



TECHNICAL UNIVERSITY OF CRETE

PETROLEUM ENGINEERING MSc COURSE

MINERAL RESOURCES ENGINEERING DEPARTMENT

Division of Mining Technology – Hydrocarbons Chemistry and Technology

«Statistical Quality Assurance and Control to Evaluate Analytical Measurement System Performance»

A Thesis submitted in partial fulfillment of the requirements for the degree of
Master of Science

BY

Aggelos Tsachouridis

Examination Committee

Nikolaos Pasadakis, Professor (Supervisor)

Dionisis Christopoulos, Professor

Nikolaos Kallithrakas, Professor

Chania, 2017

INTRODUCTION

Quality is an important factor when it comes to any product or service. With the high market competition, quality has become the main differentiator for almost all products and services. Quality control and assurance are essential in building a successful business that delivers products that meet or exceed the expectations of the customers. They also form the basis of an efficient business that minimizes waste and operates at high level of productivity. A quality control system based on a recognized standard, such as ISO 9001 published by the International Organization for Standardization, provides strong foundation for achieving a wide range of operational benefits. Therefore all manufacturers and service providers out there constantly look for enhancing their product or service quality. In order to maintain or enhance the quality of the offerings, manufacturers use quality control and quality assurance processes. These two practices make sure that the end product, or the service, meets the quality requirements and standards defined for the product or the service. There are many methods implemented by organizations to achieve and maintain a required level of quality. Some organizations believe in the concept of Total Quality Management (TQM) while others believe in internal and external standards. The standards usually define the processes and procedures for organizational activities and assist to maintain the quality in every aspect of organizational functioning. When it comes to standards for quality, there are many. ISO (International Standards Organization) is one of the prominent bodies for defining quality standards for different industries and laboratories. Therefore, many organizations try to adhere to the quality requirements of ISO. In addition to that, there are many other standards that are specific to various industries. Concluding, every organization that practices Quality Control needs a Quality Manual. The main difference between quality control and quality assurance is that quality control makes sure that the end product meets the quality requirements while quality assurance makes sure that the manufacturing process of the product does adhere to certain standards. Therefore, quality assurance can be identified as a proactive process while quality control can be noted as a reactive process.

OBJECTIVE OF THE PROJECT

The quality of data must be known and established before it can be used in any application. Data quality may be judged on the basis of its quantitative accuracy and on the confidence that can be placed in the qualitative identification of the parameters measured. In order to be able to use the results generated from a measurement process with confidence, the fitness – for – use of these results must be evaluated. This can be achieved by regular application of the measurement process to quality control samples, as well as engaging the measurement process in regular interlaboratory testing programs (round robin studies). This requires the production of data in a quality control program that permits the assignment of its statistically supported limits of uncertainty. The essential features of such a program, which consists of statistical quality control techniques, are discussed in this project that focuses on how an overall statistical quality control procedure should be conducted, fully or partially, by any engineer and in almost every laboratory environment, based on certified standards. The program that was written in Matlab was used for the fulltime evaluation of a Rock – Eval pyrolysis apparatus and a GC – MS apparatus. In order to describe as best as possible the general statistical quality control design, a description of a Ruggedness test is also included and finally, the statistical procedure with which firstly outliers can be detected in the results of an Interlaboratory – round robin test, and after their discarding, the final precision statement (in terms of repeatability and reproducibility) can be formulated, is also described. The precision statistics of a round robin study are indicatively implemented by a small Matlab code on a series of interlaboratory data results related to the different properties of an oil lubricant sample. Concluding, the aim of this thesis is the combination of simple statistical techniques into a program that enables every engineer, without having to delve deep into the world of mathematics, to statistically evaluate the data that he has at his disposal, based on certified standard techniques.

ACKNOWLEDGEMENTS

Prima facie, i am grateful to the God for the good health and wellbeing that were necessary to complete this thesis. I am also using this opportunity to express my gratitude to my family, my friends and my colleagues who supported me during the whole duration of this MSc program. I am sincerely grateful to them for their aspiring guidance, constructive criticism and friendly advice.

Heartfelt thanks and appreciation to all my professors, for their invaluable guidance and the tolerance they exhibited towards to my person all these years, with a special mention to Professors Nikos Pasadakis, Dionisis Christopoulos and Nikos Kallithrakas who also constituted my examination committee.

Last but not least, i want to thank all the faculty members of the School of Mineral Resources Engineering for their help and support and everyone who, directly or indirectly, contributed in this venture.

Table of Contents

INTRODUCTION.....	2
OBJECTIVE OF THE PROJECT	3
ACKNOWLEDGEMENTS.....	4
QUALITY CONTROL AND QUALITY ASSURANCE – GENERAL INFORMATION, BASIC DEFINITIONS AND QUALITY ASSURANCE SYSTEMS.....	8
STANDARIZATION, CERTIFICATION AND ACCREDITATION.....	9
TRADITIONAL AND MODERN PRODUCTION AND QUALITY CONTROL MODELS.....	10
QUALITY ASSURANCE SYSTEMS.....	12
ADVANTAGES AND DISADVANTAGES OF THE ISO 9000 SERIES.....	15
THE ROCK – EVAL APPARATUS	17
IDENTIFICATION OF THE KEROGEN TYPE AND THERMAL MATURITY OF THE SAMPLE	20
THE GAS CHROMATOGRAPHY – MASS SPECTROSCOPY APPARATUS.....	23
GAS CHROMATOGRAPHY	23
MASS SPECTROSCOPY	24
INJECTION TECHNIQUES.....	26
RESPONSE FACTOR AND RELATIVE RESPONSE FACTOR	27
APPLICATION OF STATISTICAL QUALITY ASSURANCE TECHNIQUES FOR THE EVALUATION OF ROCK – EVAL AND GC - MS.....	29
SCOPE	29
TERMINOLOGY USED IN QUALITY CONTROL AND QUALITY ASSURANCE PROCEDURES	29
USE OF MEASUREMENT PROCESS IN THE PETROLEUM INDUSTRY	31
MEASUREMENT SYSTEM SELF – MONITORING	32
BASIC ELEMENTS OF STATISTICS USED IN QUALITY CONTROL	32
STATISTICAL SAMPLING PROCESS	32
REFERENCE MATERIALS.....	34
THE HISTOGRAM	35
THE LEVEY-JENNINGS CONTROL CHART	38
WESTGARD RULES.....	40
PRETREATMENT OF TEST RESULTS	45
RUN CHARTS	46
CONTROL CHARTS	47
THE I CHART	49

THE MR CHART	50
EWMA OVERLAY ON I CHART	51
STUDENT'S-t test.....	52
APPROXIMATE F TEST	54
SITE PRECISION AND SITE PRECISION STANDARD DEVIATION	55
APPROXIMATE CHI - SQUARE TEST (χ^2)	55
PROCESS CAPABILITY AND PROCESS PERFORMANCE	56
PERCENTAGE OF DEFECTIVE RESULTS OUTSIDE SPECIFICATIONS.....	57
BENCHMARKING	58
RUGGEDNESS TEST	59
PLACKETT - BURMAN DESIGN	59
PLACKETT - BURMAN Design Calculations	60
INTERLABORATORY-ROUND ROBIN TEST	62
CALCULATION AND DISPLAY OF STATISTICS.....	63
THE h AND THE k GRAPHS	65
INVESTIGATION OF INCONSISTENT RESULTS.....	66
PRECISION STATEMENT	67
ROCKEVAL PERFORMANCE EVALUATION.....	68
RESULTS AND DISCUSSION FOR TOC	68
RESULTS AND DISCUSSION FOR TMAX	73
RESULTS AND DISCUSSION FOR S1.....	76
RESULTS AND DISCUSSION FOR S2.....	80
RESULTS AND DISCUSSION FOR S3.....	83
BENCHMARK ARV SCENARIO.....	88
MEAN SCENARIO	89
UPPER L-J POINTS SCENARIO	90
LOWER L-J POINTS SCENARIO	90
GC - MS PERFORMANCE EVALUATION	92
INTERLABORATORY STATISTICS	100
GRUBB'S TEST FOR OUTLIERS.....	100
RESULTS AND DISCUSSION ABOUT MAGNESIUM m(%) - D4951 Method	101
RESULTS AND DISCUSSION ABOUT ZINC m(%) - D4951 Method	102
RESULTS AND DISCUSSION ABOUT PHOSPHORUS m(%) - D4951 Method	104
RESULTS AND DISCUSSION ABOUT CALCIUM m(%) - D4951 Method	105

RESULTS AND DISCUSSION ABOUT SULFUR m(%) – D129 Method	106
RESULTS AND DISCUSSION ABOUT DENSITY (Kg/L) – D1298 Method	107
RESULTS AND DISCUSSION ABOUT FLASH POINT – D93 Method	108
CONCLUSIONS-SUGGESTIONS	110
REFERENCES	111
APPENDIX	113
GENERAL MATLAB CODE	113
INTERLABORATORY STATISTICS CODE	122
GC – MS COMPONENTS	129

CHAPTER 1 – QUALITY CONTROL AND QUALITY ASSURANCE PRINCIPLES.

QUALITY CONTROL AND QUALITY ASSURANCE – GENERAL INFORMATION, BASIC DEFINITIONS AND QUALITY ASSURANCE SYSTEMS

Research around ancient and modern technology has proven that in all organized societies of every age, the high cultural level is always associated with an advanced technological level. This technology is based on a mechanism that ensures quality, quality control and the protection of the rights of the consumer. In ancient Babylon, capital of Mesopotamia, one of the laws that were enacted by King Hammurabi was related to the construction of monuments and is considered as the first law that regulated the construction of such sites. In ancient Greece, as many retained inscriptions reveal, a mechanism of quality control and quality assurance of all the produced materials was operating under very strict prescriptions and heavy penalties were imposed upon the transgressors. Quality is generally a subjective term and depends mainly on the eye of the beholder. Some of the most frequently used definitions about quality are the following:

- Suitability for use (Juran, 1974).
- Compliance with requirements or specifications (Crosby, 1979).
- The sum – totality – of characteristics of an entity (product or service) that determines its ability to satisfy, definitely or relatively, the needs of the user (ELOT EN ISO 8402: 1996).

Quality can be distinguished in Quality of Design and in Manufactured Quality or Quality of Conformance. The levels of quality of design are defined during the planning phase of the product and concern the type of the materials used, their sturdiness and reliability etc. Manufactured quality is the subject of quality control and quality assurance.

Analytical Quality Control includes a set of activities which measure the quality characteristics of a laboratory product, compare it with certain specifications or requirements and take the necessary actions to correct a possible discrepancy between the produced material and the one that should have been produced, if the specifications concerning the product had accurately been implemented. Quality assurance (QA) refers to the full range of practices employed to ensure that laboratory results are reliable. Quality assurance may be defined as the system of

documenting and cross referencing the management procedures of the laboratory. Its objective is to have clear and concise records of all procedures which may have a bearing on the quality of data, so that those procedures may be monitored with a view to ensuring that quality is maintained. Nowadays, quality is closely related to three terms: Standardization, Certification and Accreditation.

STANDARIZATION, CERTIFICATION AND ACCREDITATION

Standardization is a process established by consensus and approved by an international regulatory body which provides rules for common and repeated use, guidelines or attributes for certain production activities and their results, and aims at achieving the best order in a given implementation framework (ELOT EN 45020:1996). From the very first moment that a product is created, a standardization process begins and when this product is distributed to the market it is accompanied by certain specifications, the manufacturer's specifications. Usually, standardized products, or more simply standards, concern the safety of the users of the products and do not interfere with manufacturing or designing processes so as not to affect the creativity, inventiveness and technological training of the manufacturer. What is simpler is the process by which standards are produced, setting the necessary rules for the production, composition and properties that a product should have. The standards that are established are distinguished into three categories: classical, state and national. The standardization regulatory body for Greece is the Hellenic Standardization Organization, ELOT, which maintains a big library with constantly updated files of all the Greek, European and international standards.

Certification is the process by which the conformity or compliance of a product with specified requirements is certified (ELOT EN45020:1996). The certification of compliance of a product with a standard is optional or compulsory depending on the legal framework of each country. Generally, there are two kinds of certifications. The certification of a product so as to evaluate its compliance with a standard, and the certification of a product based on its quality. In the first category, the product undergoes a series of certified tests in a laboratory and according to the test results its compliance with a standard is evaluated. If the level of "agreement" is considered satisfying then a certificate of compliance for the product is granted. In the certification of a product based on the quality, a manufacturer who wishes for his product to be certified and be characterized by repeatable quality, resorts to the European certification ELOT EN ISO 9000 according to which the manufacturer develops a system of activities so as to confirm that his product and the process he implements in order to produce it, have certain quality. Every country has established certification committees. In Greece, the official and certified certification body is ELOT.

Accreditation is the process with which a competent regulatory body provides official recognition that enables another entity or person to carry out specific projects (ELOT EN45020: 1996). The accreditation regulatory body in Greece is the National Accreditation Council of the ESYD which was established in 1994 and operates in the Ministry of Development. ESYD has established criteria and guidelines for evaluating the compliance with the accreditation criteria. All certification bodies must be accredited and acquire an Accreditation Certificate.

Even though quality and quality control as terms existed for centuries, only recently they emerged as primary elements of the success of a company and were adopted by the science of Management. The evolution of the production and quality control models that took place and evolved during the past 50 years is described below.

TRADITIONAL AND MODERN PRODUCTION AND QUALITY CONTROL MODELS

After the Second World War, an abrupt economic development resulted in high demand for goods and services. With every passing year, the different social classes increased their demands and, subsequently, the quality of the produced goods and services increased so as to sufficiently satisfy the increased needs of the consumers. Every social group mainly focused on the raw production, so as to satisfy these needs, and neglected the high quality that these products and services should display. For the evaluation of the quality of the products and services, a final and massive quality control was conducted and all those products that were regarded as acceptable were channeled into the market while the rest that were regarded as non-acceptable were rejected, and were either reprocessed (so as to achieve certain quality standards) or completely discarded. This production model was broadly accepted and disseminated and was labeled as the Traditional Production Model. Nowadays this production model starts being considered obsolete but there are still many companies throughout the world that still implement it.

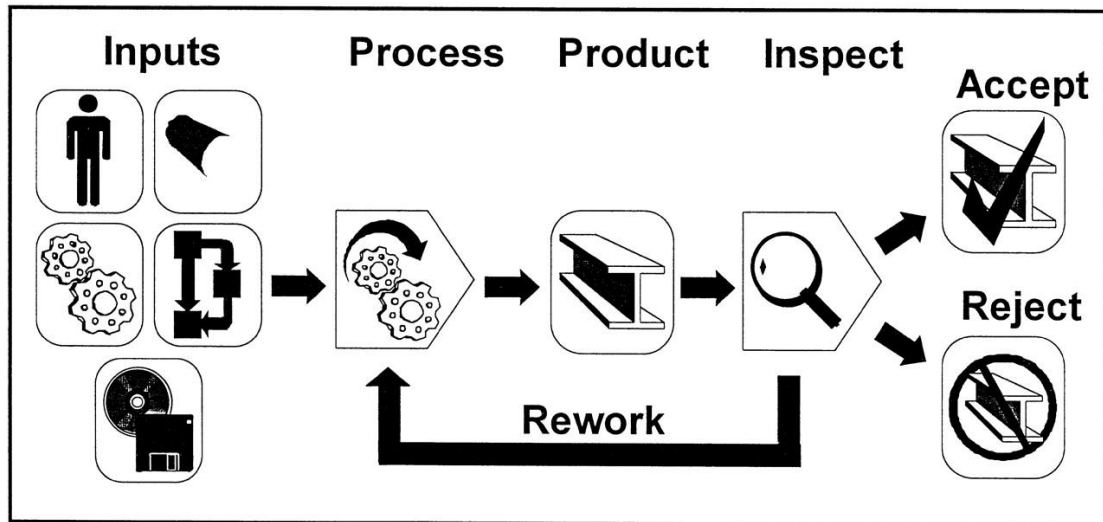


Image 1: Traditional production model.

Despite its broad acceptance and dissemination, the traditional production model has some disadvantages. These disadvantages are presented below.

- It does not take under consideration the human factor, the quality, the security and the compliance with the environmental rules that the production process of every material or service should follow.
- It creates a relatively big quantity of useless products – services.
- The final and massive quality control that this model implements is not always feasible and effective. Again human factors restrict the effectiveness of this process (knowledge, experience, work pressure, reduced interest etc.) Furthermore, the final quality control detects the defective products but does not solve the production problem that lead to the defection. In essence, this massive inspection can detect a quality problem but does not offer any reason for this problem or possible solutions.

All these disadvantages of the traditional production model lead to an increased production cost which, in turn, lead Europe and North America to an economic dead-end by the end of the 1950's. At the same time, Japan implemented new production methods that were primarily based on quality and resulted to rapid economic improvement. This mindset gradually affected the way that many companies used to function and produce, and lead them to focus on the following objectives prior and during the production process:

- Satisfaction of the needs of the consumers.
- Improvement of quality.
- Reduce of the production and processing cost.

- Improvement of the communication of the employees.
- Investment in technology.
- Development and constant training of the working personnel.

These changes that were gradually developed over the previous 50 years were importantly boosted by the theories of a group of business executives who were stern supporters of quality, and finally resulted in the creation of a new production model that was labeled as the Modern Production Model. The evolution of quality over this time period encompasses 4 phases – periods: 1) Control and Inspection, 2) Statistical Quality Control, 3) Quality Assurance, 4) Overall Quality Management. As far as the first phase is concerned, G.S Radford suggested that control and inspection of quality is a responsibility of those who exercise administration. The creation of more than one production phases lead to the need of a more detailed control of production and of the quality of the produced materials. 1931 has been a landmark for the second phase, the statistical quality control, when W.A Shewhart published his famous book “Economic Control of Quality of Manufactured Product” in which he provided an exact and “measurable” definition of quality control in production, he developed new techniques for daily supervision of the production and suggested new methods for the improvement of quality. During the quality assurance phase quality, which was a more restricted and focused in production term, expanded and affected the administration system. Essential role in the quality assurance phase plays the term of the cost of quality and the sustenance of quality in all the spectrum of activities (design, production, quality control, market distribution). Given that in every production process the defective products are always costly, the main issue that had to be addressed at that point of time was what should have been the level of quality according which a product should be considered as acceptable or not. Another landmark for the evolution of the quality control and quality assurance has been the book of J. Juran “Quality Control Handbook” in which he addressed all the foretold issues. Finally, in 1956 A. Feigenbaum introduced the term of overall quality management. He remarked that all the processes related to production would only be effective, in terms of quality, if the different departments of a company are in direct collaboration. In different case, any error that would possibly occur during the initial stages of production, would have been impossible to be identified before the final product is channeled to the market and subsequently reaches the consumers.

QUALITY ASSURANCE SYSTEMS

Nowadays more than ever, quality has evolved into a basic precondition for the survival of the products within the demanding international market limits. Within

this new globalized environment, there is a dire need for introduction of quality assurance systems, for the configuration of a new business environment which will be based on the new mentality that the consumer has developed and the demand for new products. So as to avoid market related problems and to ensure the protection of the rights of the consumer, it is of vital importance to develop mutually accepted criteria with which the quality of a product or a service will be certified. The totality of all these criteria consist what we call today as Quality Assurance Systems. Quality systems belong to a category of management systems of a process. As management system of a company is defined the summation of standard specifications under which a company operates. In these specifications, the organizational structure of a company, the business activities, the required equipment for the effective operation of the company and the required personnel for implementation of the work, are included. Depending on the objectives of a system, they can be distinguished into different categories such as quality systems, systems of environmental protection, internal operating regulations, internal control systems etc. Quality systems aim at the adoption of operating specifications so as to improve the quality of provided products and services. The most applicable and generally widespread guide for the development of such quality systems is the international standard ISO 9000.

So as to ensure the quality of production and to establish an international production language, special series of standards were developed from the International Standards Organization so as to evaluate the operation of all companies. In 1987 the introduction and development of quality control and quality assurance occurred in an international level with the development of the ISO 9000 series which described the structure that a quality control system should exhibit. The first edition of 1987 was based on three standards known as ISO 9001 – 87, ISO 9002 – 87 and ISO 9003 – 87. In 1994, the first update of the original series took place. In this update, five standards were included – the three previous, original, standards of the 1987 version, and two additional guide standards, ISO 9000 and ISO 9004. The next update occurred in 2000 and induced 2 important changes. First of all, the three initial standards of the first version are unified into one, the ISO 9001, which was officially adopted by the European Union and subsequently by ELOT under the code name ELOT EN ISO 9001:2000. The second change was the demand for constant improvement of its performance. With the implementation of this new version of the standard, a new contemporary and effective model of quality control was introduced that set the consumer at the center of every business activity, emphasized on the human potential of a company and determined that the process is the basic fundamental element of every company. The last update of this standard occurred in 2008. This new revised version, known as EN ISO 9000:2008, is the

current active standard which is generally adopted by all companies throughout the world and is composed from the following standards:

- EN ISO 9000:2008 (Fundamentals & Vocabulary) – Describes the basic principles of the quality control systems and defines the terminology that they implement.
- EN ISO 9001:2008 (Quality Management Systems Requirements) – Defines the requirements that a quality control system should display for use, where the ability of an organization to provide products and services that satisfy the demands of the market is evaluated.
- EN ISO 9004:2008 (Quality Management Systems, Guidelines for performance improvements) – Provides guidelines regarding the improvement and efficiency of the quality control system. The objective of the standard is the improvement of the efficiency and the satisfaction of the needs of the customers.

The basic philosophy of the ELOT EN ISO 9000:2008 standard is essentially based on the constant implementation of an improvement cycle. The figure below depicts that improvement cycle.

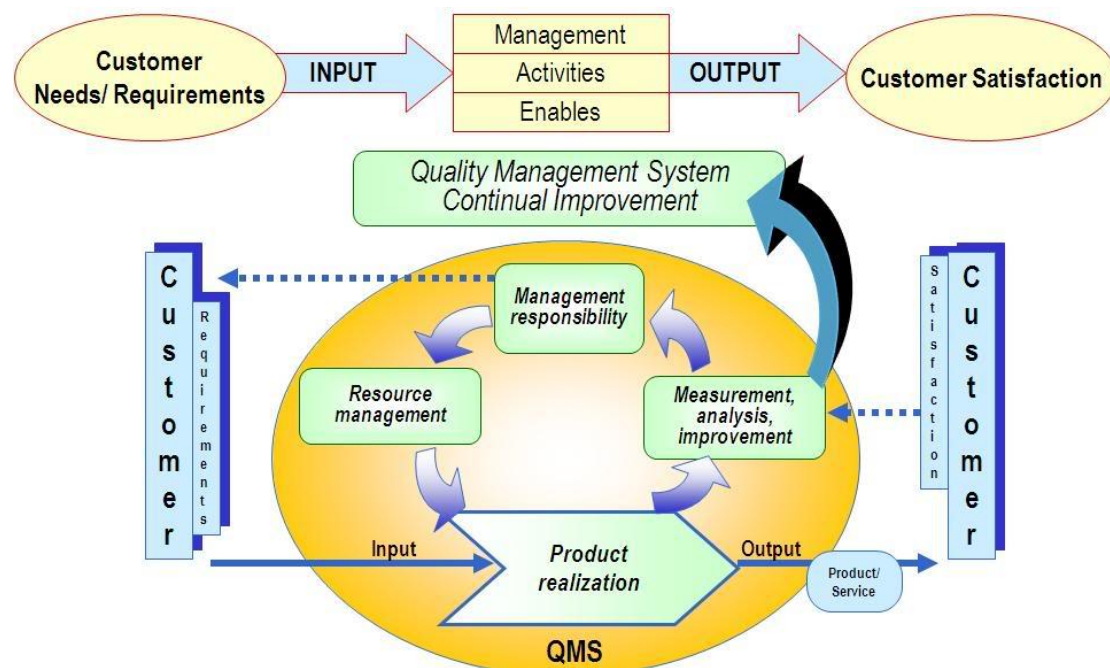


Image 2: Quality management system according to ELOT EN ISO 9000:2008.

ELOT EN ISO 9001:2008 standard (an improvement of the previous standard) defines the requirements of a quality control system, when a company or an organization is called to prove its ability to provide products and services that satisfy the demands

of the consumers, according to the implemented quality framework. More specifically, an organization is obliged to follow a series of requirements that include:

- Compliance with the relative to the product legislation.
- Determination and implementation of quality control charts.
- Active participation of the business management.
- Appointing of a person or a team that will be in charge of supervising the quality control system.
- Constant monitoring and recording of the operating elements.
- Constant monitoring and recording of the satisfaction of the customers.

The ISO 9000 series has certain advantages, disadvantages and difficulties of implementation that are described below.

ADVANTAGES AND DISADVANTAGES OF THE ISO 9000 SERIES

A relatively new standard series like ISO 9000 displays both advantages and disadvantages, with the disadvantages being the main subject of consideration for improvement in the upcoming, future editions. Advantages can be distinguished into internal and external advantages.

Internal Advantages:

- It is a new method of administration that promotes the competitiveness.
- The sensitivity and awareness of the personnel in matters of quality is increased
- The systematic approach in work functions and the sufficiently defined and revised connection of the different functions has as a result the increased productivity of the working personnel.
- Decrease in instability and improvement of the quality of the products and services.
- Obvious reduction of the defective and reprocessed products.
- Implementation of the ISO 9000 series has the most stable background for every business and organization for the promotion of an overall quality control system.

External Advantages

- Provides satisfaction of demands or market pressure.
- Constant quality creates satisfied customers and attracts new ones.

- Offers strong competitive advantage.
- Improves the reliability, fame and trust in the products of a company.
- Minimizes the discarding of defective products from the customers.
- Supports and facilitates the exporting to new markets and increases the overall market capital.

Disadvantages – Difficulties in implementation

- Extensive capital investment and continuous training of the working personnel is required.
- The implementation of the series creates excessive bureaucracy.
- It requires important investments and the consumption of many men – hours during the preparation of certification.
- It does not necessarily guarantee that the quality of the products or services will be improved.
- Some of the abrupt changes that a revised version may involve could possibly create discomfort to the employees.

Concluding, the installation and operation of a quality control system is vital for every company and organization nowadays and it requires that both the company executives and the working personnel fully comprehend the importance of its development and preservation. Quality control in general is a vital process that, according to ISO 9000 and the revised ISO 9001 series, emphasizes on three main aspects. Elements such as control, job management, well defined processes, performance and integrity criteria. Competence, such as knowledge, skills, experience and qualifications and finally, soft elements such as personnel, confidence, organizational structure, motivation and team spirit.

Inspection is another vital component of quality control, where the products are examined visually and a list with a description of unacceptable and defective products is created by the product inspector. There is a tendency for individual consultants and organizations to name their own unique approaches to quality control—a few of these have ended up in widespread use and universal acceptance. Namely, some of them are the statistical quality control (SQC), the total quality control (TQC), the statistical process quality control (SPC), the company wide quality control (CWQC), the Six Sigma (6σ) and the lean Six Sigma ($L6\sigma$) quality control. In this specific project, a combination of the SQC and SPC principles was implemented so as to evaluate statistically the operation of an analytical measurement system.

CHAPTER 2: EVALUATION OF THE ROCK – EVAL AND GC – MS APPARATUSES WITH THE INVOLVEMENT OF QUALITY ASSURANCE TECHNIQUES

THE ROCK – EVAL APPARATUS

The Rock – Eval pyrolysis method has been extensively employed by the industry for hydrocarbon exploration for decades. By pyrolysis, we refer to the decomposition of organic matter by extensive heating in the absence of oxygen. Pyrolysis is implemented so as to evaluate the thermal maturity of an organic sample and to determine the petroleum potential of the sedimentary basin from which the sample was obtained. The term thermal maturity is used in the characterization of the chemical reactions that convert organic matter into petroleum and depend on heat processes. Two Rock – Eval systems are most frequently used in the industry nowadays, Rock – Eval II and Rock – Eval VI, with the later gradually replacing the former as the predominant apparatus. In Rock – Eval, II the sample is pulverized and 70 – 80 mg are weighted into stainless steel crucibles which have screened top and bottom so as to allow the movement of carrier gas through the sample. The carrier gas is Helium. The sample is initially heated under an inert atmosphere of Helium at 300 °C for 4 minutes, and then progressively heated to 600 °C with a rate of 25 °C per minute. In the first thermal stage (-300 °C), the soluble compounds – Bitumen – already present in the sample are extracted, while during the second thermal stage (300 °C – 600 °C), the insoluble organic compounds – Kerogen – are cracked down into pyrolytic products. The two major differences In Rock – Eval VI are, firstly, that the oven of the system is able to reach a terminal Temperature of 800 °C (which means that the second thermal stage starts at 300 °C and ends at 800 °C) and secondly, that the carrier gas used is Nitrogen and not Helium. Another important deviation between the two apparatuses is the fact that that the S1 peak calculated by Rock – Eval II differs slightly from the S1 calculated by Rock – Eval VI and that is because the real sample temperature is different from 300 °C (variations up to 20 °C can be observed). This is due to the fact that in Rock – Eval VI, the probe is in direct contact with the sample, so the sample temperature measurement is much more precise compared to the one that the II apparatus provides, where the probe is inserted into the oven and does not come into direct contact with the crucible.

The released hydrocarbons are monitored by a flame ionization detector (FID) and a thermal conductivity detector (TCD). More specifically, in the first thermal stage the carrier gas sweeps the volatile products out of the oven and into a splitter. The first half of the split effluent is sent to a CO₂ trap and the second split is directed into a

hydrogen FID where the detection and quantitative measurement of the hydrocarbons takes place. After the terminal temperature is achieved (600 °C and 800 °C for both systems respectively) and the pyrolysis is over, the CO₂ trap is led into a TCD where the quantitative measurement of the evolved CO₂ takes place. Finally, the results of the procedure are collected and sent into a recorder. The parameters measured and visually depicted as peaks by the system, are the **S1**, **S2**, **S3**, **T_{MAX}** and **TOC**. These parameters are explained below.

S1 – Represents the amount of free hydrocarbons already present in the sample before the pyrolysis. When S1 is large relative to S2, possible contamination should be suspected.

S2 – Represents the amount of hydrocarbons generated through thermal cracking of nonvolatile organic matter. S2 is an indication of the quantity of hydrocarbons the sediments could potentially produce should burial and maturation continue. It also represents the amount of hydrogen in the kerogen.

S3 – Represents the amount of CO₂ produced during the pyrolysis of kerogen. S3 is an indication of the amount of oxygen contained in the kerogen.

T_{MAX} – Represents the temperature at which the maximum release of hydrocarbons from thermal cracking of kerogen occurs during pyrolysis. T_{MAX} is an indication of the maturation stage of the organic matter.

S4 – Represents the amount of carbon that comes from graphite, in other words the residual carbon content of the sample. The residual carbon content has no potential to generate hydrocarbons due to the absence of hydrogen and the overall chemical structure of the molecule.

TOC (Total Organic Carbon) – Represents the amount of residual and pyrolyzed organic carbon and can be determined by oxidizing the organic matter remaining in the sample after the pyrolysis is over.

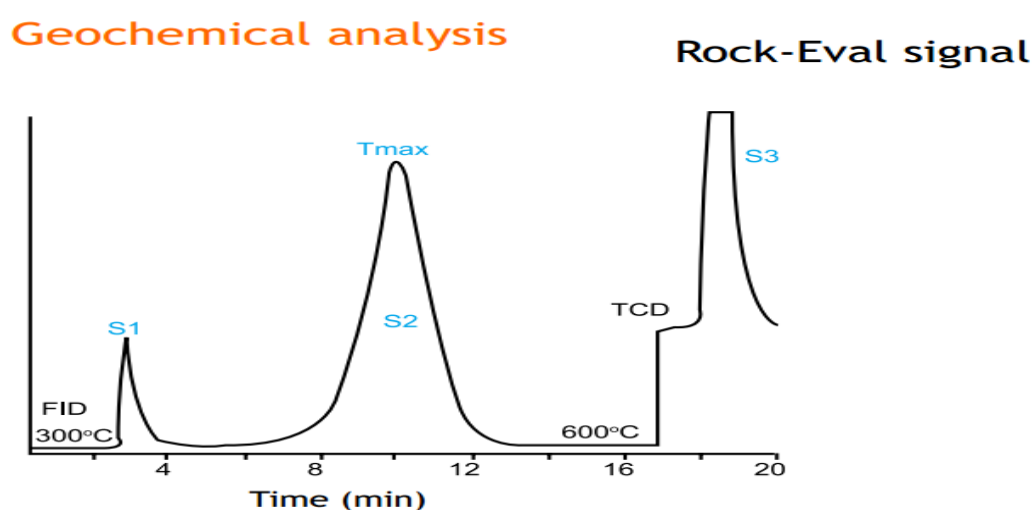


Image 3: A typical Rock – Eval signal.

There are also some quantities that are directly derived from the five foretold results that the Rock – Eval system provides. These quantities are the **Productivity Index**, the **Pyrolyzed Carbon**, the **Hydrogen Index**, the **Oxygen Index**, the **Organic type indicator** and the Vitrinite Reflectance. The overall geochemical evaluation of the sample is based on these indexes, which are explained below:

Productivity Index (PI) – $[S1 / (S1+S2)]$ characterizes the evolution level of the organic matter. In an ideal situation with the burial depth increasing, S1 should increase while S2 should decrease resulting in PI increasing with depth and maturation.

Pyrolyzed Carbon or Petroleum Potential (PC) – $[0.83 * (S1 + S2)]$ represents the carbon content, the maximum quantity of hydrocarbons capable of being produced from the source rock, given sufficient depth and time.

Hydrogen Index (HI) – $[(100 * S2) / \text{TOC}]$ represents a parameter used to characterize the origin of the organic material and indirectly correlates the hydrogen to carbon ratio. Marine organisms and algae have higher hydrogen to carbon ratios compared to land plants, because they are not so exposed to oxidization. HI typically ranges from 100 to 600 in geological samples.

Oxygen Index (OI) – $[(100 * S3) / \text{TOC}]$ represents a parameter that indirectly correlates the oxygen to carbon ratio. OI typically ranges from 0 to 150.

Organic Type Indicator – $[S2/S3]$ represents another index who essentially indicates the depositional environment (marine or terrestrial) from which the sample was obtained, but also the hydrogen richness of the kerogen.

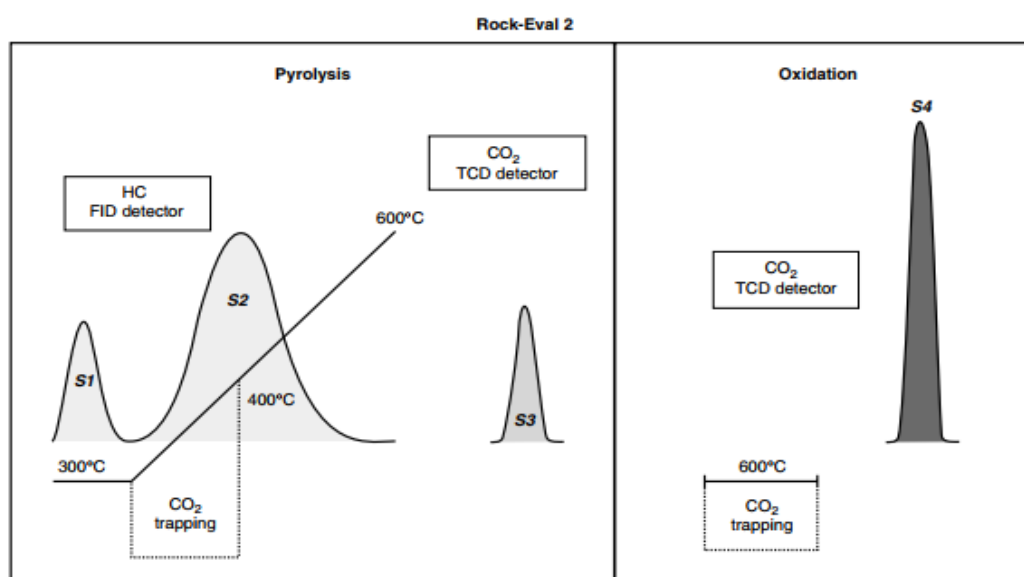


Image 4: Rock – Eval II pyrolysis conditions.

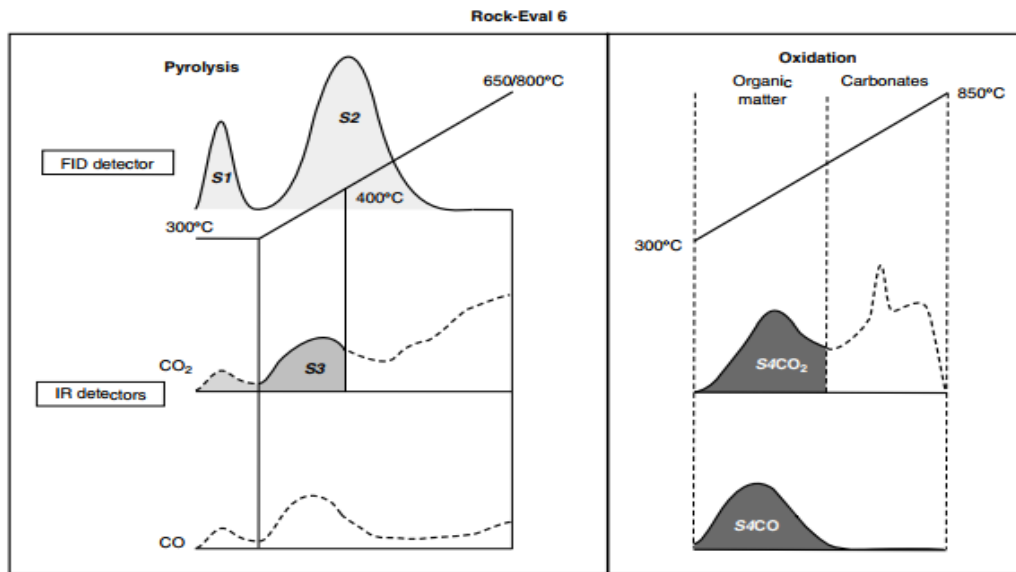


Image 5: Rock – Eval VI pyrolysis conditions.

IDENTIFICATION OF THE KEROGEN TYPE AND THERMAL MATURITY OF THE SAMPLE

After the pyrolysis is completed and the relative peaks are collected, the characterization of the kerogen type and the thermal maturity of the sample is conducted. The basic tool which is traditionally implemented for this task in the industry is the Van Krevelen diagram. Originally developed by Dirk Willen Van Krevelen, this diagram was initially used to classify coals and predict compositional evolution during thermal maturation, utilizing the oxygen to carbon and hydrogen to carbon ratios.

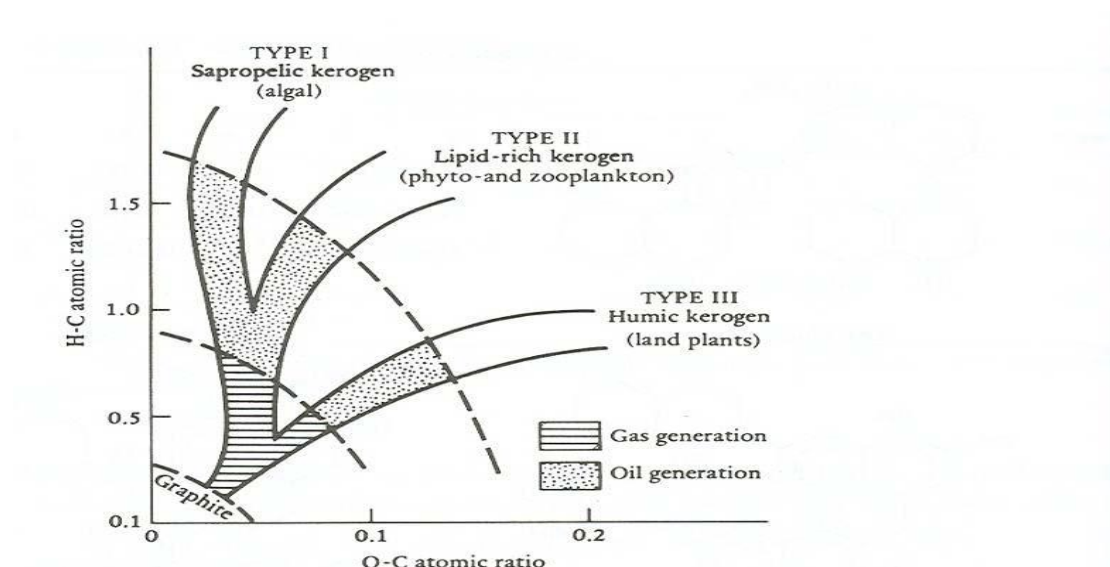


Image 6: Original Van – Krevelen Diagram.

During the 1970's, Tissot et al modified the original diagram by substituting the O/C and H/C ratios with the Oxygen Index and Hydrogen Index produced by the pyrolysis, so as to identify the kerogen of sedimentary rocks based on their maturity. Three major kerogen families were identified and classified as kerogen type I, II and III. Kerogen type I is oil prone, kerogen type II is usually oil prone and rarely gas prone, while kerogen type III is mainly gas prone. Obviously, the first two types have higher hydrogen to carbon ratios, whereas the third type has higher oxygen to carbon ratio. Thus, the first two kerogen types are connected with lacustrine and marine environments, as far as the origin of the sample is concerned, and the third type is connected with terrestrial environments.

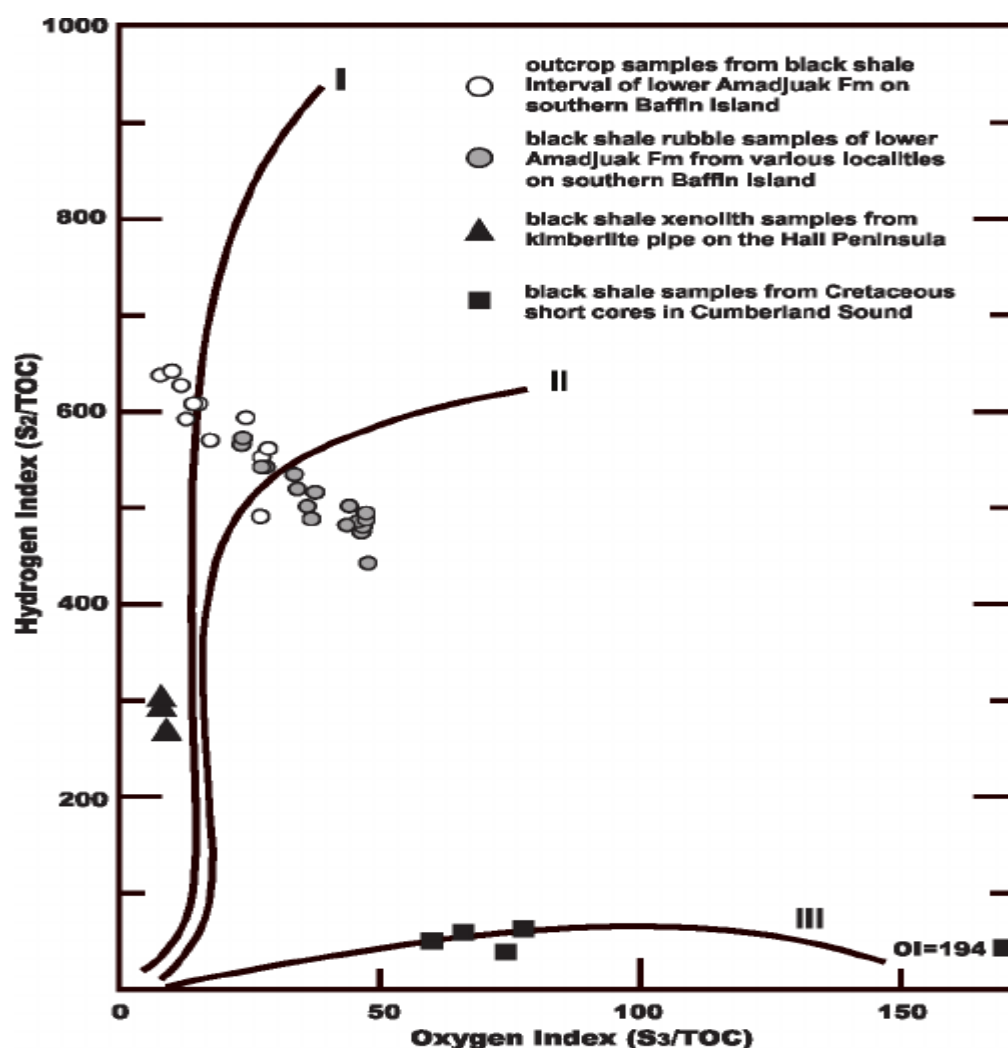


Image 7: Modified Van – Krevelen Diagram of samples from the Hall Peninsula.

Finally, the overall geochemical characterization of the samples based on their petroleum potential, the predominant kerogen type and their thermal maturity, is conducted by comparing the data results with certain limit values which are presented on the table below.

Evaluation of the Rock-Eval data

Potential (quantity)	TOC (wt.%)	Rock-Eval (mg/g rock)		Bitumen (ppm)	Hydrocarbons (ppm)
		S1	S2		
Poor	<0.5	<0.5	<2.5	<500	<300
Fair	0.5–1	0.5–1	2.5–5	500–1000	300–600
Good	1–2	1–2	5–10	1000–2000	600–1200
Very good	2–4	2–4	10–20	2000–4000	1200–2400
Excellent	>4	>4	>20	>4000	>2400

(a)

Kerogen (quality)	Hydrogen index (mg hydrocarbon/g TOC)	S2/S3	Atomic H/C	Main product at peak maturity
I	>600	>15	>1.5	Oil
II	300–600	10–15	1.2–1.5	Oil
II/III	200–300	5–10	1.0–1.2	Oil/gas
III	50–200	1–5	0.7–1.0	Gas
IV	<50	<1	<0.7	None

(b)

Maturity	Maturation			Generation		
	R _o (%)	T _{max} (°C)	TAI	Bitumen/TOC*	Bitumen (mg/g rock)	Production index (S1/(S1+S2))
Immature	0.20–0.60	<435	1.5–2.6	<0.05	<50	<0.10
Mature						
Early	0.60–0.65	435–445	2.6–2.7	0.05–0.10	50–100	0.10–0.15
Peak	0.65–0.90	445–450	2.6–2.7	0.15–0.25	150–250	0.25–0.40
Late	0.90–1.35	450–470	2.9–3.3	–	–	>0.40
Postmature	>1.35	>470	>3.3	–	–	–

(c)

Image 8: Limit values for the evaluation of Rock – Eval data results.

THE GAS CHROMATOGRAPHY – MASS SPECTROSCOPY APPARATUS

Gas chromatography – mass spectroscopy (GC – MS) is an analytical method that combines the features of gas chromatography and mass spectroscopy to separate a mixture into its compounds. The separation is based on the different distribution of the compounds between two phases. One of the phases remains stationary and the other (mobile) sweeps through or above the stationary phase and causes the separation. The combination of these two techniques creates a very effective system that allows a much finer degree of substance identification than either technique used separately.

GAS CHROMATOGRAPHY

The equipment used for gas chromatography generally consists of an injection port at one end of a column packed with substrate material and a detector at the other end of the column. An inert carrier gas, usually He or N, propels the sample down the column which is placed within a thermostable oven, while various flow meters and pressure gauges warrant a constant gas flow. To ensure proper separation, the sample must enter the column in a discreet, compact packet. Normally the sample is injected into the injection port with a hypodermic needle and syringe capable of measuring the specimen amount. The needle is stuck into a replaceable neoprene or silicone rubber septum that covers the injection port. The injection port is maintained at a temperature at which the sample vaporizes immediately. Ideally, the sample spreads evenly along the cross section of the column, forming a plug. The column is a several meters long tube, often packed with a sand-like material to promote maximum separation. Columns are commonly obtained pre-packed by vendors. As the sample moves through the column, the different molecular characteristics determine how each substance in the sample interacts with the column surface and packing. The column allows the various substances to partition themselves. Different compounds elute the column at different times. The amount of time that a compound is retained within the GC column is known as the retention time and it is directly analogous to the boiling point of the compound. This means that the heavier compounds, which have higher boiling points, are going to elute the column later than the lighter compounds. This comparison of the retention times of the different compounds is what gives the GC its analytical usefulness. Finally, the GC apparatus uses a detector to measure the different compounds as they emerge from the column. Among the available detectors are 1) the argon ionization detector, 2) the flame ionization detector, 3) the flame emission detector, 4) the thermal conductivity detector and 5) the electron capture detector. Choosing the proper detector depends upon the use. Each component ideally produces a specific spectral peak that may be recorded on a paper chart or electronically. The size of the peaks is proportional to the quantity of the corresponding substances in

the specimen analyzed. The peak is measured from the baseline to the tip of the peak. A typical GC apparatus is depicted below.

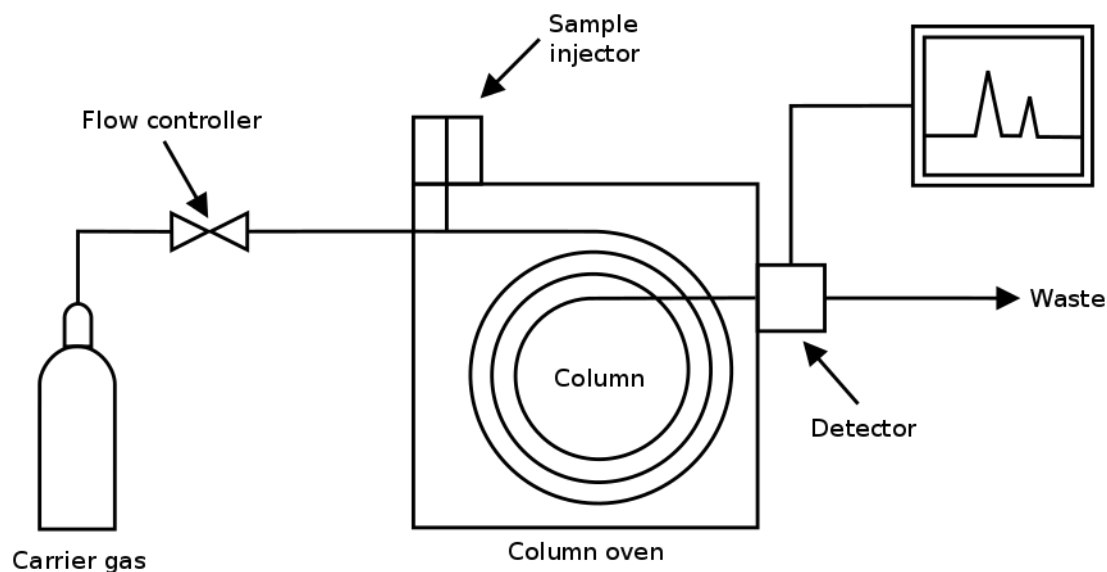


Image 9: Gas Chromatography Apparatus.

MASS SPECTROSCOPY

The most common type of mass spectrometer (MS) associated with a gas chromatograph (GC) is the quadrupole mass spectrometer, sometimes referred to as "Mass Selective Detector" (MSD). The quadrupole consists of four parallel metal rods in which the opposing rods are connected electrically with each other. Ions travel down the quadrupole between the rods. Only ions of a certain mass – to – charge ratio will reach the detector for a given ratio of voltages. This permits selection of an ion with a particular m/z or allows the operator to scan for a range of m/z -values by continuously varying the applied voltage. Mathematically this can be modeled with the help of the Mathieu differential equation. A mass spectrometer consists of three components: an ion source, a mass analyzer / filter, and a detector. The molecules of the different compounds that leave the GC enter the Mass Spectrometer and they are ionized by various methods with typically only one method being used at any given time. The ionization happens when the molecules of the sample collide with a beam of high voltage electrons and they are shattered into well – defined fragments. Each fragment is charged and travels a curved path through the quadrupole and towards the detector. When an individual charged particle collides with the detector surface, several electrons (also charged particles) emit from the detector surface. Next, these electrons accelerate towards a second surface, generating more electrons, which bombard another surface. Each electron carries a

charge. Eventually, multiple collisions with multiple surfaces generate thousands of electrons which emit from the last surface. The result is an amplification of the original charge through a cascade of electrons arriving at the collector. At this point the instrument measures the charge and records the fragment mass as the mass is proportional to the detected charge. The MS instrument produces the output by drawing an array of peaks on a chart, the "mass spectrum." Each peak represents a value for a fragment mass. A peak's height increases with the number of fragments detected with one particular mass. As in the case of the GC detectors, a peak may differ in height with the sensitivity of the detector used. A typical assembly of a GC-MS system is presented below.

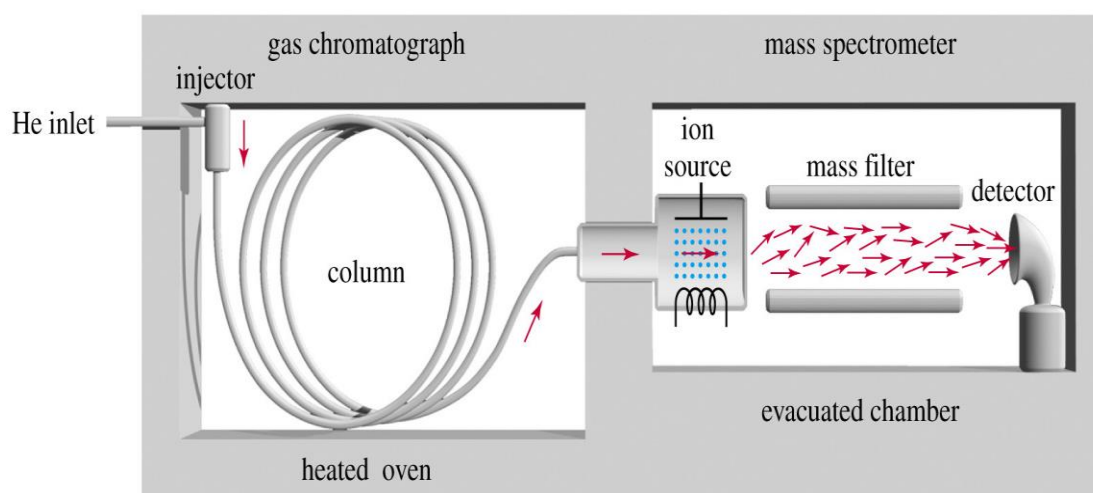


Image 10: GC – MS Apparatus.

A typical mass spectrum that the GC – MS system delivers at the end of the analysis is depicted below.

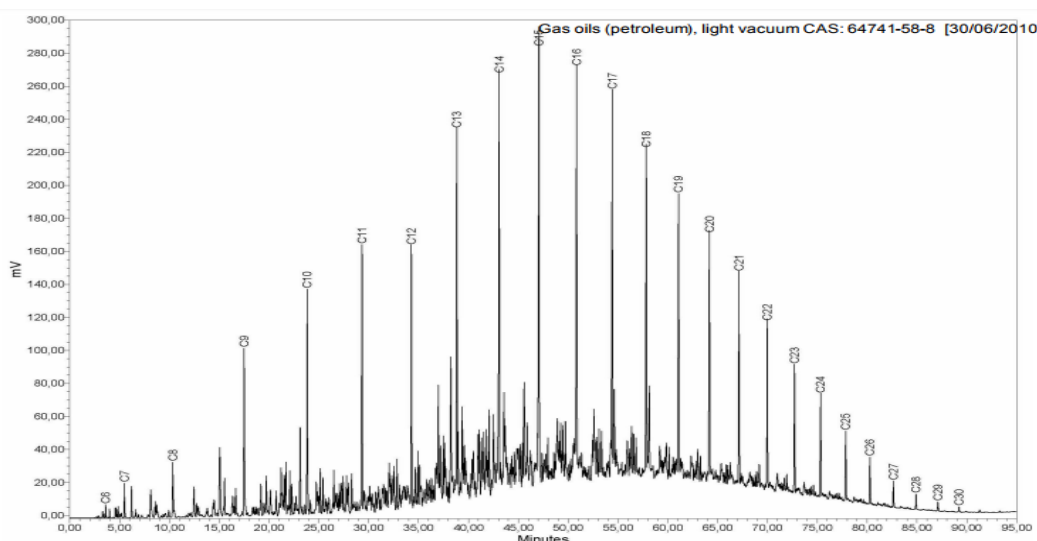


Image 11: Mass Spectrum.

INJECTION TECHNIQUES

There are three major capillary injection techniques – split, splitless and on – column. Nearly every standard capillary injector is capable of split and splitless injection, while on column injection requires a more dedicated capillary injector. Split injection is the technique that is most frequently employed since it provides the highest resolution and system efficiency. Split injections are used for highly concentrated samples with typical per component concentrations of $0.1 - 10 \mu\text{g}/\mu\text{L}$. The volume of the injected sample is usually $1 \mu\text{L}$ as was already mentioned above. The sample is instantly vaporized upon injection and rapidly mixed with the carrier gas. A small amount of the carrier gas enters the column and a much larger amount leaves the injector via the split vent. Since the carrier gas introduces the vaporized sample into the column, only a small amount of the sample also enters the column. The total gas flow at the moment of the injection is quite high (the sum of the column and the split vent flows). The sample is rapidly swept into the column which accounts for the high efficiency of the split injections. By measuring the column flow and the split vent flow, the amount of sample “splitting” can be calculated as the ratio of the two flows. Thus, a split ratio of 1:50 indicates that one part of the sample enters the column while 50 parts of it are discarded out of the split vent.

The split / splitless injector

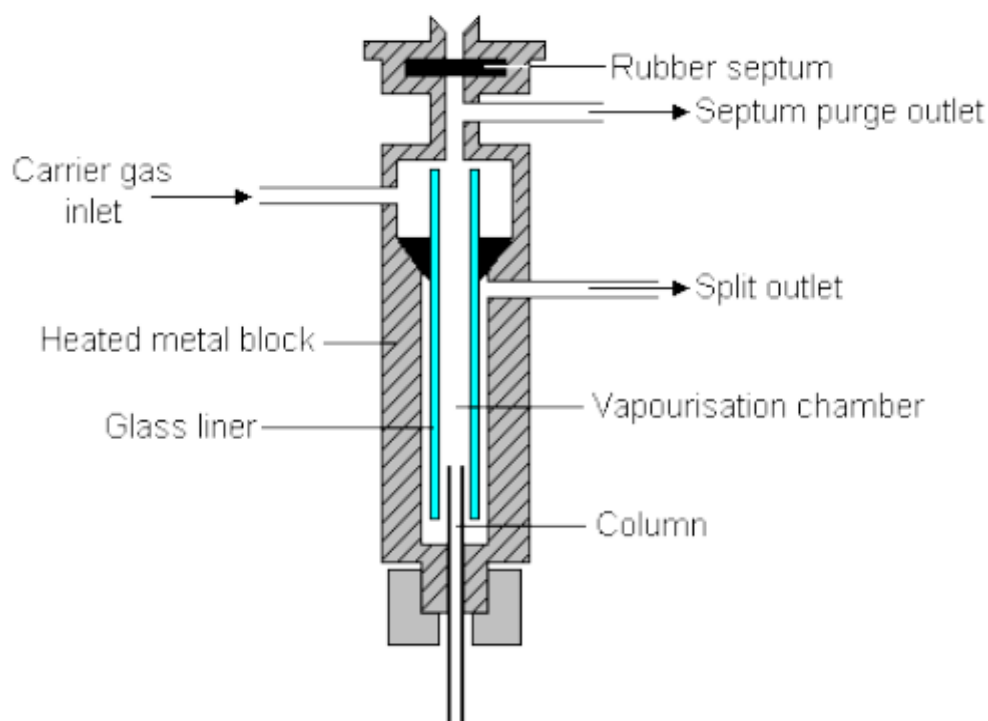


Image 12: Split – Splitless injector.

RESPONSE FACTOR AND RELATIVE RESPONSE FACTOR

In order to quantitate our analyte, a relationship between peak area and concentration must be defined. In chromatography, a response factor is defined as the ratio between the concentration of a compound being analyzed and the response of the detector to that compound. One of the main reasons to use response factors is to compensate for the irreproducibility of manual injections into a gas chromatograph. Injection volumes for GCs can be 1 microliter (μL) or less and are difficult to reproduce. Differences in the volume of injected analyte leads to differences in the areas of the peaks in the chromatogram and any quantitative results are suspect. To compensate for this error, a known amount of an internal standard (a second compound that does not interfere with the analysis of the primary analyte) is added to all solutions (standards and unknowns). This way, if the injection volumes (and hence the peak areas) differ slightly, the ratio of the areas of the analyte and the internal standard will remain constant from one run to the next.

This comparison of runs also applies to solutions with different concentrations of the analyte. The area of the internal standard becomes the value to which all other areas are referenced.

A chromatogram will show a response from a detector as a peak. While there are several ways to quantify the peak, one of the most common is peak area, thus:

$$\text{Response Factor} = \frac{\text{Peak Area}}{\text{Concentration of the compound}}$$

Quantitative analysis is the determination of a compound's concentration in a sample (analyte). Response factors are important when using GC for quantitative analysis. Each analyte has a unique Response factor under given conditions. Therefore, for repeatability in sample measurements, a method of eliminating variability in the response factor must be used. One of the easiest ways to eliminate variation in the response factor is to use relative response factors and an internal standard to calibrate the GC. The response factors calculated for each analyte are then used to establish the Relative Response Factor between the two analytes. Using RRF's is beneficial when one of the samples is an internal standard - a calibration then allows multiple analyses to be run on different samples. The formula that calculates the relative response factor for an unknown compound and an internal standard is given below:

$$\text{Relative Response Factor} = \frac{A_x * [S]}{[x] * A_s}$$

Where A_x and $[x]$ is the peak area and the concentration of the unknown compound and A_s and $[S]$ is the peak area and concentration of the internal standard used.

The internal standards are compounds with similar physical and chemical properties with the sample that was analyzed. The internal standards that were used in this particular GC – MS analysis were the NC12 – d26 for the calibration of components C10 to C15 and Nc16 – D34 for components C16 to C35. With this way a calibration curve is being created. A calibration curve is simply a graph where concentration is plotted along the x-axis and area (Response) is plotted along the y-axis. Several internal standards are run at different concentrations and the points where they intersect in the chart are connected with a straight line which represents the calibration curve. For every different analyte, a different calibration curve is being constructed.

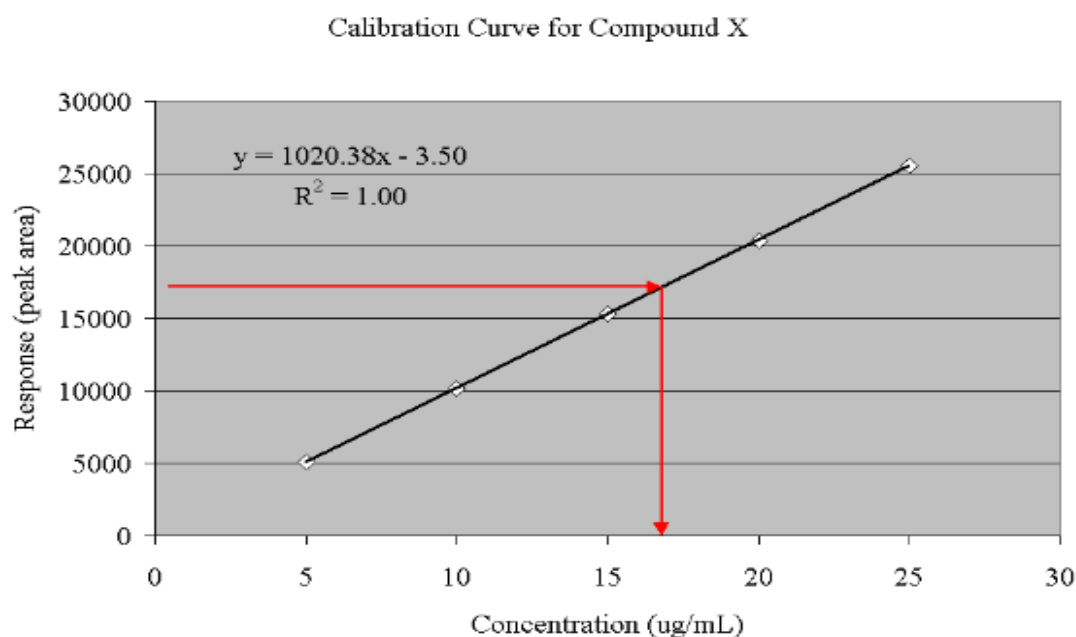


Image 13: Typical Calibration curve.

Apart from the Rock – Eval results, the thermal maturity of the petroleum can be evaluated from certain substances, the Biomarkers, which can directly be calculated from the GC – MS results. Biomarkers are complex organic compounds composed of carbon, hydrogen and other secondary elements such as oxygen, nitrogen or sulfur which can be found in petroleum source rocks. Simply put, biomarkers are complex carbon based compounds derived from formerly living organisms. Their interpretation leads to conclusions about the thermal maturity of the sample and the depositional environment from which the sample originated. Some of the most frequently used biomarkers are the Pristane / Phytane ratio, the Carbon Preference Index and the Odd to Even Predominance.

The Carbon Preference Index is defined as:
$$CPI = \frac{2 \cdot (C_{23} + C_{25} + C_{27} + C_{29})}{C_{22} + 2 \cdot (C_{24} + C_{26} + C_{28}) + C_{30}}$$

The Odd to Even Predominance is defined as:

$$OEP(1) = \frac{C21+6*C23+C25}{4*C22+4*C24} \text{ and } OEP(2) = \frac{C25+C27+C29}{4*C26+4*C28}$$

APPLICATION OF STATISTICAL QUALITY ASSURANCE TECHNIQUES FOR THE EVALUATION OF ROCK – EVAL AND GC - MS

SCOPE

The scope of the following practice is to provide information for the design and operation of a program to monitor the ongoing stability and precision and bias of an analytical measurement system using a collection of generally accepted statistical quality control procedures and tools. This program should be generally applicable to a big variety of laboratory test methods, to validate process stream analyzers and to monitor the difference between two similar analytical measurement systems which are measuring the same property, using the same method.

TERMINOLOGY USED IN QUALITY CONTROL AND QUALITY ASSURANCE PROCEDURES

Accepted reference value – ARV: A specific value that serves as a benchmark for comparison and it is derived based on scientific principles. Alternatively, it can be an assigned or consensus value derived from the experimental work of a scientific institute or international organization.

Accuracy: The proximity of agreement between a test result and an accepted reference value.

Bias: A systematic error that contributes to the difference between a population mean of the measurement or test results and an accepted reference value.

Assignable Cause: A factor that contributes to deviation and that is feasible to detect and identify.

Precision: The proximity of agreement between test results obtained under prescribed procedures, from the same homogeneous sample.

Repeatability conditions: Conditions under which, in the same laboratory, with the same operator who is using specific equipment, independent test results are obtained by following the same practice method, using random specimens from the same sample of material.

Reproducibility conditions: Conditions under which, in different laboratories, test results are obtained with the same practice method, using random specimens from the same sample of material.

Analytical measurement system: An assortment of laboratory equipment such as samplers, instruments, display devices, data handlers and transmitters that is used to produce quantitative results of a specific property for an unknown sample.

Check Standard: A material with a fixed reference value which is used in the determination of an analytical measurement system's accuracy.

QC sample: Stable and homogeneous material having physical or chemical properties similar to those of the typical samples tested by the measurement system. They are used in quality assurance programs to monitor the precision and stability of the system.

Random cause: A factor, usually of minor importance, that in quality control and quality assurance procedures contributes to variation and that is difficult to identify.

In – Statistical – Control: An experimental procedure or analytical measurement system that displays variations which can only be associated with random causes.

Quality Control samples: Homogeneous materials that are used in quality control and quality assurance programs to evaluate the stability and precision of the analytical measurement system.

Expected Value: A theoretical value that the average of results should approximate if the number of results was infinite.

Site Precision Conditions: Conditions under which test results are obtained in a single laboratory, by one or more operators who are implementing the same practice method, utilizing the same equipment, on a single analytical measurement system, using random specimens from the same sample of material.

Site Precision – R' : A value below which the absolute difference between two test results obtained under site precision conditions may be expected to occur with a probability of 95%. It equals to 2.77 (or roughly 2.8) times the standard deviation of the test results.

Site precision Standard deviation: The standard deviation of the test results that were obtained under site precision conditions.

Published reproducibility value – R : A benchmark published value, used in site precision evaluation.

Validation sample: A Quality Control sample or a check standard that is used to validate system precision and bias during quality assurance testing.

Systematic Errors: Errors that shift measurements from their true value by the same amount or fraction and in the same direction all the time. They affect the accuracy but not the reliability of a result and they are usually related to problematic or incorrectly used laboratory equipment.

Random Errors: Errors that shift measurements from their true value by a random amount and in a random direction. They affect the reliability but not the overall accuracy of a test result.

Outlier: A datum that according to a statistical test does not belong to the distribution of the rest of the data.

Grubb's limit: The statistical limit of the Grubb's test that determines if a datum is an outlier

USE OF MEASUREMENT PROCESS IN THE PETROLEUM INDUSTRY

In a quality conscious and cost competitive environment, proficiency and statistical control of measurement processes used for product certification are expected between suppliers and customers. The measurement process is a sub – process integral to all manufacturing processes. Like any other process, it has inputs and outputs. Inputs to the measurement process are usually samples taken at various stages of the manufacturing process, while outputs are numerical values. In the petroleum refining industry, two types of measurement processes are commonly encountered. One is the traditional laboratory – based process where a small sample is extracted from the main process streams and analyzed off – line using standard test methods. The other, commonly referred to as a continuous process analyzer system, is a field – deployed, fully automated instrumentation system designed to provide analytical information on representative samples continuously extracted from the main process streams.

The use of measurement process outputs (numerical values) can be categorized into three common applications:

- Manufacturing process control: To support decisions on appropriate manipulations (known as control actions) to key manufacturing process variables in order to meet required performance criteria.
- Product property conformance to specification: to test the hypothesis that the batch of product from which the test sample is taken meets the required quality specifications.
- Measurement process self – monitoring: To test the hypothesis that the test results generated by the measurement process are fit – for – use.

MEASUREMENT SYSTEM SELF – MONITORING

In order to be able to use the results generated from a measurement process with confidence, the fitness – for – use of these results must be evaluated. This can be achieved by regular application of the measurement process to quality control samples, as well as engaging the measurement process in regular interlaboratory testing programs (round robin studies). Through application of control chart techniques and round robin results, the performance (stability/precision, bias, over time variation) of the measurement process can be monitored and assessed. The set of activities used to achieve this objective is known as measurement process quality assurance program.

BASIC ELEMENTS OF STATISTICS USED IN QUALITY CONTROL

Statistics is the science that deals with the collection, description and analysis of data with main objective the extraction of conclusions that enables us to better comprehend the phenomena of the world that surrounds us. Statistical methods are the basic mathematical tools implemented in all quality control procedures. The three basic cognitive objects of statistics which are employed in quality control are the Sampling Process, the Descriptive Statistics and the Statistical Inference.

STATISTICAL SAMPLING PROCESS

Sampling is the sector of statistics that refers to the different techniques of data gathering. As statistical data are considered all the observations that are collected with a random or a well-defined manner, and consist the elements around which, after a statistical processing, conclusions will be extracted about the population under study. The data collected from such a population make up the sample. In order for these conclusions to be valid, it is imperative that the sample is representative of the population from which it originates. To understand the concept of sampling, someone should be first able to understand the definition of a population. A population can be thought of as a group of items about which information should be obtained. According to ISO 2589 – 1, a population is “a collection of units of a product from which a sample shall be drawn and inspected to determine conformance with the acceptability criteria”. Although the idea of population is easy to comprehend, the sampler must first ascertain what questions he wishes to answer as without this basic concept, the correct population with which the answers will be provided will be impossible to establish. The sample is simply a group of items from the population of interest. As it is invariably impossible or simply not cost effective to sample the whole population, the size of the sample is an important factor that has to be evaluated. Taking a small sample can often lead to

improved precision regarding the parameter of interest as more care can be invested into the measurement of each sample unit, thereby improving accuracy. Taking measurements from the whole population though, is likely to be time consuming and will ultimately affect the precision of the parameter of interest but it will be more representative of the general population. Ultimately, the sample should be as large or small as the sampler wishes. In principle a larger sample provides more information about the population. The size of the original population does not play any role in the accuracy provided by the sample, but it is the size of the sample that dictates its own accuracy. If it is achievable, it is advisable to take a larger sample to provide more precision for the parameter of interest and therefore more comfort to the decision maker. Careful documentation during sampling is required so that all relevant information on the nature of the sample (when it was taken, where it was taken and under what conditions it was taken) are clearly recorded on site at the time of sampling by the person conducting the sampling. This is necessary because variations in sampling procedures can have a marked effect on the results of analysis. It is very difficult to quantify these effects and, therefore, the most practical way to control this stage of the analytical process is to document sampling conditions as fully as possible. Quality assurance of sampling can be achieved in the following ways:

- Strictly adhere to standard operating procedures for sampling.
- Ensure all equipment is in working order.
- Record all conditions which applied during sampling.
- Take strict precautions to avoid contamination.
- Following those simple procedures should help to ensure that the quality of samples matches the quality of analysis.

In designing a sampling scheme, it is critical that the sample is representative of the population for which answers are sought, as it was already explained. The basic methods of sampling are the following:

- Systematic Sampling – Samples of defined size are obtained with a random manner, and during fixed periodic intervals, from the population under study.
- Stratified Sampling - Stratification is the process of dividing members of the population into homogeneous subgroups before sampling. No population element should be excluded and every element must be assigned to only one stratum. After stratification, systematic sampling is applied within each stratum.
- Directional Sampling – It is a special subcategory of stratified sampling with the main difference that after the dividing of the population into subgroups, the sampling is carried out not randomly but in a specific manner.

REFERENCE MATERIALS

QC samples are used to establish and monitor the precision of the analytical measurement system. They are stable and homogeneous materials having physical or chemical properties, or both, similar to those of typical samples tested by the analytical measurement system. Before the analysis, the quantity of the material needed for each specific lot of QC sample is estimated so as to accommodate the number of analytical measurement systems for which it is to be used and provide determination of QC statistics for a desirable period of time. Initially, the material is collected into a single isolated container. Subsequently it is thoroughly mixed so as homogeneity is ensured and the necessary tests are conducted in order to validate that the QC sample meets the characteristics of its intended use. Finally, the QC samples are stored to ensure that all analyses of samples from a given lot are performed on essentially identical material. For volatile samples, storage in one container which is repeatedly opened may result in loss of light ends. This problem can be avoided by chilling and splitting the bulk sample into smaller containers, each with sufficient quantity to conduct the analysis. Similarly, for samples prone to oxidation, the bulk sample can be splitted into smaller containers that can be blanketed with inert gas prior to being sealed and remain sealed until the sample is needed.

Check standards are used to estimate the accuracy of an analytical measurement system. A check standard may be a commercial standard reference material when such material is available in appropriate quantity, quality and composition, or alternatively it can be prepared from a material that is analyzed under reproducibility conditions (from various similar measurement systems). In an ideal world, a check standard should have a specified composition that is representative of routinely produced material and have an ARV produced with zero error, something which is unrealistic and unattainable. However, ARV with error that is small relative to the test method precision can be achieved by averaging the results from multiple testing systems. Even for those test methods that do have limited check standard testing requirements, the check standards are typically pure compounds or a blend of pure compounds which are typically not representative of the products tested. Therefore unknown composition effects prelude the extrapolation of performance statistics related to the system. Materials circulated as part of an interlaboratory exchange program can offer a cost – effective alternative to site – specific check standards. In order for an exchange sample to be usable as a check standard, the standard deviation of interlaboratory exchange program should not be greater than the reproducibility standard deviation of the test method. In general it is recommended that a minimum of 16 non – outlier results be used in calculating the ARV to reduce the uncertainty of the ARV by a factor of 4 relative to the measurement system single value precision.

Basic objective of the statistics as it was already mentioned is the study of the sample that will eventually lead to some conclusions concerning the population from which it was obtained. After the data sampling, a statistical processing is implemented and results in the construction of certain tables and charts that enables us to better understand and evaluate the qualities of the data results that concern us and the performance of the system. The analytical quality control (AQC), refers to all those statistical processes designed to ensure that the test results of a laboratory analysis are consistent, comparable, accurate and within specified limits of precision. The qualitative and quantitative data generated can then be used for decision making. This stage includes the construction of a series of charts that enable us to visually detect possibly erroneous data that fall outside certain statistical limits and can be associated with either systematic or random errors. Furthermore, a series of statistical tests are carried out during this stage and more specifically the Anderson – Darling normality test, the Student's t test, the Chi – square test and the F test. Quality control begins with sample collection and ends with the reporting of data. Another important part of a statistical quality control that defines the precision of an analytical test method is the interlaboratory study. An interlaboratory (round – robin) study should be conducted so as to compare the results of an individual measurement system with the results that similar systems of other laboratories provide, by implementing the same experimental method and subsequently, framing a precision statement about the method. The round – robin study is the final stage of an overall statistical quality control. Prior to this study, sensitivity – ruggedness test should also be performed so as to detect which experimental factors influence the test results and to determine how thoroughly these factors must be controlled. The ruggedness test should always precede the interlaboratory studies so as to restrict the allowable ranges of the critical variables of the test method, differently the precision statement provided by the round robin test will be inaccurate.

THE HISTOGRAM

The histogram is one of the seven basic tools of quality as was already mentioned before. A histogram is a simple and powerful quality control tool and an accurate representation of the distribution of numerical data. To construct a histogram, the first step is to “bin” the range of values – divide the entire range of values into a series of intervals – and then count how many values fall into each interval. The data is represented by columns on a graph which vary in height, depending on the frequency a specific datum (or range of data) occurred in a study. A histogram is used as a quality tool for a variety of reasons:

- Display of data in an easy to interpret graphical manner.
- Depiction of frequency of occurrence of data values.
- Reveals the centering and variation of data values.
- Provides a rough idea about the underlying distribution of the data.
- Enables future prediction of process performance.
- Enables identification in changes in process parameters.
- It provides visual evidence concerning the deviation of the data from the accepted reference value.

Histograms should not be confused with Bar Charts. Bar charts measure the frequency of categorical data, and the classes for a bar chart are these categories. On the other hand histograms are used for data that is at least at the ordinal level of measurement.

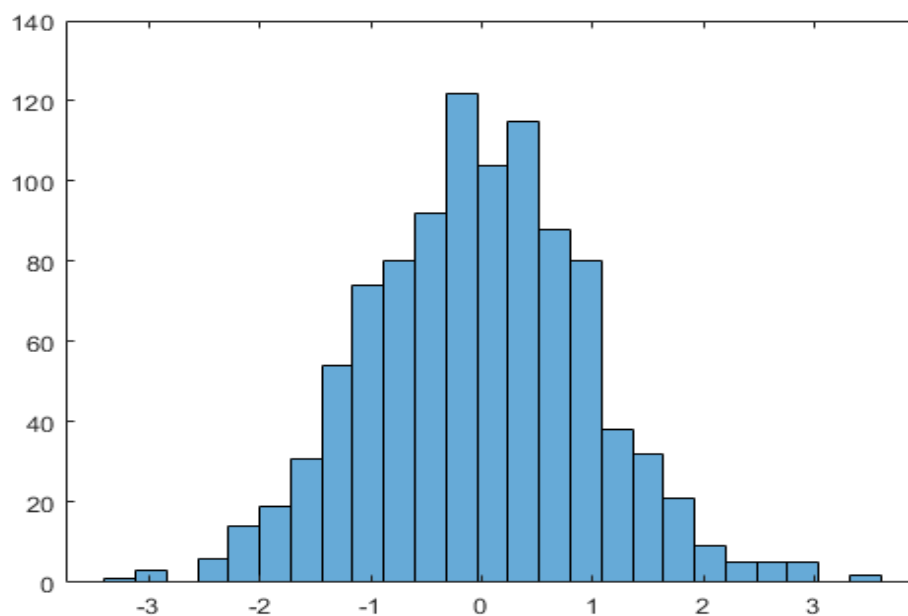


Image 14: A typical Histogram.

ANDERSON-DARLING STATISTIC

The Anderson-Darling test is a statistical test invented in 1952 by Theodore Wilbur Anderson and Donald A. Darling. It is usually applied to test if a set of data is adequately described by the normal distribution and generally it is considered as a powerful statistical tool, widely used to detect departures from normality. Today, most frequentists throughout the world consider that the normal distribution is one of the basic factors that divulge if a process is in statistical control state in small sized samples. It is accepted that data should be moderately normal. The assumption of normality has to be checked for many statistical procedures, namely parametric tests and control charts because their validity depends on it. Minor departures from normality do not significantly affect the results of these procedures but severe departures can increase the number of false out – of – control signals and subsequently, the results of the parametric tests may be misleading about the performance of the process. Smaller sized samples should always be tested for normality because presumably they have more control over the process. Many laboratory textbooks suggest that in larger samples ($N > 50$) the rejection of normality does not importantly affect the overall process, since the sampling distribution tends to be normal regardless of the shape of the data, an assumption that derives from the Central Limit Theorem. This should not be generally adopted as a “de facto” doctrine and should be either accepted or rejected based on the character of the process.

The Anderson – Darling test involves the following steps:

Array the test results in ascending order such as:

$$x_1 \leq x_2 \leq \dots x_n$$

Obtain standardized variate from the x_i as follows:

$$W_i = \frac{(X_i - X_{avg})}{s}$$

where s is the sample standard deviation and X_{avg} is the average of the test results

Convert the W_i values to standard normal cumulative probabilities P_i by utilizing the cumulative probability table for the standardized normal variate z .

$$P_i = \text{Probability}(z < W_i)$$

Compute A^2 as follows:

$$A^2 = \frac{\sum_{i=1}^n (2i - 1) * [\ln(P_i) + \ln(1 - P_{n+1-i})]}{n} - n$$

Compute A^{2*} as follows:

$$A^{2*} = A2 * \left(1 + \frac{0.75}{n} + \frac{2.25}{n^2}\right)$$

Upon finishing the calculations described above, the hypothesis of normality will be revisited and will be accepted or rejected, based on the value of A^{2*} . If the value is less than 0.752 then the hypothesis of normality is accepted at the 95% confidence level. By decreasing this specific value – 0.752 – the accepted confidence level increases above 95%. Similarly, if it is increased above 0.752, the accepted confidence level is being decreased below 95%. Although the normal distribution is the basis of following control charts, they are considered relatively robust and reliable in the face of non – normally distributed quality characteristics also.

THE LEVEY-JENNINGS CONTROL CHART

The Levey – Jennings control test in essence produces another statistical chart that is widely used in laboratory quality control. It was named after S. Levey and E.R Jennings who in 1950 suggested the use of control charts in the clinical laboratory. The basic concept of the Levey Jennings control chart is almost identical to the one that Shewhart applied on his Individual control chart, with the main difference being traced in the way that the standard deviation is estimated. More specifically, the Levey – Jennings control chart uses the long term estimation of sigma while the Shewhart Control charts use the short term estimation. As the long term estimation of sigma it is indicated the overall amount variation that the sample in its totality exhibits, while with the short term estimation of sigma it is indicated the amount of variation displayed within subgroups of the overall sample. The long term variation of sigma is always wider than the short term variation. In order to produce a Levey – Jennings chart, the standard deviation is used. The chart is employed to graphically depict successive (run to run) quality control data values. The first step includes the calculation of the decision-control limits. These limits are the $m \pm s$, $m \pm 2s$, $m \pm 3s$. The first range – $m \pm s$ – adheres to 68% of all the quality control data values. Likewise, $m \pm 2s$ adheres to 95.5 % of all the data values while the remaining 4.5% data results fall outside the $m \pm 2s$ range. Finally, 99.7% of all data values are found within the $m \pm 3s$ range and, subsequently, only 0.3% of the data will fall outside this limit. The $m \pm 3s$ ranged is considered in general and throughout the world as the “limit” of acceptance for a value in statistical quality control. Some laboratories incorrectly consider that all quality control data values that fall outside the $m \pm 2s$ range area are invalid and out of control. So it is almost certain that laboratories that accept these limits reject good data points.



Levey-Jennings Chart

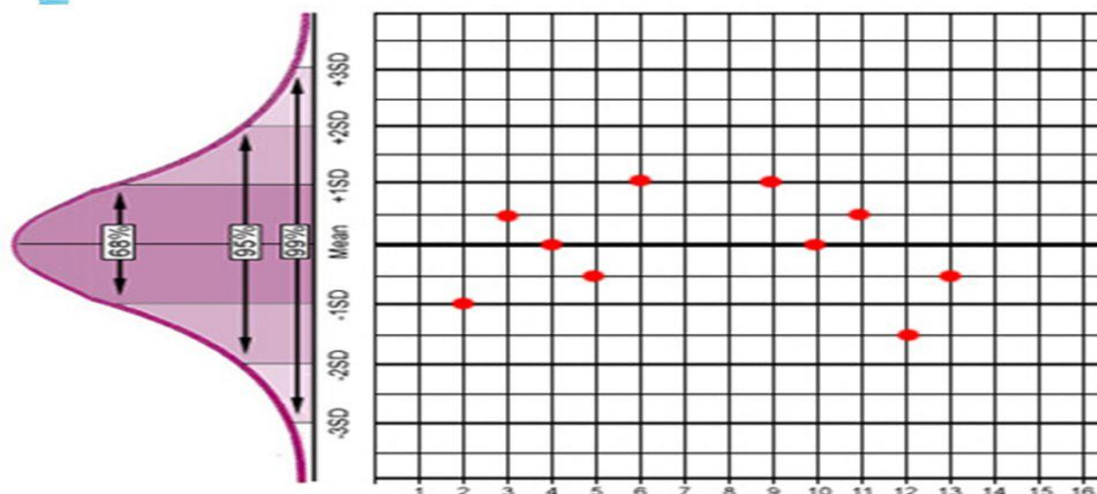


Image 15: A typical Levey – Jennings Chart.

The Levey – Jennings chart that was developed can also be overlaid onto a bell shaped curve to further illustrate, and confirm the normal distribution of the test values. In cases where the normality assumption is not valid, the interpretation of the Levey – Jennings chart is not reliable.

The errors in this practice can be either systematic or random. A systematic error is detected by a change in the mean of the values which may be gradual and be depicted as a trend, or it may be abrupt and be depicted as a shift in the values. The random error, for quality control results, is essentially any deviation (positive or negative) from the calculated mean. As was already mentioned, any data points that fall outside the $m \pm 3s$ range (99.7% of the data values) are considered as unexpected or unacceptable random errors.

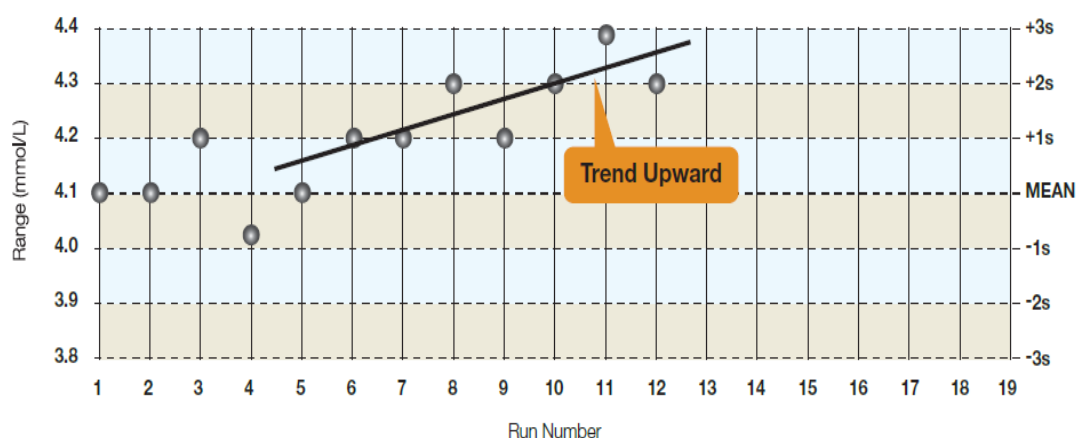


Image 16: A Levey – Jennings Chart with an upward trend.

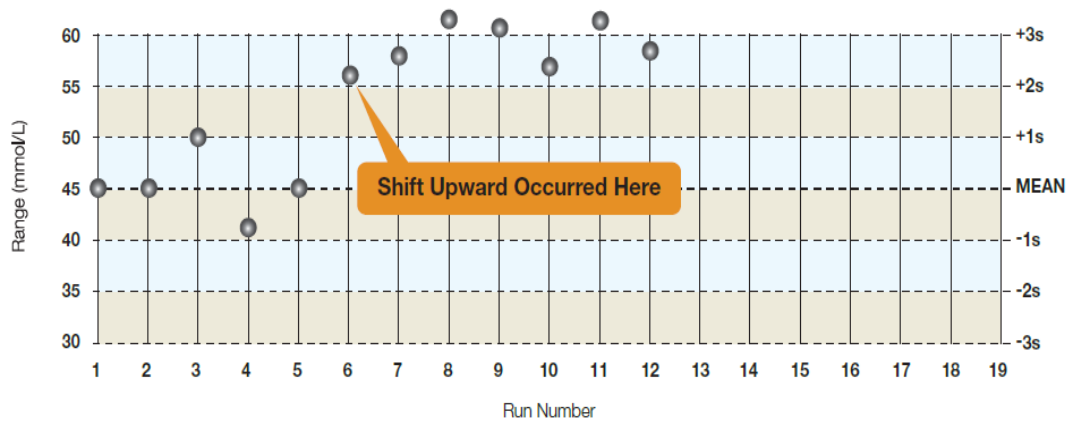


Image 17: A Levey – Jennings Chart with an upward shift.

WESTGARD RULES

In 1981, Dr. James Westgard, established a set of rules that eventually became the basis for evaluating analytical quality control result data. His basic intention was to come up with a basic set of rules that would enable anyone to detect possibly erroneous data and flag inconsistent results in his analytical measurement system. His work was based on principles of statistical control that had already been used in the industry from the 1950's. This set is consisted of nine rules that are used to assess the quality of an analytical run. Dr. Westgard also developed a specific notation. More specifically, his control rules are expressed as N_L where N denotes the number of control observations while L denotes the statistical limit for accepting or rejecting the control observation. Below, these eight rules are explained in detail.

RULE 1_{2s}

This rule alerts that a single data point falls outside the $m \pm 2s$ range. This is simply a warning rule which indicates that a systematic or random error may be present in the analytical measurement system, so the connection between this point and other data results must be reassessed. If not outward reason can be identified for this result to be outside the $\pm 2s$ limit, then it should be accepted as a random error.

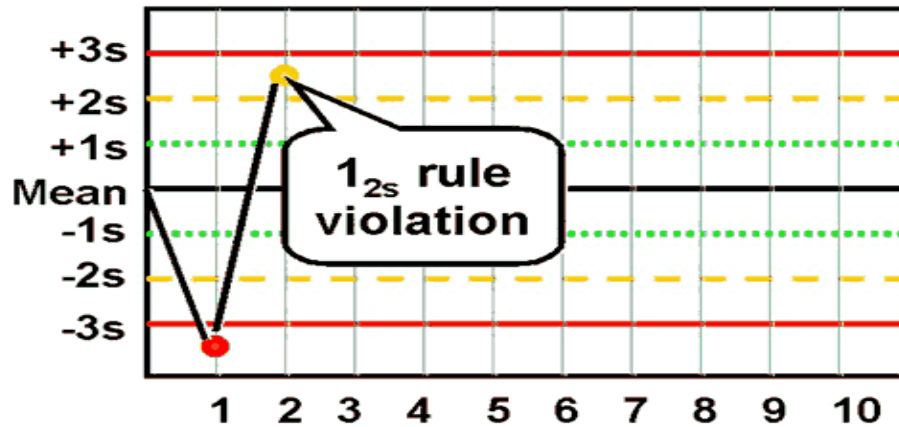


Image 18: Violation of the 1_{2s} rule.

RULE 1_{3s}

This rule indicates an unacceptable random error or the beginning of a large systematic error. Obviously, any data result that falls outside the $m \pm 3s$ limits violates this rule. A run is rejected when a single control measurement exceeds the mean plus 3s or the mean minus 3s control limit.

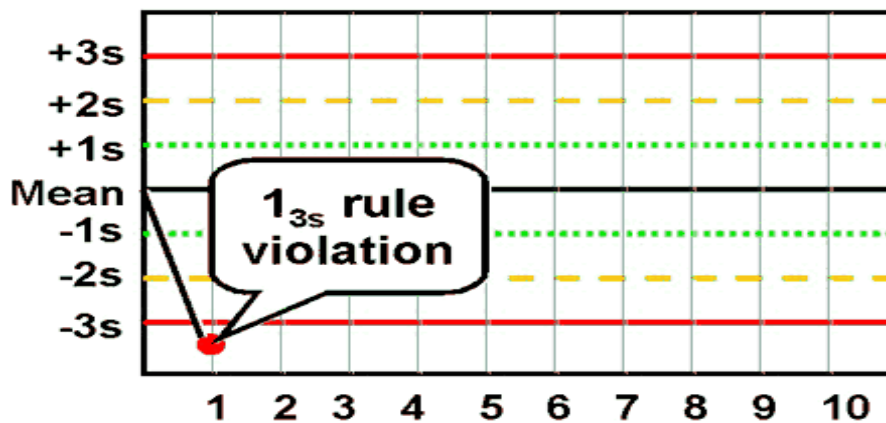


Image 19: Violation of the 1_{3s} rule.

RULE 2_{2s}

This rule indicates that two consecutive data points fall outside the $m \pm 2s$ boundaries and identifies a systematic error only. This rule is violated when two consecutive data results, on the same side of the mean, exceed the $\pm 2s$ limit.

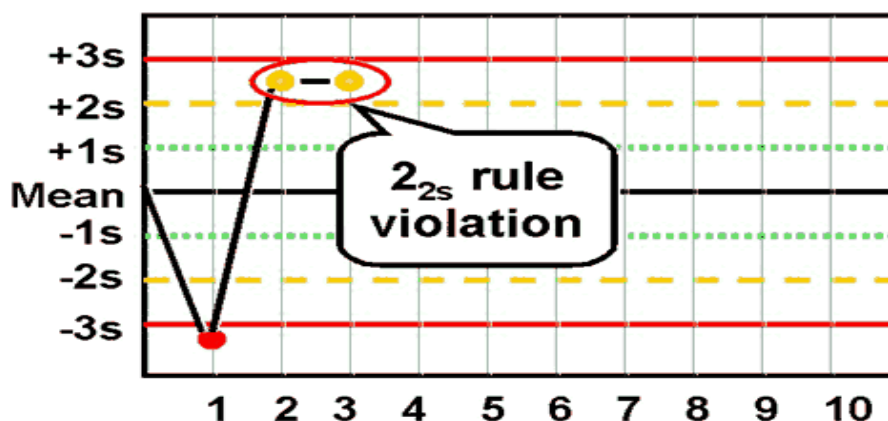


Image 20: Violation of the 2_{2s} rule.

RULE R_{4s}

The R_{4s} rule detects random error only and should be interpreted only within-run. As the name suggests, the rule is violated when 2 consecutive data results abstain from each other distance $4s$ or greater.

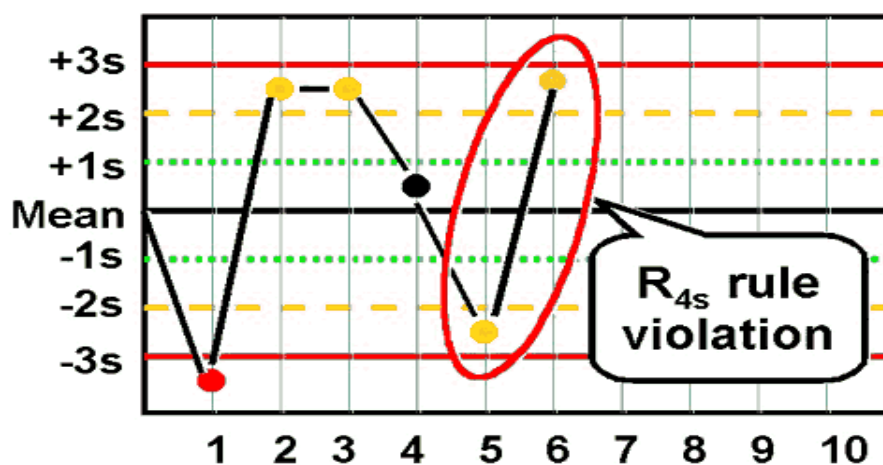


Image 21: Violation of the R_{4s} rule.

Any violation of the following five Westgard rules does not necessarily indicate direct rejection of the analytical run. On the contrary, these violations are related to smaller systematic errors or analytical bias which usually is insignificant or irrelevant and can be addressed by performing maintenance or calibration to the analytical system.

RULE 3_{1s}

The 3_{1s} rule is violated when three consecutive data points, on the same side of the mean, are greater than $1s$.

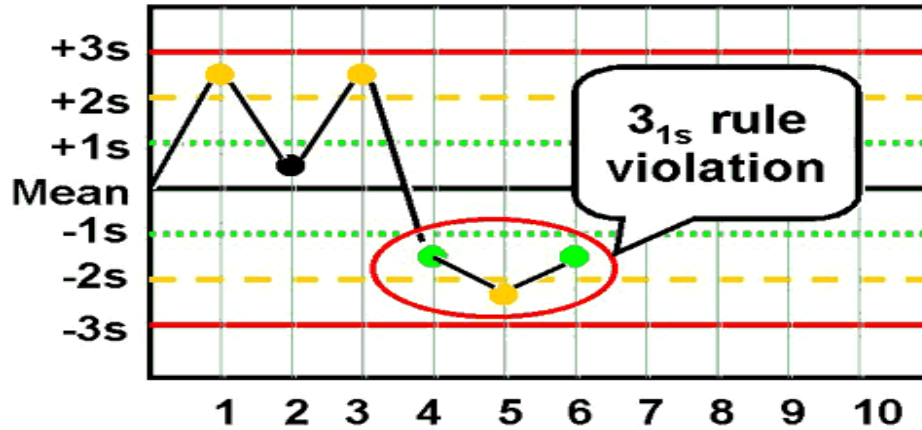


Image 22: Violation of the 3_{1s} rule.

RULE 4_{1s}

The 4_{1s} rule is violated when four consecutive data points, on the same side of the mean, are greater than $1s$.

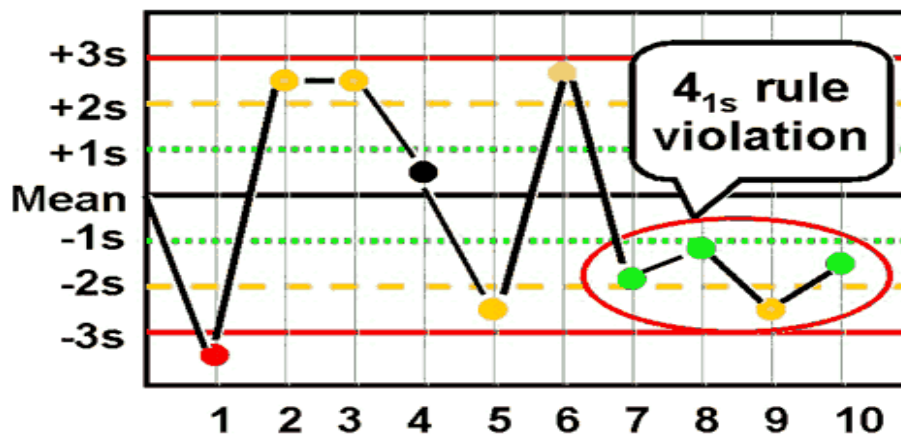


Image 23: Violation of the 4_{1s} rule.

Violation of the 3_{1s} and/or 4_{1s} rules is indicative of systematic error in a single area of the chart and the specific values that belong to that area should be further investigated.

8_x, 10_x AND 12_x RULES

These three rules are violated when eight, ten and twelve consecutive data points respectively are located on the same side of the mean regardless of the specific standard deviation in which they are located. Again, as was the case with the previous rules, their violation is indicative of systematic error in a single area of the chart.

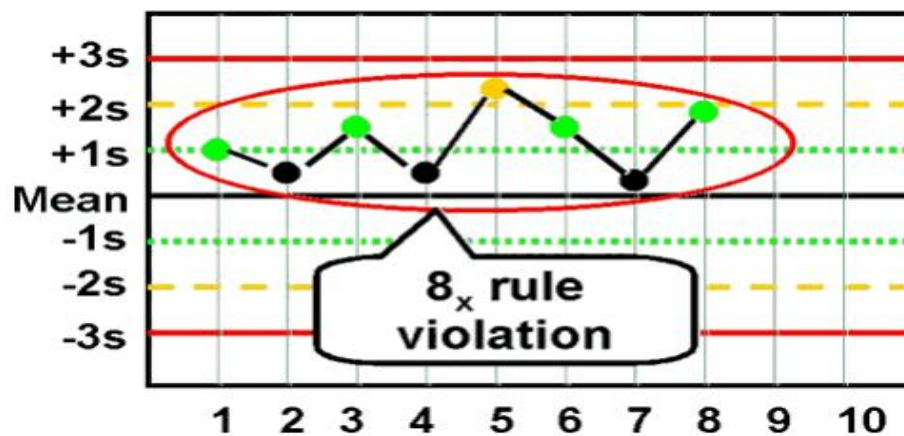


Image 24: Violation of the 8_x rule.

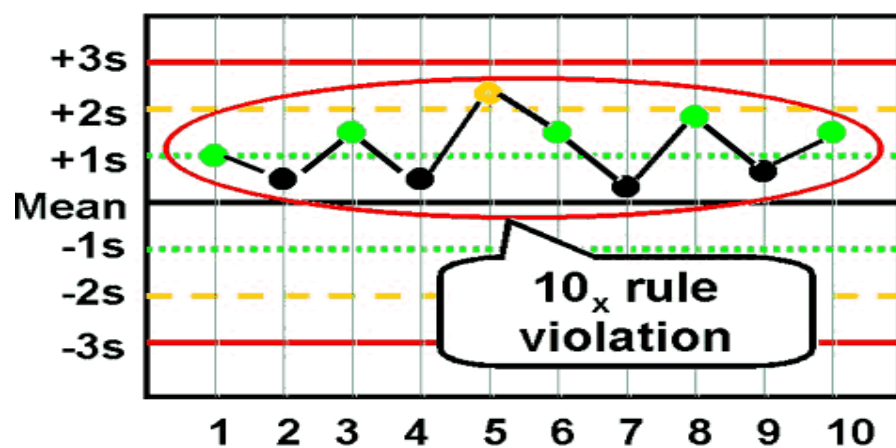


Image 25: Violation of the 10_x rule.

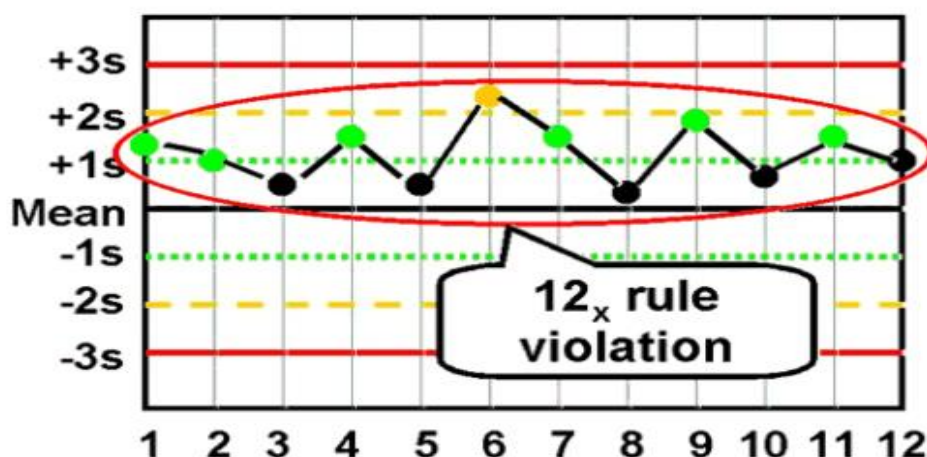


Image 26: Violation of 10_x rule.

The evaluation of data results based on the Levey – Jennings chart and the Westgard rules is generally a more lenient evaluation in comparison to the techniques that will be described below, because as it was already mentioned the long term estimation of sigma is used for its construction and not the short term that the following charts use.

PRETREATMENT OF TEST RESULTS

Assessment, control charting and evaluation of the data produced by an analytical system should be applied after the test results are appropriately pretreated. The purpose of pretreatment is to standardize the control chart scales so as to allow for data from multiple check standards (if more than one are used during the process) to be compared on the same chart. For Quality Control sample test results, no data pretreatment is typically applied since results for different Quality Control samples are generally not plotted on the same chart. For check standard sample test results and depending on the measurement system precision, two cases apply:

Case 1- If either all of the check standard results are from one or more lots of check standard material having the same ARV (reference value), or the precision of the measurement system is constant across all levels, then the pretreatment procedure consists of calculating the difference between the raw test results and the ARV:

$$\text{Pretreated Result} = \text{Raw Test Result} - \text{ARV}$$

Case 2- Test results are for multiple lots of check standards with different ARV's and the precision of the analytical measurement system varies with level.

$$\text{Pretreated Result} = \frac{[\text{Raw Test Result} - \text{ARV}(\text{of the sample})]}{\text{Standard deviation at the ARV level}}$$

Statistical techniques are applied to the pretreated data results in order to a) Identify erroneous data, b) Assess initial results, c) Construct, interpret and maintain Run and Control Charts and d) Quantify long term measurement precision and bias.

RUN CHARTS

A Run Chart is a line graph of data plotted over time and it is used as a quick test of the analytical system's performance. Preferably, pretreated results are plotted. By collecting and charting data over time, it is possible to detect "abnormal" trends or patterns in the process. Because they do not use control limits, run charts cannot determine if a process is stable. However, they can indicate how the process is running. The run chart can be a valuable tool at the beginning of a project, since it displays process performance over time. If 15 data points or more are available, a run chart can be constructed to detect special causes – something beyond the usual variability of the process – acting on the process. These special causes include 1) Shifts, 2) upward and downward Trends and 3) unusual Patterns that may be spotted and investigated further. These special causes can also be detected in the Individual Chart which in essence is a Run Chart with control limits and its construction will be explained later.

- **Shifts:** Eight or more consecutive points on one side of the center line, something which indicates that a special cause has influenced the process. Points on the center are not taken under consideration since they don't break the string, not add to it.
- **Trends:** Six consecutive jumps on the same direction indicate that a special cause is acting on the process to cause a trend. Flat line segments don't count, either to break a trend or to count towards it.
- **Pattern:** A pattern that recurs eight or more times in a row is indicative of a special cause.

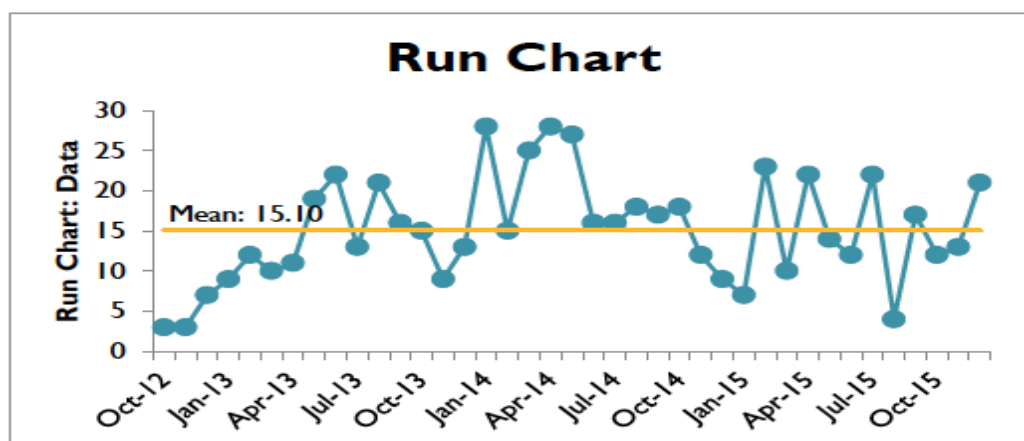


Image 27: Typical Run Chart.

For more robust monitoring of an analytical system's performance and more detailed information about when your process is exhibiting significant deviation from the expected, the construction of Control Charts succeeds the Run Chart. With the Control charts it is possible to detect special causes more quickly and with more elaborated accuracy.

CONTROL CHARTS

Control charts, also known as Shewhart charts (after Walter A. Shewhart) or process-behavior charts, are a statistical process control tool used to determine if a manufacturing or business process is in a state of control. As mentioned above, the control chart was invented by Walter Shewhart who was working for the Bell Labs in the 1920's. The main objective of the company's engineers at that time was the improvement of the reliability of their telephony transmission systems and since the amplifiers and other relative equipment had to be buried underground, there was a strong business need to reduce the frequency of failures and repairs. By that time, the engineers have already realized the importance of reducing variation in a manufacturing process and had also realized that frequent process-adjustment in reaction to non-conformance actually increased variation and resulted in quality degradation. So in May 1924, Shewhart addressed the problem in terms of Common and Special causes of variation and introduced the Control Charts as a tool for distinguishing between the two. In his sort report, he pointed out that bringing and maintaining an analytical system into a state of statistical control where there is only common-cause variation, is vital both for the economic management of the system and to forecast future output. While conducting his experiments, Shewhart concluded that data from physical processes naturally produce a normal-Gaussian distribution curve, widely known as "bell curve". Finally, he observed that while every process exhibits variation, some processes exhibit controlled variation that is natural to the process, while others exhibit uncontrolled variation which sometimes is absent from the process causal system. In the decades that followed his work became widespread known and his Control charts served as a strong statistical tool that helped the manufacturing industry progress and expand significantly. The control chart is considered as one of the seven basic tools of Quality control.

The seven basic tools of quality control is a designation that describes a specific set of graphical techniques that provide flexible and easy solutions in troubleshooting issues related to quality control. They are called "basic" because they can be even used by people who are not specialized in this specific field of mathematics and they provide solution to a vast variety of quality related problems. These basic tools are:

- The Cause and Effect or “Ishikawa” diagram
- Check sheet
- Histogram
- Run Chart
- Control Chart
- Pareto chart
- Scatter diagram

Control Chart

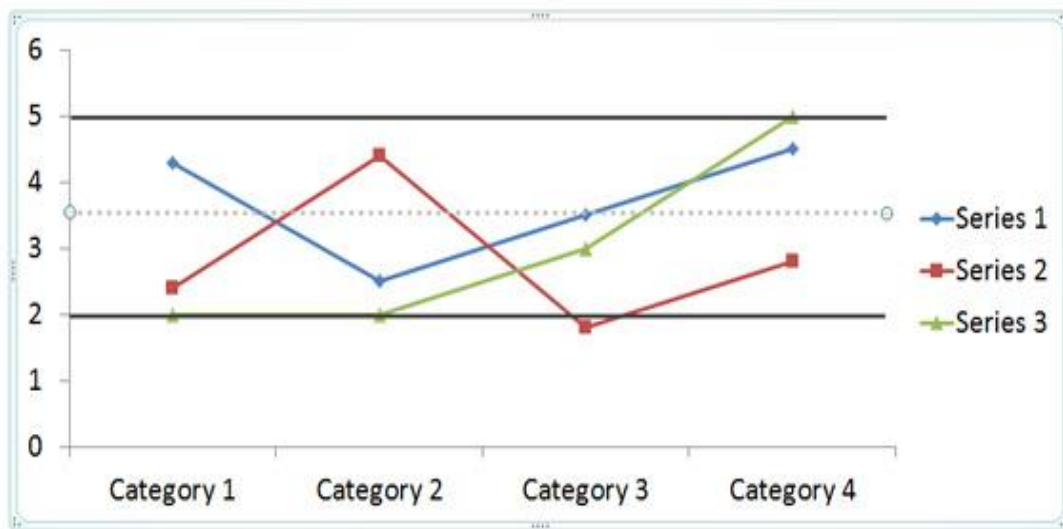


Image 28: Typical Control Chart.

In this specific project, the Control Charts that were constructed after the Run chart were 1) the I (Individual) control chart, 2) the MR (Moving Range) control chart and 3) The Exponentially weighted moving range (EWMA) chart and 4) the exponentially weighted moving average (EWMA) overlay on the Individual chart so as to enhance its detection power for small level shifts.

These techniques are recommended tools for **(a)** routine recording of Quality Control samples and check standard test results and **(b)** forthright assessment of the “in statistical control” status of the analytical system that generated the data. All control charts have three main objectives:

- Monitoring the stability of a process (Even in the most stable of processes, some variations may be detected and a further attempt to “fix” minor fluctuations in the process may result into instability.)
- Detecting whether a process is stable and ready to be improved.
- Demonstrating process performance.

THE I CHART

The I or Individual chart is a run chart to which upper and lower control limits and a specific horizontal center line have been incorporated. It is produced only after a minimum of 15 preprocessed data results have been obtained from the analytical measurement system. The horizontal center line is introduced in the run chart at the level of the mean of all the results:

$$I_{avg} = \frac{\sum_1^n I_i}{n}$$

Upper and lower control limits are added also in the run chart computed from the MR_{AVG} :

$$MR_{avg} = \frac{\sum_1^{n-1} |I_{i+1} - I_i|}{n - 1}$$

$$\text{Upper Control Limit} = I_{avg} + 2.66 * \frac{\sum_1^{n-1} |I_{i+1} - I_i|}{n - 1}$$

and

$$\text{Lower Control Limit} = I_{avg} - 2.66 * \frac{\sum_1^{n-1} |I_{i+1} - I_i|}{n - 1}$$

where 2.66 is obtained by dividing 3 by the sample size specific anti – biasing constant for $n=2$.

Individual data points that fall outside the “borders” defined by the upper and lower control limits are indications of instability, reveal shifts in the process that alter the mean or variance of the measured statistic and denote that the system goes out – of – control. Efforts should be made to determine the cause. More specifically, the following occurrences should be considered as potential signs of instability and should be revisited further:

- Two out of three consecutive data points on the I chart that are more than $1.77 * MR_{AVG}$ distant from the center line in the same direction.
- Five consecutive data points on the I chart that are more than $0.89 * MR_{AVG}$ from the center line in the same direction.
- Eight or more consecutive data points on the I chart that fall on the same side of the center line.

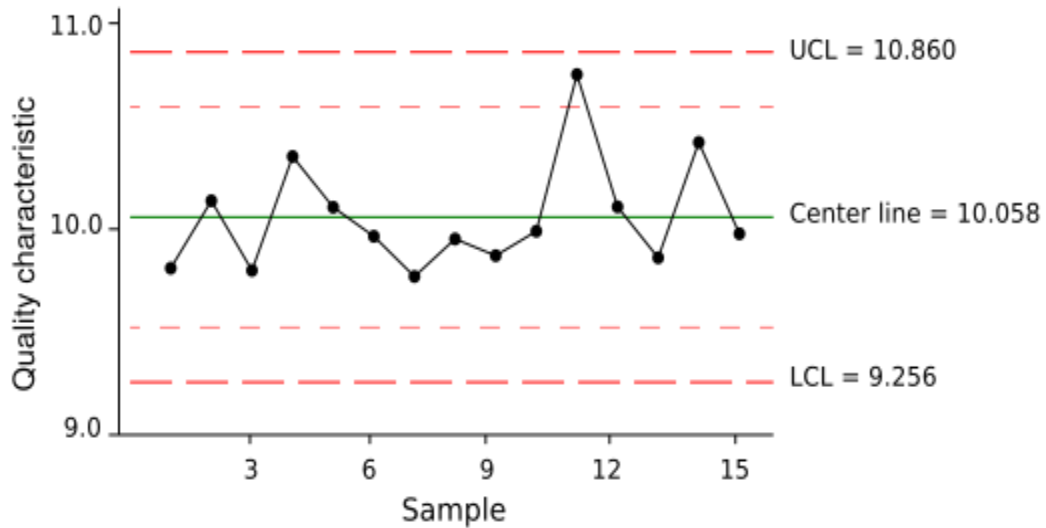


Image 29: Typical Control Chart with Upper and Lower Control Limits.

THE MR CHART

The Moving Range chart is a control chart that, as the name indicates, is constructed by plotting the values derived from the time-ordered sequential data points. Each moving range point is being calculated as $X_n - X_{n-1}$ and hence the MR chart is going to have one data point less than the Individual control chart.

Subsequently, MR_i is defined as:

$$MR_i = |I_i - I_{i-1}|$$

It is important to mention that there is not lower control limit for the MR chart and the upper control limit is given by:

$$UCL_{MR} = 3.27 * \frac{\sum_{i=1}^{n-1} |I_{i+1} - I_i|}{n - 1}$$

where 3.27 is taken from the sample size – specific D_4 anti – biasing constant for $n=2$.

The MR chart contributes in assessing the stability of a process in terms of precision caused by the over time variation between consecutive, individual data points. As was the case with the Individual chart, data points that fall outside the “boundaries” defined by the control limits indicate certain instability in the process. An assignable cause that triggered the instability of the system should be determined and if possible removed from the affected data points of both the Individual and the Moving Range chart.

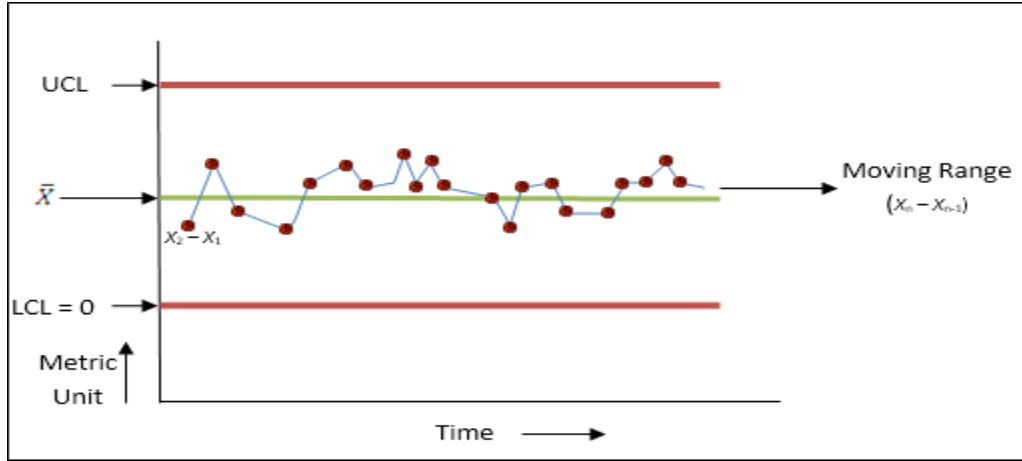


Image 30: Typical Moving Range Chart.

Here it is important to mention that the normal distribution is not assumed not required in the calculation of the control limits of the both the Individual and the MR chart as was demonstrated by Wheeler recently, thus making the two charts a very robust statistical tool.

EWMA OVERLAY ON I CHART

The EWMA trend line is typically overlaid on the Individual chart to enhance its sensitivity in detecting mean shifts that are small and difficult to detect. Each EWMA value is a weighted average of the current preprocessed result and previous results, with the weights decreasing exponentially with the age of the reading. While other control charts treat rational subgroups of samples individually, the EWMA chart tracks the exponentially weighted moving average of all prior sample means. EWMA weights samples in geometrically decreasing order so that the most recent samples are weighted most highly while the most distant samples contribute very little. More specifically, a sequence of $EWMA_i$ values are calculated and overlaid on the I chart, using the following recursion equation:

$$EWMA_i = I_i$$

$$EWMA_i = (1 - \lambda) * EWMA_{i-1} + \lambda * I_i$$

where λ is the exponential weighting factor.

The parameter λ determines the importance of older data for the calculation of EWMA. Smaller values of λ give more weight to recent data while larger values give more weight to older data. A value of 0.4 was used for the application of this

practice because it closely emulates the run rule effects of conventional control charts. Conveniently, a λ value of 0.4 places the control limits to 1.5 – sigma for the I chart. The control limits for the EWMA chart are calculated with the use of the exponential weighting factor λ as follows:

$$UCL_{\lambda} = I_{avg} + 2.66 * \frac{\sum_1^{n-1} |I_{i+1} - I_i|}{n-1} * \sqrt{\frac{\lambda}{2-\lambda}}$$

$$LCL_{\lambda} = I_{avg} - 2.66 * \frac{\sum_1^{n-1} |I_{i+1} - I_i|}{n-1} * \sqrt{\frac{\lambda}{2-\lambda}}$$

It is important to mention that the above control limits are “tighter” than the control limits of the Individual chart thus, it can detect outliers that the Individual chart did not.

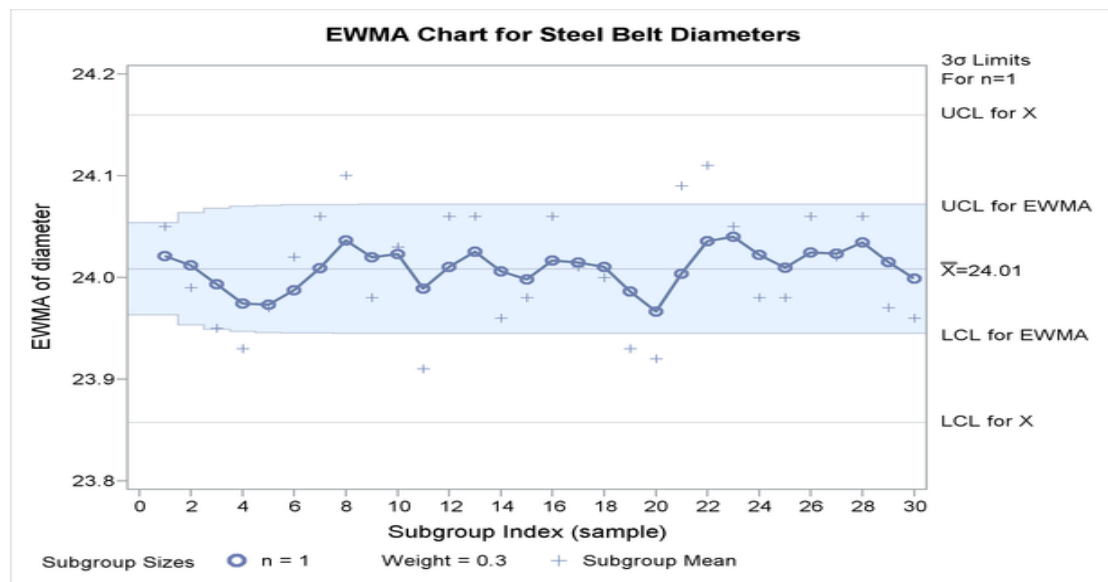


Image 31: Typical EWMA Chart with λ factor of 0.3.

Although the normal distribution is the basis of the EWMA chart, it is also robust and reliable in the face of non – normally distributed quality characteristics also.

STUDENT'S-t test

Among the most commonly used statistical significance tests applied to small data sets (population samples) is the series of Student's tests. One of these tests is used for the comparison of two means, which is commonly applied to many cases. The outcome of these tests is the acceptance or rejection of the null hypothesis (H_0). The null hypothesis generally states that: "Any differences, discrepancies, or suspiciously

outlying results are purely due to random and not systematic errors". A Student's t test is used to check if a sample value comes from a population with a mean different from a hypothesized value, μ_0 . It is commonly applied when the test statistic would follow a normal distribution if the value of a scaling term in the test statistic was known. When the scaling term is unknown and is replaced by an estimate based on the data, the test statistics (under certain conditions) follow a Student's t distribution.

In this project, a t test was performed on pretreated check standard test results to check for bias relative to the ARV. Since during pretreatment, accepted reference value has been subtracted from the raw results, the hypothesized mean value equals to zero ($\mu_0=0$). For the purpose of performing the t test, two methods can be implemented, the root-mean square method and the MR approach.

By the root –mean square method, the standard deviation of the pretreated results is calculated as:

$$S_I = \sqrt{\frac{\sum (I_i - I_{avg})^2}{n - 1}}$$

And the t value of the test is calculated as:

$$t = \frac{\sqrt{n} * |I_{avg} - \mu_0|}{S_I}$$

Alternatively, the t value can be calculated by the MR approach. Following this method, the alternate t value was computed as follows:

$$t_{MR} = \frac{\sqrt{n} * |I_{avg} - \mu_0|}{\left(\frac{\sum_{i=1}^{n-1} |I_{i+1} - I_i|}{n - 1} \right) \cdot 1.128}$$

where μ_0 is the hypothesized mean value, which is zero as already explained above.

The computed absolute t value is being compared with the published in the literature critical t values, for $(n-1)$ degrees of freedom for the root-mean square method, or for $(n-1)/2$ degrees of freedom if the MR approach is implemented. If the absolute value of the calculated t is less than or equal to the published critical value, then μ_0 is statistically undistinguishable from the mean of the distribution, a fact that for check standard testing indicates no statistically identifiable bias in the measurement system. On the other hand, if the computed absolute t value is greater than the published value, then μ_0 is statistically distinguishable from the mean of the distribution with 95% level of confidence. For check standard checking this would indicate statistically identifiable bias in the system. Here it is important to mention

that the result of the Student's test gives a relative statement about the bias present in the system. There may be situations where the statistical test is important but the magnitude of bias detected is fairly small relative to the business application need. In these cases the bias is deemed mainly of statistical significance and less of practical.

APPROXIMATE F TEST

The F test of equality of variances was initially created as a variance ratio in 1925 by Sir Ronald Fisher. In statistical quality control an F -test of equality of variances tests for the null hypothesis that two normally or approximately normally distributed populations have the same variance. In this project, an approximate F test is used to compare the variation exhibited by an analytical measurement system over two different time periods. Alternatively, it can also be implemented to compare the site precision estimated from a series of results from one check standard with that estimated using a different check standard. In order to implement the F test the following procedure is followed:

The F value is being computed as follows:

$$F = \frac{MR_{avg}(1)}{MR_{avg}(2)} = \frac{\sum_{i=1}^{n-1} |I_{i+1} - I_i|}{n-1} (1) / \frac{\sum_{i=1}^{n-1} |I_{i+1} - I_i|}{n-1} (2)$$

Where $MR_{avg}1$ represents the larger of the two moving average ranges and $MR_{avg}2$ the smaller.

The computed F value is being compared with the Critical F value which is obtained from literature tables, with $(n_1 - 1)/2$ degrees of freedom for the numerator and $(n_2 - 1)/2$ for the denominator.

If the computed F value is greater than the value obtained from the table, the two precisions are statistically distinguishable and this fact leads to the conclusion that there is 95% of confidence that the analytical process performed to produce $MR_{avg}1$ has larger site precision, in other words is less precise, than the analytical process that produced $MR_{avg}2$. A significant F – test should trigger an investigation for assignable causes.

On the other hand, if the computed F value is smaller than the table value, then the precisions of the two samples are statistically indistinguishable.

Finally, if the two precisions are statistically indistinguishable, they can be combined-“pooled” into a single estimate. The pooled precision estimate is computed as follows:

$$MR_{pooled} = \frac{(n_1 - 1) * MR_1 + (n_2 - 1) * MR_2}{n_1 + n_2 - 2}$$

SITE PRECISION AND SITE PRECISION STANDARD DEVIATION

The site precision is a value below which the absolute difference between two test results obtained under site precision conditions may be expected to occur with a probability of 95%. **Site precision and the site precision standard deviation** of the measurement system can eventually be calculated from the MR_{avg} of the MR chart for that specific lot as:

$$R' = 2.46 * \frac{\sum_1^{n-1} |I_{i+1} - I_i|}{n - 1}$$

And the site precision standard deviation as:

$$\sigma_{R'} = \frac{(\frac{\sum_1^{n-1} |I_{i+1} - I_i|}{n - 1})}{1.128}$$

Alternatively, R' can be calculated using the root – mean square formula for standard deviation:

$$\sigma_{R'} = \sqrt{\frac{\sum (I_i - I_{avg})^2}{n - 1}}$$

and

$$R' = 2.77 * \sigma_{R'}$$

APPROXIMATE CHI – SQUARE TEST (χ^2)

The Chi – square test is specifically employed here to compare the site precision of an analytical measurement system with a published reproducibility value. The Chi – square test includes the following steps:

Calculation of the chi – square statistic as follows:

$$\chi^2 = \frac{(n-1)*R'^2}{2R^2}$$

where R' is the estimated site precision, which equals to $2.46*MR_{avg}$, and R is the published reproducibility value.

Comparison of the computed X^2 with the critical X^2 value available in the literature, for $(n-1)/2$ degrees of freedom.

If the computed value is greater than the literature value then the site precision transcends the published reproducibility with 95% of confidence, while if it is less than or equal to the literature value then the site precision is statistically indistinguishable from the published reproducibility of the method.

PROCESS CAPABILITY AND PROCESS PERFORMANCE

The process capability index is the ratio of two quantities that relates the natural tolerance limits with the specification limits. In essence it is an indication of the production capabilities of a process in accordance to the specification limits, the Lower Specification Limit (LSL) and the Upper Specification Limit (USL). The older and most common process capability index is the C_p index who quantitatively expresses the connection between the USL – LSL range and the 6σ range.

$$C_p = \frac{USL - LSL}{6\sigma}$$

C_p describes the capabilities of a process under the assumption that the mean of the population is similar with the average of LSL – USL range. When the mean of the process is different from the average of the specifications limits then the capabilities of the process are expressed with the C_{pk} index that equals to:

$$C_{pk} = \min\left\{\frac{m - LSL}{3\sigma}, \frac{USL - m}{3\sigma}\right\}$$

- Negative values of C_{pk} indicate that the average quality is outside the specification limits.
- If C_{pk} equals to zero, the average quality equals to one of the two specification limits.
- If $0 < C_{pk} < 1$, then the average quality is within the specification limits but one part of the process is not.
- If $C_{pk} = 1$, then one of the control limits is equal to one of the specification limits
- If $C_{pk} > 1$, then the process is within the specification limits.

Similar to the process capability index, the process performance index P_p rate utilizes the long term estimation of sigma and is defined as:

$$P_p = \frac{USL - LSL}{6s}$$

Similarly, the process performance rate is defined as:

$$P_{pk} = \min\left\{\frac{USL - m}{3s}, \frac{m - LSL}{3s}\right\}$$

Per ASTM E2281 standard, process capability indices (C_p , C_{pk}) compare the variability of a process quality measure against product specifications or tolerances and assume that the process is in state of statistical control. In different case, the conclusions derived from the results of the indices will be inaccurate and misleading about the process. Process performance indices (P_k , P_{pk}) are useful during the initial set – up of a process but also in situations when the process is not in a state of statistical control.

The process performance indices essentially denote how the process has actually performed in the past. The main difference with the process capability indices is that C_p indices describe what the process is capable of doing in the future, assuming that it remains under statistical control state. On the other hand, process performance indices cannot be used for future prediction because the process is not in state of control and therefore the conclusions will be unreliable. C_{pk} and P_{pk} will converge to almost the same value when the process is in a state of statistical control.

- Negative values of P_{pk} indicate that the process falls outside the specification limits and the process is producing a large proportion of defective output.
- If $0 < P_{pk} < 1$, then the average quality is within the specification limits but one part of the process is not.
- If $P_{pk} = 1$ then the tolerance variation equals to the process variation.
- If $P_{pk} > 1$ then the process variation is less than the tolerance variation.

PERCENTAGE OF DEFECTIVE RESULTS OUTSIDE SPECIFICATIONS

To determine the percentage of results that falls outside the specification limits, it is necessary to find how many estimated standard deviations exist between the overall average and each specification limit. The number of standard deviations is known as the Z value. Z values are used to determine the percentage of output that is outside the specification limits. This allows the conversion of the data distribution to a normal and standardized distribution while adding the probabilities of failure above the USL and below the LSL. The calculation is as follows:

$$(Above\ the\ USL)\ Z_{upper} = \frac{USL - m}{s}$$

and

$$(Below\ the\ LSL)\ Z_{lower} = \frac{m - LSL}{s}$$

By utilizing the normal distribution table, the estimated proportion of output that is above the upper specification and below the lower specification can be found. In order to convert this proportion to percentage, multiply it by 100. In this way, the percentage of defective products (based on the specification limits) can be estimated.

BENCHMARKING

Benchmarking is the process of comparing the cost, cycle time, productivity or quality of a specific product or method to another that is widely considered to be an industry standard or best practice. In project management benchmarking can also support the selection, planning and deliver of products. In the process of best practice benchmarking, someone can identify the best product or process of his industry/laboratory and compare all products and processes to those he identified as benchmarks. Essentially, benchmarking provides a snapshot of the business performance of a system and assists in understanding the level of the product in relation to a particular standard. The result is often a case for making changes in order to make improvements in the process and eventually in the quality of the products. Other important advantages of benchmarking are that it improves process effectiveness, helps in cost reduction and indicates the weak parameters of an overall measurement or production process.

RUGGEDNESS TEST

While conducting an analytical test method, it is vital to take under consideration that the environmental conditions under which the test is conducted affect importantly the final results and the general accuracy of the analytical measurement system. A ruggedness test should be performed so as to detect which experimental factors influence the results provided by the method and to determine how thoroughly these factors must be controlled. The ruggedness test should always precede the interlaboratory studies (which usually are the final tests for evaluating the precision of a test method) so as to restrict the allowable ranges of the critical variables of the test method, differently the precision statement provided by the round robin test will be inaccurate. A ruggedness test is conducted by making systematic changes in the variables associated with the test method and observing the size of the associated change in the test method results. The experimental designs that are most frequently employed in ruggedness testing are called “Plackett – Burman” designs. The advantage of these designs over other designs is based on the fact that they are particularly easy to implement and very efficient in improving a test method. These designs require the simultaneous change of the levels of all the experimental variables and subsequently they arbitrate the effect of each variable on the test results. For a two level “Plackett – Burman” design, the two levels are set so as not to be greatly different between each other and the effect of any variable in the test results is not importantly affected by changes in the level of the other variables. Ruggedness tests should be conducted within a single laboratory because it is easier to monitor the effects of the different experimental factors on the test results. Indicative variables that should be monitored during such a test are Temperature, Pressure, mass %, relative humidity, duration of a measurement, etc.

PLACKETT – BURMAN DESIGN

There are two basic assumptions made in ruggedness testing. The first assumption is that the simultaneous change of a number of variables can be described as the simple addition of the fixed effects for each variable. The second assumption is that the effect for each variable is not influenced by the effects of other variables. The effects derived based on this assumption are labeled as “main effects”. On the other hand, if strong dependence is observed between the effects of the variables, additional factors should be recognized and labeled as “interactions”.

A Plackett – Burman design requires N measurements where N must be an integer multiple of four. This kind of P – B designs are employed to estimate up to $N-1$ main effects but these main effects will be “polluted” by the interactions. A typical P – B design for $N = 8$ measurements, and therefore for 7 factors, is depicted below.

	Factor							Test Result
	A	B	C	D	E	F	G	
1	+	+	+	-	+	-	-	1.1
2	-	+	+	+	-	+	-	6.3
3	-	-	+	+	+	-	+	1.2
4	+	-	-	+	+	+	-	0.8
5	-	+	-	-	+	+	+	6.0
6	+	-	+	-	-	+	+	0.9
7	+	+	-	+	-	-	+	1.1
8	-	-	-	-	-	-	-	1.4

Image 32: Plackett – Burman design for N = 8.

Every column of the design has equal number of (+) and (-), where (+) for a factor declares that the measurement was taken with that factor set at the high level, whereas (-) declares that the factor was set at low level during the measurement. The four (+) and (-) of the experimental factor A, are directly associated with an equal number of (+) and (-) of the B factor, a fact that leads to the orthogonality of the main effects. With the term orthogonality it is implied that the A factor is not affected by the B factor. This orthogonality along with the forbearance of the fact that the main effects are polluted by the interactions, are the two main characteristics of all ruggedness tests.

PLACKETT – BURMAN Design Calculations

The calculations of a Plackett – Burman design include the following steps:

Calculation of the effect that each factor has on the analytical measurement, as the average of the measurements taken at high level minus the average of the measurements taken at low level. So for the random factor B:

$$Effect\ of\ B = \frac{\sum B(+)}{\frac{N}{2}} - \frac{\sum B(-)}{\frac{N}{2}}$$

Calculation of the standard deviation for an effect. So for the random factor B:

$$\sigma_{effect\ B} = \sqrt{(2/N)^2\ Variance[\sum B(+) - \sum B(-)]} = \frac{2s}{\sqrt{N}}$$

Here it is important to mention that since the standard deviation of an effect is inversely proportional to the number of measurements, it would seem like a good practice to construct large P – B designs. In practice though, smaller to moderate size P – B designs provide better results since they require less time to be completed and

the interactions in larger designs become more complicated. If more than 8 experimental factors have to be monitored, another P – B design can be employed.

Calculation of the statistical significance of an effect by utilizing a *t* test. So for *m* supplementary measurements (where *m* is not necessarily equal to *N*), for the random factor B:

$$t_{m-1} = \frac{\text{Effect of } B}{\frac{2s}{\sqrt{N}}}$$

If the statistical significance of the effect is important then the test method specification should be reconsidered.

In the case that the laboratory has conducted two or more ruggedness tests then the effects of each experimental factor should be investigated separately and, subsequently, an average effect of every factor and therefore an average standard deviation and statistical significance should also be estimated. Finally, and before proceeding to the interlaboratory study, a distinction between controllable and uncontrollable factors should be made. Controllable or fixed factors, as the name implies, are those experimental factors that can be directly controlled and be set at a specific high or low level. On the other hand, uncontrollable or random factors are those that cannot be directly controlled, such as the operator or the instruments used during the test. Uncontrollable factors cannot be efficiently processed by the ruggedness test design since the two levels required for the design cannot be specifically defined as high or low. During a round robin test, the imprecision that derives from uncontrollable factors can be measured and expressed with the repeatability standard deviation (*Sr*). If the combined imprecision of these factors is negligible, they should be treated as a single factor and their combined imprecision should be observed. A different approach to face this problem is to maintain all experimental factors constant and repeatedly sample one random factor. A final Plackett – Burman design for two sets of data is depicted below.

Factor	Level	First Data Set		Second Data Set		Difference, (d) between Effects
		Average	Effect	Average	Effect	
A	25	3013	...	3007
A	30	2972	+ 41	2959	+ 48	-7
B	Y	2992	...	2980
B	N	2993	-1	2987	-7	+ 6
C	Y	2996	...	2989
C	N	2990	+ 6	2978	+ 11	-5
D	1	3006	...	2990
D	3	2979	+ 27	2976	+ 14	+ 13
E	Y	3007	...	2995
E	N	2979	+ 28	2972	+ 23	+ 5
F	Y	3031	...	3026
F	N	2954	+ 77	2941	+ 85	-8
G	10	2992	...	2985
G	5	2993	-1	2982	+ 3	-4

Image 33: Plackett – Burman design for two sets of data.

INTERLABORATORY-ROUND ROBIN TEST

Analytical tests conducted on presumably same materials and in presumably identical conditions do not necessarily produce the same results. This fact is attributed to random errors inherent in all test procedures. In most of the cases, the factors that affect the final test result cannot be controlled, a fact that has to be taken under consideration when interpreting the test data. A variety of factors may contribute to the variability in application of a test method. These factors are: *a)* the operator, *b)* the equipment used, *c)* the calibration of the analytical system and *d)* the environmental conditions under which the measurement was taken (humidity, temperature, etc.). Generally, when the same analytical method is applied in different laboratories, all of the factors mentioned above are different. The general term for expressing the agreement between a test result value and an accepted reference value is accuracy. Ultimately, any analytical procedure should be tested both in terms of bias and in terms of precision. The practice that was implemented in this project is a procedure for determining the **precision** of an analytical test method.

When assessing a test method, precision is expressed in two terms, repeatability and reproducibility. Under repeatability conditions the four factors that were previously mentioned, and affect the result of the analytical measurement, are kept constant and their contribution to the variability is minimal. On the other hand, under reproducibility conditions these four factors differ and contribute importantly to the variability of the test result. Therefore, conducting a test under repeatability conditions is a practice implemented individually by each laboratory and it is not a recommended part of an interlaboratory test. The reproducibility measure indicates the level of precision expected when random portions of a homogeneous sample are sent to different laboratories. In order to obtain reasonable estimates of reproducibility precision, it is vital to avoid excessive sanitization of the data and to recognize and approach possibly erroneous results, which may have unacceptable assignable causes such as deviancy from the prescribed procedure. Important questions that arise when conducting an interlaboratory test and affect the confidence in the precision statements following the study, include how consistent are the data used in the study, which is the number of the laboratories involved, the number of test results per laboratory, as well as the number, range and type of the materials used in each study. Furthermore, in order to check the consistency of the data obtained in an interlaboratory test, the *k-value* statistic (within-laboratory precision) and the *h-value* statistic (laboratory to laboratory precision) should be implemented. The final objective of a round – robin test, after the calculation of the *h* and *k* statistics, is to frame a precision statement (*R*) related to the precision of the analytical measurement system used which, in cases of large scale tests, can be used as a benchmark for monitoring the precision of an individual system in a within laboratory test.

Generally, a complete interlaboratory test consists of three basic procedures. The planning of the interlaboratory study, the guiding of the phase of the study and finally, the calculation and visual display of the statistics related to the results provided by every laboratory. The evaluation of the data results is based on graphical and statistical tools which are applied to monitor for unusual and inconsistent results. This basic methodology is also implemented by the ASTM E 691 – 99 Standard (practice for conducting an interlaboratory study to determine the precision of a test method), a practice which was – partially – applied on this project, since the planning of the study and the guiding of the phase of the study were already performed by the laboratories that conducted the study. More specifically, the aim of the project was to frame a precision statement for a report of results of Engine Oil Lubricants (ASTM Committee D – 2 Interlaboratory Crosscheck Program) whose basic properties were analyzed by a number of laboratories. These properties include the mass (%) of Magnesium, Calcium, Phosphorus, Sulfur and Zinc, the density and the flash point of the lubricant sample. The sample material used was a monograde oil, therefore viscosity tests were not included. Since the report included only the final results of the study, only the third step of the procedure mentioned above was practically implemented through a Mat Lab code and the calculation of the precision statistics was depicted on relative tables and graphs. Furthermore, there is also a divergence from the standard ASTM practice, related to the repeatability (within laboratory) deviation – and by extension to the *k value* statistic of the results which was zero – since the report contained only a single final value as a test result from every laboratory. Following, the calculation of the statistics and the arrangement of the results into tables and graphs is described.

CALCULATION AND DISPLAY OF STATISTICS

The calculation of the cell statistics includes the following steps:

Calculation of the cell average \bar{x} for each laboratory as follows:

$$\bar{x} = \sum_{i=1}^n \frac{x_i}{n}$$

where x is the test result in one cell and n is the number of test results in one cell.

Calculation of the cell standard deviation S as follows:

$$S = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{(n - 1)}}$$

Calculation of the Average of the Cell Averages X_{avg} :

$$X_{avg} = \sum_{1}^P \frac{x_{avg}}{P}$$

where x_{avg} is the cell average and P is the number of laboratories involved.

Calculation of the Cell Deviation d :

$$d = x_{avg} - X_{avg}$$

Calculation of the Standard Deviation of the Cell Averages Sx_{avg} :

$$Sx_{avg} = \sqrt{\frac{\sum_1^P d^2}{(P-1)}}$$

Calculation of the Repeatability Standard Deviation Sr :

$$Sr = \sqrt{\frac{\sum_1^P s^2}{P}}$$

Calculation of the Reproducibility Standard Deviation SR :

$$SR = \sqrt{(Sx_{avg})^2 + (Sr)^2 * \frac{(n-1)}{n}}$$

Finally, after performing the above calculations, the consistency h and k statistics should be estimated and the relative graphs should be constructed so as to graphically evaluate the differences between the laboratories and monitor for inconsistent results. The flagging of the inconsistent results is based on critical values attributed to h and k statistics at the 0.5% confidence level (which crudely corresponds to the 3s range of the Student's test). The h critical values depend on the number of the laboratories participating in the study, while the k critical values depend on the number of the test results, per material, that each laboratory provides. More specifically, these critical values are obtained after performing a Student's – t test and an F – ratio test on the data results.

The h (between laboratory consistency statistic) and the k (within laboratory consistency statistic) are calculated as:

$$h = \frac{d}{Sx_{avg}}$$

and

$$k = \frac{S}{Sr}$$

THE h AND THE k GRAPHS

As far as the h graph (laboratory to laboratory precision) is concerned there are three general patterns for these plots. In the first one, all laboratories provide both positive and negative values. In the second one, the number of positive laboratories equals to the number of negative laboratories. Both of the aforementioned patterns are usual and do not require further investigation. In the last pattern, one laboratory provides positive (or negative) h values which are opposed to the negative (or positive) values provided by the other laboratories. Obviously, this is an unusual pattern and the accuracy of the measurement performed by this laboratory should be questioned.

In the k graph, there is one predominant pattern in which one laboratory provides very large or very small k values for all the materials. Very high k values indicate within laboratory imprecision while very small values indicate a very insensitive measurement scale.

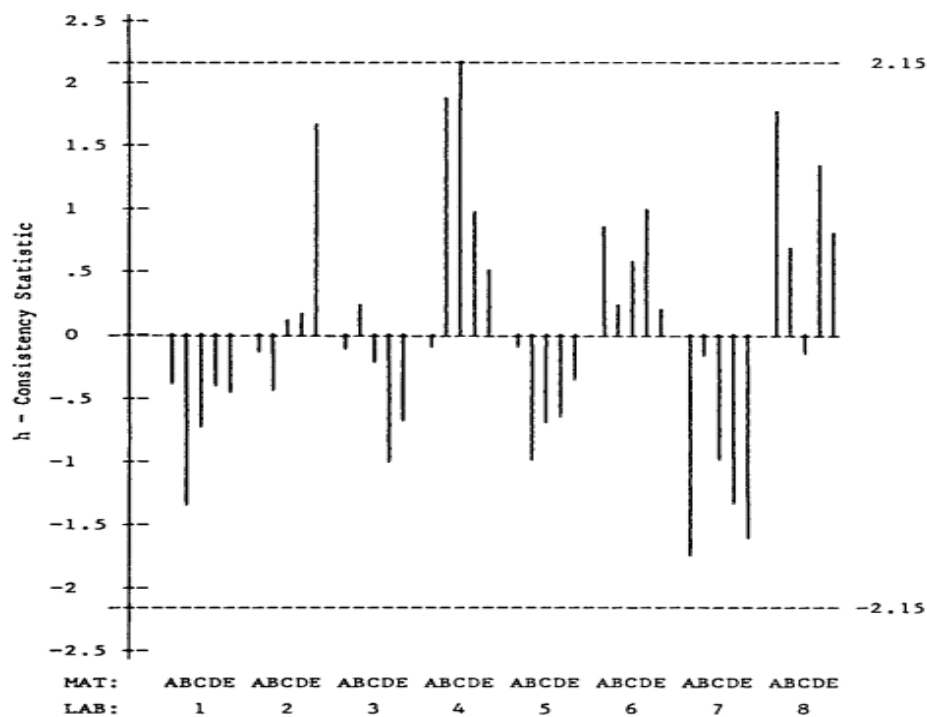


Image 34: A typical h graph – Materials within Laboratories.

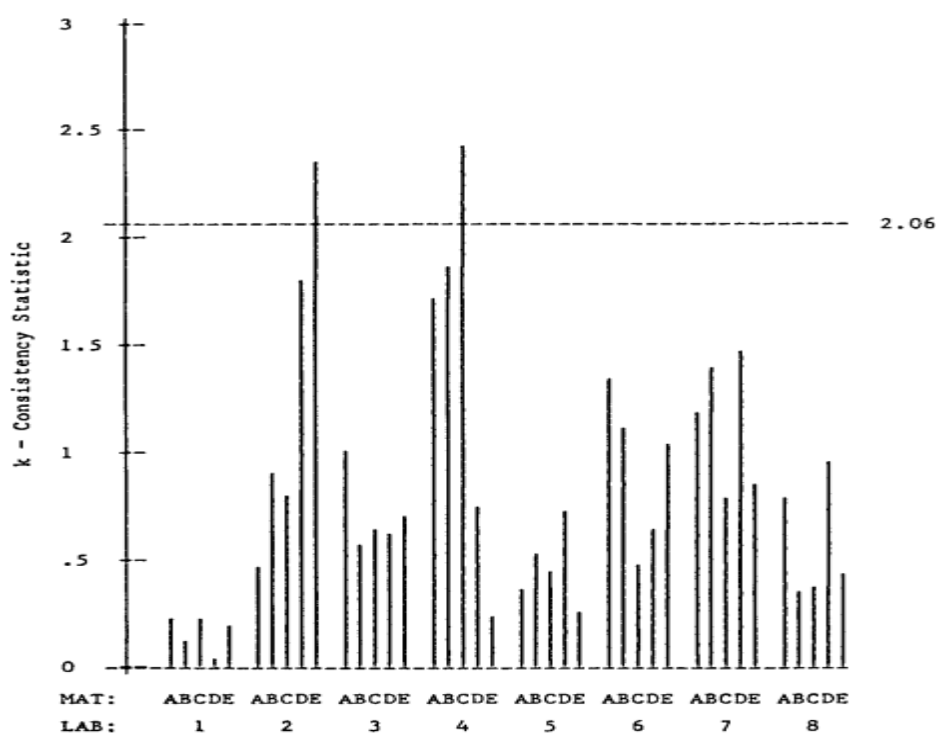


Image 35: A typical k graph – Materials within Laboratories.

INVESTIGATION OF INCONSISTENT RESULTS

After having constructed the h and k graphs, an examination should be conducted so as to locate the test results that exceed the critical values but also those that deviate significantly from the others. Both deviations should be explained with proper reasoning. Usually, inconsistent results are due to either sampling and clerical or procedural errors. The sampling and clerical errors include errors in the calculations, prematurely rounded data and possible mislabeling of the test results of one or more materials. Procedural errors are associated with any deviation from the protocol or the standard test method that should have been implemented. Big quality difference in the laboratory equipment and shifts in the number of significant digits are the two main reasons that usually lead to procedural errors. If the investigation reveals no errors, then the inconsistent results should be retained and the precision statistics related with them should be published. On the contrary, if a reason for the deviation was determined, then a series of options should be considered as possible solutions. First of all, if the deviation between a test result provided by a laboratory and the values provided by the rest is illogically large, then the value can be neglected and deleted from the study or, alternatively, the laboratory that provided it can be asked to repeat the measurement. Moreover, if the interlaboratory study includes a large number of laboratories and if the investigation stage does not reveal any obvious cause for some unusual cell values,

then this cell must also be removed from the study. Here it is important to mention that any action that will ultimately lead to more than 5% discarding of the test results will ultimately affect the precision statement of the study. Loss of precision may also be a result of wide range interpretation due to opacity of the standard practice method but also due to poor maintenance practices performed by some laboratories.

PRECISION STATEMENT

After the investigation stage has been concluded and the inconsistent results have been either discarded or reintroduced into the study (the laboratory that provided an inconsistent result repeated the measurement), the statistical procedure already described has to be re-implemented and produce the “corrected” statistics. With the corrected statistics, the final statement that summarizes the precision of the test method (based on the repeatability and reproducibility limits) is given. The repeatability (within laboratory) and reproducibility (laboratory to laboratory) limits, for 95% probability, are calculated as follows:

$$r = 2.8 S_r$$

$$R = 2.8 S_R$$

Concluding, the corrected precision statistics are organized and presented into a table as shown below.

Material	\bar{x}	$s_{\bar{x}}$	s_r	s_R	r	R
A	0.4048	0.1131	0.0150	0.1137	0.04	0.32
B	0.8841	0.0447	0.0322	0.0519	0.09	0.14
C	1.1281	0.1571	0.1429	0.1957	0.40	0.55
D	1.2686	0.0676	0.0375	0.0742	0.11	0.21
E	1.9809	0.0538	0.0396	0.0628	0.11	0.18
F	4.1814	0.2071	0.0325	0.2088	0.09	0.58
G	5.1843	0.2172	0.1330	0.2428	0.37	0.68
H	10.4010	0.5630	0.1936	0.5848	0.54	1.64
I	16.3610	1.0901	0.2156	1.1042	0.60	3.09

Image 36: Presentation of the corrected precision statistics of an ILS study.

CHAPTER 3 – EVALUATION OF THE ROCK – EVAL AND GC – MS SYSTEMS AND INTERLABORATORY STUDY

The statistical techniques described in Chapter 2 were incorporated in a Mat lab code which was implemented for the evaluation of a series of data results produced by a Rock – Eval pyrolysis apparatus and a GC – MS apparatus. Subsequently, the precision and consistency statistics also described in Chapter 2 were applied on a series of data results of properties coming from an interlaboratory, round – robin study, for a sample of a monograde oil, and after the construction of the required graphs, a precision statement about the different analytical methods was formulated. Relative MATLAB codes were written for both tasks. Since this study was a “historical” overview of the performance of the two systems, an assignable cause for the possible outliers could not be determined since the required information was missing. The purpose of the statistical testing was to determine how often they appear. The statistical procedure described above was implemented in a series of S1, S2, S3, TOC and T_{MAX} results for the Rock – Eval apparatus while for the GC – MS, the code that was written in MATLAB first calculated the Relative Response Factors of the different compounds of different samples and then implemented the statistical procedures described, with them as an input. The relative response factors were evaluated instead of the standard concentrations so as to prove the usefulness of keeping a common calibration equation or the need of adopting a new one for each study. Once again, for the evaluation of both systems it will be considered that the process goes out – of – control when there are outliers in the control charts and when there are major departures from the normality at the 95% confidence level. Given that the outliers could not be associated with normal or non - assignable sources of variation, it will be considered that every outlier causes instability in the process and leads to an out – of – control state. The relative results are presented below.

ROCKEVAL PERFORMANCE EVALUATION

RESULTS AND DISCUSSION FOR TOC

Anderson - Darling A2* Value	0.2926	
Student's t(a) Value	7.2893	
MRavg	0.0264	
Upper Limit for I Chart	0.0372	
Lower Limit for I Chart	-0.1181	

Upper Limit for EWMA Overlay	-0.0021	
Lower Limit for EWMA Overlay	-0.0792	
Upper Limit for MR Chart	0.0862	
m / s (Levey – Jennings)	3.2392	0.0290
m+S / m-S (Levey – Jennings)	3.2683	3.2102
m+2S / m-2S (Levey – Jennings)	3.2973	3.1812
m+3S / m-3S (Levey – Jennings)	3.3264	3.1522
F Value	1.0795	
MRpooled	0.0274	
Site Precision R' for TOC	0.0649	
X ² Value	-	
Cp / Cpk	-	-
Pp / Ppk	1.5789	1.1198
Zupper / Zlower	6.1143	3.3593

Table 1: TOC results.

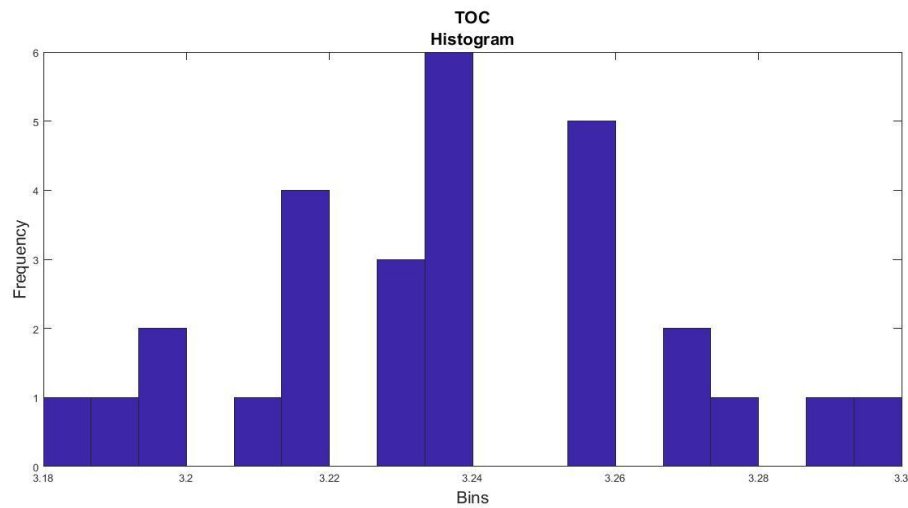


Figure 1: Histogram of the TOC results.

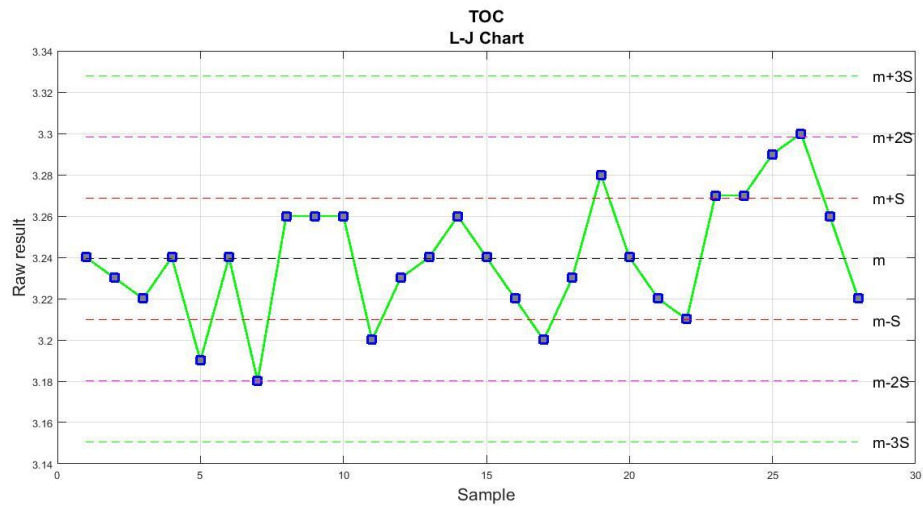


Figure 2: Levey – Jennings of the TOC results.

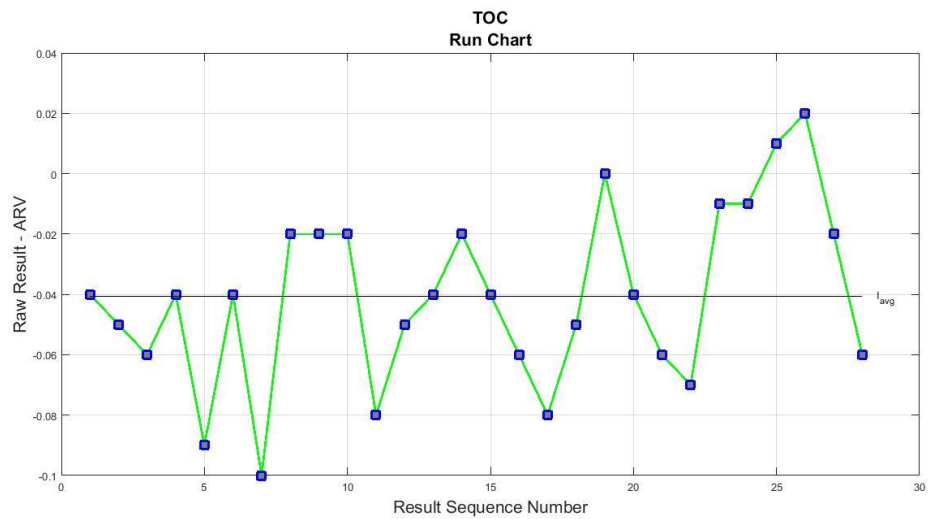


Figure 3: Run Chart of the TOC results.

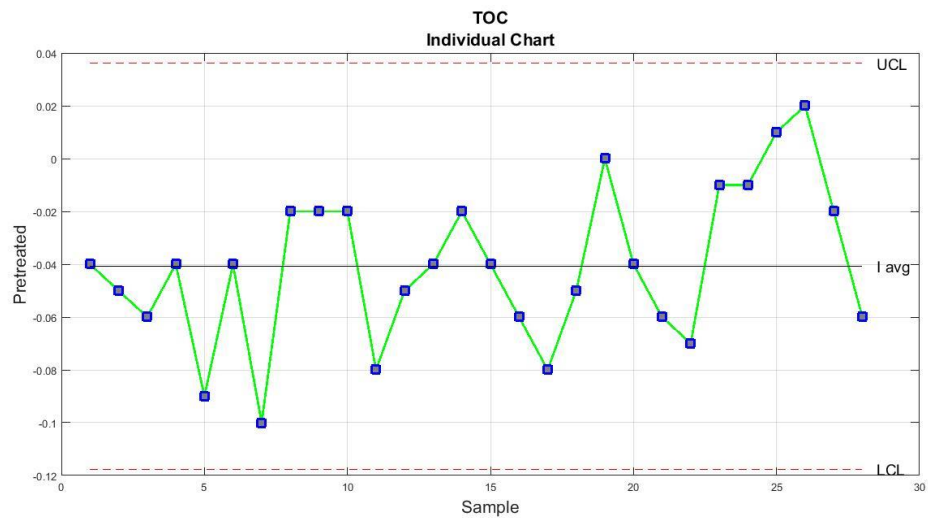


Figure 4: Individual Chart of the TOC results.

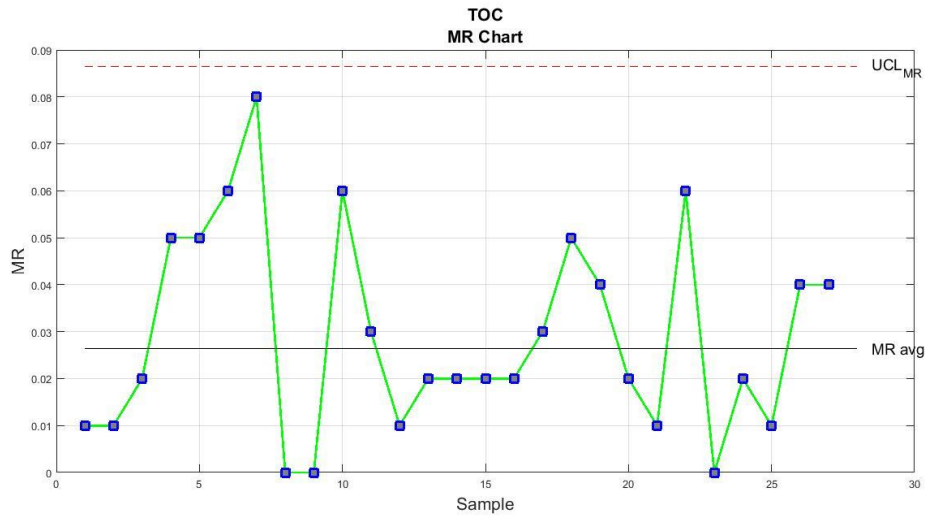


Figure 5: MR Chart of the TOC results.

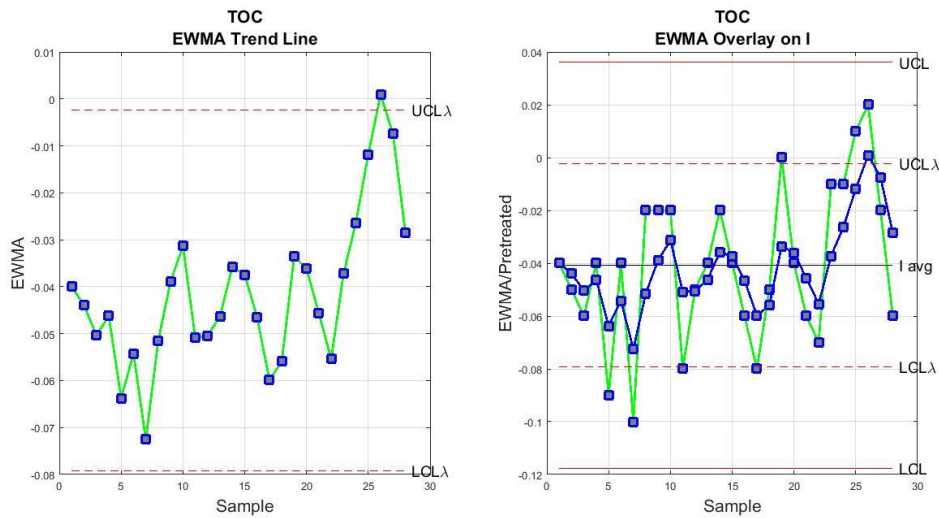


Figure 6: EWMA trend line and EWMA Overlay on the Individual Chart.

The histogram of the TOC results provides a rough idea about the distribution of the data results which indeed seems to approximate the Gaussian. The Anderson Darling test confirmed the initial visual observation made by the Histogram about the distribution since it resulted in an A_2^* value of 0.2926 and therefore smaller than 0.752, which is the marginal value for the normality hypothesis at the 95% confidence level. In the Levey – Jennings control chart, no data point is located above or below the $m \pm 3s$ limits respectively, a fact that indicates that no unacceptable random or large systematic errors are present in the system. As far as the Westgard rules are concerned, 2 violations were detected by the program. Initially, violation of the 12_s rule for data points 7 and 26, and 41_s rule violation for data points 23, 24, 25 and 26. The first violation consists a warning rule and was accepted as a random error. The second violation is related to smaller systematic errors which are usually insignificant and can be addressed by performing

maintenance or calibration to the analytical measurement system. By observing the Run Chart and the Individual Chart of the TOC results no unusual patterns and no special causes (shifts or trends) beyond the usual variability of the process can be discerned. Additionally, no outliers can be visually screened since all the preprocessed results fall within the upper and lower limits of the Individual chart. In the Moving Range control chart all the sequential points are located below the upper limit of the chart, a fact that indicates relative precision and that the process variation is in control. The EWMA trend line though indicates that data point 26 results above the upper limit of the chart and it is connected with a upward shift in the process that begins from point 23. With its overlay on the Individual Chart, some smaller shifts can also be detected (points 5, 7, 19, 25 and 26) but this kind of shifts do not affect the general stability of the system importantly since, in most cases, they are corrected internally. The Student's test resulted in an absolute t (a) value of 7.2893 which is greater than 2.1448 (the critical t value for $(n-1)/2$ degrees of freedom). This indicates that the mean of the sample is statistically distinguishable from the hypothesized mean value, with 95% confidence, and therefore there is statistically identifiable bias in the system relative to the ARV. The observation made from the Student's test is logical if someone considers that the test results had a mean equal to 3.2392 and the ARV of the standard was equal to 3.28. Subsequently, for the purpose of conducting an F test, the MRavg of these TOC results was compared with the MRavg of the TOC results of another set of data so as to compare the variation exhibited by the analytical measurement system over two different time periods. The F test resulted in an F value of 1.0795 which is smaller than the critical F value of 3.11 (for a population of this size), a fact that reveals good agreement between the results of the two different time periods and that the two precisions are statistically undistinguishable. The two different MR's (0.0264 and 0.0285) were combined into a single MR_{pooled} which was equal to 0.0274. The site precision concerning the TOC results was equal to 0.0649. Since the EWMA chart revealed an outlier (point 26), the process is considered to be out – of – control and therefore the Process Capability indices are not reliable. The Process performance indices can be safely calculated though, since they can be used even when the process is not in state of control. P_p resulted in a value of 1.5789 and P_k was equal to 1.1198, something which indicates that the process variation is less than the tolerance variation. Subsequently, Z_{upper} was equal to 6.1143, something which denotes that virtually 0% of the past results ended above the upper certification limit, while Z_{lower} was equal to 3.3593, something which denotes 0.39% of the overall past results were below the specification limit. No published reproducibility value exists in the literature for the TOC thus the Chi – square test could not be directly applied by the code written in Matlab.

RESULTS AND DISCUSSION FOR TMAX

Anderson - Darling A2* Value	0.7087	
Student's t(a) Value	0.1200	
MRavg	1.2857	
Upper Limit for I Chart	3.4292	
Lower Limit for I Chart	-3.4107	
Upper Limit for EWMA Overlay	1.7291	
Lower Limit for EWMA Overlay	-1.6998	
Upper Limit for MR Chart	4.2042	
m / s (Levey – Jennings)	416.0357	1.5465
m+S / m-S (Levey – Jennings)	417.5822	414.4892
m+2S / m-2S (Levey – Jennings)	419.1287	412.9428
m+3S / m-3S (Levey – Jennings)	420.6751	411.3963
F Value	1.1367	
MRpooled	1.3712	
Site Precision R' for T _{MAX}	3.1628	
X ² Value	-	
Cp / Cpk	-	-
Pp / Ppk	0.4233	0.4158
Zupper / Zlower	1.2473	1.2926

Table 2: T_{MAX} results.

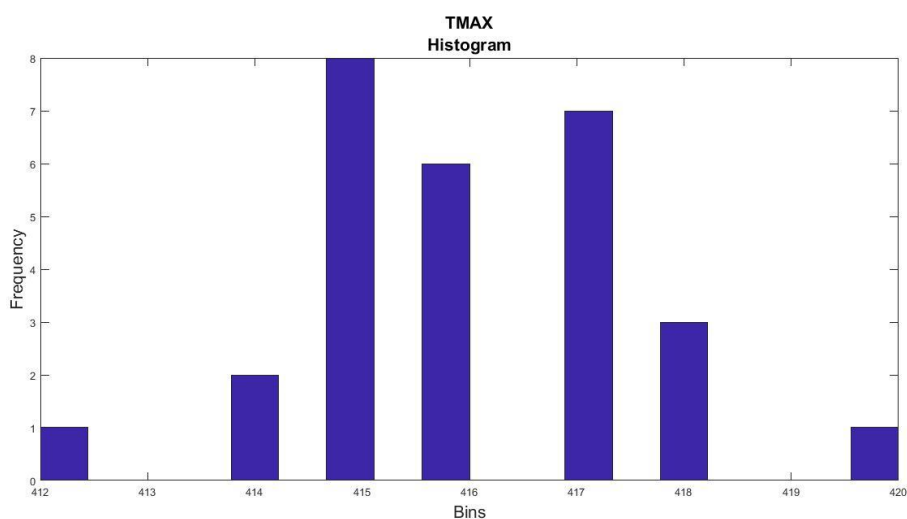


Figure 7: Histogram of the TMAX results.

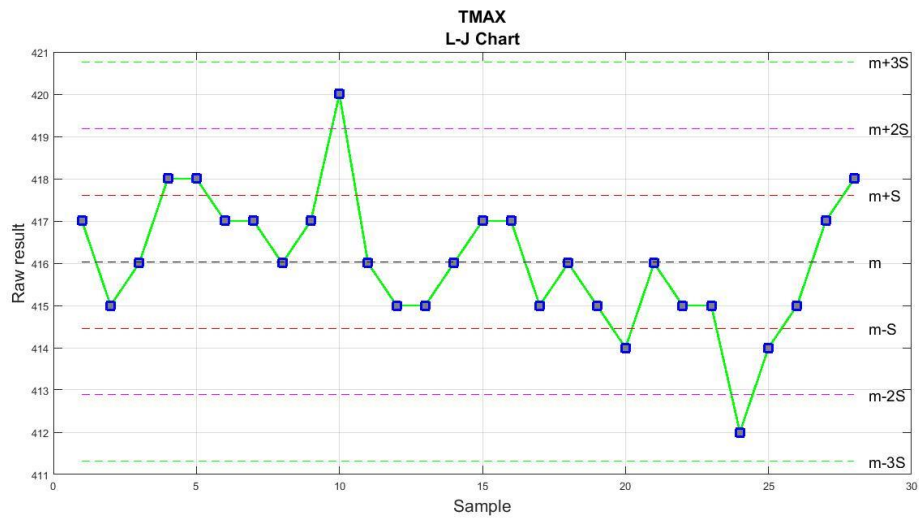


Figure 8: Levey – Jennings of the TMAX results.

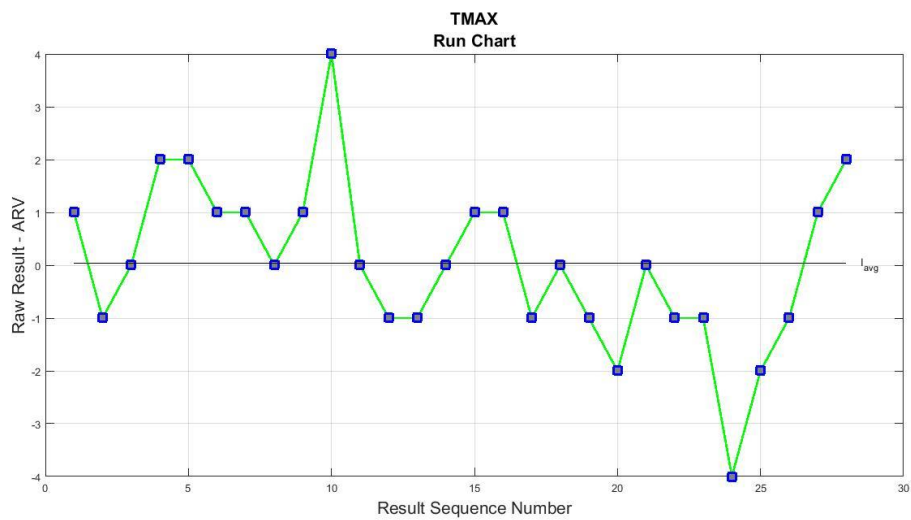


Figure 9: Run Chart of the TMAX results.

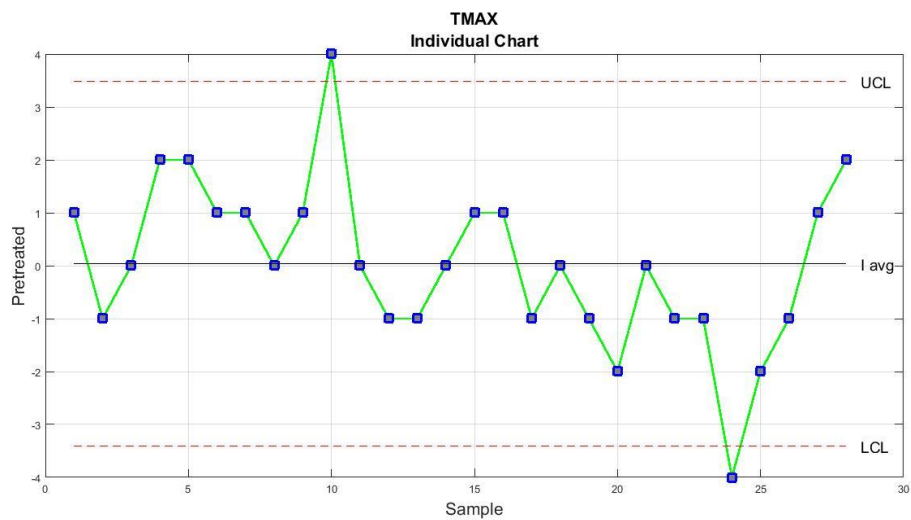


Figure 10: Individual Chart of the TMAX results.

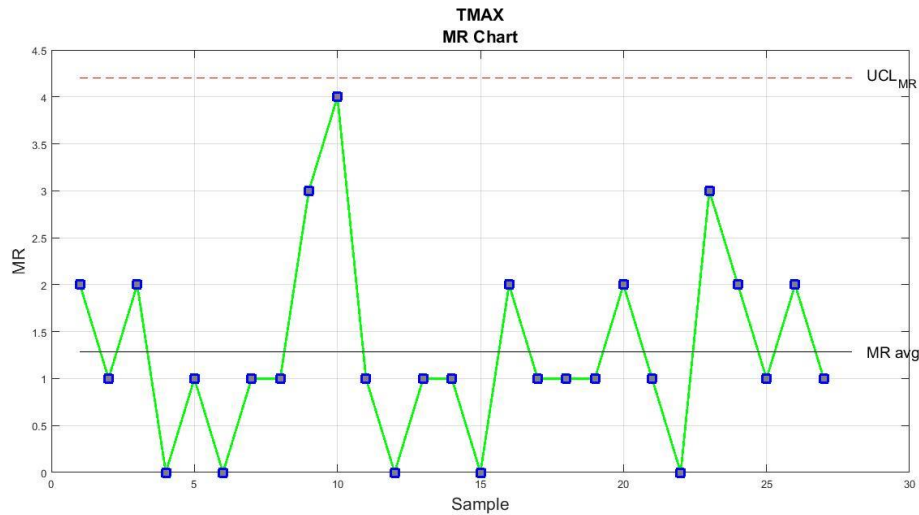


Figure 11: MR Chart of the TMAX results.

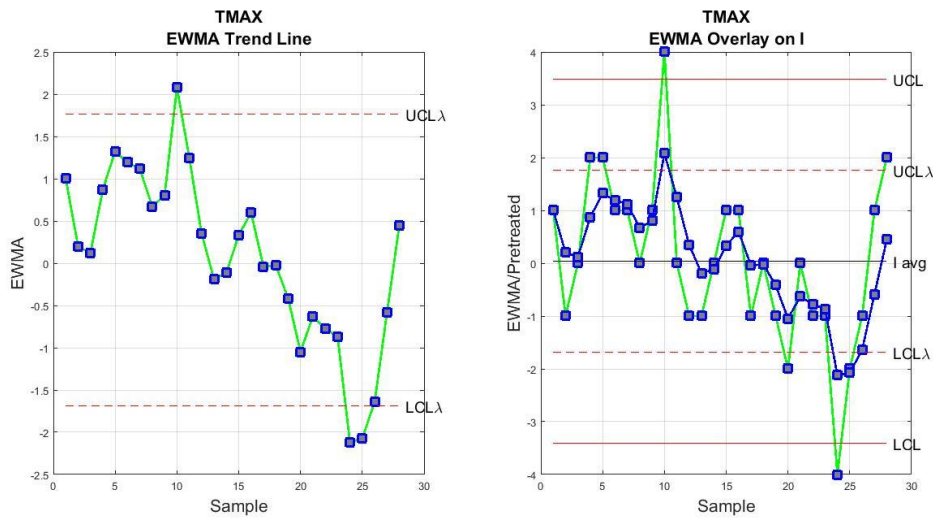


Figure 12: EWMA trend line and EWMA Overlay on the Individual Chart.

The histogram of the TMAX gives a visual indication that the Gaussian distribution adequately describes the results. The normality assumption was accepted with 95% confidence as the Anderson – Darling test resulted in an A2* value that was equal to 0.7087 and therefore lower than 0.752, the marginal normality hypothesis value. As far as the Levey – Jennings chart and the Westgard rules are concerned, the program detected 2 violations. Points 10 and 24 violate the 12_s rule, since they both exceed the $m \pm 2S$ range limit, and points from 17 to 26 violate the $10x$ rule. The first violation was accepted as a random error while the second violation indicates smaller systematic error. By observing the Run Chart of the T_{MAX} results, the 2 points that were mentioned above seem to exhibit a certain discrepancy related to the other points. Individual Chart confirms this observation since both data points fall above and below the upper and lower limits of the chart and they cause an upward and downward shift respectively. As was explained in chapter 2, individual values

that fall outside the control limits of the I Chart are indications of instability in the process and denote that the system goes out of control. The Moving Range Chart divulges that no points exceeded its upper limit. The EWMA trend line confirms that points 10 and 24 cause large upward and downward shifts respectively. Furthermore, the effect of point 24 on point 25 can also be observed, since it “attracts” point 25 below the lower limit of the chart, a fact that is logical since the calculation of the EWMA values is based upon the recursion equation that was described in chapter 2. The EWMA overlay on the Individual Chart reveals some minor shifts as well (points 4, 5, 20, 25) but as it was explained above, they do not affect the process importantly. The Student’s test resulted in a $t(a)$ value of 0.1200 and since it was lower than the critical published t value, which for a sample of this size equals to 2.1448, the mean of the distribution of the T_{MAX} values is less than the hypothesized mean value and there is no bias relative to the ARV. The F test resulted in an F ratio value of 1.1367 which is smaller than the critical F value of 3.11, a fact that reveals good agreement between the results of the two different time periods and that the two precisions are statistically undistinguishable, a fact that further advocates to the general stability and repeatability of the values. The two different MR_{avg} (1.2857 and 1.4615) were combined into a single “pooled” MR which was equal to 1.3712. Concluding, the site precision concerning the T_{MAX} results was equal to 3.1628. Since there are outliers, the process is not in a state – of – control and therefore the Capability indices are not reliable. As far as the performance indices are concerned P_p was equal to 0.4233 and P_k was equal to 0.4158 something which indicates that the average quality is within the specification limits but one part of it is not. Concluding, Z_{upper} was equal to 1.2473 and Z_{lower} was equal to 1.2926. They both indicate a large amount of deficiency and more specifically 10.56% above the USL and 9.85% below the LSL, as far as the history of the system is concerned. No published reproducibility value exists in the literature for the TMAX thus the Chi – square test could not be directly applied by the code written in Mat lab.

RESULTS AND DISCUSSION FOR S1

Anderson - Darling A2* Value	0.8255	
Student's $t(a)$ Value	22.6630	
MRavg	0.0157	
Upper Limit for I Chart	0.0861	
Lower Limit for I Chart	0.0168	
Upper Limit for EWMA Overlay	0.0682	
Lower Limit for EWMA Overlay	0.0338	
Upper Limit for MR Chart	0.0467	

m / s (Levey – Jennings)	0.1907	0.0118
m+S / m-S (Levey – Jennings)	0.2025	0.1789
m+2S / m-2S (Levey – Jennings)	0.2143	0.1671
m+3S / m-3S (Levey – Jennings)	0.2261	0.1553
F Value	1.0194	
MRpooled	0.0155	
Site precision R' of S1	0.0386	
X ² Value	-	
Cp / Cpk	-	-
Pp / Ppk	0.1971	-1.2306
Zupper / Zlower	-3.6917	4.8741

Table 3: S1 results.

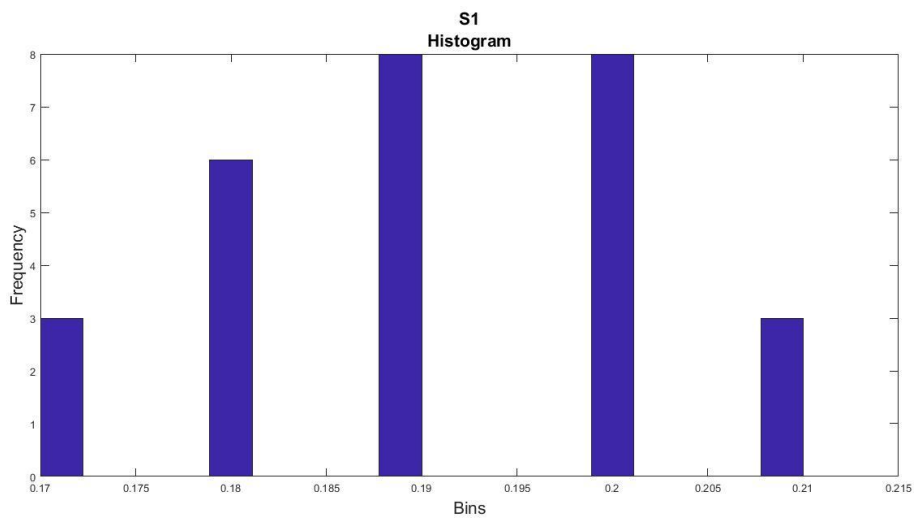


Figure 13: Histogram of the S1 results.

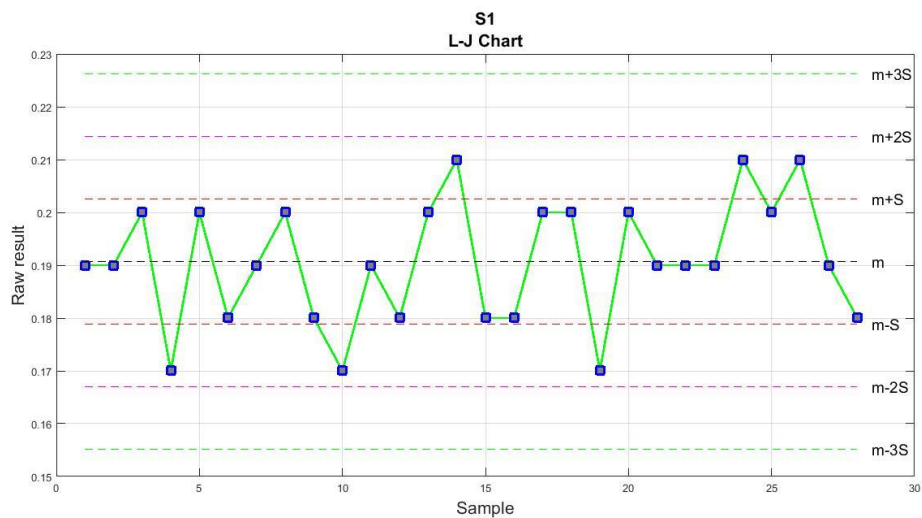


Figure 14: Levey – Jennings Chart of the S1 results.

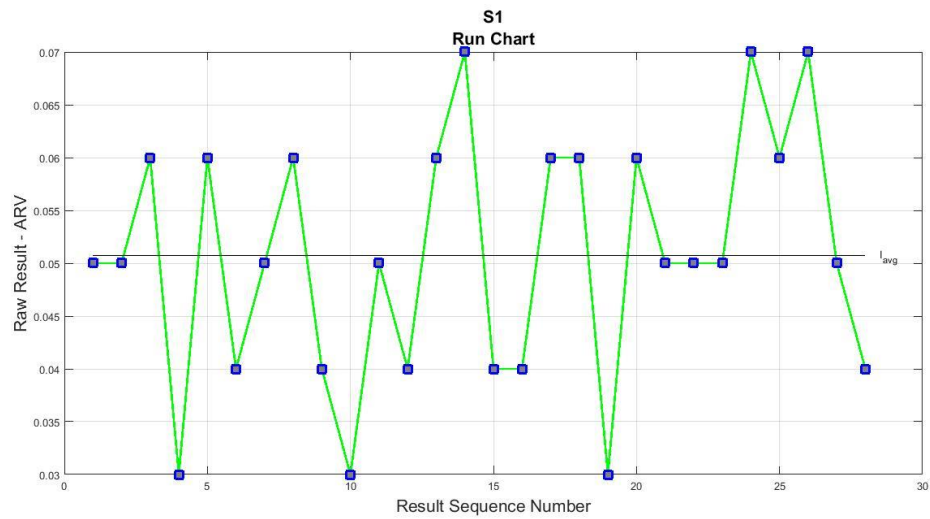


Figure 15: Run Chart of the S1 results.

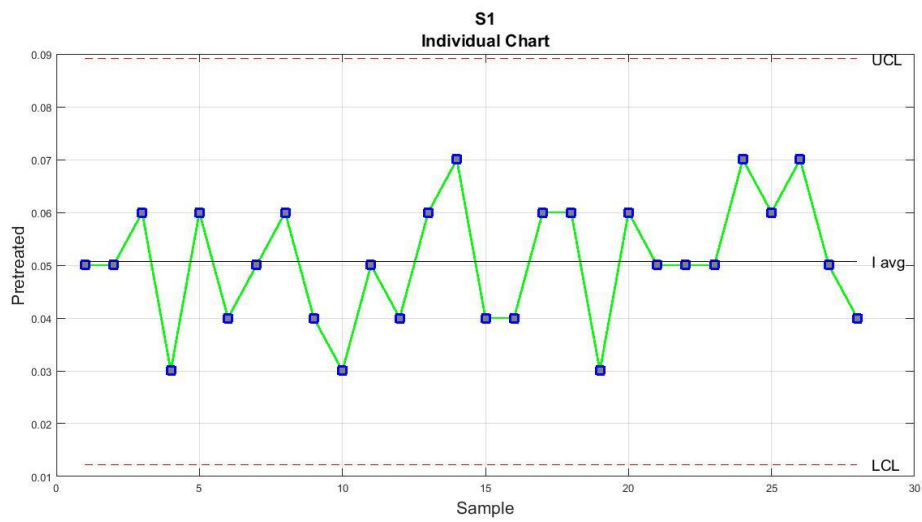


Figure 16: Individual Chart of the S1 results.

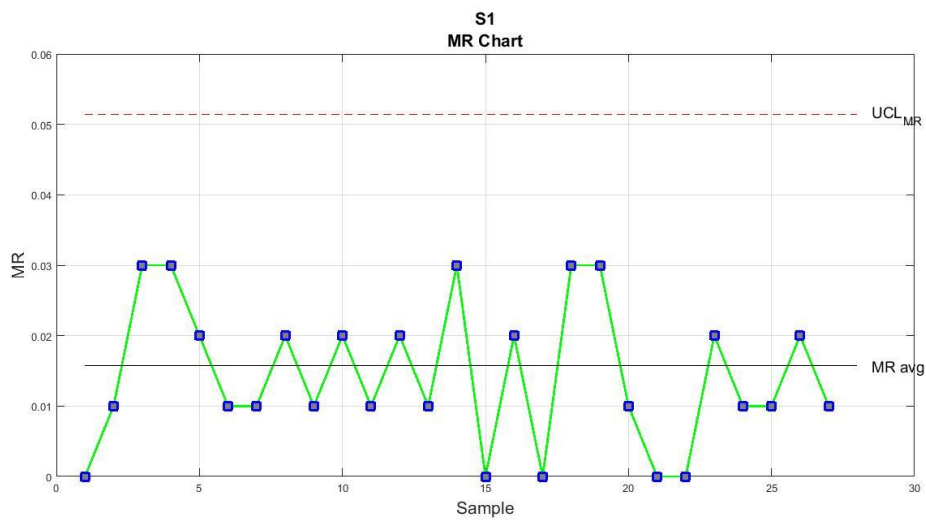


Figure 17: MR Chart of the S1 results.

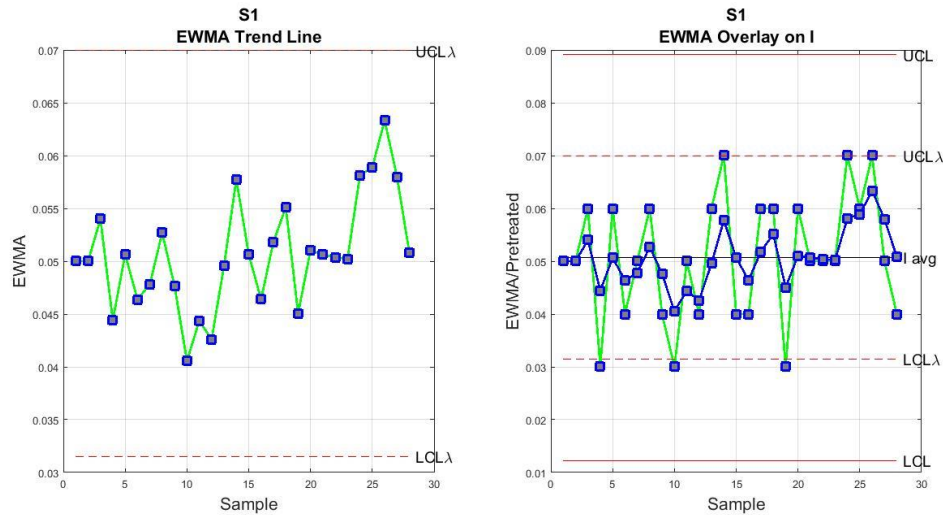


Figure 18: EWMA trend line and EWMA Overlay of S1.

The histogram of the S1 results gives the impression that the normal distribution is slightly skewed to the right. Indeed, the Anderson – Darling test resulted in an $A2^*$ value of 0.8295 and therefore the normality assumption is rejected at the 95 % confidence level. It should be mentioned here that the degree of divergence from the normality causes an analogous degree of divergence in the accuracy of the results of the parametric tests from their true value. In this case, the error in the results of the parametric tests was considered acceptable, and will be taken under consideration in the evaluation of the S1 results. Given that the normality assumption was rejected, the interpretation of the Levey – Jennings chart does not lead to reliable conclusions and it is omitted. By observing the Run Chart and the Individual Chart of the S1 results, no unusual patterns and no special causes (shifts or trends) beyond the usual variability of the process can be discerned, as well as no point with obvious discrepancy related to the others. In the individual chart, all data points fall within the upper and lower limits, a fact that reveals stability in the process and repeatability between the values that the system produces. On the Moving Range chart also no point exceeds the upper limit a fact that further advocates in the precision of the system and the repeatability of the values. The EWMA overlay on the Individual chart reveals some small process shifts (points 4, 10, and 19). The Student's test indicated that a bias relative to the ARV is present in the system and that the mean of the samples is statistically distinguishable from the hypothesized mean value. The calculated $t(a)$ value equals to 22.6630 and is significantly greater than 2.1448. This result is logical if someone considers that the mean of the data results was equal to 0.1907 and the ARV was equal to 0.14. The F test resulted in an F value of 1.0194 which is smaller than the critical F value of 3.16, a fact that reveals good agreement between the results of the two different time periods and furthermore advocates in the stability of the system. The two MR_{avg} (0.0157 and 0.0154) were pooled into a single one that was equal to 0.01557.

Concluding, the site precision concerning the S1 results was equal to 0.0386. Given that the normality hypothesis was rejected, the calculation of the process capability indices once again leads to unreliable results and therefore their values are not included. Furthermore, P_k index results in a negative value equal to -1.2306, something which indicates that the process falls outside the specification limits and produces a large amount of defective results. No published reproducibility value exists in the literature for the S1 thus the Chi – square test could not be directly applied by the code written in Mat lab.

RESULTS AND DISCUSSION FOR S2

Anderson - Darling A2* Value	0.3088	
Student's t(a) Value	1.4008	
MRavg	0.1929	
Upper Limit for I Chart	0.5711	
Lower Limit for I Chart	-0.4551	
Upper Limit for EWMA Overlay	0.3045	
Lower Limit for EWMA Overlay	-0.1985	
Upper Limit for MR Chart	0.6307	
m / s (Levey – Jennings)	12.4800	0.1854
m+S / m-S (Levey – Jennings)	12.6654	12.2945
m+2S / m-2S (Levey – Jennings)	12.8509	12.1090
m+3S / m-3S (Levey – Jennings)	13.0364	11.9236
F Value	1.2701	
MRpooled	0.2208	
Site Precision R' for S2	0.4745	
X ² Value	-	
Cp / Cpk	0.9748	0.8773
Pp / Ppk	0.8824	0.7942
Zupper / Zlower	2.3825	2.9120

Table 4: S2 results.

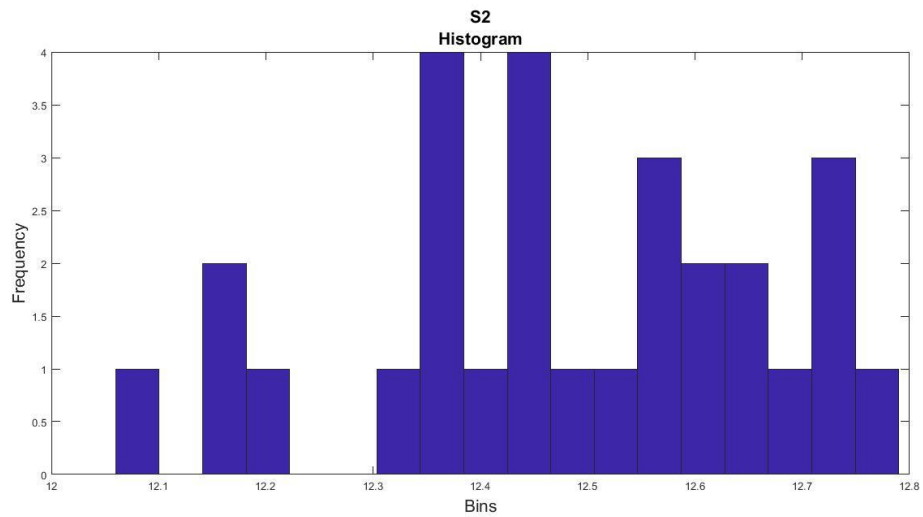


Figure 19: Histogram of the S2 results.

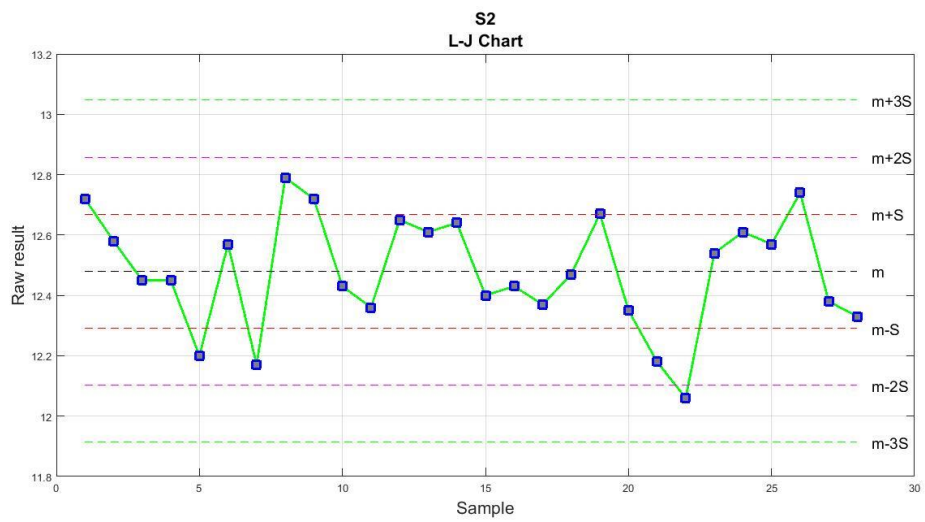


Figure 20: Levey – Jennings Chart of the S2 results.

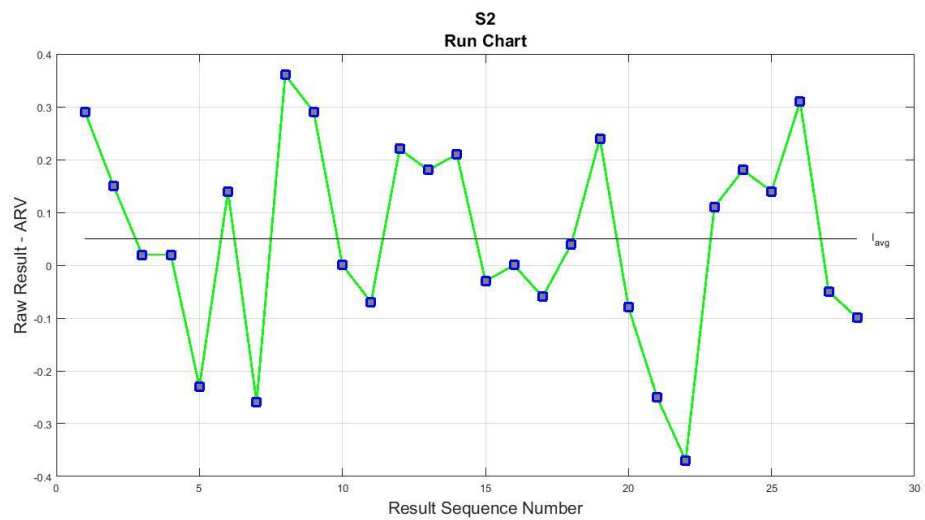


Figure 21: Run Chart of the S2 results.

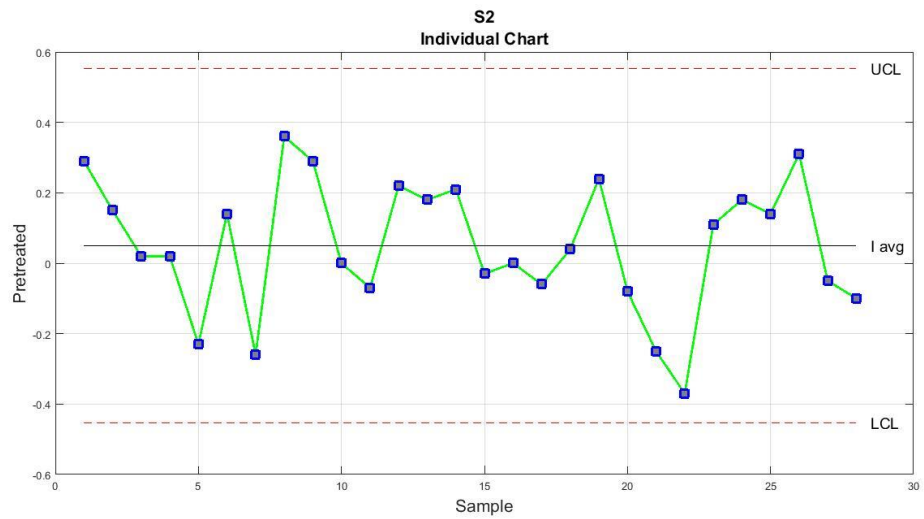


Figure 22: Individual Chart of the S2 results.

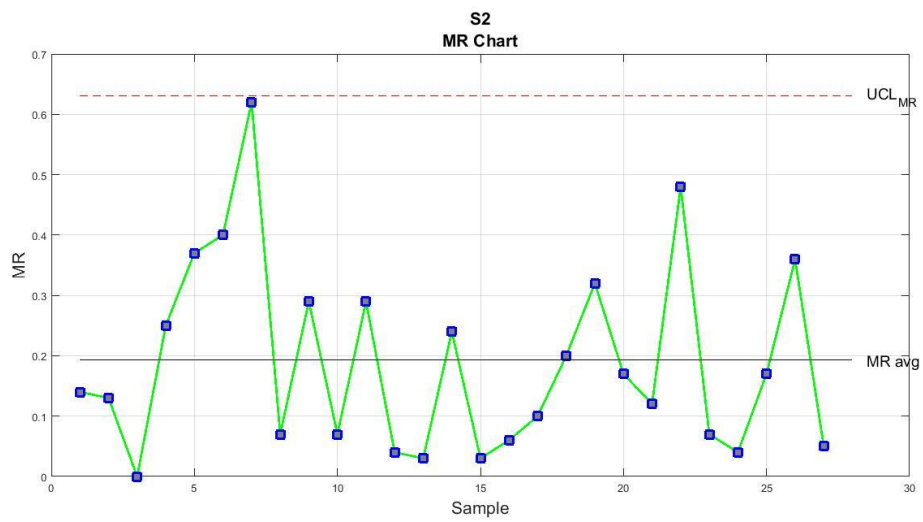


Figure 23: MR Chart of the S2 results.

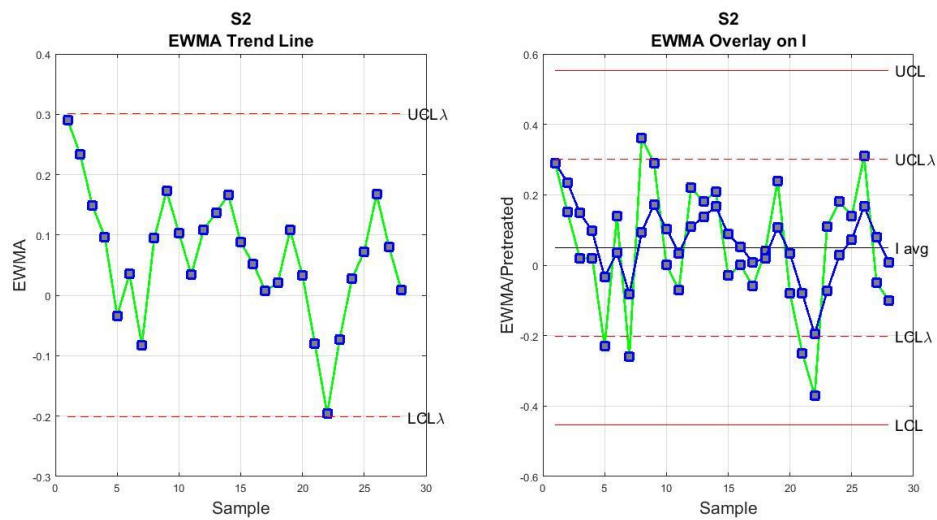


Figure 24: EWMA trend line and EWMA Overlay on the Individual Chart.

As was the case with the previous test results, here also the histogram indicates normal distribution for the data, a hypothesis that is confirmed at the 95% confidence level by the Anderson – Darling test which resulted in an $A2^*$ value of 0.3088 and therefore it was smaller than the marginal 0.752 value. As far as the Levey – Jennings chart and its evaluation by the Westgard rules is concerned, the program indicated violation of the 12_s rule by point 22 since it results below the $m - 2S$ range limit. No outward reason could be determined for this violation and thus it was accepted as a random error. By observing the Run Chart and the Individual Chart of the S2 results, no unusual patterns and no special causes (shifts or trends) beyond the usual variability of the process can be discerned. In the Moving Range Chart, again no points exceeded its upper limit, an indication of system stability as well. The enhanced sensitivity of the EWMA Overlay on the Individual chart indicated that there are some smaller shifts present in the process (points 5, 7, 8, 21, 22 and 26). Additionally, no bias relative to the ARV was detected by the Student's test since the calculated $t(a)$ value was equal to 1.4008 and subsequently smaller than 2.1448. This fact reveals that the mean of the samples is statistically undistinguishable from the hypothesized mean value and no bias relative to the ARV is present in the system. The F test resulted in a value of 1.2701, smaller than 3.16, and as was the case with the previous measurements, the two different MR_{avg} (0.1929 and 0.2423) were pooled into a single MR with a value of 0.2208. Concluding, the site precision concerning the S2 results was equal to 0.4745. Since the normality assumption was accepted by the Anderson – Darling test and no outlier was identified by the control charts, the process is considered in a state of statistical control and therefore both process capability and process performance indices are reliable. C_p resulted in a value of 0.9748 and C_{pk} in a value of 0.8773, a fact which denotes that the average quality is within the specification limits. P_p and P_{pk} resulted in 0.8824 and 0.7942 respectively. As far as the percentage of the defective results is concerned Z_{upper} was equal to 2.3825, something which corresponds to 0.87% of past results above the USL, and Z_{lower} was equal to 2.9120 which corresponds to 0.18% of past results below the LSL. No published reproducibility value exists in the literature for the S2 thus the Chi – square test could not be directly applied by the code written in Mat lab.

RESULTS AND DISCUSSION FOR S3

Anderson - Darling $A2^*$ Value	0.3985	
Student's $t(a)$ Value	9.9165	
MR_{avg}	0.0325	
Upper Limit for I Chart	0.1684	
Lower Limit for I Chart	0.0062	

Upper Limit for EWMA Overlay	0.1252	
Lower Limit for EWMA Overlay	0.0389	
Upper Limit for MR Chart	0.1062	
m / s (Levey – Jennings)	0.8725	0.0432
m+S / m-S (Levey – Jennings)	0.9157	0.8293
m+2S / m-2S (Levey – Jennings)	0.9589	0.7861
m+3S / m-3S (Levey – Jennings)	1.0021	0.7429
F Value	1.449	
MRpooled	0.0276	
Site Precision R' of S3	0.0799	
X ² Value	-	
Cp / Cpk	-	-
Pp / Ppk	1.5144	0.8847
Zupper / Zlower	2.6691	6.4172

Table 5: S3 results.

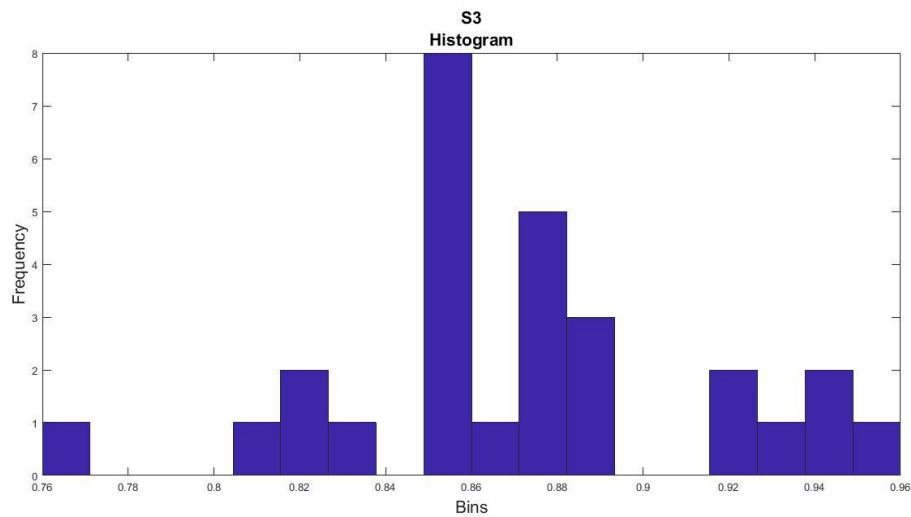


Figure 25: Histogram of the S3 results.

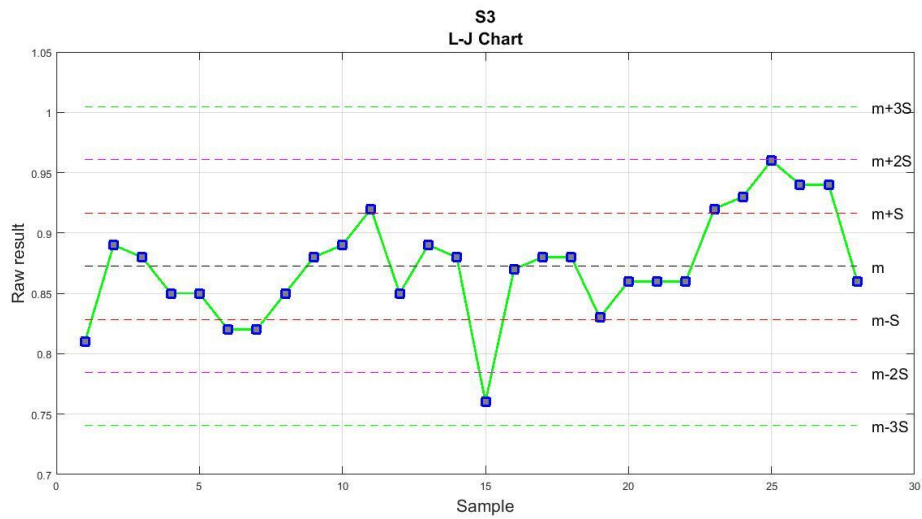


Figure 26: Levey – Jennings Chart of the S3 results.

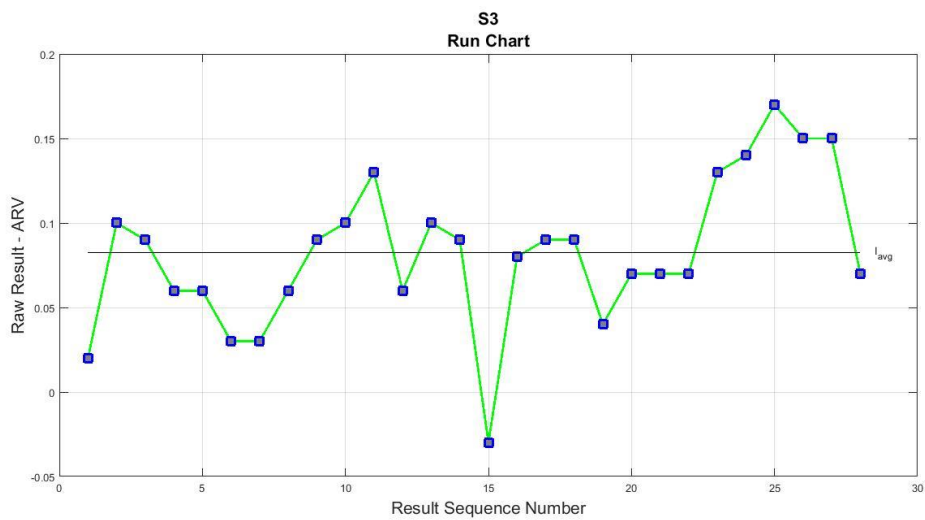


Figure 27: Run Chart of the S3 results.

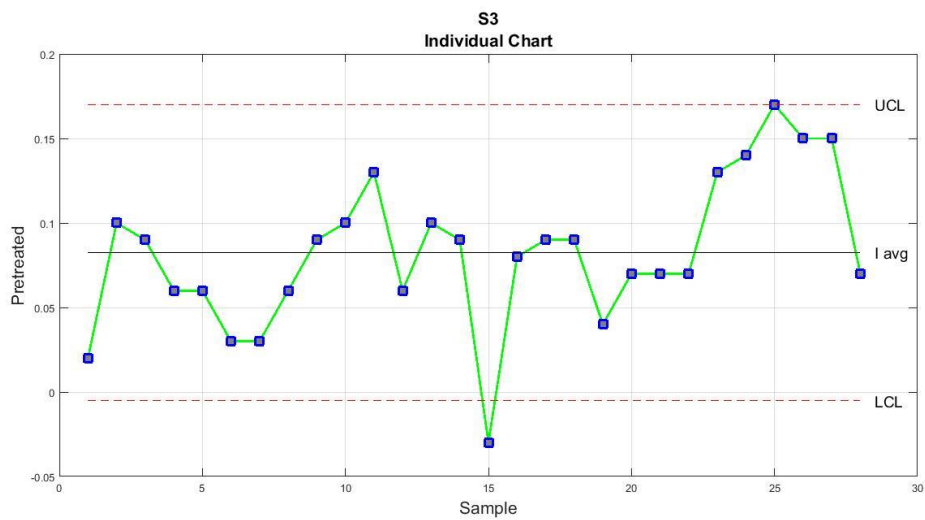


Figure 28: Individual Chart of the S3 results.

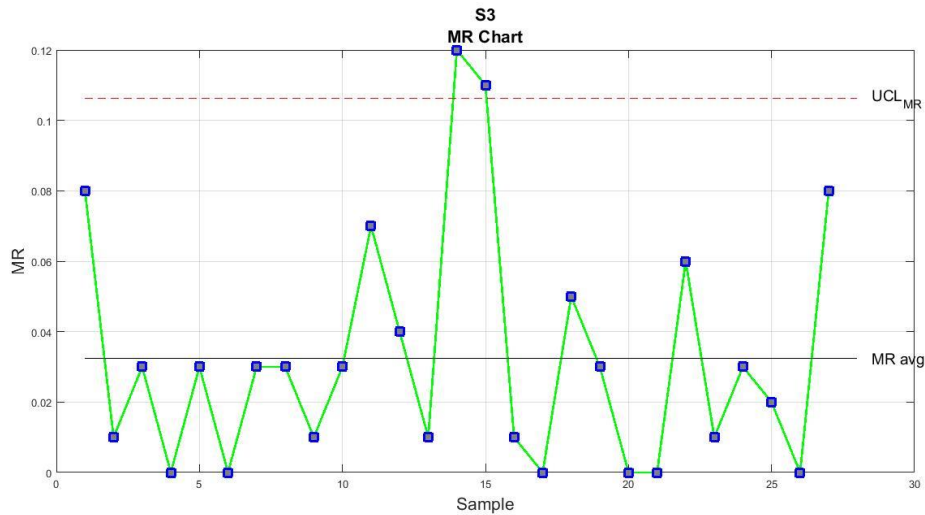


Figure 29: MR Chart of the S3 results.

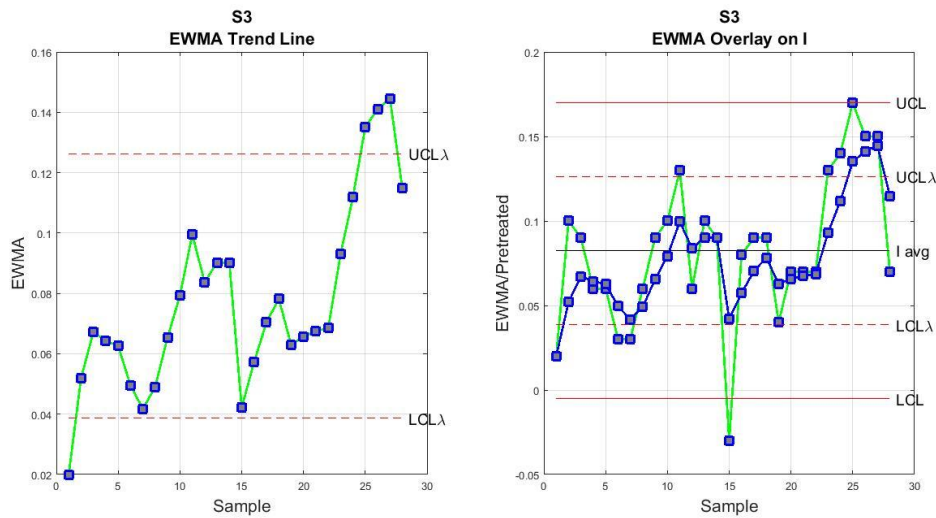


Figure 30: EWMA trend line and EWMA Overlay on the Individual Chart.

The normality assumption is confirmed by both the visual inspection of the histogram and the result of the Anderson – Darling test at the 95% confidence level. The test resulted in an A_2^* value of 0.3985 and therefore it was smaller than the marginal 0.752 value. As far as the Levey – Jennings Chart for the S3 results and the Westgard rules are concerned, the program indicates that point 15 falls below the $m - 2S$ range limit, therefore violates the 12_s Westgard rule and was accepted as a random error. Furthermore, points 23 to 26 violate the 41_s rule. In the Run Chart of the S3 results, the same point seems to exhibit a more intense fluctuation than the others and Individual Chart confirms the visual observation about point 15, since it is located below the lower control limit and furthermore causes an abrupt downward shift in the process. In the Moving Range Chart, both points 14 and 15 result above the upper limit and denote that the process goes out of control. Here it is important to mention that this kind of instability also affects the control limits of the Individual

Chart as it was explained in Chapter 2 but no assignable cause that triggered this instability could be determined and removed from the affected data points. In the EWMA trend line, Point 1 results below the lower control limit while points 25, 26 and 27 result above the upper control limit (larger shifts in the process). An upward trend can also be detected here, starting from point 19 and ending to point 27. As was the case with the previous measurements, smaller shifts in the process that do not affect importantly the system are also detected by EWMA Overlay on the Individual Chart. The Student's t test indicated that a bias relative to the ARV is present in the system and that the mean of the samples is statistically distinguishable from the hypothesized mean value. The calculated t (a) value equals to 9.9165 which greater than 2.1448. This result is logical if someone considers that the mean of the data results was equal to 0.8725 and the ARV was equal to 0.79. The F ratio that was calculated from the F test was equal to 1.449, smaller than 3.16, and the two individual MR_{avg} (0.0325 and 0.0224) were pooled into a single which was equal to 0.0276, a fact that further advocates to the stability and repeatability of the results over time. Concluding, the site precision concerning the S3 results was equal to 0.0799. Given that the system was out – of – control the inclusion of the C_p and C_{pk} values is once again omitted. The P_p and the P_{pk} indices were equal to 1.5144 and 0.8897 respectively a fact that reveals acceptable average quality. Finally Z_{upper} was equal to 2.6691 and Z_{lower} was equal to 6.4172. The first value indicates that the process has produced 0.39% of defective results above the USL and virtually 0% below the LSL. No published reproducibility value exists in the literature for the S3 thus the Chi – square test could not be directly applied by the code written in Matlab.

All things considered, there are some special causes (shifts and trends) in the control charts that reveal local instabilities in the process, with most important violations those of points 10 and 24 of the T_{MAX} results and point 15 of the S3 results but the Rock – Eval system generally seems to operate in an acceptable manner in terms of precision, given that even the most inherently stable manufacturing processes can exhibit smaller shifts and trends. The basic unanswered question though is if the accuracy of the test results provided by the system is acceptable or not. As was already mentioned, the business application need is what primarily defines the accuracy limits of a process and less so the result of a statistical test which has a relative importance. So subsequently, the geochemical evaluation of the results and an accuracy statement about the apparatus was conducted based on the benchmarking concept. This concept was based on how much the actual quantities calculated from the test results (Productivity Index, Oxygen Index, Hydrogen Index, Petroleum Potential, Organic Type indicator and Vitrinite Reflectance) will deviate from the “ideal” calculated quantities, if the ARV of the S1, S2, S3, TOC and TMAX results was used for their calculation. The Van – Krevelen diagram and the Peter's

table, which contains the limit values for the evaluation of the Rock – Eval results, were employed for this purpose so as to quantify the difference between the two cases. With this comparison, a statement about the general “business” accuracy of the system was formulated. More specifically, four different scenarios that are presented below were conducted:

- Productivity Index, Petroleum Potential, Hydrogen Index, Oxygen Index, Organic type Indicator and Vitrinite Reflectance calculated with the ARV’s.
- Productivity Index, Petroleum Potential, Hydrogen Index, Oxygen Index, Organic type Indicator and Vitrinite Reflectance calculated with the mean of the actual results.
- Productivity Index, Petroleum Potential, Hydrogen Index, Oxygen Index, Organic type Indicator and Vitrinite Reflectance calculated with the point that is located further above the central line (m) of the Levey – Jennings Chart.
- Productivity Index, Petroleum Potential, Hydrogen Index, Oxygen Index, Organic type Indicator and Vitrinite Reflectance with the point that is located further below the central line (m) of the Levey – Jennings Chart.

The first scenario, which uses the ARV’s S1, S2, S3, TOC and TMAX, will be evaluated geochemically based on the Van – Krevelen diagram and the Peters table with the limit values that was presented above and this evaluation will be considered as the benchmark with which the other three scenarios will be compared. If the geochemical evaluation of the data of the other 3 scenarios will produce a geochemical statement about the organic matter that deviates from the one that the ARV scenario produces, then the actual “business” accuracy of the system will be deemed poor. The results of the benchmark scenario are summarized below:

BENCHMARK ARV SCENARIO

	Values	Peter’s Table	Van Krevelen Diagram
S1	0.14	Poor	
S2	12.43	Very good	
S3	0.79		
T _{MAX}	416	Immature	
TOC	3.28	Very good	
Productivity Index	0.011	Early Mature	
Petroleum Potential	10.43		
Hydrogen Index	378.96	Kerogen Type II	
Oxygen Index	24.08		
Organic Type Indicator	15.73		
Vitrinite Reflectance	0.328	Immature	
Kerogen Type			Kerogen Type II

Table 6: Benchmark scenario geochemical characterization.

The Peters table with the limit values of the Rock – Eval results classifies the organic matter geochemically based on the level of thermal maturation of the organic matter, the petroleum potential and the kerogen (quality of the organic matter) which is predominant in the sample. More specifically, the **S1** value of the ARV scenario reveals that the sample is poor (quantitatively) in free hydrocarbons and volatile compounds. The **S2** value on the other hand, which represents the amount of hydrocarbons that were generated through thermal cracking, characterizes the sample as very good, again in terms of quantity. The **S3** value, which represents the amount of the Oxygen contained in the organic sample, equals to 0.79 and will be utilized along with the S2 in the organic type indicator formula in order to determine the depositional environment (marine or terrestrial) from which the organic matter originated. The **TOC** value (3.28) which represents the total amount of carbon in the sample (pyrolyzed + residual) is also characterized as very good. As far as the maturation of the organic matter is concerned, a **Tmax** value of 416 °C characterizes the sample as late mature something which in accordance with the Van – Krevelen diagram will suggest the type of kerogen (quality) of the organic matter. The **productivity index** value of the ARV scenario is compared with the table value and indicates early mature organic matter, in terms of the generation, since it falls within the 0.10 – 0.15 range. As far as the type of the kerogen (which determines the source rock quality) is concerned, the **Hydrogen index** indicates that we have kerogen type 2, and thus mainly oil prone organic matter. The **organic type indicator** on the other hand, equals to 15.7 and therefore indicates kerogen type 1, something which again suggests oil prone organic matter. Finally, the Van – Krevelen diagram confirms that a Tmax of this magnitude is probably related to kerogen type 2 organic matter and the combination of the Hydrogen Index and Oxygen Index further advocates in the Kerogen type 2 observation. Based on the Organic Type indicator and the amount of S3 it can be stated that the organic matter was not exposed to intense oxidization conditions something which probably indicates marine origin.

The other three scenarios that were explained above were also materialized and the results are summarized in the following tables:

MEAN SCENARIO

	Values	Peter's Table	Van Krevelen Diagram
S1	0.1907	Poor	
S2	12.48	Very good	
S3	0.8725		
T_{MAX}	416	Immature	
TOC	3.24	Very good	
Productivity Index	0.015	Early Mature	

Petroleum Potential	10.51		
Hydrogen Index	385.27	Kerogen Type II	
Oxygen Index	26.93		
Organic Type Indicator	14.31		
Vitrinite Reflectance	0.328	Immature	
Kerogen Type			Kerogen Type II

Table 7: Table 8: Mean scenario geochemical characterization.

UPPER L-J POINTS SCENARIO

	Values	Peter's Table	Van Krevelen Diagram
S1	0.21	Poor	
S2	12.79	Very good	
S3	0.96		
T _{MAX}	420	Immature	
TOC	3.29	Very good	
Productivity Index	0.016	Early Mature	
Petroleum Potential	10.79		
Hydrogen Index	388.75	Kerogen Type II	
Oxygen Index	29.18		
Organic Type Indicator	14.32		
Vitrinite Reflectance	0.4	Immature	
Kerogen Type			Kerogen Type II

Table 8: Upper L-J points scenario geochemical characterization.

LOWER L-J POINTS SCENARIO

	Values	Peter's Table	Van Krevelen Diagram
S1	0.17	Poor	
S2	12.06	Very good	
S3	0.76		
T _{MAX}	412	Immature	
TOC	3.18	Very good	
Productivity Index	0.014	Early Mature	
Petroleum Potential	10.15		
Hydrogen Index	379.24	Kerogen Type II	
Oxygen Index	23.9		
Organic Type Indicator	15.86		
Vitrinite Reflectance	0.26	Immature	
Kerogen Type			Kerogen Type II

Table 9: Lower L-J points scenario geochemical characterization.

Even though the Student's test result indicated that in the cases of S1, S3 and TOC there was a bias in the system relative to the ARV and subsequently poor accuracy, this benchmark concept reveals that in all three alternative cases, the geochemical characterization of the organic matter remains the same, a fact that indicates acceptable business accuracy. A general conclusion that can also be elicited about the system is that the general accuracy of the S1 results is lower because S1 peak is connected with the more volatile compounds already present in the sample which can be "lost" easier during the pyrolysis and therefore not be detected during the analysis. Similarly, the relative poor accuracy of the S3 results probably indicates that the TCD detector does not work ideally and gives inaccurate information about the CO₂ that was produced during the pyrolysis. Since both S1 and S3 exhibit poor accuracy, this fact affects also the TOC accuracy, which as was already mentioned directly includes S1 in its formula.

GC - MS PERFORMANCE EVALUATION

The evaluation of the performance of the GC – MS system was based on the behavior of the relative response factors and not on the direct measured quantities of the system (concentrations/areas). By statistically evaluating the relative response factors instead of the concentrations, the performance of the system in time can be monitored and it is easier to identify the need for a new calibration curve and a possibly ill prepared quality control sample. For a time period of almost four years (January 2013 – September 2016) the relative response factors of the different compounds were calculated and statistically evaluated. It is important to mention that the system was not operating constantly but within specific time intervals and that what was used as an “ARV” for the implementation of the statistical techniques was in reality a calculated mean value of the R.R.F’s, that can be considered as the Site Expected Value (SEV). The parametric tests that were used for the evaluation of the Rock – Eval system cannot be performed here since the result of the Student’s test does not correspond to the real accuracy of the system, there is no published reproducibility value for the R.R.F’s and no F test could also be performed since the objective was to evaluate the general behavior of the R.R.F’s in time, in their totality. The process stability and process performance indices could also not be implemented since no specification limits were available for an indirect measurement (R.R.F’S) and the evaluation was solely based on interpretation of the charts. For 55 samples, a code was written that was first calculating the relative response factor of each compound and then implementing the charting techniques that were used for the evaluation of the Rock – Eval system. The evaluation was based on the behavior of the Relative Response factor of the different compounds during this time period but also on the behavior of the R.R.F of each compound within the specific time intervals that the system was operating, thus two arrangements of charts were developed for this purpose. The size of the sample was adequately large (>50) therefore any deviations from the normality were considered to have no important impact on the process and given that the control charts that were described are reliable and robust for non – normally distributed characteristics also, a direct comparison between non normal and normal relative response factors can be conducted safely, with the exception of the Levey – Jennings chart in cases where the normality was rejected. Below, four random sets of charts of relative response factors are indicatively presented, starting from a lighter compound (C13) and progressing towards the heavier (C29) so as to reach to some conclusions concerning the general operation of the system.

Component	A ^{2*}	MRavg	Site Precision
C12	2.1863	0.0728	0.1787
C13	1.5228	0.0864	0.2137
C14	2.3312	0.0869	0.2107
C15	2.6480	0.1103	0.2691
C16	1.4959	0.0708	0.1737
C17	2.0951	0.0732	0.1785
Pr	1.1618	0.0588	0.1449
C18	2.0858	0.0734	0.1807
Ph	0.2353	0.0810	0.2033
C19	2.6023	0.0825	0.2029
C20	0.9090	0.0918	0.2275
C21	0.4782	0.0966	0.2404
C22	0.5702	0.1090	0.2707
C23	0.6088	0.1152	0.2860
C24	0.5203	0.1169	0.2894
C25	0.5932	0.1115	0.2764
C26	0.4589	0.1299	0.3213
C27	0.5140	0.1317	0.3261
C28	0.7049	0.1209	0.2987
C29	0.3992	0.1281	0.3183
C30	0.3700	0.1274	0.3164

Table 10: A^{2*} and Site precision of the analytes.

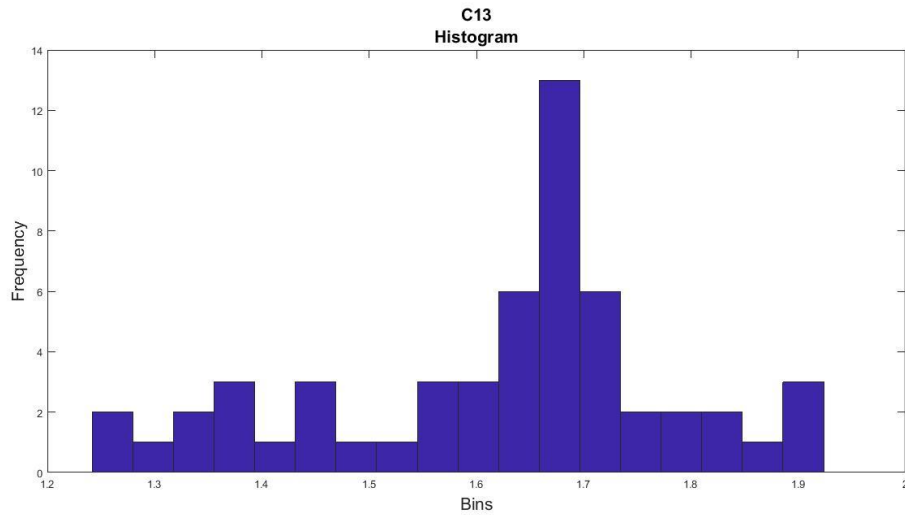


Figure 31: Histogram of C13 R.R.F.

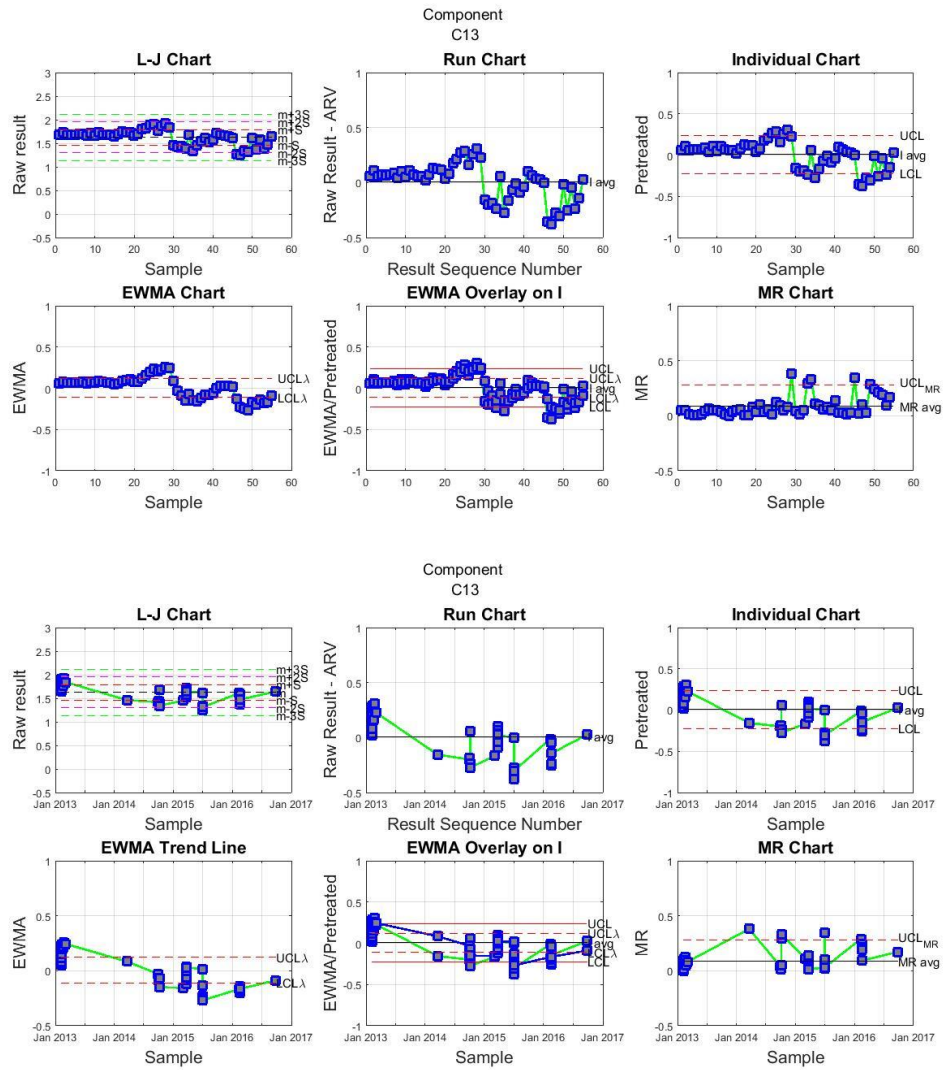


Figure 32: R.R.F over time behavior of C13.

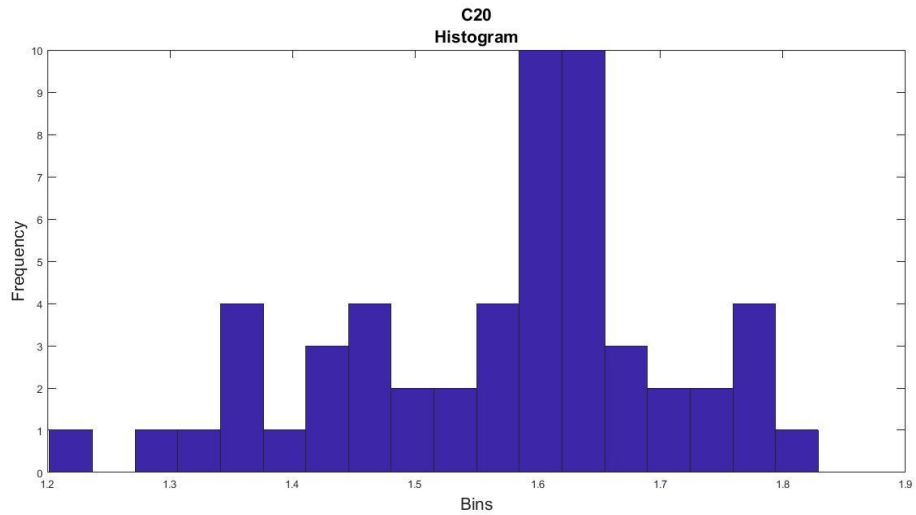


Figure 33: Histogram of C20 R.R.F.

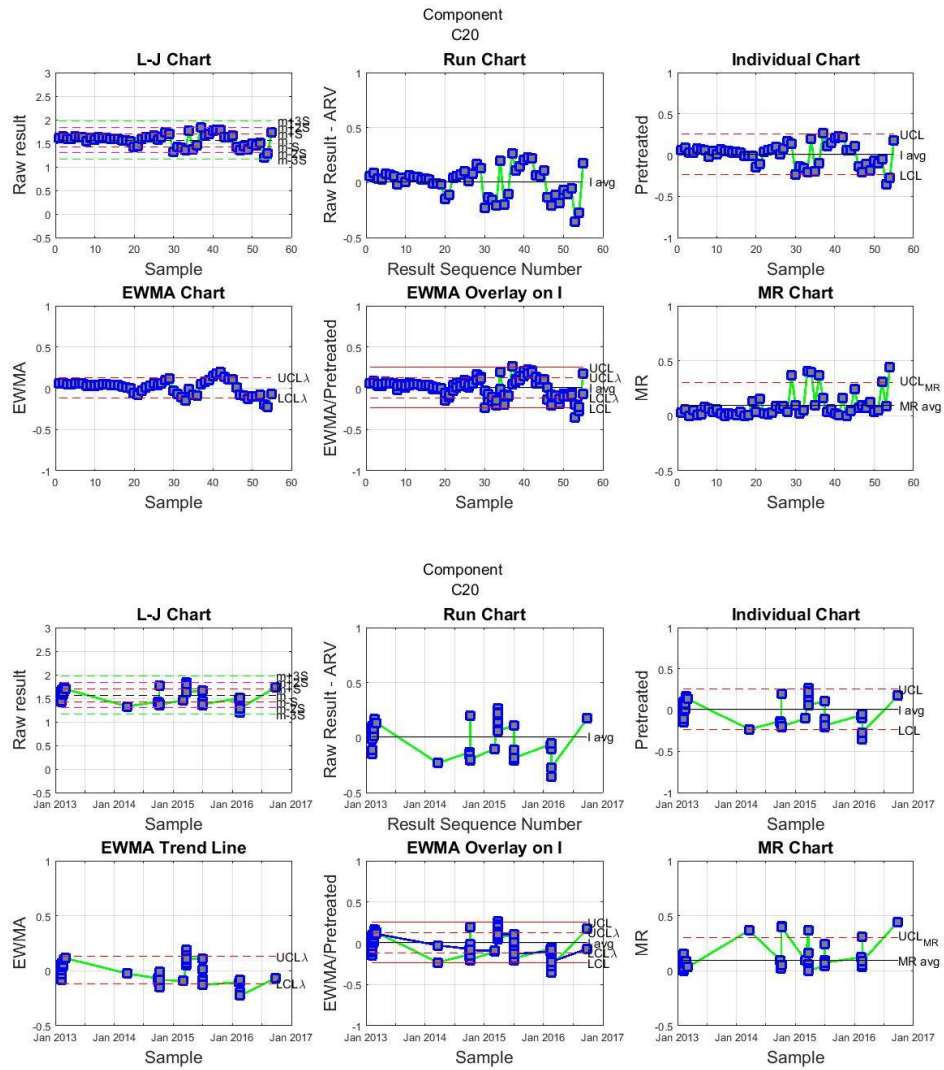


Figure 34: R.R.F. over time behavior of C20.

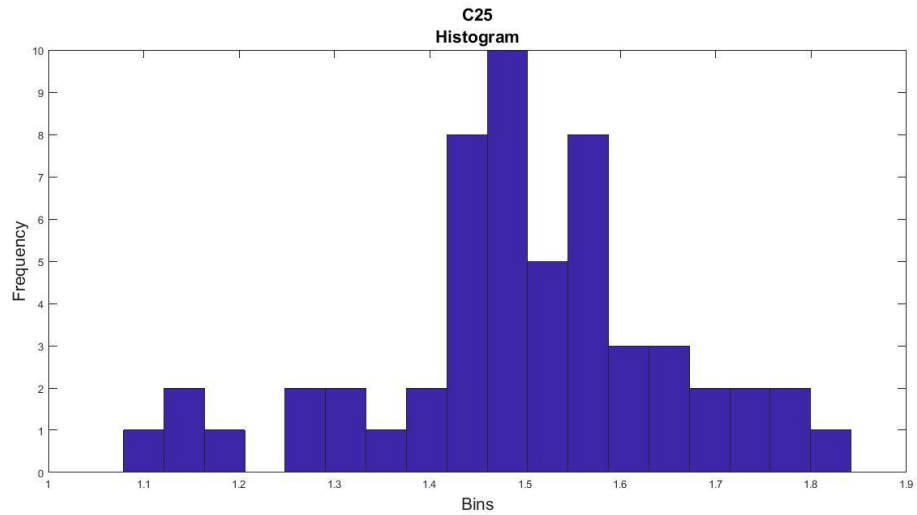


Figure 35: Histogram of C25 R.R.F.

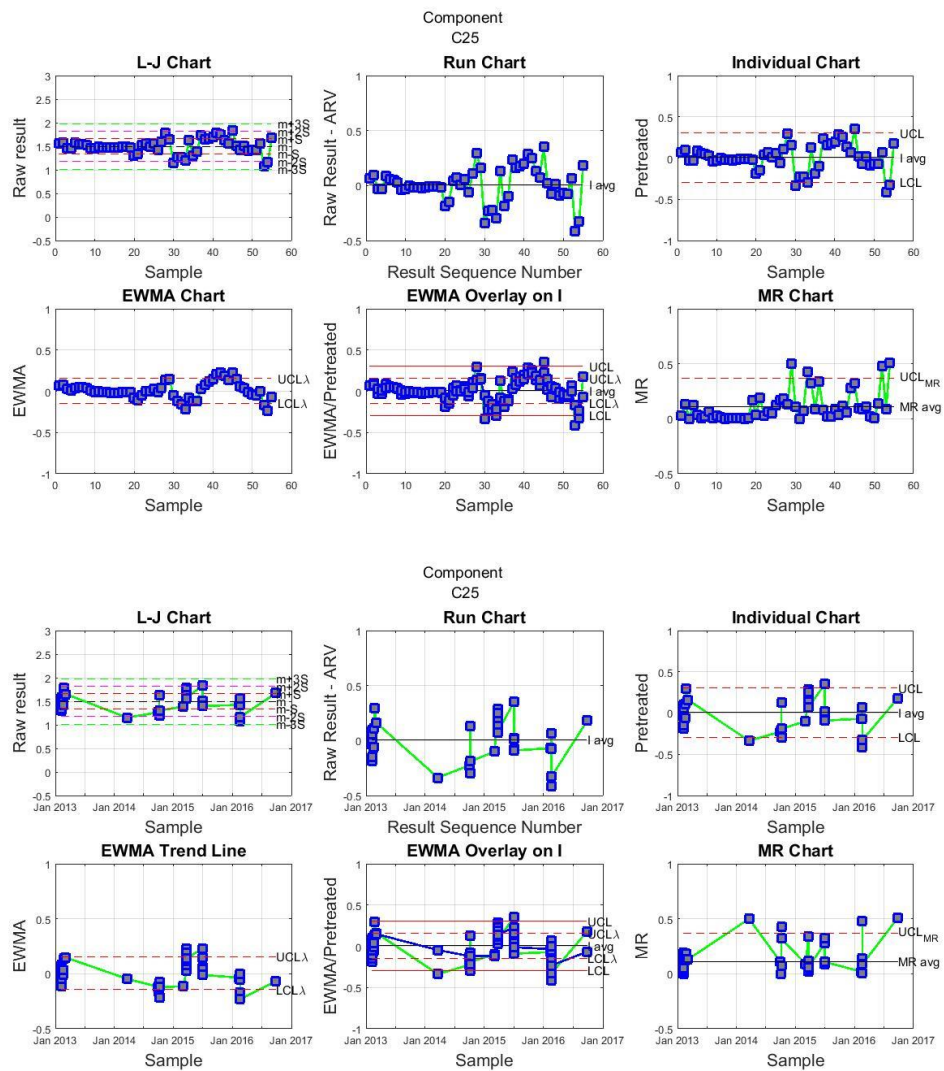


Figure 36: R.R.F over time behavior of C25.

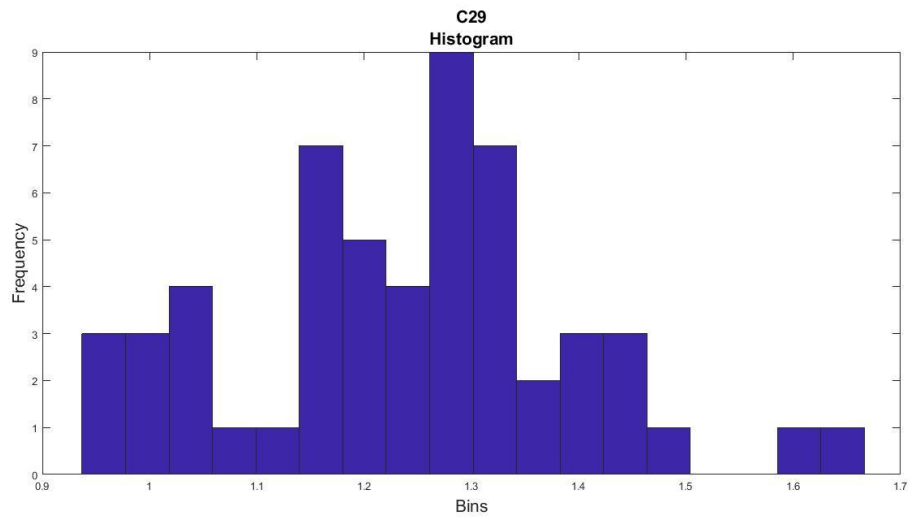


Figure 37: Histogram of the C29 R.R.F.

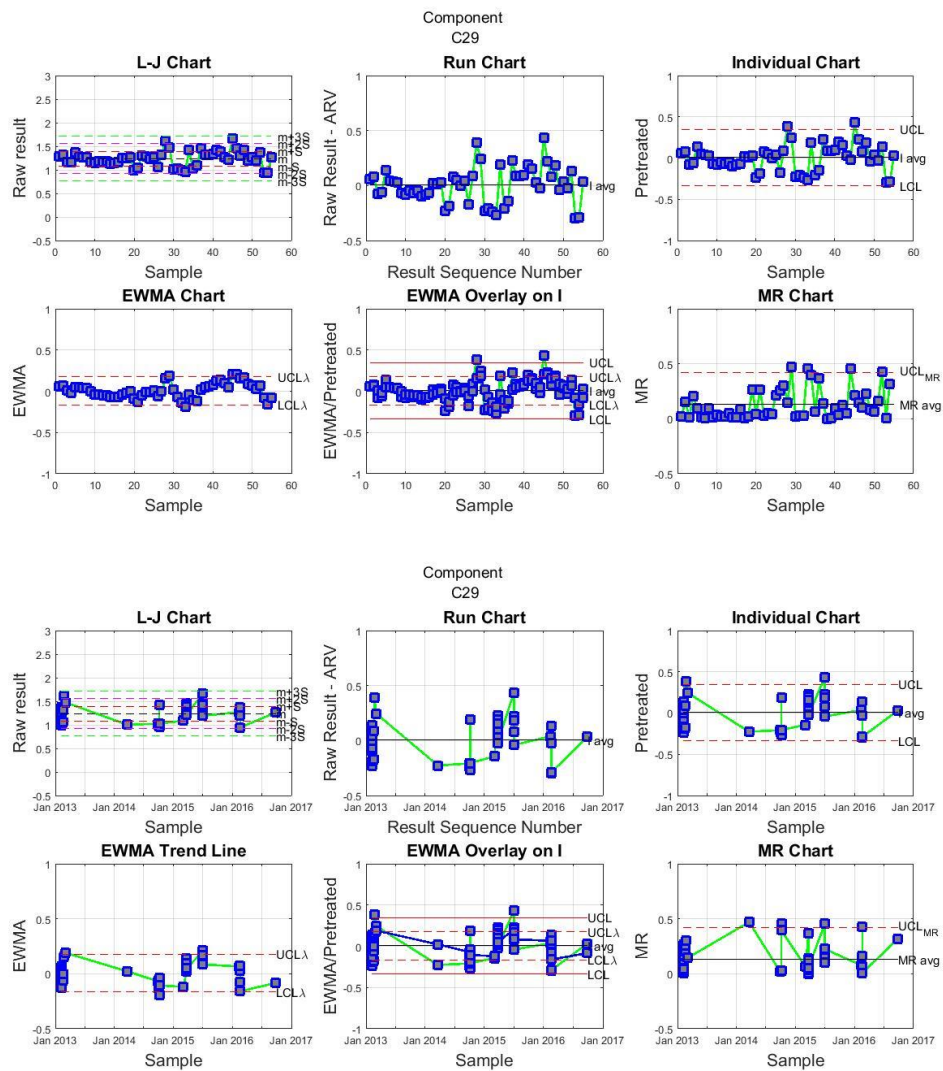


Figure 38: R.R.F. over time behavior of C29.

The general behavior of the R.R.F of the different compounds in time followed similar trends something which was to be expected. The values of the R.R.F of the different compounds were different due to the different participation of the C_6H_{13} ion in each of them. The evaluation of the performance of the system in terms of precision/repeatability was based on the behavior of the R.R.F between the lighter to heavier compounds during the overall operating period but also within the same time intervals as it was already explained. In general, the system operates poorly and given that many outliers are present (about which there is no option to assign specific - assignable causes), an attempt to define the reasons behind the poor performance of the system will be made. Generally the normality assumption is rejected by the Anderson – Darling test for the majority of the lighter compounds and that means that the conventional Levey – Jennings chart does not describe the precision of the system with absolute reliability. On the other hand, the normality hypothesis is accepted in the C21 – C30 range so general assumptions based on the Levey – Jennings chart can be made for the heavier compounds. By observing the above set of charts, two basic conclusions can be made about the performance of the system. Firstly, two periods of operation, as far as the stability and the precision of the relative response factors is concerned, can be discerned. In the first period, the system seems to operate with relative stability while in the second period, intense instability can be detected and the system goes obviously out of control. This trend is similar for all the components. This can be attributed either to the measurement system or to the injected standard that was used. Under the assumption that the injected internal standard exhibited correct concentration, it can be concluded that the measurement system exhibits different signal intensity through time, therefore we cannot rely on an average calibration curve for each analysis and a new one should be adopted for each study (for every analyte). Also, on the same basis, since the system was not operating constantly but within specific time intervals, the disparity of the response factors can also be attributed to the current state of the measurement system, with obvious reasons the poor maintenance practices and possibly the ageing of the apparatus. A common problem that affects the response and usually emerges with the ageing of the system is the normal degradation of the stationary phase of the column. This happens when the stationary phase is exposed to high temperatures and in the presence of oxygen for a long period of time. This gradual degradation of the column may cause retention of some of the compounds within the column and subsequently result to poor precision between the relative response factors. If the system is assumed to work properly, and based on the first – stable – period of operation, the grouped data of the second arrangement of charts indicates that the precision of the relative response factors is decreasing as we progress from lighter to heavier compounds. This can be attributed to the incorrect preparation of the quality control samples and their incorrect injection. Heavier compounds require a finer pretreatment than the lighter and it is

possible that many of them were not heated properly and therefore were not introduced into the system for the analysis. The rapid flow velocity may create a phenomenon where the less volatile compounds do not have sufficient time to fully vaporize and as a result condense at the rubber septum or exit the split outlet without being introduced with the correct split ratio into the column. This phenomenon is called discrimination of the heavier – less volatile compounds. Discrimination describes the phenomenon during which only a portion of the heavier analytes enters the column from the inlet mainly due to fast sample injection or due to the wrong position of the liner packing which cannot act as a surface for the evaporation / condensation of the heavier analytes. It is possible that the decline in the precision from the lighter to the heavier compounds is partially due to this discrimination, so the precision of the system may be improved under different injection conditions. Concluding, the precision of the system may improve after injector and column maintenance. During injector maintenance, the old liners are replaced with new ones so as to diminish the discrimination phenomenon that was described above. During column maintenance, a part of the column is replaced so as to improve and restore analyte response. Usually 5-15 cm of the column is removed and reinstalled. As far as the general “business” accuracy of the system is concerned, a general conclusion that can be made is that for the heavier fraction (C₂₁ – C₃₀) no ¹³S violation can be detected in the Levey – Jennings chart, something which would indicate an excessive random or systematic error. Given that the Carbon Preference Index and the Odd to Even Predominance Index are calculated with the C₂₀+ compounds, the fact that the Levey – Jennings chart reveals no outliers for the heavier fraction indicates certain reliability as far as the conclusions about the thermal maturity of the samples (based on these biomarkers) is concerned.

INTERLABORATORY STATISTICS

In the Interlaboratory study, the sample material used was a monograde oil that was distributed in a number of laboratories so as to obtain estimates about basic properties of the sample and, after a statistical processing and charting of the results, formulate a precision statement about the test method. These properties were the mass (%) of Magnesium, Calcium, Zinc, Phosphorus and Sulfur in the sample, the density (Kg/L) at 15 degrees Celsius and the Flash Point of the oil. In the charts presented below, the results from the various laboratories are presented and based on the statistical procedure described in Chapter 2, some results are discarded before the final precision statement is formulated for each one of the properties. As it was already mentioned, since in the ASTM report every laboratory provided only one final value as a result, the within – laboratory repeatability statistics cannot be estimated and, subsequently, the k statistic with its relative graph cannot also be produced (although the overall procedure was included in the Mat lab Code). It is also important to mention that the final report of the results of the samples already defined some test results as outliers, based on the 98% Grubb's Limit. These outliers were not taken under consideration during the implementation of this standard practice method and eventually, in some cases, more data results were labeled as outliers after the implementation of the ASTM E 691 – 99 standard practice. The discarding of these additional points resulted into a difference between the reproducibility limit R that was calculated with this practice and the reproducibility limit R that was included into the initial report. So an indirect comparison between the 98% Grubb's Limit evaluation and the specific standard's evaluation was conducted.

GRUBB'S TEST FOR OUTLIERS

Grubb's test (named after Frank E. Grubbs) also known as extreme studentized deviate test, is a statistical test used to detect outliers in a univariate data set assumed to come from a normality distributed population. Grubb's test identifies one outlier at a time and subsequently expunges it from the data set. The test is iterated until no outliers can be detected. Given that multiple iteration's change the probabilities of detection, the test should not be used for small sample sizes. Grubb's test is defined for the hypothesis where H_0 : There are no outliers in the data set and H_a : There is one outlier in the data set. The Grubb's test is briefly described below.

The Grubb's test statistic is defined as:

$$G = \frac{\text{Max } |Y_i - Y_{avg}|}{S}$$

Where Y_{avg} is the sample mean and S is the standard deviation.

The one sided version of the test which determines if the maximum or the minimum value of the data set is an outlier includes:

$$G = \frac{Y_{avg} - Y_{Min}}{S}$$

$$G = \frac{Y_{Max} - Y_{avg}}{S}$$

For the min and max value respectively.

For the two sided test, the hypothesis of no outliers is rejected at a significance level of α if:

$$G > \frac{N-1}{N} * \sqrt{\frac{t_{\frac{\alpha}{2N}, N-2}^2}{N-2 + t_{\frac{\alpha}{2N}, N-2}^2}}$$

With $t_{\frac{\alpha}{2N}, N-2}^2$ denoting critical value of the t – distribution with N-2 degrees of freedom and a significance level of $\alpha / (2N)$.

RESULTS AND DISCUSSION ABOUT MAGNESIUM m(%) – D4951 Method

P (Number of laboratories)	29	
X – Avg of the cell Avg's	0.0751	
Sx – Standard Deviation	0.0061	
Sr – Repeatability St. Dev.	-	
SR – Reproducibility St. Dev.	0.0061	
Critical h statistic value	2.64	-2.64
Reproducibility Estimate (Grubb's)	0.0168	
Reproducibility Estimate R	0.0168	

Table 9: Statistical results for Magnesium.

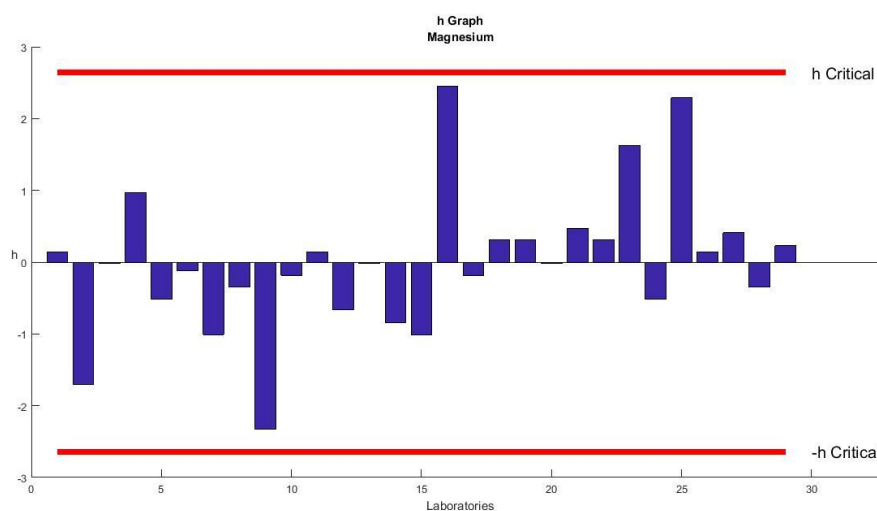


Figure 39: Consistency graph for Magnesium.

Initially, 34 laboratories provided a test result about the m (%) concentration of the Magnesium in the sample using the D4951 Method but the Grubbs' Limit identified that five of them (Laboratories No 50, 61, 67, 86 and 90) were outside the statistical limits and labeled them as outliers. The implementation of the standard practice method revealed no additional outliers and therefore the precision statement was formulated without having to discard more data results than those the initial report already discarded. The reproducibility value R that was calculated, was equal to the one that the report also provided ($R = 0.0168$)

RESULTS AND DISCUSSION ABOUT ZINC m(%) – D4951 Method

P (Number of laboratories)	32	31
X – Avg of the cell Avg's	0.1179	
Sx – Deviation	0.0054	
Sr – Repeatability St. Dev.	-	
SR – Reproducibility St. Dev.	0.0054	
X – Corrected Avg of the cell Avg's		0.1173
SR – Corrected Reproducibility St. Dev.		0.0044
Critical h statistic value	2.64	-2.64
Adjusted critical h value	2.64	-2.64
Reproducibility Estimate (Grubb's)	0.0168	
Reproducibility Estimate R	-	0.0126

Table 10: Statistical results of Zinc.

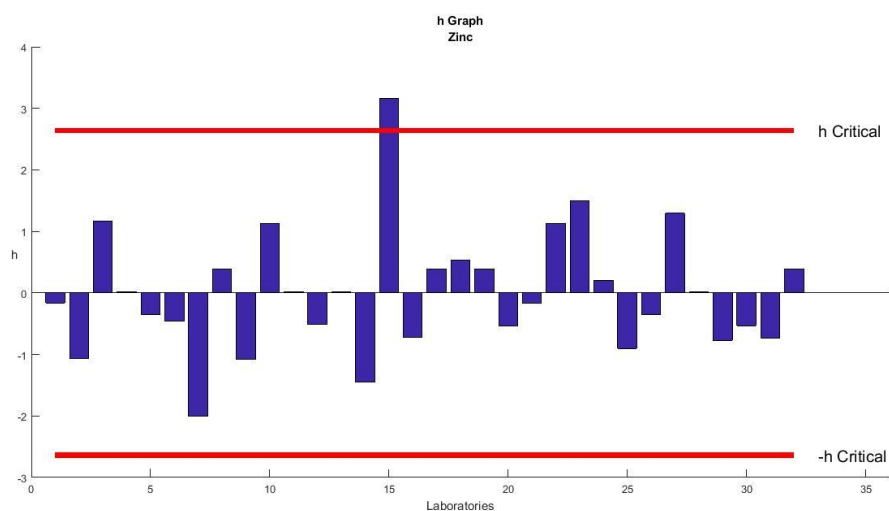


Figure 40: Consistency graph for Zinc.

Initially, 33 laboratories provided a test result about the m (%) concentration of the Zinc in the sample using the D4951 Method but the Grubbs' Limit identified that one of them (Laboratory 90) was outside the statistical limits and was labelled as an outlier. The implementation of the standard practice method revealed that apart from Laboratory 90, Laboratory 50 also provided a test result that exceeds the critical h value. Subsequently, this value was discarded and was not taken under consideration for the formulating of the precision statement. The final reproducibility value (R) calculated (0.0126) was different from the precision reproducibility R that the report provided (0.0168), due to the discarding of this data result. Below the graph with the corrected h statistic for Zinc is depicted.

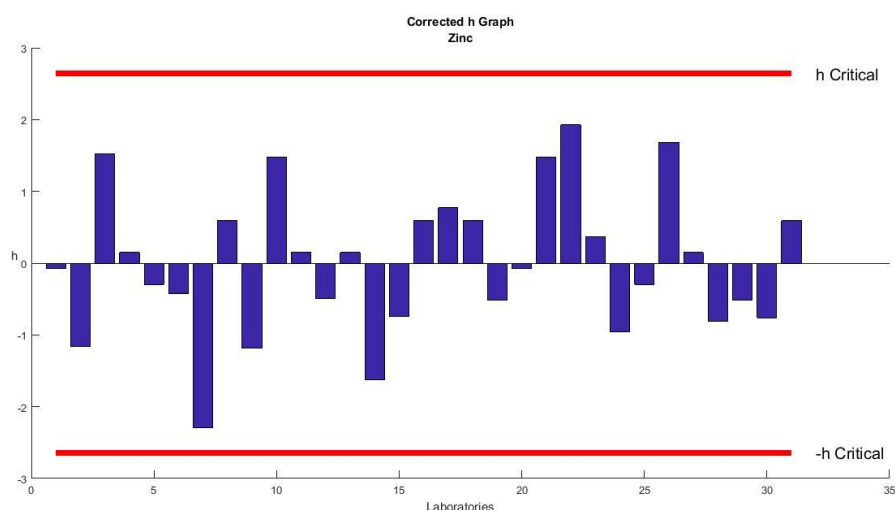


Figure 41: Corrected consistency graph for Zinc.

RESULTS AND DISCUSSION ABOUT PHOSPHORUS m(%) – D4951 Method

P (Number of laboratories)	30	28
X – Avg of the cell Avg's	0.1059	
Sx – Deviation	0.0096	
Sr – Repeatability Standard Dev.	-	
SR – Reproducibility Standard Dev.	0.0096	
X – Corrected Avg of the cell Avg's		0.1039
SR – Corrected Reproducibility St.Dev.		0.0062
Critical h statistic value	2.64	-2.64
Adjusted critical h value	2.63	-2.63
Reproducibility Estimate (Grubb's)	0.0267	
Reproducibility Estimate R	-	0.0174

Table 11: Statistical results of Phosphorus.

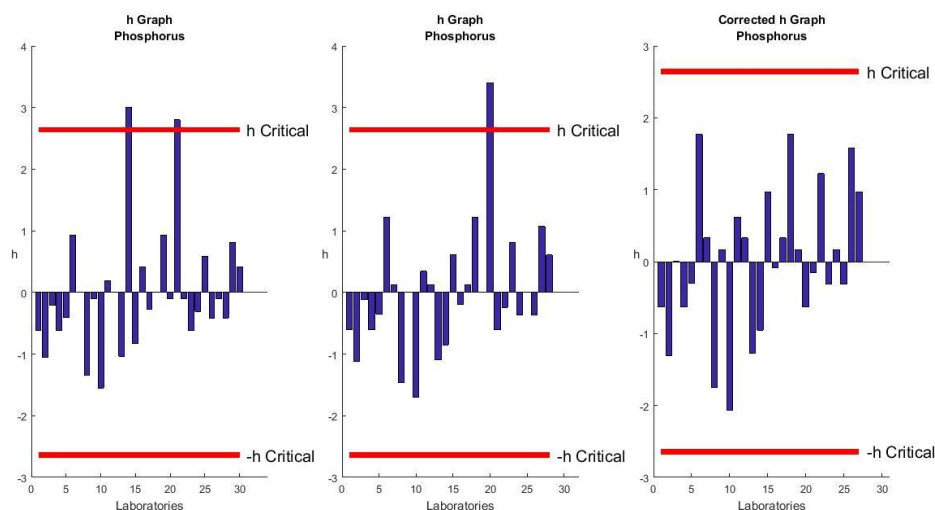


Figure 42: Consistency graph for Phosphorus.

In the initial report of the round – robin test, 30 laboratories provided a result about the m (%) of the Phosphorus in the sample without any result being identified as an outlier by the statistical control performed based on the Grubbs' statistical limit. After implementing the iterative statistical process of the E 691 – 99 Standard though, Laboratories 50 and 61 violated the critical h value, were discarded as outliers and were not taken into account for the final precision statement. The final reproducibility value (R) calculated (0.0174) was different from the precision reproducibility R that the report provided (0.0267) due to the discarding of these 2 data results.

RESULTS AND DISCUSSION ABOUT CALCIUM m(%) – D4951 Method

P (Number of laboratories)	32	
X – Avg of the cell Avg's	0.1118	
Sx - Deviation	0.0054	
Sr - Repeatability St. Dev.	-	
SR - Reproducibility St. Dev.	0.0054	
Critical h statistic value	2.64	-2.64
Reproducibility Estimate (Grubb's)	0.0149	
Reproducibility Estimate R	0.0149	

Table 12: Statistical results for Calcium.

From the 33 laboratories that provided a result for the m (%) of the Calcium in the sample, one laboratory – No 90 – violated the Grubbs' statistical limit and was discarded as an outlier. The implementation of the E 691 – 99 standard practice method revealed no additional outliers and therefore the precision statement was formulated without having to discard more data results than those the initial report already discarded. The reproducibility limit value R that was calculated, was equal to 0.0149, equals to the one that the Grubb's limit provided.

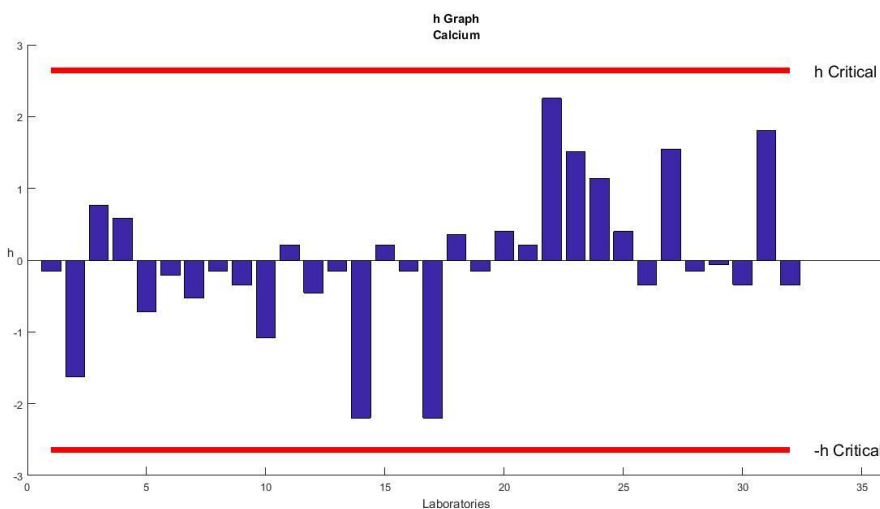


Figure 43: Consistency graph for Calcium.

RESULTS AND DISCUSSION ABOUT SULFUR m(%) – D129 Method

P (Number of laboratories)	14	
X – Avg of the cell Avg's	0.4193	
Sx - Deviation	0.0409	
Sr - Repeatability St. Dev.	-	
SR - Reproducibility St. Dev.	0.0409	
Critical h statistic value	2.44	-2.44
Reproducibility Estimate (Grubb's)	0.1133	
Reproducibility Estimate R	0.1133	

Table 13: Statistical results for Sulfur.

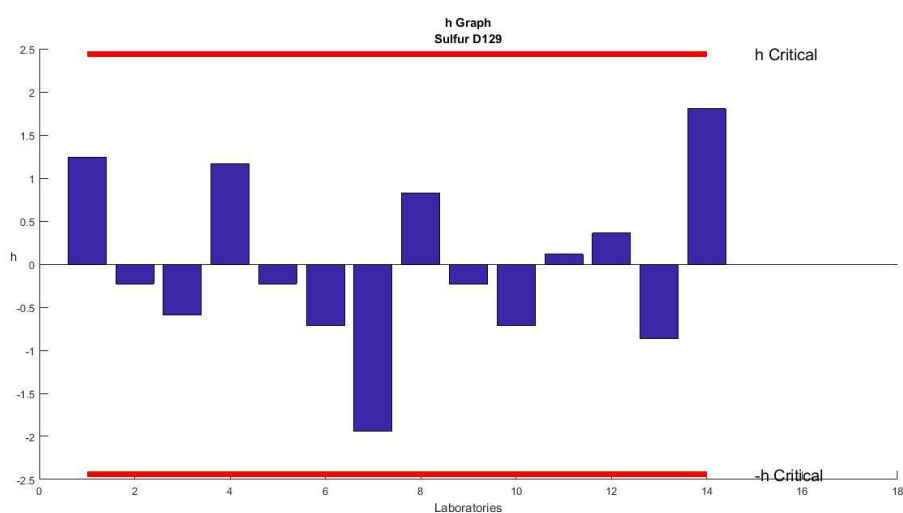


Figure 44: Consistency graph for Sulfur m(%) .

In the round – robin study, 16 laboratories provided a result value for the m(%) concentration of Sulfur in the sample using the D129 Method. From these 16 values the results that laboratories 24 and 92 provided were labeled as outliers from Grubbs' statistical limit and were discarded from the study. The implementation of the E 691 – 99 standard practice method revealed no additional results that should be treated as outliers and the calculation of the precision statistics resulted into a reproducibility value R equal to 0.1133.

RESULTS AND DISCUSSION ABOUT DENSITY (Kg/L) – D1298 Method

P (Number of laboratories)	50	48
X – Mean of the Cell Avg's	0.9023	
Sx – Deviation	0.0015	
Sr – Repeatability Standard Dev.	-	
SR – Reproducibility Standard Dev.	0.0015	
X – Corrected Avg of the cell Avg's		0.9026
SR – Corrected Reproducibility St.Dev.		0.0011
Critical h statistic value	2.64	-2.64
Adjusted critical h value	2.64	-2.64
Reproducibility Estimate (Grubb's)	0.0041	
Reproducibility Estimate R	-	0.0029

Table 14: Statistical results of Density.

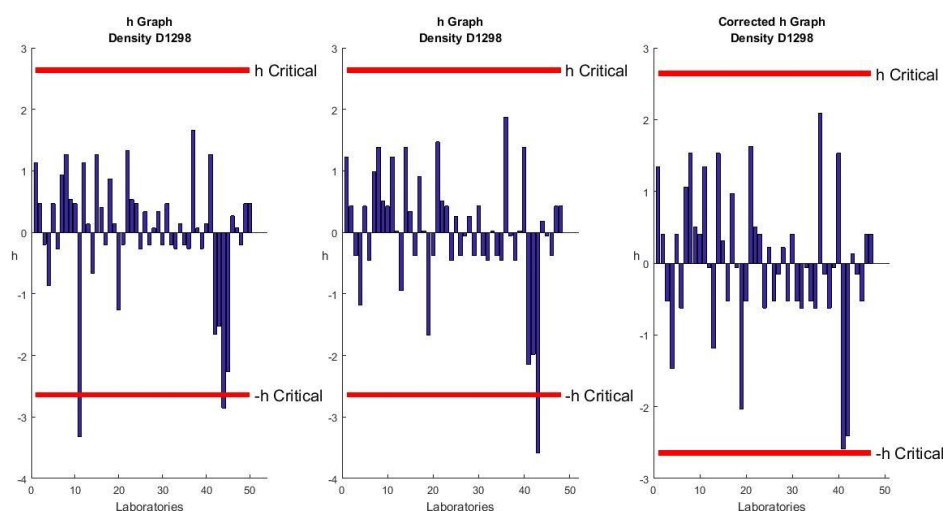


Figure 45: Consistency graph for Density (Kg/L).

Initially, 52 laboratories provided test result values for the density (Kg/L) of the monograde oil sample using the D1298 Method. The Grubbs' statistical limit identified that results from laboratories 41 and 84 violate its limits and therefore discarded them as outliers. The iterative statistical process of E 691 – 99 standard practice method revealed that laboratories 24 and 86 should also be considered as outliers since they both fall below the critical h limit and the calculation of the reproducibility value R was conducted without them. The reproducibility value R was found equal to 0.0029 while the reproducibility value based on the Grubb's Limit provided in the report was equal to 0.0041.

RESULTS AND DISCUSSION ABOUT FLASH POINT – D93 Method

P (Number of laboratories)	46	45
X – Mean of the Cell Avg's	204.585	
Sx – Deviation	5.9471	
Sr – Repeatability Standard Dev.	-	
SR – Reproducibility Standard Dev.	5.9471	
X – Corrected Avg of the cell Avg's		204.9822
SR – Corrected Reproducibility St.Dev.		5.3610
Critical h statistic value	2.64	-2.64
Adjusted critical h value	2.64	-2.64
Reproducibility Estimate(Grubb's)	16.47	
Reproducibility Estimate R	-	15.0107

Table 15: Statistical results of Flash Point.

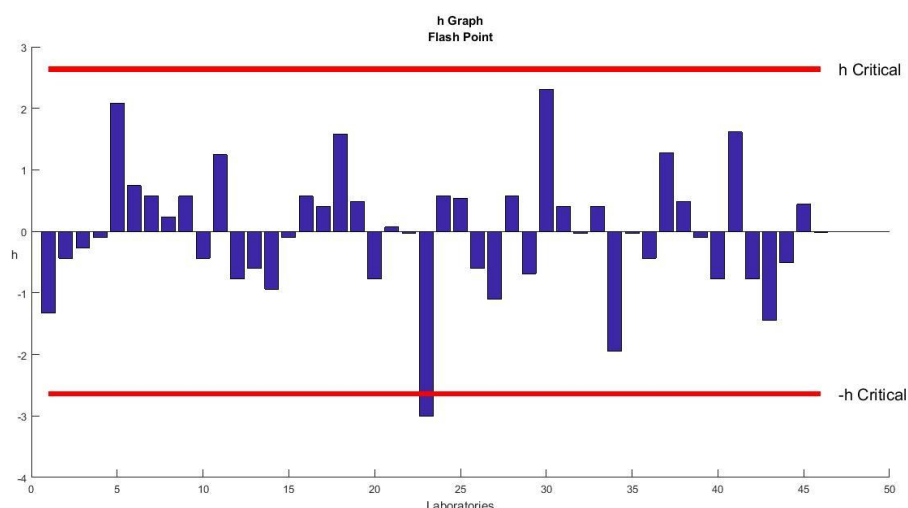


Figure 46: Consistency graph for Flash Point (degrees Celsius).

As far as the flash point of the sample is concerned 53 laboratories participated in the study and provided an estimate for its value based on the D93 Method. Seven of them, laboratories 3, 11, 21, 43, 67, 78, and 79 resulted in a value that was rejected as an outlier from the Grubbs' statistical limit. Furthermore, the implementation of the E 691 – 99 standard practice method revealed that laboratory 47 resulted in a value that violates the critical h limit and therefore it was discarded by the iterative process of the code and was not taken under consideration for the calculation of the reproducibility value R, which was found equal to 15.0107. Below the graph with the corrected h statistic for the Flash Point of the sample is depicted.

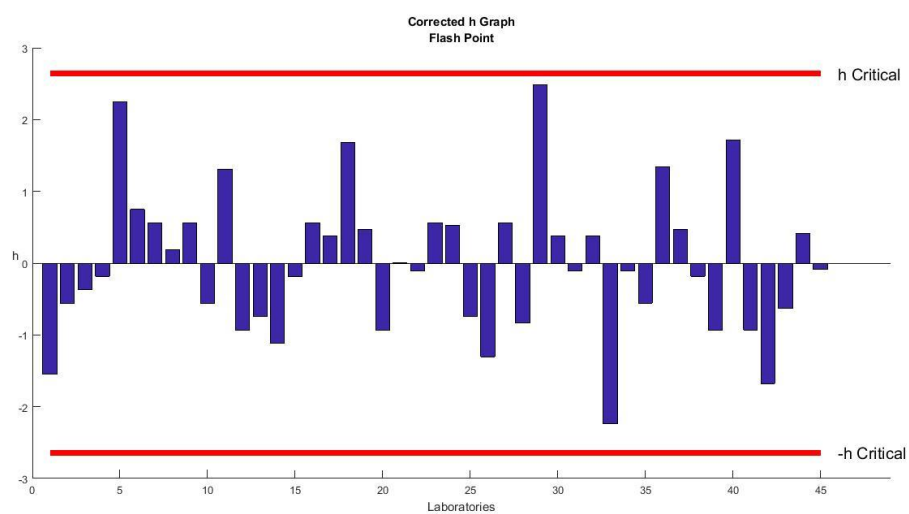


Figure 47: Corrected consistency graph for Flash Point (degrees Celsius).

Property	Xavg	Sx	Sr	SR	r	R
Magnesium	0.0751	0.0061	-	0.0061	-	0.0168
Zinc	0.1173	0.0044	-	0.0044	-	0.0126
Phosphorus	0.1039	0.0062	-	0.0062	-	0.0174
Calcium	0.1118	0.0054	-	0.0054	-	0.0149
Sulfur	0.4193	0.0409	-	0.0409	-	0.1133
Density	0.9026	0.0011	-	0.0011	-	0.0029
Flash Point	204.9822	5.3610	-	5.3610	-	15.0107

Table 16: Summary of Precision Statistics.

CONCLUSIONS-SUGGESTIONS

Statistical process control is an important aspect of an overall quality control procedure which ensures that a process operates efficiently and within the desired production requirements. In analytical measurement processes in laboratories, all data should be produced under a quality control program which will ensure that the results generated from a measurement process are statistically accurate and precise before using them in chemical and geochemical evaluations. The main objective of this work was to evaluate the performance of a Rock – Eval and a GC – MS apparatus based on historical data results produced by the two systems and, subsequently, act as the first step of a small laboratory manual which describes general statistical quality processes followed in laboratories (Ruggedness testing – interlaboratory statistics). Rock – Eval was evaluated based on the direct results produced by the system but also on the fact that even though the system does not operate as optimal as possible, this fact does not significantly change the geochemical characterization of the samples. GC – MS was evaluated not based on the direct measured quantities that the system produces (peak areas / concentrations) but on the behavior of an indirect measurement, the relative response factors. Also it should be mentioned that the samples that were used in GC MS were quality control samples and not check standards so the system was evaluated in terms of precision mainly. The relative response factor is by itself a method of eliminating variability in the response factors and a way to ensure repeatability in sample measurements. The conventional SPC processes that were implemented by the Mat lab code (control charts, parametric tests) theoretically describe the performance of an analytical measurement system sufficiently but the basic assumption underlying is that there is no correlation between the results generated by the process. However, when there is significant autocorrelation in a process, traditional control charts are not that effective and robust. Many refinery and smelting operations have been shown to have auto correlated observations and new – regression – control charts have been developed recently to monitor such processes. A regression based control chart derives from the combination of conventional control charts and a regression analysis and assumes that the values of the dependent variable are linearly related with the values of the independent variable. In light of what was said, a good suggestion for future work would be to monitor for possible autocorrelation between the results produced by the two systems – via a neural networks approach – and accordingly either include a regression control chart in the code or try to “correct” the conventional charts already used in the evaluation of the performance of the two systems.

REFERENCES

1. Standard N. Standard Practice for Applying Statistical Quality Assurance Techniques to Evaluate Analytical Measurement System Performance 1. 2002;5(January 2001):1-29.
2. Cooper G. Basic Lessons in Laboratory Quality Control - QC Workbook. *Bio-Rad Lab*. 2008:62.
3. Karaoglan AD. Graduate School of Natural and Applied Sciences Regression Control Chart for Regression Control Chart for. 2010.
4. Vardeman SB, Jobe JM. : Basics, Measurement, Control, Capability, and Improvement. 2015.
5. ASTM. Standard Guide for Conducting Ruggedness Tests 1. *Astm*. 2015;14(January 1990).
6. Astm. E691 - 12, Standard Practice for Conducting an Interlaboratory Study to Determine the Precision of a Test Method. 2012.
7. Behar F, Beaumont V, De HL, Penteado B. Rock-Eval 6 Technology: Performances and Developments. *Oil Gas Sci Technol – Rev IFP*. 2001.
8. Nuñez-Betelu L, Baceta JI. Basics and application of Rock-Eval/TOC pyrolysis: an example from the uppermost Paleocene/lowermost Eocene in the Basque Basin, Western Pyrenees.
9. Agilent. Fundamentals. *Agil Present*. 2014;42(11):2555-2568.
10. Troubleshooting GC. GC Troubleshooting. *J Infect Dis*. 2013.
11. Pasadakis N. Instrumental methods in the analysis of petroleum fluids (II). (li).
12. Michael Galetakis. *Έλεγχος Ποιότητας Ορυκτών Πρώτων Υλών.*; 2017.
13. Montgomery, Douglas (2005), [Introduction to Statistical Quality Control](#), [Hoboken, New Jersey: John Wiley & Sons](#), pp. 348–349.
14. [Wheeler, Donald J.](#) (2009-05-26), ["When Can We Trust the Limits on a Process Behavior Chart?"](#)
15. Grubbs, Frank E. (1950). "Sample criteria for testing outlying observations".
16. Chambers, John; William Cleveland; Beat Kleiner; Paul Tukey (1983). Graphical Methods for Data Analysis. Duxbury.
17. A Poots, T Woodcock (2012). ["Statistical process control for data without inherent order"](#). BMC Medical Informatics and Decision Making. [London, UK: BioMedCentral](#).

18. "Seven Basic Quality Tools". The Quality Toolbox. Milwaukee, Wisconsin: American Society for Quality. T
19. Westgard JO, Barry PL, Hunt MR, Groth T. A multi-rule Shewhart chart for quality control in clinical chemistry. Clin Chem 1981;27:493-501.
20. "Non-Normality and Tests on Variances". Biometrika. 40(3/4): 318–335.
21. Pasadakis N. "Γεωχημεία Πετρελαίου" (2014) - Εκδόσεις ΤΖΙΟΛΑ.

APPENDIX

GENERAL MATLAB CODE

```
function [] = quality(raw, ARV, R, flag, USL, LSL)
    % Pretreatment
    pretreated = raw - ARV ;
    avg_pretreated = mean(pretreated) ;
    MR_sum = 0 ;
    for i = 2 : size(pretreated,2)

        MR_sum = MR_sum + abs(pretreated(i) - pretreated(i-1)) ;

    end
    MR_avg = MR_sum/(size(pretreated,2) - 1) ;
    R_tonos = 2.46*MR_avg;
    UCL = avg_pretreated + 2.66*MR_avg ;
    LCL = avg_pretreated - 2.66*MR_avg ;

    avg_pretreated_vec = repmat(avg_pretreated, 1, size(pretreated,2)) ;
    UCL_vec = repmat(UCL, 1, size(pretreated,2)) ;
    LCL_vec = repmat(LCL, 1, size(pretreated,2)) ;

    % Plot Run Chart
    figure(1)
    subplot(2,3,1)
    hold on
    box on
    plot(pretreated, '-gs', 'LineWidth',2, 'MarkerSize',10,
'MarkerEdgeColor','b', 'MarkerFaceColor',[0.5,0.5,0.5]) ;
    plot(avg_pretreated_vec, '-k')
    xlabel('Result Sequence Number', 'fontsize',16);
    ylabel('Raw Result - ARV', 'fontsize',16);
    title({flag ; 'Run Chart'}, 'fontsize',16);
    grid on
    hold off
    text(size(raw,2)+0.5,avg_pretreated_vec(1),'I_{avg}')

    % Plot Individual Chart
    figure(2)
    subplot(2,3,2)
    plot(pretreated, '-gs', 'LineWidth',2, 'MarkerSize',10,
'MarkerEdgeColor','b', 'MarkerFaceColor',[0.5,0.5,0.5]);
    hold on
    plot(avg_pretreated_vec, '-k')
    plot(UCL_vec, '--r')
    plot(LCL_vec, '--r')
    xlabel('Sample', 'fontsize',16);
    ylabel('Pretreated', 'fontsize',16);
    title({flag ; 'Individual Chart'}, 'fontsize',16);
    hold off
    grid on
    text(size(raw,2)+0.5,avg_pretreated_vec(1),'I_{avg}')
```

```

text(size(row,2)+0.5,UCL_vec(1),'UCL')
text(size(row,2)+0.5,LCL_vec(1),'LCL')

lambda = 0.4 ;
EWMA = zeros(1, size(pretreated,2)) ;
EWMA(1,1) = pretreated(1,1) ;
for i = 2 : size(EWMA,2)

    EWMA(1,i) = (1 - lambda)*EWMA(1,i-1) + lambda*pretreated(1,i) ;

end
UCL_l = avg_pretreated + 2.66*MR_avg*sqrt(lambda/(2 - lambda)) ;
LCL_l = avg_pretreated - 2.66*MR_avg*sqrt(lambda/(2 - lambda)) ;

UCL_l_vec = repmat(UCL_l, 1, size(pretreated,2)) ;
LCL_l_vec = repmat(LCL_l, 1, size(pretreated,2)) ;

% Plot EWMA Trend Line
figure(3)
subplot(2,3,3)
plot(EWMA, '-gs', 'LineWidth',2, 'MarkerSize',10, 'MarkerEdgeColor','b',
'MarkerFaceColor',[0.5,0.5,0.5]);
hold on
plot(UCL_l_vec, '--r')
plot(LCL_l_vec, '--r')
xlabel('Sample', 'fontsize',16);
ylabel('EWMA', 'fontsize',16);
title({flag ; 'EWMA Trend Line'},'fontsize',16);
hold off
grid on
text(size(row,2)+0.5,UCL_l_vec(1),'UCL \lambda')
text(size(row,2)+0.5,LCL_l_vec(1),'LCL \lambda')

% Plot EWMA Overlay
figure(4)
subplot(2,3,4)
plot(pretreated, '-gs', 'LineWidth',2, 'MarkerSize',10,
'MarkerEdgeColor','b', 'MarkerFaceColor',[0.5,0.5,0.5]);
hold on
plot(avg_pretreated_vec, '-k')
plot(UCL_vec, '-r')
plot(LCL_vec, '-r')
plot(EWMA, '-bs', 'LineWidth',2, 'MarkerSize',10, 'MarkerEdgeColor','b',
'MarkerFaceColor',[0.5,0.5,0.5]);
plot(UCL_l_vec, '--r')
plot(LCL_l_vec, '--r')
xlabel('Sample', 'fontsize',16);
ylabel('EWMA/Pretreated', 'fontsize',16);
title({flag ; 'EWMA Overlay on I'},'fontsize',16);
hold off
grid on
text(size(row,2)+0.5,avg_pretreated_vec(1),'I_{avg}')

```

```

text(size(row,2)+0.5,UCL_vec(1),'UCL')
text(size(row,2)+0.5,LCL_vec(1),'LCL')
text(size(row,2)+0.5,UCL_l_vec(1),'UCL \lambda')
text(size(row,2)+0.5,LCL_l_vec(1),'LCL \lambda')

MR = zeros(1, size(pretreated,2)) ;
MR(1,1) = pretreated(1,1) ;
for i = 2 : size(MR,2)

    MR(1,i) = abs(pretreated(1,i) - pretreated(1,i-1)) ;

end
MR_avg = mean(MR) ;
UCL_MR = 3.27*MR_avg ;

MR_avg_vec = repmat(MR_avg, 1, size(MR,2)) ;
UCL_MR_vec = repmat(UCL_MR, 1, size(MR,2)) ;

% Plot MR Chart
figure(5)
subplot(2,3,5)
plot(MR(2:size(MR,2)), '-gs', 'LineWidth',2, 'MarkerSize',10,
'MarkerEdgeColor','b', 'MarkerFaceColor',[0.5,0.5,0.5]);
hold on
plot(MR_avg_vec, '-k')
plot(UCL_MR_vec, '--r')
hold off
xlabel('Sample', 'fontsize',16);
ylabel('MR', 'fontsize',16);
title({flag ; 'MR Chart'},'fontsize',16);
grid on
text(size(row,2)+0.5,MR_avg_vec(1),'MR_{avg}')
text(size(row,2)+0.5,UCL_MR_vec(1),'UCL_{MR} ')

[h, p ,A, cv] = adtest(row') ;

upper68 = mean(row) + std(row) ;
upper955 = mean(row) + 2*std(row) ;
upper997 = mean(row) + 3*std(row) ;
lower68 = mean(row) - std(row) ;
lower955 = mean(row) - 2*std(row) ;
lower997 = mean(row) - 3*std(row) ;

raw_avg_vec = repmat(mean(row), 1, size(row,2)) ;
upper68_vec = repmat(upper68, 1, size(row,2)) ;
upper955_vec = repmat(upper955, 1, size(row,2)) ;
upper997_vec = repmat(upper997, 1, size(row,2)) ;
lower68_vec = repmat(lower68, 1, size(row,2)) ;
lower955_vec = repmat(lower955, 1, size(row,2)) ;
lower997_vec = repmat(lower997, 1, size(row,2)) ;

```

```

% Plot L-J Chart
figure(6)
subplot(2,3,6)
plot(raw, '-gs', 'LineWidth',2, 'MarkerSize',10, 'MarkerEdgeColor','b',
'MarkerFaceColor',[0.5,0.5,0.5])
hold on
plot(raw_avg_vec, '--k')
plot(upper68_vec, '--r')
plot(lower68_vec, '--r')
plot(upper955_vec, '--m')
plot(lower955_vec, '--m')
plot(upper997_vec, '--g')
plot(lower997_vec, '--g')
hold off
xlabel('Sample', 'fontsize',16);
ylabel('Raw result', 'fontsize',16);
title({flag ; 'L-J Chart'}, 'fontsize',16);
grid on
text(size(raw,2)+0.5,raw_avg_vec(1), 'm')
text(size(raw,2)+0.5,upper68_vec(1), 'm+S')
text(size(raw,2)+0.5,lower68_vec(1), 'm-S')
text(size(raw,2)+0.5,upper955_vec(1), 'm+2S')
text(size(raw,2)+0.5,lower955_vec(1), 'm-2S')
text(size(raw,2)+0.5,upper997_vec(1), 'm+3S')
text(size(raw,2)+0.5,lower997_vec(1), 'm-3S')

% Plot Histogram
figure(7)
hist(raw,18)
title({flag ; 'Histogram'}, 'fontsize',16);
xlabel('Bins', 'fontsize',16);
ylabel('Frequency', 'fontsize',16);

% Process Control and Process Capability
sigma = MR_avg/1.128
mew = mean(raw);
es = std(raw);

c_p = (USL-LSL)/(6*sigma)
c_pk = min((USL-mew)/(3*sigma), (mew-LSL)/(3*sigma))
p_p = (USL-LSL)/(6*es)
p_pk = min((USL-mew)/(3*es), (mew-LSL)/(3*es));
z_up = (USL-mew)/es
z_low = (mew-LSL)/es
sigma_R = sqrt((sum((pretreated-
avg_pretreated).^2))/(size(pretreated,2)-1))

[h, p, ci, stats] = ttest(raw', ARV) ;
t = stats.tstat ;

n = size(pretreated,2) ;
X = (n-1)*(2.46*MR_avg)^2/(2*R^2) ;

```



```

A
t
MR_avg
R_tonos
X

% L - J WESTGARD RULES -----
-----

fprintf('%s\r\n',' ')

% RULE 1_2_S

rule_name = 'Rule 1-2S';
fprintf('%s\r\n', rule_name)

n_clean = 0;
for i = 1:size(raw,2)
    if raw(1,i)>upper955_vec(i)
        fprintf('%s','Positive violation at sample number ',num2str(i))
        fprintf('%s\r\n',' ')
        n_clean = n_clean + 1;
    end
    if raw(1,i)<lower955_vec(i)
        fprintf('%s','Negative violation at sample number ',num2str(i))
        fprintf('%s\r\n',' ')
        n_clean = n_clean + 1;
    end
end
if n_clean == 0
    fprintf('%s\r\n', 'No violations')
end
fprintf('%s\r\n',' ')

% RULE 1_3_S

rule_name = 'Rule 1-3S';
fprintf('%s\r\n', rule_name)

n_clean = 0;
for i = 1:size(raw,2)
    if raw(1,i)>upper997_vec(i)
        fprintf('%s','Positive violation at sample number ',num2str(i))
        fprintf('%s\r\n',' ')
        n_clean = n_clean + 1;
    end
    if raw(1,i)<lower997_vec(i)
        fprintf('%s','Negative violation at sample number ',num2str(i))
        fprintf('%s\r\n',' ')
        n_clean = n_clean + 1;
    end
end
if n_clean == 0

```

```

        fprintf('%s\r\n', 'No violations')
    end
    fprintf('%s\r\n', ' ')

    % RULE 2_2_S

    rule_name = 'Rule 2-2S';
    fprintf('%s\r\n', rule_name)

    n_clean = 0;
    for i = 1:(size(raw,2)-1)
        if raw(1,i)>upper955_vec(i) && raw(1,i+1)>upper955_vec(i+1)
            fprintf('%s','Positive violation at sample numbers',num2str(i),' ',num2str(i+1))
            fprintf('%s\r\n', ' ')
            n_clean = n_clean + 1;
        end
        if raw(1,i)<lower955_vec(i) && raw(1,i+1)<lower955_vec(i+1)
            fprintf('%s','Negative violation at sample numbers',num2str(i),' ',num2str(i+1))
            fprintf('%s\r\n', ' ')
            n_clean = n_clean + 1;
        end
    end
    if n_clean == 0
        fprintf('%s\r\n', 'No violations')
    end
    fprintf('%s\r\n', ' ')

    % RULE R_4_S

    rule_name = 'Rule R-4S';
    fprintf('%s\r\n', rule_name)

    n_clean = 0;
    for i = 1:(size(raw,2)-1)
        if abs(raw(1,i)-raw(1,i+1))>abs(upper955_vec(i)-lower955_vec(i))
            fprintf('%s','Violation at sample numbers ',num2str(i),' ',num2str(i+1))
            fprintf('%s\r\n', ' ')
            n_clean = n_clean + 1;
        end
    end
    if n_clean == 0
        fprintf('%s\r\n', 'No violations')
    end
    fprintf('%s\r\n', ' ')

    % RULE 3_1_S

    rule_name = 'Rule 3-1S';
    fprintf('%s\r\n', rule_name)

    n_clean = 0;

```

```

    for i = 1:(size(raw,2)-2)
        if raw(1,i)>upper68_vec(i) && raw(1,i+1)>upper68_vec(i+1) &&
raw(1,i+2)>upper68_vec(i+2)
            fprintf('%s','Positive violation at sample numbers
',num2str(i),' ',num2str(i+1),' ',num2str(i+2))
            fprintf('%s\r\n',' ')
            n_clean = n_clean + 1;
        end
        if raw(1,i)<lower68_vec(i) && raw(1,i+1)<lower68_vec(i+1) &&
raw(1,i+2)<lower68_vec(i+2)
            fprintf('%s','Negative violation at sample numbers
',num2str(i),' ',num2str(i+1),' ',num2str(i+2))
            fprintf('%s\r\n',' ')
            n_clean = n_clean + 1;
        end
    end
    if n_clean == 0
        fprintf('%s\r\n', 'No violations')
    end
    fprintf('%s\r\n',' ')

% RULE 4_1_S

rule_name = 'Rule 4-1S';
fprintf('%s\r\n', rule_name)

n_clean = 0;
for i = 1:(size(raw,2)-3)
    if raw(1,i)>upper68_vec(i) && raw(1,i+1)>upper68_vec(i+1) &&
raw(1,i+2)>upper68_vec(i+2) && raw(1,i+3)>upper68_vec(i+3)
        fprintf('%s','Positive violation at sample numbers
',num2str(i),' ',num2str(i+1),' ',num2str(i+2),' ',num2str(i+3))
        fprintf('%s\r\n',' ')
        n_clean = n_clean + 1;
    end
    if raw(1,i)<lower68_vec(i) && raw(1,i+1)<lower68_vec(i+1) &&
raw(1,i+2)<lower68_vec(i+2) && raw(1,i+3)<lower68_vec(i+3)
        fprintf('%s','Negative violation at sample numbers
',num2str(i),' ',num2str(i+1),' ',num2str(i+2),' ',num2str(i+3))
        fprintf('%s\r\n',' ')
        n_clean = n_clean + 1;
    end
end
if n_clean == 0
    fprintf('%s\r\n', 'No violations')
end
fprintf('%s\r\n',' ')

% RULE 8_X

rule_name = 'Rule 8-X';
fprintf('%s\r\n', rule_name)

n_clean = 0;

```

```

    for i = 1:(size(row,2)-7)
        if row(1,i)>row_avg_vec(i) && row(1,i+1)>row_avg_vec(i+1) &&
row(1,i+2)>row_avg_vec(i+2) && row(1,i+3)>row_avg_vec(i+3) &&
row(1,i+4)>row_avg_vec(i+4) && row(1,i+5)>row_avg_vec(i+5) &&
row(1,i+6)>row_avg_vec(i+6) && row(1,i+7)>row_avg_vec(i+7)
            fprintf('%s','Positive violation at sample numbers
',num2str(i),' ',num2str(i+1),' ',num2str(i+2),' ',num2str(i+3),'
',num2str(i+4),' ',num2str(i+5),' ',num2str(i+6),' ',num2str(i+7))
            fprintf('%s\r\n',' ')
            n_clean = n_clean + 1;
        end
        if row(1,i)<row_avg_vec(i) && row(1,i+1)<row_avg_vec(i+1) &&
row(1,i+2)<row_avg_vec(i+2) && row(1,i+3)<row_avg_vec(i+3) &&
row(1,i+4)<row_avg_vec(i+4) && row(1,i+5)<row_avg_vec(i+5) &&
row(1,i+6)<row_avg_vec(i+6) && row(1,i+7)<row_avg_vec(i+7)
            fprintf('%s','Negative violation at sample numbers
',num2str(i),' ',num2str(i+1),' ',num2str(i+2),' ',num2str(i+3),'
',num2str(i+4),' ',num2str(i+5),' ',num2str(i+6),' ',num2str(i+7))
            fprintf('%s\r\n',' ')
            n_clean = n_clean + 1;
        end
    end
    if n_clean == 0
        fprintf('%s\r\n', 'No violations')
    end
    fprintf('%s\r\n',' ')

    % RULE 10_X

    rule_name = 'Rule 10-X';
    fprintf('%s\r\n', rule_name)

    n_clean = 0;
    for i = 1:(size(row,2)-9)
        if row(1,i)>row_avg_vec(i) && row(1,i+1)>row_avg_vec(i+1) &&
row(1,i+2)>row_avg_vec(i+2) && row(1,i+3)>row_avg_vec(i+3) &&
row(1,i+4)>row_avg_vec(i+4) && row(1,i+5)>row_avg_vec(i+5) &&
row(1,i+6)>row_avg_vec(i+6) && row(1,i+7)>row_avg_vec(i+7) &&
row(1,i+8)>row_avg_vec(i+8) && row(1,i+9)>row_avg_vec(i+9)
            fprintf('%s','Positive violation at sample numbers
',num2str(i),' ',num2str(i+1),' ',num2str(i+2),' ',num2str(i+3),'
',num2str(i+4),' ',num2str(i+5),' ',num2str(i+6),' ',num2str(i+7),'
',num2str(i+8),' ',num2str(i+9))
            fprintf('%s\r\n',' ')
            n_clean = n_clean + 1;
        end
        if row(1,i)<row_avg_vec(i) && row(1,i+1)<row_avg_vec(i+1) &&
row(1,i+2)<row_avg_vec(i+2) && row(1,i+3)<row_avg_vec(i+3) &&
row(1,i+4)<row_avg_vec(i+4) && row(1,i+5)<row_avg_vec(i+5) &&
row(1,i+6)<row_avg_vec(i+6) && row(1,i+7)<row_avg_vec(i+7) &&
row(1,i+8)<row_avg_vec(i+8) && row(1,i+9)<row_avg_vec(i+9)
            fprintf('%s','Negative violation at sample numbers
',num2str(i),' ',num2str(i+1),' ',num2str(i+2),' ',num2str(i+3),'

```

```

',num2str(i+4),' ',num2str(i+5),' ',num2str(i+6),' ',num2str(i+7),'
',num2str(i+8),' ',num2str(i+9))
    fprintf('%s\r\n',' ')
    n_clean = n_clean + 1;
end
end
if n_clean == 0
    fprintf('%s\r\n', 'No violations')
end
fprintf('%s\r\n',' ')

% RULE 12_X

rule_name = 'Rule 12-X';
fprintf('%s\r\n', rule_name)

n_clean = 0;
for i = 1:(size(raw,2)-11)
    if raw(1,i)>raw_avg_vec(i) && raw(1,i+1)>raw_avg_vec(i+1) &&
raw(1,i+2)>raw_avg_vec(i+2) && raw(1,i+3)>raw_avg_vec(i+3) &&
raw(1,i+4)>raw_avg_vec(i+4) && raw(1,i+5)>raw_avg_vec(i+5) &&
raw(1,i+6)>raw_avg_vec(i+6) && raw(1,i+7)>raw_avg_vec(i+7) &&
raw(1,i+8)>raw_avg_vec(i+8) && raw(1,i+9)>raw_avg_vec(i+9) &&
raw(1,i+10)>raw_avg_vec(i+10) && raw(1,i+11)>raw_avg_vec(i+11)
        fprintf('%s','Positive violation at sample numbers
',num2str(i),' ',num2str(i+1),' ',num2str(i+2),' ',num2str(i+3),'
',num2str(i+4),' ',num2str(i+5),' ',num2str(i+6),' ',num2str(i+7),'
',num2str(i+8),' ',num2str(i+9),' ',num2str(i+10),' ',num2str(i+11))
        fprintf('%s\r\n',' ')
        n_clean = n_clean + 1;
    end
    if raw(1,i)<raw_avg_vec(i) && raw(1,i+1)<raw_avg_vec(i+1) &&
raw(1,i+2)<raw_avg_vec(i+2) && raw(1,i+3)<raw_avg_vec(i+3) &&
raw(1,i+4)<raw_avg_vec(i+4) && raw(1,i+5)<raw_avg_vec(i+5) &&
raw(1,i+6)<raw_avg_vec(i+6) && raw(1,i+7)<raw_avg_vec(i+7) &&
raw(1,i+8)<raw_avg_vec(i+8) && raw(1,i+9)<raw_avg_vec(i+9) &&
raw(1,i+10)<raw_avg_vec(i+10) && raw(1,i+11)<raw_avg_vec(i+11)
        fprintf('%s','Negative violation at sample numbers
',num2str(i),' ',num2str(i+1),' ',num2str(i+2),' ',num2str(i+3),'
',num2str(i+4),' ',num2str(i+5),' ',num2str(i+6),' ',num2str(i+7),'
',num2str(i+8),' ',num2str(i+9),' ',num2str(i+10),' ',num2str(i+11))
        fprintf('%s\r\n',' ')
        n_clean = n_clean + 1;
    end
end
if n_clean == 0
    fprintf('%s\r\n', 'No violations')
end
fprintf('%s\r\n',' ')

end

```

INTERLABORATORY STATISTICS CODE

```
function [] = praxis(A, h_crit, name)

flag = 0;
hist_counter = 0;
while (flag == 0)

n = size(A,2);
P = size(A,1);
hist_counter = hist_counter+1;
counter = 0;
cell_avg = zeros(P,1);
cell_std = zeros(P,1);

pos(1:P) = h_crit;
neg = -pos;

for i = 1:P
    cell_avg(i) = sum(A(i,:));
    cell_std = sqrt(sum((A(i,:)-cell_avg(i)).^2)/(n-1));
end

cell_std(isnan(cell_std)) = 0;
tot_avg = mean(cell_avg);

d = cell_avg-tot_avg

S_tot_avg = sqrt(sum(d.^2)/(P-1))
S_r = sqrt(sum(cell_std.^2)/P);
S_R = sqrt(S_tot_avg^2+(S_r^2)*((n-1)/n))
h = d./(S_tot_avg);

figure(hist_counter)
hold on
bar(h)
plot(pos,'LineWidth',5,'color',[1 0 0])
plot(neg,'LineWidth',5,'color',[1 0 0])
hold off
title({'h Graph' ; name})
xlabel('Laboratories')
ylabel('h')
xlim([0 P+4])
set(get(gca,'YLabel'),'Rotation',0)
text(P+1,h_crit,'h Critical','FontSize',15)
text(P+1,-h_crit,'-h Critical','FontSize',15)
for i = 1:P
    if h(i)>h_crit || h(i)<-h_crit
        A(i) = [];
        pos(i) = [];
    else
        counter = counter + 1;
    end
end
end
if counter == P
```

```

        flag = 1;
    end
    end
    if hist_counter ~= 1
        title({'Corrected h Graph' ; name})
    end
    k = cell_std/S_r;
    r = 2.8*S_r;
    R = 2.8*S_R
end

```

TABLE A1.10 95th Percentile of Student's *t* Distribution

Degrees of Freedom	<i>t</i>
1	12.7062
2	4.3027
3	3.1824
4	2.7764
5	2.5706
6	2.4469
7	2.3646
8	2.3060
9	2.2622
10	2.2281
11	2.2010
12	2.1788
13	2.1604
14	2.1448
15	2.1314
16	2.1199
17	2.1098
18	2.1009
19	2.0930
20	2.0860
21	2.0796

Table 19: 95th Percentile of Student's Distribution.

	n																	
Order	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
1	-1.83	-1.86	-1.89	-1.91	-1.94	-1.96	-1.98	-2.00	-2.02	-2.04	-2.05	-2.07	-2.09	-2.10	-2.11	-2.13	-2.14	-2.15
2	-1.28	-1.32	-1.35	-1.38	-1.41	-1.44	-1.47	-1.49	-1.51	-1.53	-1.55	-1.57	-1.59	-1.61	-1.63	-1.64	-1.66	-1.68
3	-0.97	-1.01	-1.05	-1.09	-1.12	-1.15	-1.18	-1.21	-1.23	-1.26	-1.28	-1.30	-1.32	-1.35	-1.36	-1.38	-1.40	-1.42
4	-0.73	-0.78	-0.82	-0.86	-0.90	-0.93	-0.97	-1.00	-1.03	-1.05	-1.08	-1.10	-1.13	-1.15	-1.17	-1.19	-1.21	-1.23
5	-0.52	-0.58	-0.63	-0.67	-0.72	-0.76	-0.79	-0.83	-0.86	-0.89	-0.92	-0.94	-0.97	-0.99	-1.01	-1.04	-1.06	-1.08
6	-0.34	-0.40	-0.46	-0.51	-0.55	-0.60	-0.64	-0.67	-0.71	-0.74	-0.77	-0.80	-0.83	-0.85	-0.88	-0.90	-0.93	-0.95
7	-0.17	-0.24	-0.30	-0.36	-0.41	-0.45	-0.50	-0.54	-0.58	-0.61	-0.64	-0.67	-0.70	-0.73	-0.76	-0.78	-0.81	-0.83
8	0.00	-0.08	-0.15	-0.21	-0.27	-0.32	-0.37	-0.41	-0.45	-0.49	-0.52	-0.56	-0.59	-0.62	-0.65	-0.67	-0.70	-0.72
9	0.17	0.08	0.00	-0.07	-0.13	-0.19	-0.24	-0.29	-0.33	-0.37	-0.41	-0.45	-0.48	-0.51	-0.54	-0.57	-0.60	-0.63
10	0.34	0.24	0.15	0.07	0.00	-0.06	-0.12	-0.17	-0.22	-0.26	-0.31	-0.34	-0.38	-0.41	-0.45	-0.48	-0.51	-0.53
11	0.52	0.40	0.30	0.21	0.13	0.06	0.00	-0.06	-0.11	-0.16	-0.20	-0.24	-0.28	-0.32	-0.35	-0.39	-0.42	-0.45
12	0.73	0.58	0.46	0.36	0.27	0.19	0.12	0.06	0.00	-0.05	-0.10	-0.15	-0.19	-0.23	-0.26	-0.30	-0.33	-0.36
13	0.97	0.78	0.63	0.51	0.41	0.32	0.24	0.17	0.11	0.05	0.00	-0.05	-0.09	-0.13	-0.17	-0.21	-0.25	-0.28
14	1.28	1.01	0.82	0.67	0.55	0.45	0.37	0.29	0.22	0.16	0.10	0.05	0.00	-0.04	-0.09	-0.13	-0.16	-0.20
15	1.83	1.32	1.05	0.86	0.72	0.60	0.50	0.41	0.33	0.26	0.20	0.15	0.09	0.04	0.00	-0.04	-0.08	-0.12
16		1.86	1.35	1.09	0.90	0.76	0.64	0.54	0.45	0.37	0.31	0.24	0.19	0.13	0.09	0.04	0.00	-0.04
17			1.89	1.38	1.12	0.93	0.79	0.67	0.58	0.49	0.41	0.34	0.28	0.23	0.17	0.13	0.08	0.04
18				1.91	1.41	1.15	0.97	0.83	0.71	0.61	0.52	0.45	0.38	0.32	0.26	0.21	0.16	0.12
19					1.94	1.44	1.18	1.00	0.86	0.74	0.64	0.56	0.48	0.41	0.35	0.30	0.25	0.20
20						1.96	1.47	1.21	1.03	0.89	0.77	0.67	0.59	0.51	0.45	0.39	0.33	0.28
21							1.98	1.49	1.23	1.05	0.92	0.80	0.70	0.62	0.54	0.48	0.42	0.36
22								2.00	1.51	1.26	1.08	0.94	0.83	0.73	0.65	0.57	0.51	0.45
23									2.02	1.53	1.28	1.10	0.97	0.85	0.76	0.67	0.60	0.53
24										2.04	1.55	1.30	1.13	0.99	0.88	0.78	0.70	0.63
25											2.05	1.57	1.32	1.15	1.01	0.90	0.81	0.72
26												2.07	1.59	1.35	1.17	1.04	0.93	0.83
27													2.09	1.61	1.36	1.19	1.06	0.95
28														2.10	1.63	1.38	1.21	1.08
29															2.11	1.64	1.40	1.23
30																2.13	1.66	1.42
31																	2.14	1.68
32																		2.15

FIG. A1.4 z-Values

Table 17: Z – Values used in Anderson – Darling test.

	-0.09	-0.08	-0.07	-0.06	-0.05	-0.04	-0.03	-0.02	-0.01	0.00
-3.5	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002
-3.4	0.0002	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003
-3.3	0.0003	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0005	0.0005	0.0005
-3.2	0.0005	0.0005	0.0005	0.0006	0.0006	0.0006	0.0006	0.0006	0.0007	0.0007
-3.1	0.0007	0.0007	0.0008	0.0008	0.0008	0.0008	0.0009	0.0009	0.0009	0.0010
-3.0	0.0010	0.0010	0.0011	0.0011	0.0011	0.0012	0.0012	0.0013	0.0013	0.0013
-2.9	0.0014	0.0014	0.0015	0.0015	0.0016	0.0016	0.0017	0.0018	0.0018	0.0019
-2.8	0.0019	0.0020	0.0021	0.0021	0.0022	0.0023	0.0023	0.0024	0.0025	0.0026
-2.7	0.0026	0.0027	0.0028	0.0029	0.0030	0.0031	0.0032	0.0033	0.0034	0.0035
-2.6	0.0036	0.0037	0.0038	0.0039	0.0040	0.0041	0.0043	0.0044	0.0045	0.0047
-2.5	0.0048	0.0049	0.0051	0.0052	0.0054	0.0055	0.0057	0.0059	0.0060	0.0062
-2.4	0.0064	0.0066	0.0068	0.0069	0.0071	0.0073	0.0075	0.0078	0.0080	0.0082
-2.3	0.0084	0.0087	0.0089	0.0091	0.0094	0.0096	0.0099	0.0102	0.0104	0.0107
-2.2	0.0110	0.0113	0.0116	0.0119	0.0122	0.0125	0.0129	0.0132	0.0136	0.0139
-2.1	0.0143	0.0146	0.0150	0.0154	0.0158	0.0162	0.0166	0.0170	0.0174	0.0179
-2.0	0.0183	0.0188	0.0192	0.0197	0.0202	0.0207	0.0212	0.0217	0.0222	0.0228
-1.9	0.0233	0.0239	0.0244	0.0250	0.0256	0.0262	0.0268	0.0274	0.0281	0.0287
-1.8	0.0294	0.0301	0.0307	0.0314	0.0322	0.0329	0.0336	0.0344	0.0351	0.0359
-1.7	0.0367	0.0375	0.0384	0.0392	0.0401	0.0409	0.0418	0.0427	0.0436	0.0446
-1.6	0.0455	0.0465	0.0475	0.0485	0.0495	0.0505	0.0516	0.0526	0.0537	0.0548
-1.5	0.0559	0.0571	0.0582	0.0594	0.0606	0.0618	0.0630	0.0643	0.0655	0.0668
-1.4	0.0681	0.0694	0.0708	0.0721	0.0735	0.0749	0.0764	0.0778	0.0793	0.0808
-1.3	0.0823	0.0838	0.0853	0.0869	0.0885	0.0901	0.0918	0.0934	0.0951	0.0968
-1.2	0.0985	0.1003	0.1020	0.1038	0.1056	0.1075	0.1093	0.1112	0.1131	0.1151
-1.1	0.1170	0.1190	0.1210	0.1230	0.1251	0.1271	0.1292	0.1314	0.1335	0.1357
-1	0.1379	0.1401	0.1423	0.1446	0.1469	0.1492	0.1515	0.1539	0.1562	0.1587
-0.9	0.1611	0.1635	0.1660	0.1685	0.1711	0.1736	0.1762	0.1788	0.1814	0.1841
-0.8	0.1867	0.1894	0.1922	0.1949	0.1977	0.2005	0.2033	0.2061	0.2090	0.2119
-0.7	0.2148	0.2177	0.2206	0.2236	0.2266	0.2296	0.2327	0.2358	0.2389	0.2420
-0.6	0.2451	0.2483	0.2514	0.2546	0.2578	0.2611	0.2643	0.2676	0.2709	0.2743
-0.5	0.2776	0.2810	0.2843	0.2877	0.2912	0.2946	0.2981	0.3015	0.3050	0.3085
-0.4	0.3121	0.3156	0.3192	0.3228	0.3264	0.3300	0.3336	0.3372	0.3409	0.3446
-0.3	0.3483	0.3520	0.3557	0.3594	0.3632	0.3669	0.3707	0.3745	0.3783	0.3821
-0.2	0.3859	0.3897	0.3936	0.3974	0.4013	0.4052	0.4090	0.4129	0.4168	0.4207
-0.1	0.4247	0.4286	0.4325	0.4364	0.4404	0.4443	0.4483	0.4522	0.4562	0.4602
0.0	0.4641	0.4681	0.4721	0.4761	0.4801	0.4840	0.4880	0.4920	0.4960	0.5000

Note—Probability ($z < w_i$), where w_i is the sum of the number in the left column and top row.

FIG. A1.5 p_i Values

Table 18: P_i Values used in Anderson – Darling test.

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998

FIG. A1.5 p_i Values (continued)

Table 19: P_i Values used in Anderson – Darling test (continued).

TABLE A1.11 95th Percentiles of the Chi Square Distribution	
Degrees Freedom	χ^2
7	14.1
8	15.5
9	16.9
10	18.3
11	19.7
12	21.0
13	22.4
14	23.7
15	25.0
16	26.3
17	27.6
18	28.9
19	30.1
20	31.4
21	32.7
22	33.9
23	35.2
24	36.4
25	37.7
26	38.9
27	40.1
28	41.3
30	43.8
35	49.8
40	55.8
45	61.7
50	67.5
60	79.1
70	90.5
80	101.9

Table 20: 95th Percentile of the Chi Square Distribution.

TABLE A1.12 97.5 Percentiles of the F Statistic

Denominator, degrees of freedom	Numerator													
	7	8	9	10	12	14	16	18	20	25	30	40	50	100
7	4.99	4.90	4.82	4.76	4.67	4.60	4.54	4.50	4.47	4.40	4.36	4.31	4.28	4.21
8	4.53	4.43	4.36	4.30	4.20	4.13	4.08	4.03	4.00	3.94	3.89	3.84	3.81	3.74
9	4.20	4.10	4.03	3.96	3.87	3.80	3.74	3.70	3.67	3.60	3.56	3.51	3.47	3.40
10	3.95	3.85	3.78	3.72	3.62	3.55	3.50	3.45	3.42	3.35	3.31	3.26	3.22	3.15
11	3.76	3.66	3.59	3.53	3.43	3.36	3.30	3.26	3.23	3.16	3.12	3.06	3.03	2.96
12	3.61	3.51	3.44	3.37	3.28	3.21	3.15	3.11	3.07	3.01	2.96	2.91	2.87	2.80
13	3.48	3.39	3.31	3.25	3.15	3.08	3.03	2.98	2.95	2.88	2.84	2.78	2.74	2.67
14	3.38	3.29	3.21	3.15	3.05	2.98	2.92	2.88	2.84	2.78	2.73	2.67	2.64	2.56
15	3.29	3.20	3.12	3.06	2.96	2.89	2.84	2.79	2.76	2.69	2.64	2.59	2.55	2.47
16	3.22	3.12	3.05	2.99	2.89	2.82	2.76	2.72	2.68	2.61	2.57	2.51	2.47	2.40
17	3.16	3.06	2.98	2.92	2.82	2.75	2.70	2.65	2.62	2.55	2.50	2.44	2.41	2.33
18	3.10	3.01	2.93	2.87	2.77	2.70	2.64	2.60	2.56	2.49	2.44	2.38	2.35	2.27
19	3.05	2.96	2.88	2.82	2.72	2.65	2.59	2.55	2.51	2.44	2.39	2.33	2.30	2.22
20	3.01	2.91	2.84	2.77	2.68	2.60	2.55	2.50	2.46	2.40	2.35	2.29	2.25	2.17
25	2.85	2.75	2.68	2.61	2.51	2.44	2.38	2.34	2.30	2.23	2.18	2.12	2.08	2.00
30	2.75	2.65	2.57	2.51	2.41	2.34	2.28	2.23	2.20	2.12	2.07	2.01	1.97	1.88
35	2.68	2.58	2.50	2.44	2.34	2.27	2.21	2.16	2.12	2.05	2.00	1.93	1.89	1.80
40	2.62	2.53	2.45	2.39	2.29	2.21	2.15	2.11	2.07	1.99	1.94	1.88	1.83	1.74
45	2.58	2.49	2.41	2.35	2.25	2.17	2.11	2.07	2.03	1.95	1.90	1.83	1.79	1.69
50	2.55	2.46	2.38	2.32	2.22	2.14	2.08	2.03	1.99	1.92	1.87	1.80	1.75	1.66
60	2.51	2.41	2.33	2.27	2.17	2.09	2.03	1.98	1.94	1.87	1.82	1.74	1.70	1.60
70	2.47	2.38	2.30	2.24	2.14	2.06	2.00	1.95	1.91	1.83	1.78	1.71	1.66	1.56
80	2.45	2.35	2.28	2.21	2.11	2.03	1.97	1.92	1.88	1.81	1.75	1.68	1.63	1.53
90	2.43	2.34	2.26	2.19	2.09	2.02	1.95	1.91	1.86	1.79	1.73	1.66	1.61	1.50
100	2.42	2.32	2.24	2.18	2.08	2.00	1.94	1.89	1.85	1.77	1.71	1.64	1.59	1.48

Table 21: 97.5th Percentile of the F Statistic.TABLE 5 Critical Values of h and k at the 0.5 % Significance Level^A

Critical value of h	p	Critical values of k Number of replicates, n								
		2	3	4	5	6	7	8	9	10
1.15	3	1.72	1.67	1.61	1.56	1.52	1.49	1.47	1.44	1.42
1.49	4	1.95	1.82	1.73	1.66	1.60	1.56	1.53	1.50	1.47
1.74	5	2.11	1.92	1.79	1.71	1.65	1.60	1.56	1.53	1.50
1.92	6	2.22	1.98	1.84	1.75	1.68	1.63	1.59	1.55	1.52
2.05	7	2.30	2.03	1.87	1.77	1.70	1.65	1.60	1.57	1.54
2.15	8	2.36	2.06	1.90	1.79	1.72	1.66	1.62	1.58	1.55
2.23	9	2.41	2.09	1.92	1.81	1.73	1.67	1.62	1.59	1.56
2.29	10	2.45	2.11	1.93	1.82	1.74	1.68	1.63	1.59	1.56
2.34	11	2.49	2.13	1.94	1.83	1.75	1.69	1.64	1.60	1.57
2.38	12	2.51	2.14	1.96	1.84	1.76	1.69	1.64	1.60	1.57
2.41	13	2.54	2.15	1.96	1.84	1.76	1.70	1.65	1.61	1.58
2.44	14	2.56	2.16	1.97	1.85	1.77	1.70	1.65	1.61	1.58
2.47	15	2.57	2.17	1.98	1.86	1.77	1.71	1.66	1.62	1.58
2.49	16	2.59	2.18	1.98	1.86	1.77	1.71	1.66	1.62	1.58
2.51	17	2.60	2.19	1.99	1.86	1.78	1.71	1.66	1.62	1.59
2.53	18	2.61	2.20	1.99	1.87	1.78	1.72	1.66	1.62	1.59
2.54	19	2.62	2.20	2.00	1.87	1.78	1.72	1.67	1.62	1.59
2.56	20	2.63	2.21	2.00	1.87	1.79	1.72	1.67	1.63	1.59
2.57	21	2.64	2.21	2.00	1.88	1.79	1.72	1.67	1.63	1.59
2.58	22	2.65	2.21	2.01	1.88	1.79	1.72	1.67	1.63	1.59
2.59	23	2.66	2.22	2.01	1.88	1.79	1.72	1.67	1.63	1.59
2.60	24	2.66	2.22	2.01	1.88	1.79	1.73	1.67	1.63	1.60
2.61	25	2.67	2.23	2.01	1.88	1.79	1.73	1.67	1.63	1.60
2.62	26	2.67	2.23	2.02	1.89	1.80	1.73	1.68	1.63	1.60
2.62	27	2.68	2.23	2.02	1.89	1.80	1.73	1.68	1.63	1.60
2.63	28	2.68	2.23	2.02	1.89	1.80	1.73	1.68	1.63	1.60
2.64	29	2.69	2.24	2.02	1.89	1.80	1.73	1.68	1.64	1.60
2.64	30	2.69	2.24	2.02	1.89	1.80	1.73	1.68	1.64	1.60

^AThe above critical values for the h and k consistency statistics were calculated from Student's t and the F -ratio using the following relationships:

$$h = \frac{(p-1)t\sqrt{p(t^2 + p-2)}}{\sqrt{p[1 + (p-1)F]}} \quad t \text{ with } p-2 \text{ degrees of freedom, and}$$

$$k = \frac{(p-1)t\sqrt{p(t^2 + p-2)}}{\sqrt{p[1 + (p-1)F]}} \quad F \text{ with } n-1 \text{ and } (p-1)(n-1) \text{ degrees of freedom.}$$

Table 22: h and k critical values at the 0.5% Significance Level.

Standard normal table

z	x.x0	x.x1	x.x2	x.x3	x.x4	x.x5	x.x6	x.x7	x.x8	x.x9
4.0	.00003									
3.9	.00005	.00005	.00004	.00004	.00004	.00004	.00004	.00004	.00003	.00003
3.8	.00007	.00007	.00007	.00006	.00006	.00006	.00006	.00005	.00005	.00005
3.7	.00011	.00010	.00010	.00010	.00009	.00009	.00008	.00008	.00008	.00008
3.6	.00016	.00015	.00015	.00014	.00014	.00013	.00013	.00012	.00012	.00011
3.5	.00023	.00022	.00022	.00021	.00020	.00019	.00019	.00018	.00017	.00017
3.4	.00034	.00032	.00031	.00030	.00029	.00028	.00027	.00026	.00025	.00024
3.3	.00048	.00047	.00045	.00043	.00042	.00040	.00039	.00038	.00036	.00035
3.2	.00069	.00066	.00064	.00062	.00060	.00058	.00056	.00054	.00052	.00050
3.1	.00097	.00094	.00090	.00087	.00084	.00082	.00079	.00076	.00074	.00071
3.0	.00135	.00131	.00126	.00122	.00118	.00114	.00111	.00107	.00104	.00100
2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.7	.2420	.2389	.2358	.2327	.2297	.2266	.2236	.2206	.2177	.2148
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

Table 23: Cumulative Z table.

GC – MS COMPONENTS

