ICQEM 16

"A better world with Quality!"

July 13 - 15, 2016

Guimarães, Portugal

Proceedings Book

TECHNICAL RECORD

Title

Proceedings book of the 2nd International Conference on Quality Engineering and Management, 2016

Authors/Editors

Sampaio, Paulo; Domingues, Pedro; Barros, Síria; Carvalho, André; Casadesús, Marti; Costa, Acácio; Fernandes, Ana Cristina; Marimon, Frederic; Melo, Mónica; Oliveira, Rui; Pires, António Ramos; Saraiva, Pedro.

Publisher

International Conference on Quality Engineering and Management

Date

July 2016

Cover Design

Luís Coutinho

ISBN

978-989-20-6814-5

1st Edition

200 units

2nd International Conference on
Quality Engineering and Management

A better world with Quality!

July 13-15, 2016

University of Minho, Guimarães, Portugal

This edition is published by the International Conference on Quality Engineering and Management.

Portuguese National Library Cataloguing in Publication Data

Proceedings book of the 2nd International Conference on Quality Engineering and Management edited by Sampaio, Paulo; Domingues, Pedro; Barros, Síria; Carvalho, André; Casadesús, Marti; Costa, Acácio; Fernandes, Ana Cristina; Marimon, Frederic; Melo, Mónica; Oliveira, Rui; Pires, António Ramos; Saraiva, Pedro

ISBN 978-989-20-6814-5

Publisher: International Conference on Quality Engineering and Management

Book in 1 volume, 894 pages

International Conference on Quality Engineering and Management

Paulo Sampaio - Conference Chair

School of Engineering, Systems and Production Department

Campus Gualtar, 4710-057 Braga, Portugal

http://icqem.dps.uminho.pt/

icqem@dps.uminho.pt

ISBN 978-989-20-6814-5

Organizing Committee

Chair

Paulo Sampaio, University of Minho, Portugal

Members

Acácio Costa, University of Minho, Portugal

Ana Cristina Fernandes, University of Minho, Portugal

André Carvalho, University of Minho, Portugal

António Ramos Pires, Portuguese Association for Quality, Portugal

Frederic Marimon, International University of Catalunya, Spain

Marti Casadesús, University of Girona, Spain

Mónica Melo, University of Minho, Portugal

Pedro Domingues, University of Minho, Portugal

Pedro Saraiva, University of Coimbra, Portugal

Rui Oliveira, University of Minho, Portugal

Síria Barros, University of Minho, Portugal

Scientific Committee

Alberto Á. Suárez, University of Oviedo, Spain

Alexandra Simon, Universitat Autonomous of Barcelona, Spain

Ana Cristina Fernandes, University of Minho, Portugal

Ana Fonseca, University Fernando Pessoa, Portugal

António Araújo, Stadual University of Rio de Janeiro, Brazil

António Fernando Branco Costa, Paulista Stadual University, Brazil

António Ramos Pires, Portuguese Association for Quality

Denis Devos, ASQ QMDivision Chair, Canada

Eduardo Salgado, Federal University of Alfenas, Brazil

Elizabeth Cudney, Missouri University of Science and Technology, USA

Estela Vilhena, Polytechnic Institute of Cávado and Ave, Portugal

Evandro Lorentz, Brazilian Society for Quality and Excellence in Management, Brazil

Fiorenzo Franceschini, Politecnico di Torino, Italy

Frederic Marimon, International University of Catalunya, Spain

German Arana, University of the Basque Country, Spain

Glauco Henrique de Sousa Mendes, Federal University of São Carlos, Brazil

Greg Watson, Business Excellence Solutions, Finland

Helena Alvelos, University of Aveiro, Portugal

Henriqueta Nóvoa, University of Porto, Portugal

Inaki Heras, University of the Basque Country, Spain

Isabel Lopes, University of Minho, Portugal

Jacob Kjaer Eskildsen, Aarhus University, TQM&BE Editor, Denmark

FOREWORD

Welcome to the 2nd International Conference on Quality Engineering and Management!

After the successful organization of the **1st International Conference on Quality Engineering and Management** in 2014, it is our pleasure to welcome you to the conference **2nd edition** at the University of Minho, again in the historic city of Guimarães, Portugal. This event combines two areas that are not usually brought together: Quality Engineering and Quality Management. We hope that the results of our effort will translate into a successful venture, making gradually of this conference an important scientific event in the field of Quality. As was our aim, since the beginning, the conference covers different topics related to Quality Management and Quality Engineering, including Standards, Continuous Improvement, Supply Chain Quality Management, Management Systems, Six Sigma, Quality Tools, Quality Management in Higher Education, Quality Management in Services and Total Quality Management.

In this 2nd edition the balance between Quality Management papers and Quality Engineering ones is more clear, thus accomplishing one of the fundamental goals of this conference. Approximately 120 papers have been submitted and almost 85 were accepted for presentation, after review from the Conference Scientific Committee. Additionally, some of these papers were selected by the Scientific Committee to be considered for a special issue that will be published by the International Journal of Quality and Reliability Management (SCOPUS indexed journal). Papers accepted correspond to authors from all around the world, with more than 20 countries represented at this level. Therefore, a warm acknowledgment to all speakers and authors is well deserved – Thank You! The success of this second edition derives from their efforts and participation!

We would like to thank all of our four keynote speakers, who will be with us during the two days of the event: **Eric Rebentisch**, **Jiju Antony**, **Lars Sorqvist** and **Marco Reis**. We have here the chance to listen to their contributions and new research development insights, coming from some of the most influent current Quality Academicians. Many thanks also to all the excellent work carried out by the Scientific Committee during the papers selection process. We must acknowledge as well the institutional support received from the School of Engineering of the University of Minho, University of Coimbra, University of Girona, International University of Catalunya, Portuguese Association for Quality, Algoritmi Research Centre, Luso-American Foundation, American Society for Quality, Portuguese Institute for Quality, Brazilian Association of Production Engineering, Brazilian Society of Quality and Excellence in Management, Quality for Excellence Consultancy, BQualidade, Target and Cempalavras.

Again, let's take advantage of this great opportunity and make with your contributions an event with Quality, shared and built by such a top level group of participants!

Enjoy your conference! Thank you all!

University of Minho, July 14, 2016.

Paulo Sampaio (Conference Chair)

Conference Co-founders
Paulo Sampaio, University of Minho
António Ramos Pires, Portuguese Association for Quality
Pedro Saraiva, University of Coimbra
Marti Casadesús, University of Girona
Frederic Marimon, International University of Catalunya

Measuring Final Inspectors' Discrimination Ability of Metal Structures in the Automotive Industry

Alvelos, H.[1], Gonçalves, S.[2] and Xambre, A.R.[3]

[1] CIDMA/DEGEIT - University of Aveiro, 3810-193 Aveiro, Portugal

[2] DEGEIT - University of Aveiro, 3810-193 Aveiro, Portugal

[3] CIDMA/DEGEIT - University of Aveiro, 3810-193 Aveiro, Portugal

ABSTRACT

Purpose – The main objective of this paper is to present an approach for understanding the discrimination ability of the operators responsible for the final inspection at a metal structures' production line in the automotive industry. Another important goal is to show how this type of analysis can help to improve the way quality inspectors carry out their daily activities.

Design/methodology/approach – The methodology followed consisted in selecting a sample of 30 metal structures (15 conforming and 15 non-conforming), over a time period of two months, and then asking 10 quality inspectors to classify each of the structures. Afterwards the performance of the inspectors was analysed using Receiver Operating Characteristic (ROC) Analysis.

Findings – The results indicate that inspectors are generally working efficiently and also that there is a low risk of rejecting conforming parts and of accepting non-conforming ones. These risks depend on the inspector that judges the structure so the differences between their individual performances were analysed. Improvement actions were then recommended and implemented.

Research implications – The work presented is an exploratory study and reports on the results of the proposed approach that underline the possibility of quantifying the discrimination ability of inspectors in industrial environments.

Practical implications – The quantification of the inspectors' performances of the company was completed. Afterwards, it was possible to plan and implement improvement actions regarding those performances. All the inspectors, as well as the production line responsible, felt comfortable with the results.

Originality/value – As the approach used has not been addressed by scientific literature, it is considered that the outcomes of the work are valuable for researchers on quality evaluation of inspectors. Additionally, this study may help practitioners implement the proposed method in other environments.

Keywords: Quality assessment, Inspection, Automotive industry, ROC Analysis.

Paper type: Research paper.

INTRODUCTION

Quality control is an activity that has a high level of importance throughout the supply chain in the automotive industry. This work was developed in collaboration with a second-tier supplier of various automotive manufacturers that produces, among other products, car seats and operates in the highly demanding context of the car industry.

Some of the characteristics of the car seats' metal structures are controlled by visual inspection at the end of the production line, a task that is executed by trained inspectors.

Inspection operations throughout a certain production process are used to address several issues, namely: (i) to distinguish between conforming (OK) and non-conforming (NOK) products; (ii) to determine if there were any changes in the production process; (iii) to measure a process' capability; and (iv) to categorize the product according to its quality. The most common of those alternatives is to verify if the product will be accepted or rejected by the client, whether it is an internal or external client.

The quality of the inspection is, thus, dependent on the decisions taken by the inspectors, which can be right (to detect a defect in a structure or to classify a flawless product as a good one) or wrong (to detect a defect that does not exist in a flawless product or to classify a product as a good one, when it has a defect).

Binary measurement systems are used quite often in industry and a common example is the visual inspection of products with the purpose of accepting or rejecting them (De Mast *et al.*, 2011). In this context the inspector can be regarded as a measurement instrument that visually checks the characteristics of the product and determines if it can be packed and shipped to the customer or if it should be rejected and, in that case, go through some rework or be scrapped.

Evaluating the performance of this measurement system is important in order to ensure the high quality standards required in this type of industry.

The main objective of this paper is to present an approach for measuring inspectors' performance in the final inspection workstation of a production line for car seats' metal structures, using ROC (Receiver Operating Characteristic) Analysis.

RESEARCH METODOLOGHY

The first step was to study the product and its production process with, naturally, a particular focus on the activities performed at the inspection station.

To support the work it was also necessary to study and evaluate the essential and more current literature regarding measurement systems, inspection processes and their importance, as well as methods to evaluate the discrimination ability of the inspectors, particularly ROC Analysis.

With the help of product and process specialists, and also considering problems' historic data and the workstation instructions, the sample size and the main defects to analyse, were then determined.

Additionally, the need for a specific location to storage defective and non-defective structures to serve as standards was identified, as well as the need for creating a replica of the inspection workstation. Once these physical conditions were fulfilled, it was possible to test the measurement system and so, all the inspectors were asked to analyse and classify the same set of structures twice, and the resulting data was gathered.

The final steps consisted in the data analysis, the suggestion of some improvement actions and the implementation of those actions, which resulted in the main outcomes of this work.

THEORETICAL BACKGROUND

This section is dedicated to a literature revision directed at the two main issues used in the practical study. The first sub-section focuses on concepts related to the activity of inspection, whereas the second one consists of an explanation of ROC Analysis and its fields of application.

Inspection

Historically, inspection is defined as the process or procedure of examining the attributes of a part or product in order to determine if they are in conformance with the predefined requisites (Newman, 1995 and Mital *et al.*, 1998). Testing and inspecting include measuring a certain output so as to determine if it complies with its specification (Gryna *et al.*, 2007).

According to AIAG (2010) the activity, usually called inspection, is the act of examining the parameters of the process, the components, subassemblies or final products, with the aid of adequate norms and measurement devices, in order to confirm if the process is working properly.

Inspection can be done to 100% of the parts produced or by using sampling methods. The first situation occurs when the defect or defects that are analysed can seriously affect the use of the product or the way it works. The second situation results on the inspection of a sample of the population and it's commonly used when it is impossible to check all the parts or when the inspection process results in the destruction of the part.

Planning and managing all the necessary resources for inspecting critical quality characteristics it's an important activity in any transformation industry but it consumes a lot of resources. The high associated costs are, however, compensated by the benefits of identifying faulty parts and removing them from the system.

Deciding the number and placement of inspection workstations is a responsibility shared between the people in charge of production process design and of quality control system specification (Inman *et al.*, 2013).

According to Gryna *et al.* (2007) the most common locations for inspection workstations are:

- Receiving inspection of goods from suppliers.
- After installing a production process, to ensure that defective batches are not produced.
- During critical or costly operations.
- Prior to delivery of goods from one department to the next.
- Prior to shipping finish goods to storage or to a client.
- Before performing a costly, irreversible operation.

According to Kopardekar *et al.* (1995) there are three different types of inspections: human inspection, automatic inspection and hybrid inspection.

Human inspection is performed by an inspector that evaluates the quality of a product taking into account a set of standard products and his/her own experience and know-how.

This type of inspection is used when the characteristics of the product under inspection cannot easily be measured by instruments and so, they have to be evaluated by human senses. Those are called sensory characteristics and, for evaluating them, human senses are used as measurement instruments.

Sensory characteristics can be related to the technological performance of the product (e.g. adherence of a protective coating), olfactory characteristics (e.g. smell of a perfume), taste (e.g. food), or others. An import category includes the visual quality of certain characteristics that, usually, do not have very well defined written specifications since they are difficult to quantify (Gryna *et al.*, 2007).

Sensory analysis is traditionally used in the development of new products or when there is a need to compare a product to its competitors. However, it can also be used as a tool in quality control to check certain sensory characteristics. In this context, the final step is to decide whether or not to accept the product based on the analysis of the inspectors that can consider the intensity of the defects as well as the predefined tolerances associated with the different characteristics (Baudet *et al.*, 2013).

Several authors mention that subjectivity is part of human inspection and so its performance is influenced by factors related to the task itself, environmental factors, organizational factors, motivational factors and personal factors (Mital *et al.*, 1998). Evaluating the existence of a defect can then depend, among others, on the level of knowledge of the inspector, his/her know-how and his/her perception concerning the defect (Baudet *et al.*, 2013).

The other two types of inspections are: automatic inspection where the analysis and decision is performed by a computer or some other type of automatic controller (Newman, 1995); and hybrid inspection where human inspection is complemented or aided by automatic equipment (Mital *et al.*, 1998).

Wang *et al.* (1997) mention that when it comes to human inspection, and particularly visual inspection, there are two opposite strategies: random search or systematic search. The first strategy implies that the sequence of location inspected each time changes when the inspection is repeated and in a systematic approach that sequence does not change.

Wang *et al.* (1997) also mention that when there is some memory of the locations previously inspected, it results in a better performance of the inspectors. In this context, and since the search strategy is so important, inspectors should be trained so they can use the adequate strategy (Tetteh and Jiang, 2006).

It is thus important to train inspectors so they can use a systematic search strategy adequately designed for the product being inspected (e.g. use work instructions and/or visual aids placed at the workstation).

That are several factors that influence the complexity of the inspection operation, whether it's in the visual search stage, whether it's in the decision making stage. Some of the most common complexity factors are: the number of types of defects, the complexity of the standard defects, the probability of defects occurring, and the defects' distribution.

Gallwey and Drury (1986) state that: (i) if there is a need to search for different types of defects the performance is worse than when the search is for just one type of defect; (ii) the higher the number of locations or areas inspected the worst the performance; and (iii) the complexity of the standard defects used for comparisons has no influence for the search stage but impacts the performance on the decision-making stage.

The probability of human error is higher if the working conditions are not adequate and if, during the decision-making process, subjective factors or factors such as available time and cost constraints are involved (Sylla and Drury, 1995).

ROC Analysis

The Receiver Operating Characteristic Curves Analysis (ROC Analysis) is a technique (first developed in Signal Detection Theory) that has been used in many areas, like Health Sciences (Metz, 1978; Rifkin *et al.*, 1990; Haanes *et al.*, 2015; Waterland *et al.*, 2016; Avcioglu and Sezer, 2016), Sensory Analysis (O'Mahony, 1992; Bi *et al.*, 2000; Lee *et al.*, 2007; Paredes-Olay *et al.*, 2010) or Finance (Irwin and Irwin, 2013; Liang *et al.*, 2016) to compare the performance of diagnosis equipment, or to decide in the case of dubious circumstances (Alvelos, 2002). This technique can be used when the characteristic or decision under study has a binary nature (like "approved" or "not approved" or "sweet" or "unsweet") and is evaluated using rating tests. Its output is usually represented by a curve – the ROC curve, being the area under this curve a measure of the discriminating power of the system (e.g. a diagnosis equipment, a taster or an inspector).

When applied to sensory evaluation, it provides a sensitivity measure to a stimulus or to a difference between stimuli. Ishii, Vié and O'Mahony (1992) describe and compare some of the tests and the scales that are applied more frequently in this area, and O'Mahony (1992) illustrates the use of ROC Analysis in measuring tasters' discrimination capacity between "sweet" and "unsweet" cookies.

Traditionally, in Sensory Analysis, the experiment used for applying ROC analysis is the "A-Not A" test, in which it is initially presented to each taster a sample of the product under analysis (A), followed by a series of other samples that tasters have to classify as "A" or "Not A".

As, for drawing the ROC curve, it is necessary that tasters associate to each decision ("A" or "Not A") a degree of certainty, so there are frequently used four levels of certainty. Tasters are asked to use one of the following answers for each sample: (i) "I am sure the sample is different from A", (ii) "The sample seems to be different from A, but I am not sure", (iii) "The sample seems to be equal to A, but I am not sure", or (iv)"I am sure the sample is equal to A" (Alvelos, 2002).

The ROC curve is a chart that represents the proportion of "true positives" (p_{TP}) versus the proportion of "false positives" (p_{FP}) for the various criteria, and, as Metz (1978) points out, it "(...) indicates all possible combinations of the relative frequencies of the various kinds of correct and incorrect decisions". The criterion is changed by explicitly changing the decision threshold and reinterpreting the results of the products' judging. In the case of the "A-Not A" test, when using the "strict criterion" it is assumed that the taster decides that product is "A" only in the cases he/she is sure about it, and when using the "lax criterion", it is assumed that the taster decides that the product is "A" in all the cases except when he/she is sure it is "Not A". The usual interpretation of the test (to assume that the taster decides that product is "A" in the cases he/she rates it as "sure it is equal to A" or "seems equal to A, but not sure") corresponds to the "moderate criterion". Then, for each criterion, resulting proportions (p_{TP} and p_{FP}) are calculated.

The ROC curve is obtained using the referred scale and joining the resulting points by straight lines. If the taster distinguishes at least between some of the "A" and the "Not A" products, the intermediate points on a ROC curve must be above the lower left to upper right diagonal of the ROC space. This diagonal is the ROC that corresponds to a taster rating the samples randomly. Exemplifying with the point (0.5, 0.5), it can be seen that it corresponds to the situation where the taster rates 50% of the "Not A" samples as "A" (p_{FP}), and, simultaneously, correctly rates 50% of "A" samples (p_{TP}). This is clearly a situation where the taster is simply guessing and has no sensitivity to the difference between the products. The opposite situation occurs when the taster always detects the "A" product and corresponds to a horizontal line which ordinate is equal to 1 (p_{TP} = 1).

The area under the ROC curve (A_i), when the points referred above are joined by straight line segments, has an important statistical meaning: it is a conservative estimate of taster *i* choosing the "A" sample when compared with the "Not A" one, in a single comparison between them (Alvelos, 2002). Additionally, it must be noted that this area is independent of the criterion used by the taster to "Approve" or "Not Approve" the sample. It measures the sensitivity of the taster to the difference between the two products.

PROBLEM CONTEXTUALIZATION AND RESULTS

This work was developed at a second tier supplier of the automotive industry that produces car seats' metal structures.

The production department includes four autonomous production units and, within each one, there are several production lines that manufacture products for different car's brands and models. There are also several manufacturing processes like, arch welding (MAG – Metal Inert Gas), painting, riveting, screwing, among others.

This work focused on the production of a specific metal structure (Product X) which is manufactured in four different, but identical, lines. However, just one of those lines was considered: the Model Line, where welding and assembling are the basic processes used.

The manufacturing process begins with automatic welding, done by a MAG welding robot that includes two sub processes. In the first one the main, and more expensive, components are combined and, in the second stage, smaller tubes and beams are welded to connect the rest of the components. Afterwards components are sent in batches to the painting section. The components return to the production line, after being painted, around 2.5 hours later, so the final components can be assembled.

Lastly, a final inspection is done so any eventual problem or defect is detected before the product is packed and shipped to the client.

Nonetheless, both at the welding workstation and the assembly workstation visual and/or functional control is done by the workers in order to assess the quality of the operation they have just performed.

The final inspection is the final operation and differs from the auto-controls done in the previous workstations because: (i) it focuses on the characteristics of the final product (for example, functional or dimensional controls) and not on the quality of a specific operation; and (ii) the number of points to inspect is larger so the complexity of the task is higher as is the mental strain on the inspector.

When planning the workstation dedicated to final inspection operations, ergonomic factors that affect the performance of the inspectors were considered: (i) because it is a mentally demanding task, the number of points to check were taken into account and rotation throughout the day between the inspectors was implemented so that a person is not inspecting for long periods of time (no longer than 2 hours); (ii) the movement of the part as well as of the inspector were defined so his/her posture is correct and the distance between the eyes and the part is adequate and facilitates the detection of eventual defects; (iii) environmental conditions such as noise level and temperature were controlled so as not to disturb the concentration of the inspector.

The main purpose of this work was to develop an approach for evaluating the discrimination ability of the operators responsible for the final inspection of product X at the Model Line.

The methodology adopted in the practical study consisted of five stages. First, a careful analysis of the production process of Product X was performed, in order to identify the critical quality characteristics. Following, some physical conditions were analysed and a space in the plant, where the products to be evaluated could be located, was created. The third stage consisted in the selection of the specific sample of those products, considering the defects to be detected (output of stage 1) and the available space (output of stage 2), and the fourth one comprised the planning and execution of the tests. Finally, the results were analysed and improvement actions were proposed and implemented. Following, the referred stages are detailed.

In the first stage the production process was analysed with a particular emphasis on the final inspection workstation and on the product.

Product X is a metal structure made of components united with welding seams done by automatic MAG welding and also the assembly of some components. Between those two stages, as mentioned previously, the subassembly is painted.

The main characteristics of product X that are inspected are: (i) the quality of the welding seams because those are the connecting elements that ensure the security of the structure; (ii) the dimensional requisites; (iii) the correct functioning and presence of all the assembled components; and (iv) the panting quality particularly on the visible areas.

The operations performed in the final inspection workstation are, therefore:

-control of some dimensional characteristics through the use of calibres ‘fit/does not fit’;

-control of the welding seams by evaluating their position and aspect;

-verification of the presence of all the assembled parts;

-visual analysis of the quality of the painting in the visible areas.

In the second stage a specific space was created with the purpose of storing the structures used in this study and creating a replica of the final inspection workstation similar to the one in the Model Line. This space was delimited with a red cabin which, due to space constrains, had a capacity to accommodate 30 structures (Figure 1). The use of red in the cabin was a requirement of the company since it contained defective structures and red is the colour used to signal non conformities. After being built, the structures were stored and the cabin was closed, with limited access.

Figure 1 – Test cabin.

The sample was made up of 30 units of product X, as was mentioned previously. It was defined that all of the sample parts should be analysed and classified by a specialist before they were inspected, repeatedly, by the inspectors. The sample was composed of 50% of defective units (NOK) and 50% of non-defective units (OK) as it is a good practice to have a balanced mix of OK and NOK parts, because if one type of parts is not as present it may influence the judgement regarding the inspectors’ ability to classify that type of parts (AIAG, 2010).

The defective units included several cases of parts very close to the acceptance threshold. Four types of defects could be found, 3 regarding the welding seams and 1 related to the painting.

The 3 types of defects related to the welding seams were: porous welding seam, leaked welding seam, deviated welding seam.

The porous welding seam is due to bubbles of air that are trapped inside the seam or that are formed at the surface (Figure 2 a). A porous seam leads to a weaker welding and can be caused by the presence of impurities either on the base material or the adhesive, by lack of gas protection, by excessive tension in the electric arch or by excessive distance between the torch and the part (a lot of stick-out).

A leaked welding seam means that the welding pierces one of the sections and so the seam is incomplete (Figure 2 b). This is due to excessive penetration and can result from high current intensity, an incorrect position of the torch, slow welding speed, and incorrect trajectory of the robot.

A deviated welding seam is misplaced over the areas that it should connect (Figure 2 c). This problem occurs due to the incorrect trajectory of the robot, the incorrect placement of the parts, and accessibility problems of the robot to the joint to be weld.

The defect shown on Figure 2 d) is the lack of paint on a visible area.

The defective structures were collected over a 2 months period and were the result of problems and defects that occurred at the Model Line. Some of the structures could have been reworked and repaired and so it was necessary the collaboration of the production team until all the required structures for the cabin were collected. All the structures inside the cabin were signalled so as to identify their purpose.

As explained previously, in the final inspection workstation other characteristics of product X are evaluated. However, those were not considered in this study because, for those defects, the inspectors use calibres and non-visual inspection.

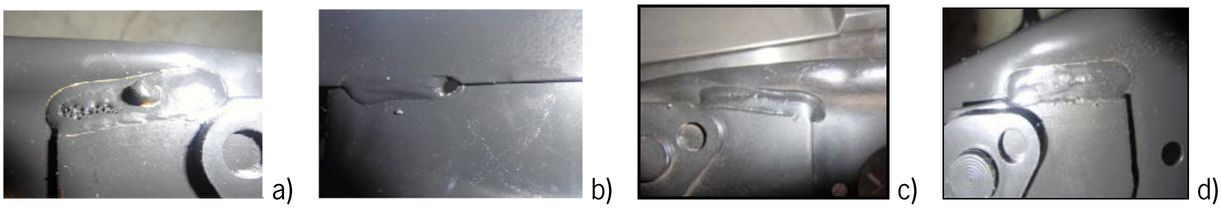

Figure 2 – Types of defects: a) porous welding seam, b) leaked welding seam, c) deviated welding seam, d) lack of paint.

The scheduling and execution of the tests were planned together with the inspectors, the line responsible, and with the rest of the production team so they could organize themselves and reduce the impact on the line whenever one element had to go into the cabin to inspect the sample. The test was done by ten inspectors of product X.

A form was created so the inspectors could register their evaluation of all of the structures in the sample. In that form they could identify themselves, identify the reference of the inspected structures, the date and type of inspection and the number of inspected structures. As far as answers they could register four possible categories: OK, I am sure; OK, but I am not sure; NOK, I am sure; NOK, but I am not sure. There was also a space to indicate the type of defect detected whenever the inspector classified a structure as NOK. Note that the OK (non-defective unit) corresponds, in the previously described "A-Not A" test, to the "A" product, while the NOK (conforming unit) corresponds to the "Not A" product.

Before any inspector did the analysis all the procedure was explained, as well as what he was expected to do and what the purpose of the study was. Also, any doubts he/she may had, were clarified.

Each inspector evaluated the structures twice and no questions were answered regarding the first inspection, before they had done the second one.

The discrimination ability of the inspectors under evaluation was then estimated using Receiver Operating Characteristic (ROC curves).

In this work inspectors are considered in the same way as tasters because they are used as measurement instruments capable of measuring characteristics that are not easily evaluated by conventional measurement instruments. So it was decided to use ROC curves for the inspectors as a way of estimating their discrimination ability between OK and NOK parts. In this context, the area under each inspector' ROC curve is an estimate of the probability of the inspector correctly identifying an OK part when it is presented together with a NOK part, in a simple test.

Table 1 summarizes the results of the tests done by two inspectors (A and B), using the scale already mentioned: NOK (I am sure the part is NOK); NOK? (I am not sure the part is NOK); OK (I am sure the part is OK), OK? (I am not sure the part is OK).

Table 1 – Results of the tests performed by two of the inspectors.

Inspector A			Inspector B		
Criteria	OK units	NOK units	Criteria	OK units	NOK units
NOK	0	28	NOK	2	23
NOK?	0	0	NOK?	1	5
OK?	0	1	OK?	2	1
OK	30	1	OK	25	1
Total	30	30	Total	30	30

The results from the previous table were then used to build Table 2 that shows the proportion of false positives (p_{FP}) and true positives (p_{TP}) for those two inspectors and for the various criteria. Notice that three 'artificial' criteria were used: (i) the more restrictive corresponds to approving only the OK parts, (ii) the intermediate one corresponds to approving the OK and OK? parts, and (iii) the broader one corresponds to approving all the parts except the NOK ones.

Table 2 – Proportion of false positives and true positives for two of the inspectors.

Inspector A			Inspector B		
Criteria	p_{TP}	p_{FP}	Criteria	p_{TP}	p_{FP}
Approved – OK Not approved - OK?, NOK?, NOK	1,00	0,03	Approved - OK Not approved - OK?, NOK?, NOK	0,83	0,03
Approved - OK, OK? Not approved - NOK?, NOK	1,00	0,07	Approved - OK, OK? Not approved - NOK?, NOK	0,90	0,07
Approved - OK; OK?, NOK? Not approved - NOK	1,00	0,07	Approved - OK; OK?, NOK? Not approved - NOK	0,93	0,23

The data presented in Table 2 correspond to the coordinates (x,y) of the points in the ROC curves for inspectors A and B. These curves, as well as the area under them, are presented in Figure 3.

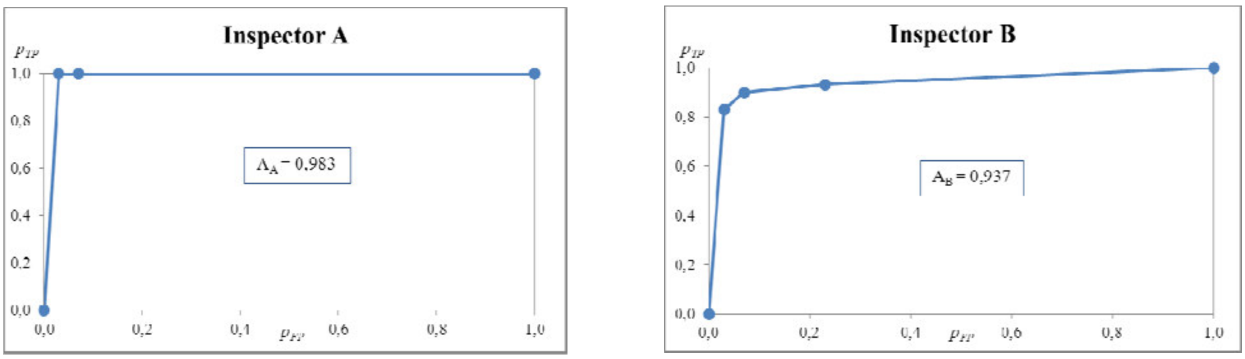

Figure 3 – ROC curves for two of the inspectors.

All the inspectors had curves above the diagonal [(0,0) (1,1)], and consequently, the respective areas are quite superior to 0,5 (all had values above 0,9) which indicates that all of them were capable of distinguish between OK and NOK units. It was also possible to detect that three inspectors (one of which is inspector A) had bigger areas. Those results can be related to the knowledge the three of them have of the entire production process since they are capable of working in any workstation of the line. This helps them to know every detail of the product and, consequently, be more sensitive to any variations in the units that are part of the sample. Product and process specialists also agreed that this knowledge of the product was beneficial for the inspectors' performance.

Besides the evaluation of the inspectors' discrimination ability, the purpose of the work was also to determine, suggest and, if possible, implement some improvement actions in order to increase their analysis capacity and autonomy. In this context, two improvement actions were suggested and implemented: (i) the first one, directed at all the inspectors, was a presentation of the global results of the study; and (ii) the second one was only directed at the inspectors with worst performance and consisted of extra training.

The first action consisted in the presentation to each individual inspector of the mistakes they had done during the test. The first step was to identify, for each inspector, which units he/she had misclassified. Then, for each of those units a comparison with other units (OK or NOK) was done and mistakes were pointed out. It was also possible to clarify their doubts and the whole procedure was carried out from a training point of view, also with the help of product and process specialists.

For the inspectors with a not so good performance (for example inspector B) additional training was identified as a need. That extra training was provided to three of the inspectors, inspector B and two others, but it should be noticed that inspector B, for example, had very little experience as inspector.

This training was based on an existing tool, the defects panel, that consists of a 2m*3m panel where there are in display: defective components and subassemblies that represent all types of different defects; components and subassemblies on the threshold between OK and NOK with indications if they were accepted or rejected by the client; and detailed indications of the visible areas of the structure. The inspectors were invited to analyse this panel more carefully and so learn more about the client's requirements.

CONCLUSIONS

This work focused on the problematic of quality inspection in the manufacturing industry, particularly the situation where visual inspection, performed by workers, is the final inspection before products are shipped to clients.

With the procedure that was implemented it was possible to analyse in detail all the inspection process. The cabin created also helped because it was possible to get a reasonably sized sample (30 units) and also install an inspection workstation.

With ROC Analysis it was possible to quantify the inspectors' performance, and it was concluded that, overall, the discrimination ability of the ten inspectors is good and they are able to distinguish between OK and NOK units. Nevertheless, two improvement actions were suggested and implemented. The first one was to present to each inspector the structures that they misclassified and to clarify why those errors happened. The second one was to organize a specific training session for those inspectors that exhibited a less satisfactory performance.

The application of the proposed methodology to other products and other production lines was suggested and implemented, which was made easier by the know-how acquired in the development of this work and also by the fact that the necessary space had already been created.

For future work, it is suggested that inspectors repeat the tests so the new results can be compared with the current ones.

It is also suggested that some of the human inspection operations, particularly when calibres are used, are moved to earlier stages of the production process and, if possible, automatized.

Finally, it is considered that the outcomes of this work are valuable not only for researchers on quality evaluation of inspectors / judges, but also for the implementation of the proposed methodology in other industrial environments.

AKNOWLEDGEMENTS

This work was supported in part by the Portuguese Foundation for Science and Technology (FCT-Fundação para a Ciência e a Tecnologia), through CIDMA - Center for Research and Development in Mathematics and Applications, within project UID/MAT/04106/2013.

REFERENCES

- AIAG (2010), Measurement System Analysis: Reference Manual, 4rd ed., Automotive Industry Action Group, Detroit (MI).
- Alvelos, H. (2002), “Análise, desenvolvimento e teste de métodos e técnicas para controlo estatístico em análise sensorial”, Faculdade de Engenharia da Universidade do Porto.
- Avcioglu, S.N., Sezer, S.D., Kucuk, M., Zafer, E., Yuksel, H., Akcan, B., Turgut, O. (2016), “Maternal serum concentrations of s-endoglin and il-6 in pregnancy complicated by preterm premature membrane rupture”, Journal of Maternal-Fetal & Neonatal Medicine, Vol. 29, No.12, pp. 1957-1962.
- Baudet, N., Maire, J. L., Pillet, M. (2013), “The visual inspection of product surfaces”, Food Quality and Preference, Vol. 27, No. 2, pp. 153-160.
- Bi, J., Templeton-Janik, L., Ennis J., Ennis, D. (2000), “Replicated Difference and Preference Tests: How to Account for Inter-trial Variation”, Food Quality and Preference, Vol. 11, pp. 269-273.
- De Mast, J., Erdmann, T. P., Van Wieringen, W. N. (2011), “Measurement system analysis for binary inspection: Continuous versus dichotomous measurands”, Journal of Quality Technology, Vol. 43, No. 2, pp. 99-112.
- Gallwey, T. J., Drury, C. G. (1986), “Task complexity in visual inspection”, Human Factors, Vol. 28, No. 5, pp. 595-606.
- Gryna, F.M., Chua, R.C.H., DeFeo, J.A. (2007), Juran's quality planning and analysis for enterprise quality, 5th ed., McGraw-Hill, Boston (MA).
- Haanes, G.G., Kirkevold, M., Hofoss, D., Eilertsen, G. (2015), “Discrepancy between self-assessments and standardised tests of vision and hearing abilities in older people living at home: an roc curve analysis”, Journal of Clinical Nursing, Vol. 24, pp. 3380-3388.
- Inman, R. R., Blumenfeld, D. E., Huang, N., Li, J., Le, J. (2013), “Survey of recent advances on the interface between production system design and quality”, IIE Transactions, Vol. 45, No. 6, pp. 557-574.
- Irwin, R.J., Irwin, T.C. (2013), “Appraising credit ratings: Does the cap fit better than the roc?”, International Journal of Finance and Economics, Vol. 18, No. 4, pp. 396-408.
- Ishii, R., Vié, A., O’Mahony, M. (1992). “Sensory Difference Testing: Ranking R-Indices are Greater Than Rating R-Indices”, Journal of Sensory Studies, Vol. 7, pp. 57-61.
- Kopardekar, P., Mital, A., Anand, S. (1993), “Manual, Hybrid and Automated Inspection Literature and Current Research”, Integrated Manufacturing Systems, Vol. 4, No. 1, pp.18-29.

- Lee, H.-S., van Hout, D., Hautus, M.J. (2007), “Comparison of performance in the A-Not A, 2-AFC, and same-different tests for the flavor discrimination of margarines: The effect of cognitive decision strategies”, Food Quality and Preference; Vol. 18, No.6, pp. 920-928.
- Liang, D., Lu, CC., Tsai, CF., Shih, GA. (2016), “Financial ratios and corporate governance indicators in bankruptcy prediction: A comprehensive study”, European Journal of Operational Research, Vol. 252, No. 2, pp. 561-572.
- Metz, C. (1978), “Basic Principles of ROC Analysis”, Seminars in Nuclear Medicine, Vol. 8, pp. 283-298.
- Mital, A., Govindaraju, M., Subramani, B. (1998), “A comparison between manual and hybrid methods in parts inspection”, Integrated Manufacturing Systems, Vol. 9, No. 6, pp. 344-349.
- Newman, T. S. (1995), “Survey of automated visual inspection”, Computer Vision and Image Understanding, Vol. 61, No. 2, pp. 231-262.
- O'Mahony, M. (1992), “Understanding Discrimination Tests: a User-Friendly Treatment of Response Bias, Rating and Ranking R-Index Tests and Their Relationship to Signal Detection”, Journal of Sensory Studies, Vol. 7, pp. 1-47.
- Paredes-Olay, C., Moreno-Fernández, M.M., Rosas, J.M., Ramos-Álvarez, M.M. (2010), “ROC analysis in olive oil tasting: A Signal Detection Theory approach to tasting tasks”, Food Quality and Preference, Vol. 21, No. 5, pp. 562-568.
- Rifkin, M., Zerhouni, E., Gatsonis, C., Quint, L., Paushter, D., Epstein, J., Hamper, U., Walsh, P., McNeil, B. (1990), “Comparison of Magnetic Resonance Imaging and Ultrasonography in Staging Early Prostate Cancer”, The New England Journal of Medicine, Vol. 323, pp. 621-626.
- Sylla, C., Drury, C. G. (1995), “Signal detection for human error correction in quality control” Computers in Industry, Vol. 26, No. 2, pp. 147-159.
- Tetteh, E. G., Jiang, S. (2006), “The effects of search strategy, task complexity and pacing on visual inspection performance”, in 9th Annual Applied Ergonomics Conference 2006 Proceedings of the international conference in Orlando, Florida.
- Wang, M. J., Lin, S. Drury, C. G. (1997), “Training for strategy in visual search”, International Journal of Industrial Ergonomics, Vol. 20, No. 2, pp. 101-108.
- Waterland, P., Ng, J., Jones, A., Broadley, G., Nicol, D., Patel, H., Pandey, S. (2016), “Using CRP to predict anastomotic leakage after open and laparoscopic colorectal surgery: is there a difference?”, International Journal of Colorectal Disease, Vol. 31, No. 4, pp. 861-868.