Information Theory
In
Quality Engineering

By

Bolennarth Svensson

A Doctoral Thesis

Design Manufactur Engineering Management
University of Strathclyde
Glasgow
UK
&
School of Engineering
University of Borås
Borås
Sweden

May 1998

Abstract.

This thesis presents the results of research into a universal theory for quality techniques.

The unique contribution that is made is twofold:

- A new quality metric is proposed.
- An integrating perspective to quality engineering is introduced through the application of information theory.

The quality metric is designed as an information distance, measuring the difference between two probability density functions. The two distributions are the actual outcome of a running process and the expected outcome, i.e. the target distribution.

The target distribution makes it possible to integrate the quality losses into the metric. The metric may be adapted to the state of knowledge of the process studied.

The new quality metric is applicable to any process, be it a product process or an administration process. The information distance metric makes the analysis procedures uniform for all types of quality characteristics.

A function based process documentation makes information theory generally applicable to quality engineering.

The function description makes it possible to visualize poor quality as a surplus of information. All quality techniques aim at minimizing the information content in the system. Quality engineering in general may be expressed as an activity to stop surplus information flow reaching the process result.

There is a natural focus on noise, i.e. influencing factors that are out of control of the user, affecting the systems. This focus is in robust design developed through a process performance perspective rather than an experimen-

tal design perspective. An effort addressing a product process subject to improvement has to be discriminated from an effort addressing the efficiency of the experimentation process used to study the product process.

The present work is of pioneering character. Thus it opens a new area of research. Areas of further research are indicated.

Keyword: Quality, Robust design, Information theory, Design, Information distance, SPC, Total design.

Acknowledgements

This research has been funded by the National Swedish Agency for Development and Research, NUTEK, and the National Swedish Administration for University Education. Apart from the financial support the author would like to acknowledge support from a number of individuals. First among them is former Professor Stuart Pugh. His encouragement, support and supervision for this work is the prime reason for the present report being presented. For his engagement in my work even through a severe illness I have the greatest admiration and gratitude.

While pursuing a work like this there are always many people around that give support such that the tasks could really be brought to an end. There is Dean Bengt Andersson of The School of Engineering at The University of Borås, in Sweden. Without his support and his strong efforts for fund raising this research had been impossible. There is also Professor Ken MacCallum and Professor Martin Gordon of Design & Manufacturing Engineering Department of University of Strathclyde, Glasgow. They took up the loose threads of my work as Professor Pugh sadly died. Their actions to get a good focus and the directions towards a target were invaluable. There is also Mr Magnus Petzäll of Uponor AB in Sweden. His support for experimental application to a real problem has been vital.

Further there are all my colleagues at Design & Manufacturing Engineering Department of The Strathclyde University and at The School of Engineering at The University of Borås. All of you have been patiently listening to my talks about entropy, quality and disorder. Pushing me forward and bringing me back to reality you have all been a great support.

Borås May 31st 1998

1. Introduction

1.1. From product control to process control.

As a background we will first give some historical remarks on the development of a statistical viewpoint of the quality issue. Then there will be some comments on the usual evolution of quality technique applications within organizations.

In the early ages of industrialization manufacturing processes were not particularly capable. Production was not run in a long series. This was due to interaction between poor process capability and market development. The situation made a 100% product control, i.e. inspection, necessary and feasible. As the market developed demand was stronger than supply. Thus the market was not a competitive one. Quality was not a priority area.

Mass production which was introduced by Henry Ford and others, put emphasis on reliable manufacturing processes. Still the market was not competitive. However the power of application of statistical methods was realized during the twenties (111), (108). At that time the fundamentals of many of the tools used in quality engineering today were developed (109), (105), (106), (107).

During World War II quality, in terms close to what we use today, was first introduced. The first really competitive "market" warfare, was experienced. Quality techniques became top secret military information (110), (104), (103). Still quality had not yet got the wide interpretation which we use today. The main emphasis was on reliability; that is, loosely interpreted as the activity of a product to reliably supply the function it had when it was shipped, whether or not that function was the customer need was not a prime issue. More strictly, the reliability of a product is now defined as the probability that the product continues to meet its specifications, (27).

Within the program for rebuilding of Japan after World War II, Juran and Deming were invited to lecture in Japan, (94). At that time, or shortly before, modern quality definitions were introduced (101). In Western countries inspection was still the main quality tool in use. At that time in the early fifties Japanese quality performance was generally very poor. However they took some very important decisions to adopt the philosophy of Juran and Deming. As a consequence of that they achieved a much better quality improvement rate than Western countries.

The philosophy of Juran and Deming included a focus on quality losses. In the early fifties a Japanese, Genichi Taguchi, introduced the concept of a quality loss function, L(y). The early history of the development of the loss function may be found in the book, "introduction to quality engineering.",(93). In that function y is the performance measure. Taguchi regarded the losses to be the sum of all costs that the product caused the society due to poor function. L(y) was further considered to be a continuous function. This concept together with a statistical approach created a new insight into the nature of quality.

With the improvement of the quality performance came problems in the application of product control. The costs of inspection were much higher than the value of the quality improvement it contributed.

Montgomery, (47), gives a nice description of the evolution of applications of different quality engineering tools as an organization is maturing in terms of quality awareness. Further Grant *et al.*,(71), have an interesting discussion on statistical quality control that has a relevance in this context. Statistical quality control is said to be the application of statistical methods to control quality. Statistical process control is most often referred to as the application of control charts as an analysis tool to observations of the outcome of a process. This is in a sense a posteriori procedure. In this thesis the use of the acronym SPC is almost exclusively used to refer to this interpretation. There is however, a further evolution as the charts are applied to critical parameters controlling the process. This latter interpretation is more related to the results of robust engineering efforts for the production process.

Looking at the evolution of the application of statistical quality control tools the first thing to happen as inspection gets outdated, is the introduction of Statistical Process Control, SPC, according to Shewhart (109), (105). This is usually done as an extension of product inspection. The prerequisites for application of control charts to process parameters are usually not present. That is to say, the detailed knowledge of the processes are not available.

The quality loss function implies a continuous improvement strategy. As it accounts for many different losses (See Chapter 3.) it also points to different processes to be controlled to minimize quality losses. Accordingly the next step in the evolution is to try to control the design process. This involves the introduction of different quality improvement tools.

An important quality improvement tool within design is robust design (93), (64). Robust design is a systematic method to exploit the nonlinear relations existing between design factors and system functions. These relations are exploited such that the influence of noise factors on the system functions is minimized. In the robust design procedure according to Taguchi, interactions between noise factors and design factors are analyzed extensively (75). This will be discussed more in detail in Chapter, 4.

Another important tool is design review. This tool is impossible to apply effectively without a drastic improvement of specifications. Specifications as seen hitherto have been very weak and sweeping. Design reviews call for a quantified specification (63). In his book Total Design Professor Pugh, (63), puts a very strong focus on the quality level of the product development specification, PDS. Design reviews have the character of product inspection in the design process.

Further progress of process control took the shape of Total Quality Control, TQC (94). Within TQC, quality planning is a major task. One of the first tools for quality planning was Quality Function Deployment, QFD (61), (46). In its original version QFD did not explicitly incorporate different quality tools. However to be able to do the QFD matrices correctly different quality tools have to be applied. Thus QFD may be regarded as a planning tool for the application of quality tools.

Looking at QFD this way, we may see that the application of the quality tools has to be made with a holistic view. In Western countries we may remind ourselves about the different waves of single quality tools (75). We have seen quality circles, value analysis (VA), value engineering (VE), (99), statistical process control (SPC), failure mode and effect analysis (FMEA), design reviews, design for manufacturing (DFM), etc. We have to conclude that these waves did not solve the problem.

Sontow *et. al., (15),* have an extension to the original QFD. Details about this extension are found in Chapter 8. This work makes the interrelationship between different quality tools very clear. The research presented in this report draws on the fact these interrelationships are made visible. The quality tools all have the same effect, i.e. minimizing information content. It is also made clear that they may not be applied in an arbitrary order.

It may for instance be argued whether FMEA can be applied early or not. In actual effect it is often applied in later stages. However as will be discussed later we would like to apply FMEA early. FTA is a good starting point for a FMEA. Sontow *et. al., (15),* give an interesting extension of the FMEA methodology. This makes FMEA more applicable in early development process stages. The results of robust design exercises will give the knowledge needed to apply control charts to process controlling parameters.

In a special report in IEEE Spectrum, (52), a comprehensive listing and description of tools necessary for concurrent engineering is given. It gives a focus on the statistical nature of quality. Accordingly methods applied have to have a statistical base. Further there is no single method to cure all problems.

1.2. Problem identification

In this section we will outline some relevant problems in the application of quality engineering tools. This will be the basis for the aims statement in Chapter 2.

The Japanese have shown that: if the different quality tools discussed in Chapter 1. are applied together they can make a great contribution. Above all the most important single factor is the awareness and engagement of all staff in quality improvement efforts, (10). In Western countries the application of the different tools have come and gone one by one, (75). Instead of a search for a unifying theoretical base the arguments have gone along the lines of superiority of one of the tools over the other, (44), (45), (42). As an example we may consider the arguments around the robust design procedure according to Taguchi. A major criticism of the Taguchi S/N–ratio metric is given by Box, (69). He demonstrates the deficiencies with some carefully designed samples of data. That way he shows that the ability to discriminate different performance levels may at times be very poor. The alternative presented is the response surface technique. This latter technique engages itself more with the efficiency of experimentation than with relation to the variability which is the source of poor quality.

The quality issue is often confused with experimental design in the area of robust design. Discussion is often centered around experimental performance, because the variability issue is getting little or no attention.

Within the application of each of the different tools, different quality metrics have been designed. They are of different foundation and are involved to a greater or lesser degree. Robust design is a good example of this, (64),(80),(42). Phadke and Taguchi make some efforts to tie the metrics used to the concept of quality loss. Still it is somewhat involved and not straight forward.

Overall the area of metrics used in robust design is an area of great interest at the present time. Many researchers devote their time to this problem. A good overview of a procedure for statistical modeling is given by Nair *et. al.* (90). This procedure does not really consider the main idea of robust design, i.e. minimizing the influence of noise. The focus is instead on the effects on the sample average. An interesting discussion over the amount of experimentation involved with the Taguchi strategy for robust design is given by Shoemaker *et. al.* (54). Their main focus is on the number of experiments. However they do consider the robust design idea. They argue for a response surface strategy to identify how to evaluate variation decreasing control factors. This is of course a feasible method but it is involved. The complexity of the analysis is increasing exponentially as the system complexity is increasing. This makes the approach less accessible for users in general. In an article, (89), Box *et. al.* address the same problem with the number of experiments needed. The results are very specialized and not generally applicable. In general we may comment that quality is tightly related to variation. A metric not taking into account the variation is of minor interest.

TQM activities are often difficult to evaluate as no quality metric usually exists for management activities. Quality systems like ISO9000 are widely accepted as a quality tool. However they are often applied due to external pressure rather than internal push. This is due to the lack of arguments for the benefits.

In summary the problems that are the focus of this thesis are:

- Quality metrics do not have a common foundation.
- Quality metrics are very difficult to design for different quality techniques, i.e. TQM.
- Quality tools are considered as individual tools rather than as a coherent and comprehensive toolbox.

As a consequence we may state some characteristics of a quality metric:

- It should be as generally applicable as possible with different quality techniques.
- It should have a monotonic correspondence with the quality level for the process studied. The quality level should in this context be based on quality losses. Ideally it should be comparable between processes. At the least it should be comparable through a sequence of modifications of one and the same process, studied.
- It should have a foundation in the statistical nature of quality.
- It should have a dynamic characteristic to show quality level improvements even if the process is far off target.
- It should be able to reflect the state of the art in the area of interest.

2. Aims

The overall objective of the research presented in this thesis is to develop the means to generate a real shift in quality level in design. The key issue is to develop some integrating factor common to most available quality tools.

The work started with the intent to study different attitudes towards application of design of experiments within quality engineering. The initial studies showed that the differences were a matter of different adaptation of design of experiments to robust design. Further, some controversies exist in the area of quality metrics, (98), (79), (80). These controversies have caused objectives in the area of robust design to be confused. Researchers have even tried to find an introduction procedure for robust design. These procedures are aimed at preventing these controversies to stop the introduction of the robust design tool. (See Goh, (38), (20))

It was thus realized that the focus had to be widened to the overall objective given above. To achieve that three different aims were established:

- to show that a useful quality metric may be designed using a concept of information distance.

- to show that the proposed metric is applicable throughout most TQC–activities and that they contribute to performance of TQM– activities.

- to show that an information theoretical viewpoint forms an integrating and theoretical base for most quality techniques.

This will be done through several tasks.

First we need to establish a common understanding of what is quality. This is done through a discussion over quality costs in Chapter 3. There we will also find a discussion on quality loss functions. This ties quality costs to quality metrics.

As said above the origin of this thesis was in robust engineering. The work still has a focus on robust design using the design of experiments. The basis for the proposal lies in this quality tool. The applicability is explored through TQM. To this end we need to clarify what is robust design using designed experiments. This is done in Chapter 4.

The leading idea of this thesis is the application of information theory to give a unifying perspective on quality engineering. Accordingly we need to introduce some basic concepts of information theory. In Chapter 5., information theory basics with relation to quality evaluation are discussed.

A quality metric based on information distance will be outlined in Chapter 6. An application procedure will be shown. The procedure will show very strong connections between the quality metric and the quality losses.

The metric will be demonstrated in applications to quality engineering examples within the area of robust design, in Chapter 7.

In Chapter 8. product development processes are analyzed. A concept of flow of information surplus is applied. This gives as a natural result what were previously empirically founded quality techniques, i.e. complexity numbers, modularization.

Thus we can see the product development process as a matter of refinement of information. Accordingly a concept of information flow is applicable to the product development process as well as to the product or the production process.

The same information flow technique may be applied to on–line quality control. Thus the information distance quality metric is introduced to SPC applications. This is done in Chapter 9.

3. Quality costs.

One of the main points in the modern philosophy of quality is the relation between quality and costs. Taguchi has defined poor quality as the sum of all costs due to malfunction of the product or process. All costs refers to all different aspects of system impact. In this section we will review different type of costs in different parts of the product delivery cycle. A short summary of these aspects may be found in Reference (47). The same reference has a summary of quality costs and is repeated in Table 1.

PREVENTION COSTS	APPRAISAL COSTS
Quality planning and engineering	Inspection and test of incoming material
New products review	Product inspection and test
Product/process design	Materials and services consumed
Process control	Maintaining accuracy of test equipment
Burn–in	
Training	
Quality data acquisition and analysis	
INTERNAL FAILURE COSTS	**EXTERNAL FAILURE COSTS**
Scrap	Complaint adjustment
Rework	Returned product/material
Retest	Warranty charges
Failure analysis	Liability costs
Downtime	Indirect costs
Yield loss	
Downgrading (Off–speccing)	

Table 1 Quality costs, (47)

Burati *et. al.* (41), makes an interesting summary of quality deviation in different steps of design and construction within civil engineering. The information breakdown there resembles the approach taken below. A relation to the product delivery process gives an insight to the quality cost generators.

Chen *et. al.* (36), analyzes the quality cost distributions in the product delivery process from an accounting perspective. This gives a good summary to support the discussion below.

Quality in the cost sense is always related to product functions. These functions are those requested by the customer. Quality costs are not always directly visible. The visibility is low for costs of being late to market, costs of badwill etc. The visibility is intermediate for excessive production costs due to poor product and process concepts. The visibility is high for quality costs concerning rework, scrap and warranty. A general estimate is that the costs with low visibility are by far the greatest. Brown *et. al.*, (112), report the invisible quality costs to be 3 or 4 times the visible quality costs. Further discussions on this subject are found in Juran's Quality Control Handbook, (78).

Mechanisms which impair quality cost in different business activities can be listed as: Customer needs assessment, Product specification, Product definition, Manufacturing process, Assembly process, Product delivery and Customer product usage. These are discussed below. In later chapters these quality costs will be related to information flow in different levels of the business activity of a corporation. This is a somewhat unusual approach. The above summaries are more traditional. The reason for the present breakdown is an attempt to track the introduction as well as the extraction of excessive information into the processes. Accordingly we see a close relation to the activities of the total design procedure.

3.1. Customer needs assessment.

In this context the marketing activity has two main objectives. The first is to supply the correct information about customer needs to the rest of the organization. The second is to sell the products made by the organization, (63). In most organizations marketing and selling are different activities. Here we put both in the same bin to emphasize the cyclic nature of the process.

The first task involves analyzing customer needs and the level of competition. The information gathered must be in a form which is digestible and requested by the rest of the company. At the same time it must contain the real voice of the customer. These demands call for a variety of competences engaged in the marketing activity. We need to get the information not only in pure commercial terms but also in technical terms of product development usage etc. The consequences of poor performance in this respect are far reaching. The first and most evident is the development costs that are spent inefficiently, trying to develop a product not requested. Secondly, the business activity gets into a spiral of negative expectations. As a reputation of poorly performing products is build up, the customers do not expect products from this company that fulfil their needs. Eventually so much market share is lost that the long term existence of the business is questioned.

The wrong information assembled here allows the development process to add its own speculation over the customer needs. This usually introduces a large amount of excessive information.

The second task, to sell, is to reassess the needs of the customer. Thus, the selling effort is a process to identify a customer need or problem that an offered product (or duty) represents a solution to. If there is a good system to keep track of the voice of the customer then the product line will fulfil the needs of the customer. Even so there may be a sales staff not sufficiently trained to offer the right product to the present need. This will either end up with a dissatisfied customer or a specialized customer adapted redesign. Either way quality losses will develop, either by the customer or by the manufacturer. In the last case high losses may in the long run develop even by the customer. That may happen as the customer tries to get a non–standard product updated or repaired, (11).

The wrong product sold to a customer allows a lot of excessive information into the process, no matter whether the product is over capable or lacks capability. In either case unstructured information gets into the process. In the first case there are unused resources, i.e. waste. In the second case there will be efforts to try to get the product to perform what it was not designed to do. This will cause failures, i.e. waste.

The marketing activity has to be organized such that the customer needs are continuously updated. Further the business process has to be integrated such that the relevant information about customer needs penetrate through the entire company.

3.2. Design.

Quality costs within design are of two kinds. These are, working on the wrong problem and working with the wrong tools, i.e. product specification and product definition.

3.2.1. Product specification.

Working on the wrong problem is a matter of specification and comprehension of the specification. The costs incurred are redesign and a loss of prototype hardware. Further, the staff is less productive trying to solve what they may or may not believe to be a problem for a second time. Productivity is also decreased by miscomprehension of the specification. Even though the right problem is specified there may be slightly different comprehensions of the problem among the staff members. This will lead to discussions and avoidable mistakes downstream. These quality hazards call for a rigorous organization of the design teams and their working process, (43), (76). As the specification is usually not very stringent the concepts vary along the project. Laboratory resources have to be reserved early. The same goes for ordering of prototype hardware. Without this, it is possible to end up with tests being carried out on a prototype of an outdated concept. Effects of focusing on the wrong problem may also be seen after the launch of a new product. An excessive number of redesigns for customer adaptation is an example of often overlooked quality costs. Redesigns also appear due to the fact that some customer needs were simply not considered in the specification. This may also be seen in the late stages of the product development process, before launch. At this stage the number of design changes should go drastically down. However in many companies the number of changes increases as launch date is approached, (14).

We note how a loose specification allows excess information to enter into the process. It may be either through speculations of what may be the real specification or through late notice of new customer needs.

3.2.2. Product definition.

Working with the wrong tools is a matter of design process organization and staff training. Organizational problems refer to issues such as controversies existing between the design office and the laboratory for physical tests. In many organizations the habits of laboratory and of design are very strong. For reasons of habit simulations may be made by physical test rather than numerical tests or vise versa, (10). Of course this may also happen for reasons of inadequate training of staff. These kinds of deficiencies lead to solutions based on too little knowledge of the system. Reliability problems, maintenance problems and production problems can be seen as a result, if staff members have inadequate knowledge of technology and/or procedures. In a situation like that they try to compensate with guesses. Again excess information is being fed into the system.

The consequences discussed above are related to direct quality costs. In addition to this indirect quality costs will also be generated. One such cost is the loss due to late market entry. Another quality cost is customer badwill as an unreliable product is experienced. Some issues of this kind are discussed by Ashley, (37), as he discusses the benefit of applying Taguchi techniques to early design stages.

3.3. Production

Quality costs in production may be structured in several ways. One way may be to divide costs into those that are impaired by production process deficiencies and those that are impaired by product documentation deficiencies. Quality costs in production can also be divided into costs of material lost, quality assessment cost, direct production costs spent on failures and loss of production capacity, (77), (47). However we chose to break quality costs in line with the phases of production, i.e. manufacturing and assembly. It is thought that thereby it is easier to identify the cost generating parts of the process.

3.3.1. Manufacturing.

In manufacturing quality costs derive from several different sources. First we have the cost for quality assessment. The assessment is made in relation to the function of the final product under production. The product function is in this context described in the product documentation generated during the design activity. The function described in the documentation may or may not be optimal in relation to the original product function given in the product specification. A trivial example here is unnecessarily tight tolerances such as fit, surface roughness, etc. These tolerances set with little relation to the product function impair quality cost through excessive production costs, i.e. quality loss.

The quality assessment results can find that some objects are not complying with the specification. These objects are either scrapped or reworked. Scrapping leads to a loss of material. It also leads to a loss of all the direct production costs spent on these objects. The rework is an excessive use of production resources. We may also note that all efforts spent on failing objects are resources taken from a limited production resource. Thus we lose the corresponding part of the production capacity.

Among the objects being passed in the quality assessment, some are still not good. Either they are not discovered as being out of specification, or the specifications are not properly set to screen out all problems. Either way these objects will cause problems downstream, i.e. assembly and usage. Even when the objects are really good they can impair quality costs. As discussed above, we have the costs of quality assessment and excessive use of production resources. Tied to the latter is of course a loss of production capacity.

The discussion above is related to a situation where the production process and the product design are fairly well matched together. This is not always the case. That is to say that the design process does not take much notice of the different limitations of the production process. In situations like that there will be excessive tool wear, machinery breakdown and excessive maintenance. Production cycle times will be longer. Not the least important factor in this area is the ruined motivation of production staff. In summary we get an inefficient use of production resources.

The excess of information in this area is obvious. One may argue that there is actually too little information as the product documentation is usually not comprehensive. However this deficiency is compensated with a lot of information entered as the process is running out of control.

3.3.2. Assembly.

What has been said about quality costs in manufacture is to a large extent valid also in assembly. During assembly however it becomes very evident that the results from previous steps in the process influence the production costs. Looking into the details of quality costs in assembly we will note clear differences and comment on smaller deviations.

Although material loss is usually less frequent in assembly, quality assessment costs are still applicable. However the assessment costs may be minimized through a properly designed statistical process control in upstream stages and a careful product specification process, (101). As the material loss is lower in assembly the loss of direct production costs spent on failing products are less. Poor manufacturing performance as well as poor recognition of design for assembly may however cause excessive use of production resources. This accounts for both direct production costs and cycle time. The latter is especially important in this stage as capital bound in the product is high. Capital costs are thus impacting with particular strength. On top of this production capacity is lost.

3.4. Customer usage.

Quality losses in connection with customer usage may be differentiated in three ways; losses to the producing company, losses direct to the customer and losses to society.

Losses to the company are of several kinds. First there are customer claims over product performance which the customer does not find corresponds to expectations, even though the product is performing at its best. This may be a function of either over selling, a poor product specification or a product development result that does not fulfil the specification. This results in an overload in the sales organization.

Secondly there are customer claims that are a result of a product function breakdown. These claims result in company warranty costs. Obviously the company loses goodwill and in the long run market share from both of these types of claims. This latter part of quality losses are usually those that are recognized. In the light of what has been said in this chapter it is obvious that this is a minor part of quality losses.

The customer is subject to several losses. The downtime of the product represents a loss of earnings. In this context earnings come either from professional use or the benefits by private use. Getting the product up and running again costs the customer both directly and indirectly; direct costs are the repair costs; indirect costs are replacement costs during downtime. In this category there are also costs to get the product to a service center or service personnel coming to the product installation site. Where a malfunction causes accidents the customer may also suffer from costs for damage to third party. Further the customer may suffer loss of income due to illness caused.

The most evident cost to society is the environmental impact that the product may have. There are also other costs such as disabilities from accidents caused by the product. A good insight to this may be found by looking at regulations and their development around product systems pressure vessels and electric transmissions.

3.5. On quality loss functions.

In the above sections we have identified a lot of different categories of quality losses. The sources of these losses are also identified to some extent. In general losses may be cast into costs to different parties. This has also been done even though it has more been in terms of putting equivalence between losses and costs. Taguchi was among the first to draw on this as he was building a quality improvement strategy, (16).

Taguchi defined poor quality as the sum of all costs to society due to malfunction of the product. In this definition he puts equivalence between loss and cost. He introduced the quality loss function. The function he introduced was very simple, $L(y) = k\,(y - m)^2$.

In the equation above and the equations below in this chapter, the following conventions have been used:

L	is the quality loss due to poor quality, caused by one individual of a system performing at level y instead of the target level m.
y	is a performance measure for the product function considered
m	target value
k	is the proportionality constant in the loss function.
Q	is the average loss in the population.
n	is the population size or the population size at each instance of signal factor M.

l	number of levels used for the signal factor M
i	is a counting index.
j	is a counting index.
y_i	is the quality performance measure for member number i in the population.
M_j	is the signal factor value at instance number j.
p	is the fraction defective.
μ	is population average.
σ	is population standard deviation

There are many studies made showing that such a target value exists. For instance in the report (97) regarding color density of television screens, it was found that at a certain color density level the customer complaints over color density were a minimum. A further illustration can be the level of comfort at different possible room temperatures which is requested from an audience. Plotting the accumulated numbers of persons feeling discomfort with a feeling of the temperature being too hot at different temperatures, raising the temperature step by step from below, one graph is generated. Then lowering the temperature step by step from above and accumulating the number of persons feeling uncomfortable due to the temperature being too low produces a second graph. If the two graphs are summed together a graph with a minimum is generated. Often the minimum is located at 20^{o} C. At this temperature there are of course some people not feeling comfortable. However it is the minimum number at this temperature and at any other temperature there would be a higher number of people feeling uncomfortable. Accordingly an air conditioning machine controlling the temperature to the optimum would still generate some loss as there are people not feeling comfortable. Loss is then not zero but minimum. The same reasoning may be applied for the color density discussed above. There are often different mechanisms generating loss at either end of the spectrum.

At the target value the loss is minimum. Whether this minimum is equal to zero or greater is of minor interest. It is a matter of definition of quality costs. At the present stage of the presentation we may regard this minimum value as a reference value, zero. Hence we may conclude that the loss function given above is the first non zero term in a Taylor expansion of the general loss function around the target value. Accordingly we know that the above expression is valid at least close to the target value.

The costs impaired affect different parties, the company, the customer and the society. At the end those costs end up as a burden to the customer. However in a quality improvement process it is rewarding to keep cost categories apart. It may for instance be very difficult to apply the loss function in a general overall sense. Most people applying quality loss function use scrapping costs or warranty costs. We know however that this is a large under–estimate of the quality costs. These costs, scrapping or warranty, are distinctive costs recognized by the organization. Even so it may be hard to apply.

There is a need to follow up the quality costs differentiated in accordance with a product system structure. In turn this system structure has to be related to a product function structure. The malfunction of different product functions is the primary source of losses. Accordingly the product function structure is needed to break down the overall quality losses. With such data the loss function methodology can be applied stringently. This is to say real economical calculations in connection with financing etc. can be used. Some Japanese companies may have already achieved this. Most Western companies, which are concerned with quality loss, use loss functions in a comparative or prioritizing manner. An interesting and formalized way of applying the loss function methodology is given by Kraslawski *et. al.*, (31). They work with an adaptive process control system based on the quality loss function and fuzzy logic.

From the above sections we find that external costs, i.e. related to customer and society, may be an order of magnitude larger than the internal costs. The external costs are generally related to the function of the product itself. However an external cost may be identified due to environmental impact of the production process. This has often been considered as an external cost, carried by society. Recently environmental legislation has been changed. Accordingly the cost for environmental impact of the production process is gradually being turned to the producing company.

Internal costs are to some small extent related to product function, i.e. warranty costs. The remaining internal costs are related to the function of the internal processes. There are four related internal processes, the product development process, the production process development process, the production process and the marketing process.

To track the quality costs in these processes the ideal performance of these processes must be known. These costs are usually not considered using a loss function approach. As has been said above a strict application of loss function methodology requires a good function description. This is usually not available for three of the internal processes referred to above. The production process is often described with function terminology. Accordingly Taguchi has designed a on–line quality control procedure based on loss function methodology, (65). The remaining three processes are lacking the prerequisites for loss function methodology. These three processes are usually not well described such that the losses in the processes may be quantified. Furthermore only recently has a process–oriented description style been adopted. With that at hand a breakdown of losses per subprocess may be carried out.

Efficiency of experimentation is an issue in the experimentation process, a subprocess of the product development process. The metric of quality level is an issue of the product development process itself. Efficiency of the experimentation process is subordinated to the product development process efficiency. Thus these two issues may be separated to a large extent.

Modern quality assurance systems rely heavily on process descriptions, (81), (82), (83), (84), (85). Thus function structures will eventually be available even for the remaining processes. Several researchers have been engaged in such descriptions for those processes. Different process descriptions have been presented. The concept of Total design has been advocated by Pugh (76). A similar approach is given by Pahl & Beitz, (74) . The procedure of QFD has also won a great deal of attention during the last decade, (61). It has been elaborated first into EQFD, by Pugh & Clausing (49) and later into RCFA, by Sontow & Clausing (15) . Further discussions on this subject will be given in chapter 8.

3.6. Taguchi loss function and signal to noise ratio

The quality loss function as proposed by Taguchi emphasizes variability. This is because it is a continuous function with a global minimum. The variability in this context is measured as variance around the target value and not only around population mean. The loss function as referenced above gives the loss for individuals. As we are mainly working with populations in the quality improvement work we would like to carry this function over in terms of average loss, for the different types of quality characteristics (64).

Some quality characteristics *e.g.* wear are such that the less there is the better it is. The target value is zero. Those are called "Smaller is better". The loss function (average loss) for a population in this case is given by:

$$Q = k\left[\frac{1}{n}\sum_{i=1}^{n} y_i^2\right]$$

Taguchi proposed taking the decibel value of this average loss as a quality metric. The constant k was taken out of the formula as relative improvement was studied. Decibel value was argued to improve additivity of factor effects in quality improvement activities. The negative of the decibel value was called the signal to noise ratio, S/N. Hence S/N for smaller is better is:

$$S/N = -10Log_{10}\left[\frac{1}{n}\sum_{i=1}^{n} y_i^2\right]$$

"Nominal is best" is a type of quality characteristic where a distinctive optimum target value m exists. The room temperature example discussed above is a good example. This is a very common type of characteristic. The average loss for a population with mean μ and standard deviation σ is:

$$Q = k((\mu - m)^2 + \sigma^2)$$

When improving a characteristic of the type nominal is best there are two types of action to take. First the mean μ may be adjusted to target m. Thereafter standard deviation may be decreased. The adjustment also affects the standard deviation. If mean is adjusted with a factor m/μ the standard deviation becomes $\sigma m/\mu$. Hence the average loss for nominal is best after adjustment is:

$$Q_a = km^2\left(\frac{\sigma^2}{\mu^2}\right)$$

Again taking the negative of the decibel value of the average loss disregarding the constant factor the expression for S/N is:

$$S/N = 10Log_{10}\left(\frac{\mu^2}{\sigma^2}\right)$$

The handling of these kinds of characteristics usually involves a two step procedure, average adjustment and variation minimization. This recognition has lead to the introduction of a special class of performance measures, performance measures independent of adjustments, PerMia. With such measures the two steps can be performed independently. Alternatively use a one step procedure for process improvement. The proposals in this thesis is along this latter line.

Some quality characteristics are of the type that the higher the value the better it is. The target value is infinity. In reality it is very hard to find characteristics that are like that. Sometimes strength may be argued to be of that kind. This type is usually called "Larger is better". Taguchi used the inverse of that characteristic y, to form the loss function. The average for a population is then:

$$Q = k\left[\frac{1}{n}\sum_{i=1}^{n}\frac{1}{y_i^2}\right]$$

The negative of the decibel value gives the S/N–ratio:

$$S/N = -10Log_{10}\left[\frac{1}{n}\sum_{i=1}^{n}\frac{1}{y_i^2}\right]$$

Dynamic characteristics are measures where the input signal factor (see Figure 1 in chapter 4.) may vary and control the function signal accordingly. In this case the loss is a function of the deviation from the ideal performance, $y_j = M_j$. This loss may be expressed as:

$$Q = \frac{k}{ln}\left[\sum_{j=1}^{l}\sum_{i=1}^{n}(y_{ij} - M_j)^2\right]$$

For the dynamic system the ideal performance (target performance) is $y = M$. With linear regression a function $\underline{y} = \beta M$ may be fitted to the observed data y_{ij}. To minimize the average loss from the population y_{ij} is to adjust β to be equal to 1.0 and thereafter to minimize the standard deviation σ_e to form the line $y = M$. In analogy with what was done for nominal is best the S/N–ratio expression may be derived as:

$$S/N = 10Log_{10}\left(\frac{\beta^2}{\sigma_e^2}\right)$$

The above discussion is based on a linear regression. It is noted that it is a matter of a least square deviation exercise. It may also be the case that the ideal response is a nonlinear function $\underline{y}(M)$. In that case the analysis becomes very involved, a task for a specialist.

With a characteristic of the type "Fraction defective" you have two possible outcomes, acceptable or defective. For this kind of characteristic the arguments to design a S/N–ratio expression may be cast in the following way. If the fraction defective is p then you need to produce $1/(1-p)$ units to produce one accepted unit. The production cost for the $(1/(1-p)-1)$ excessive units is lost. Thus the average loss per accepted unit is proportional to $p/(1-p)$. Hence:

$$Q = k\left(\frac{p}{1-p}\right)$$

The negative of average loss gives an expression for the S/N–ratio:

$$S/N = -10Log_{10}\left(\frac{p}{1-p}\right)$$

It is noted that the reasoning for the last expression accounts for losses due to production costs. This may be alright if the system of prime interest is a production system. We will however come back to this as we later will discuss information content. Then we will consider acceptable results in general rather than strictly results from a production process.

Table 2 which summarizes the S/N–ratio expressions shows that for each different type of quality characteristic there is a separate formula for the S/N–ratio. For further details refer to Phadke, (64).

In later sections we will revert to the relation between quality metric and average quality losses. The new metric proposed in this thesis will present a unification in the way to analyze the different cases.

Problem type	**Adjustment**	**S/N–ratio expression** (comments)
Smaller is better	None	$-10Log_{10}((\Sigma_i\, y_i^{\,2})/n)$
Nominal is best	Scaling	$10Log_{10}(\mu^2/\sigma^2)$
Larger is better	None	$-10Log_{10}((\Sigma_i\, (1/y_i^{\,2}))/n)$
Fraction defective	None	$-10Log_{10}(p/(1-p))$
Ordered categorical		(Use accumulating analysis, see chapter 5.)
Continuous–Continuous dynamic	Scaling	$10Log_{10}(\beta^2/\sigma^2)$
Digital – Continuous dynamic		Divide the problem into two separate problems of C–C or Nominal the best type.
. . .		
. . .		

Table 2 S/N–Ratio formulas for some Different problem types.

4. Robust design

In Chapter 1. the concept of robust design was introduced briefly. A more detailed discussion of the philosophy will be given in the present chapter. In this section we will outline and comment on the procedure for robust design according to Dr Taguchi. He categorizes the different quality measures which can be used with the S/N–ratio metrics discussed in Chapter 3.. A review of the different categories is given below. A procedure description relevant for traditional experimental design exercise would differ particularly in steps a., c., e., g. and h. according to Table 3. The emphasis in traditional experimental design is on analysis efficiency while the emphasis in the robust design procedure is on system functions. In section 13.1. we will reproduce an application from Bell Labs. Comments on this example is given in section 4.2. In chapter 8. the robust design procedure will be put into the context of an information flow approach to quality engineering. In that context the reasoning for the different steps of the procedure becomes more clear.

Below a procedure for application of robust design is described. It is important to recognize that the procedure definitely calls for the work to be organized in a working group. The working group represent the present state of knowledge in the company of the problem area.

4.1. Robust design procedure

The procedure is divided into a number of different steps as summarized in Table 3.

This structured summary has been compiled by the author on the basis of the presentation by Phadke, (64).

Step	Description
a.	Quantifying system functions.
b.	System factor listing.
c.	Categorizing system factors.
d.	Design space and factor levels.
e.	Experimentation planning.
f.	Experimentation.
g.	Analysis.
h.	Verification.

Table 3 Robust design procedure.

a. Quantifying system functions.

Robust design is a quality improvement exercise. Quality is always related to a function of a system. The function has to be quantified for us to be able to judge the quality of the system.

The first step in the procedure is thus to decide what to measure. To be able to do this a function description of the system is needed. This could be made in accordance with the function analysis procedure proposed by Akiyama, (48). With the function description at hand, different measures to quantify the final function may be proposed. Thereafter a choice of the most efficient measure is made. The most efficient measure would be that with the most information content. We will later return to how to determine the information content of a signal (or measure). In general we will here state that continuous valued data contains more information than does categorized data. We conclude that the measures have to be classified. Further we note that the more information we get in the chosen measure the fewer the replications needed for each experiment.

In his book "System of experimental design.", (80), Taguchi makes a listing of different measure categories. These are summarized in Table 4 below. Some indications are also given as to what analysis tool to use.

Category	**Relevant analysis**
Simple discrete values (simple enumerative values).	S/N–ratios
Simple continuous values.	S/N–ratios
Fixed marginal enumerative values.	Accumulating analysis Frequency analysis
Multi–fractional values.	Accumulating analysis Frequency analysis
Multi–enumerative values.	Accumulating analysis Frequency analysis
Multi–variable values.	Accumulating analysis
Dynamic characteristics.	S/N–ratios

Table 4 Quality measure categories.

(1) Simple discrete values (simple enumerative values).

This refers to data that could be counted such as 1, 2, and so on, and could be for example the number of items sold, or number of particles. Analysis tools are usually S/N–ratios as discussed above. It has been shown that the S/N–ratio has a different definition depending on the appearance of the characteristic. It may be Smaller is better, Nominal is best, or Larger is better.

(2) Simple continuous values.

This term refers to continuous values such as weight, length, time, hardness, etc. The number of items sold in a day is discrete but the average number of items sold in a day is a continuous variable. As long as we stick to single values at each instant it does not matter from an analysis point of view whether the values are continuous or discrete. The same analysis tools are used.

(3) Fixed marginal enumerative values.

This is the case when a fixed number (known) of observations (data) are in some way classified into a limited number of classes. One may distinguish several different appearances of this kind of data.

i) *Ranked data* are the kind of data often created in a process of judgement by humans, i.e. grading as good, fair, normal, poor and unacceptable.

ii) *Off–scaled data* are the kind of data created when classifying continuous data by stratification in a non–proportional way. In a fatigue test we may classify life as follows, I) samples flawed from start, II) samples lasting 1 – 100 hours, III) samples lasting 101 – 500 hours, IV) samples lasting more than 500 hours.

iii) *Gauge values* are generated when stratifying continuous data on the basis of size in a proportional fashion. End classes may be unlimited as in the off–scale data example above.

iv) *Ordered Data*, this could be compared with ranked data but on a floating scale whereas ranked data are on a fixed scale. This kind of data is generated in the different evaluations in a beauty contest. If possible it is much better to use ranked data. This could be achieved through the use of one reference observation.

v) *Pure categorical* values; this is classified data where order between classes is of no importance. As an example of this kind of data consider the guesses in the game Paper–Stone–Scissors.

Data grouped according to (i), (ii), (iii) and (iv) are analyzed using accumulating analysis. Data according to (v) can be analyzed with frequency analysis. The analysis tools are not tied to the loss function in a straightforward way. Any attempt that may be tried to construct such a relation will be a challenge to the imagination.

(4) Multi–fractional values.

Multi–fractional values are very close to Fixed marginal enumerative data. Here the total number of observations is unknown, i.e. infinite. Hence instead of the number of individual observations in each class, the fraction of observations falling into each class is used. The data from a thickness gauging procedure of a steel mill can be used as an example. If these data are stratified, the fraction of observations falling into each interval gives the multi fractional values. The example outlined here gives ordered classes. Of course data may be considered to be multi fractional even though the classes are not ordered. For example in a household budget spending may be divided into fractions of spending in three classes 1) food, 2) other essential items and 3) miscellaneous expenses. In this case classes are not ordered and analysis is carried out using frequency analysis. With ordered classes accumulating analysis is used. A relation to the loss function may not be easy to design.

(5) Multi–enumerative values.

This is a case where a number of grades, i.e. ordered classes, exist. A certain number of observations are classified into these grades. The number of observations are unknown beforehand but countable. An example is the number of small defects, medium defects and large defects on a manufactured sample. The analysis type used is accumulating analysis. Classes that are not ordered may also be encounter. If in the example above you use defect type instead of defect size is used, frequency analysis instead of accumulating analysis is applied. Analysis tools are not founded on a common base.

(6) Multi–variable values.

This is very similar to the multi–enumerative values; the difference being that you have continuous valued data. As an example you may take the harvest of apples from a tree. The apples may be graded as first, second, third and fourth grade. Out of a harvest of 50 kg you may get 12 kg first grade, 18 kg second grade, 20 kg third grade and 10 kg fourth grade. These data (i.e. 12, 18, 20 ,10) are multi–variable values. Accumulating analysis is used. The base for this analysis is as pointed out above not tied to loss function or any common foundation.

(7) Dynamic characteristics.

For systems having a varying signal factor the function may still be categorized as above, see Figure 1. The difference here is that S/N has to be formulated with consideration of the whole range of system usage. The base for this S/N–ratio is in regression analysis. The connection to the loss function is fairly clear. Linear or non–linear regression is used depending on whether a linear characteristic of the system is wanted or not.

In the application of the Taguchi procedure it has come to be very common that the system is optimized to get a linear response relation. The relation may be written $y = \alpha_0 + \beta_0 M$. Here y is the function system signal and M is the input signal. α_0 and β_0 are intercept and proportionality constants respectively in the relation. Using the minimum square deviation approach, expressions for the optimum values for α_0 and β_0 may be derived. The data used in these expressions are the raw data for each control factor setup. When analyzing experiments carried out for this kind of system characteristics response tables have to be calculated for S/N, α_0 and β_0. The optimization is then a trade off choice of control factor levels to get maximum S/N and the closest possible agreement for the response relation, see Phadke (64).

b. System factor listing.

The second step is to list the different factors that are influencing the function of the system. As a guidance in this work there is the function description derived as above. An efficient way to organize the factor listing is the Ishikawa diagrams (94). A brainstorming activity with the objective to bring the relevant factors forward may be appropriate here. It is important to save this listing together with the priorities set later in the process. This listing is a source of information when the result of the experiment is that the present knowledge in the company is not good enough. This situation will be commented on further under the heading Verification.

A general system subject of study for robust design may be illustrated as a process flow shown in Figure 1.

c. Categorizing system factors.

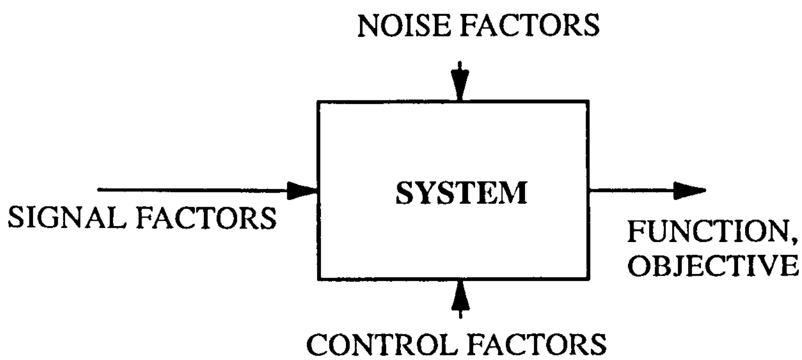


Figure 1 General dynamic system description

In *the third step* you have to categorize the different factors are categorized as control factors, signal factors or noise factors, see Figure 1. A very good discussion on this way of observing the system is given by LeÓn *et. al.,* (88).

Control factors are those factors, over which there is full control of a product or process by a designer. They may for example, be dimensions on a drawing or heat treatment prescribed. Signal factors are those factors that the user of a system uses to vary the function of the system. Sometimes the term design factors is used for these factors. This kind of factor may be exemplified by the steering angle on the steering wheel in the steering system of a car. There may not always be signal factors present. When signal factors are present we call the system dynamic. Notice the difference between the term dynamic here within the area of robust engineering on the one hand and in control engineering on the other hand.

Finally there are noise factors. These are factors that are not controlled. The reason why the noise factor could not be controlled may be purely economical. It is common, but not necessary, to split the noise factors into three groups: environmental noise factors, variation between individuals (i.e. production variations for example) and deterioration. Examples in the first group are ambient temperature, and humidity of air. The second group may be represented by production capabilities. I.e. the deviation of a control factor from the specified target value may be considered as noise. Deterioration is more obvious. It is wear and tear.

d. Design space and factor levels.

When the different factors have been listed and categorized we need to decide what values these variables may take. Thus *the fourth step* is to set up the boundary of the design space and decide how accurately it should be searched. At this stage it is also very important that the different factors are prioritized. In the scope of application of robust design the boundary may or may not be the ultimate limits. I.e. by improvement of an existing design or process the risk of total failure during experimentation may put more strict limits than would otherwise be the case.

As the first estimate of boundaries is established the level of knowledge concerning the effect of each of the factors can be evaluated. Factors that are well known may be treated with more factor levels than those which are less well known. Thus, the more levels there are, the more information is obtained. In this process decisions on the different factor levels for each factor is taken.

With regard to the different functions considered for the system under study the different factors may interact. The level of interaction among design factors has to be evaluated in detail. According to Taguchi those interactions that are certain to exist should be taken into the experimental design. Otherwise, less likely interactions should be left out. The term "certain to exist" is used to indicate the general agreement in the working group.

Interactions amongst noise factors are considered but never taken into the planning. The reason for this is that we are not interested in detailed knowledge of noise factors. The prime interest is to get the noise factor effects low. To illustrate that, consider the effect on maximum power from an internal combustion engine from a 5 degree centigrade shift of the ambient temperature. To what use is a knowledge of an effect of 3 units shift in the maximum power if the ambient temperature cannot be controlled. The engine system may not benefit from this knowledge.

The only time when interactions amongst noise factors may be of interest is when compound noise factors are designed. Compound factors are experimentation factors that are a mixture of several physical factors and can be an efficient means to reduce the experimentation while still achieving the robustness. Interactions between noise factors and control factors are the heart of robust design. They are handled by the use of an inner and an outer array. (See Figure 2)

Referring to Figure 1 the robust design study of such a system is carried out using two combined experimentation plans laid out as shown in Figure 2. The two experimentation plans are given as the inner array for control factors and the outer array for noise factors. Thus for each setup of control factor levels one system function result is observed for each setup of noise factor levels. Accordingly the interactions between control factors and noise factors is studied extensively. The outer array design is focused on getting the most possible influence of noise into the experimental results. That way the maximum effect of control factor settings on noise influence may be analyzed. The outer array design may be rationalized using compound factors as sketched above. A further discussion on compound factors is given below.

Figure 2 represents one way of taking care of the central issue in robust design. Minimization of sensitivity to noise is made through intensive exposure to noise during experimentation, using orthogonal arrays. Alternative ways may be used, as long as the central issue is respected. In reference (18) a comparative study between Monte Carlo simulations and the concept with orthogonal arrays is made. They found the concept with orthogonal arrays sound and the same conclusion was drawn for the use of SN–ratios. Similar investigations were done by Otto *et. al.* (26) and Liou *et. al.*, (24).

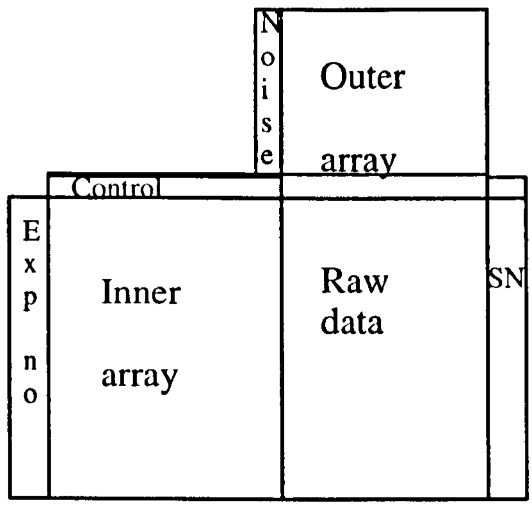


Figure 2 Experiment plan with inner and outer array.

Freeny and Nair, (39), introduced an interesting way of handling noise factors even though they are not controlled under experimentation. The mathematics of this approach is however very involved. In addition it does not give a direct connection to quality losses. This may be achieved by a proper choice of quality metric. That issue is not considered in this thesis.

Nedler *et. al.*, (55) and Pregibon, (96) have been investigating generalized linear models to produce a metric for quality. When it comes down to a mathematical, statistical rigor that is often applied on a trial and error base. I.e. there is no general procedure available to design a metric for each case application. The response surface modeling approach has recently also been advocated by McGovern, (7), (8) and Hamada, (25). This way of designing analysis tools is hard to distribute in a large community of average mathematically skilled users. A common framework which is more generic is needed. Alternatively the involved parts have to be embedded in a computerized support system where the complexity becomes transparent to the user.

Sandvik, (42), and Myers *et. al.* (40), argues in favor of including the noise factors in the same matrix as the control factors. This technique is more efficient if the level of interaction between the noise factors and the control factors is low. In such a case however the potential of robust design is low. In a situation where there is a high level of interaction the design needs to be adapted, resulting in an expanding matrix. It is with a high level of interaction that robust design has a major potential. This approach is a useful alternative but it is also more demanding on the user.

In many areas this discussion over efficiency is no longer so important. By computer simulations it is often possible to incorporate massive noise exposure at low additional cost. Welch *et. al.* (58), illustrates this situation through a simulation of a VLSI circuit design. They applies a quality loss function metric.

In the area of electronic circuit design the simulation tools incorporating noise exposure seem to develop rapidly. Liu *et. al.* (59) describe a comprehensive package for design and simulation. Even other areas such as production planning are subject to computer simulations using Taguchi strategies, (22), (23).

To complete this step, a priority list over the different design factors is made up. When setting priorities between the different factors the relation between factor and function measure should be considered. This relation, should according to Taguchi be of an energy character in a broad sense. This is very hard to grasp for many applications. Phadke discusses this in Chapter 6 of his book Quality Engineering using Robust Design, (64). He brings up the issue of monotonicity of the quality characteristics used. Monotonicity in this context has also been studied by Wilde, (28). The author believes that what is meant is an information relation. We will come back to this in later chapters.

e. Experimentation planning.

In *the fifth step* a decision has to be taken about the size of the experimentation plan. With the priorities at hand a first proposal of an experimentation plan is made. Control factors are fitted into an inner orthogonal array. The assignments of control factors to columns of orthogonal arrays may be done in either way according to Box, (98), or Taguchi, (80). Some researchers try to get those two alternatives closer together than they already are, (21).

The noise factors are fitted into an outer orthogonal array. This outer array is to control the number of repetitions of each control factor setting governed by the inner array. Further the outer array ensures a big enough influence of noise on the experimental results. A good way to keep the number of experiments down is to generate compound noise factors. I.e. a single noise factor combination that gives extreme noise stress to the system function is put into one (or two) compound noise factors. These compound factors may typically have three, four or five levels. You should at this point be very observant to the treatment of the noise factors in the outer array. It is as you see very ruff due to the fact that we are not particularly interested in the interactions amongst the noise factors themselves. This is so because in actual use of the system studied the designer can not control the noise factors. In the case of development experimentation you try to use controlled levels of the noise factors to be able to analyze the effect of control factor settings on their influence.

If the first experimentation plan does not fit into the economical limits present, it may be necessary to reassess the number of factor levels, the factor priorities and to make a new plan.

f. Experimentation.

The sixth step is to carry out the experiments and keep a careful log of results and factor settings. Preferably special data collection forms are used.

g. Analysis.

Analysis and prediction is *the seventh step* of the procedure. The data collected during experimentation are analyzed. An analysis method is chosen according to the category of measure used for the system function. (See discussions about the first step of the procedure and references (80), (64).) The analysis results are summarized in a recommended optimal setting of design factor levels. The optimality takes into account a trade off between different system functions. Further a prediction of system function performance is made.

h. Verification.

The most important step is *the eighth step.* A verification experiment is carried out with the recommended design factor levels. This experiment is replicated using the noise factors in the same way as the total plan. The result of this experiment is analyzed and compared with the prediction. In case of confirmation the design model proposed in the plan is confirmed. In other cases if the knowledge of the present system is not complete within the organization, factor listing and priorities have to be reassessed.

4.2. A robust engineering example

Appendix A an example of an application of robust engineering in production process development is given. The example is given as an illustration of robust engineering according to Taguchi as it is applied in many industries today.

The example is taken from Bell Labs, (64). It is an example of a process improvement in the manufacturing of VLSI chips. This example includes several different characteristics such as the number of surface defects and the thickness deposition rate. The part of the example that we consider is that one that deals with the number of surface defects.

The control factors for the manufacturing process used in the experiment are:

Table 5 Control factors.

Factor
A. Deposition temperature (C^o)
B. Deposition pressure (mtorr)
C. Nitrogen flow (sccm)
D. Silane flow (sccm)
E. Settling time (min)
F. Cleaning method

The control factors are assigned to the inner array. The noise factors are treated as compound factors in this example. The different noise factor levels are position of an individual wafer in the furnace and position of a cut chip on an individual wafer. This compounds the variation in gas flow, gas temperature and gas concentration. Thus the outer array is more or less collapsed into a row of observations.

The example illustrates several different ways to analyze experiments using different quality characteristics. It demonstrates the difficulties to find a consistent uniform quality metric. The accumulating analysis results are close to the results from the S/N–ratio analysis. However it is shown that the accumulating analysis is in some situations subject to subjective judgements. According to the recommended analysis procedure (see above) accumulating analysis and S/N–ratio are equally preferred procedures, for this kind of data, "number of defects".

The example also demonstrates the procedure of robust engineering. In particular the way that noises are handled are shown quite clearly. The way of compounding noises is often overlooked by researchers in their criticism of the amount of experiments involved, (42). The discussion about noises in connection with the example in reference (64), illustrates the difficulty with the concept of energy relations between system factors and system functions as proposed by Taguchi. There is a need for a general foundation that generates a better perception of these relations.

The recommendation for process setting from the robust design exercise in the example at hand, is a trade off between different relevant characteristics. This is a common situation. Looking at the recommendation of analysis methods there is no common quality metric. This is a problem as you come to a trade off situation. As the results from different quality characteristics are judged using different metric the trade off is very subjective. At this time it is worth noting that the traditional experimental design, (98), does not either give any assistance in this respect.

The data from this example will be used in chapter 7. to show the first application of a new quality metric.

We conclude some observations from this example:

1. Accumulating analysis includes a subjective judgement
2. Trade off between characteristics with different analysis methods is very hard.
3. A clear conection between the used analysis methods and the quality loss is not available.

5. Information basics in quality engineering.

In the previous sections we have discussed some tools used in quality engineering. Some of the controversies particularly around robust engineering have been brought up. In that area the quality metric used by Taguchi, i.e. SN–ratio, has been particularly a focus for criticism. This and the following Chapter will give a deeper analysis of one of the approaches to generate a consistent quality metric. That approach, information theory, was indicated by Taguchi (80). This section will be engaged with basic observations in this field together with some illustrative examples compiled by the author.

5.1. Information

What is information? In general information is associated with something read in a book or a newspaper. That is of course true but it is not a very precise definition. The question posed is a very hard one to answer. Some philosophical discussions are found in two books of Bateson, (87), (51). Bateson says, "information is a reference that makes a difference". The terms in which information are presented varies between different domains. For example in digital computing a series of "1" and "0" is a representation while in a play with two dices, the pair of numbers showing the number of eyes on the faces of the dices are another representation.

Looking at a digital signal we may subjectively reason about the amount of information in a signal. We consider two cases. The first case is characterized by the probability of a "1" being 1.0. The second case is characterized by the probability of a "0" being 1.0. In both of these cases it is true that if we know the value of one bit we with certainty know the value of the next bit. Accordingly after the first observation nothing new is to be found by further observations.

Now we look at a case where the probability of "1" and "0" are the same and equal to 0.5. Then we know nothing about the value of the next bit from the knowledge of the value of the present bit. Accordingly we learn something new with each observation. This case represents a situation of maximum information for a system with two possible outcomes.

Next consider a play with two dice. We make observations of the pair of numbers indicating the number of eyes on each individual dice. We get 36 different unique pairs of numbers that all have the probability 1/36 to appear. Thus we have no possibility to judge in advance of a throw which two numbers will appear next. In a number system, i.e. like the binary, a certain number, n_d, of digits is needed to describe the number of possible outcomes of our play of dices. This number n_d is proportional to *ln(36).*

Shannon, (102), introduced the number $E = -p\,ln(p) - (1-p)\,ln(1-p)$ as a measure of information in a binary signal, see Figure 3.

$$E = -p\,ln(p) - (1-p)\,ln(1-p)$$

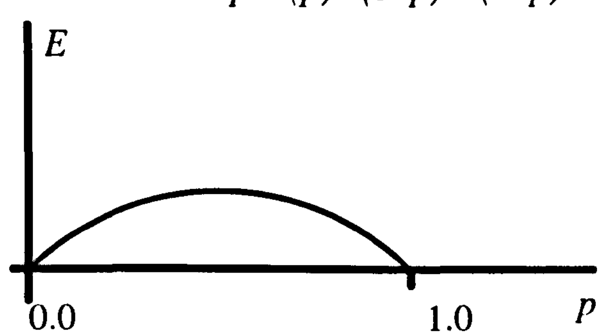


Figure 3 Entropy of a binary signal, with the probability p for a bit to be "1".

In general that number is written $E = -\sum_{i=1}^{n} p_i \ln(p_i)$ for a signal or observation with n possible outcomes. The number E is called entropy (or information). The number p_i is the probability of getting outcome number i.

Applying that expression to our play of dice above gives $E_{pair} = -\sum_{(i=1)}^{36} \frac{1}{36} ln(\frac{1}{36}) = ln(36)$. From that result we may conclude that the information in the case of number pairs in our play with dice is the number of digits you need in a number system to describe the number of possible outcomes.

Why should that complicated formula be used if just the logarithm of the number of possible outcomes could be used? There has to be something more to it.

To get some more reasoning for the more complex expression with probabilities we look a little more on our play of dice. Let us make some restriction on the observations. We do not observe the different pairs of numbers as individuals. Instead we take as our observation the sum of the two numbers in a pair. This restrict our outcomes to single numbers i.e. $\{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$. As different pair of numbers may give the same sum, the probabilities of the eleven different outcomes are not equal. In fact they are $\{\frac{1}{36}, \frac{2}{36}, \frac{3}{36}, \frac{4}{36}, \frac{5}{36}, \frac{6}{36}, \frac{5}{36}, \frac{4}{36}, \frac{3}{36}, \frac{2}{36}, \frac{1}{36}\}$. This gives us $E_{sum} = 2.270$ if we evaluate the general Shannon entropy from the previous page, using the eleven probabilities above. In comparison we have $E_{pair} = 3.584$. We note that our restriction on the observations has lowered the information by $\Delta E = 1.314$.

Our restriction is a structure that we apply to our observations. From the reasoning above we may conclude that this structure represent an information $\Delta E = 1.314$. Further we observe that the structure is documented as a probability density function. For a system we see that the less structure the higher the information content in the system.

Once again going back to our discussion on the binary signal we introduce a known pattern in the appearance of the "1" and "0". It may be that there is always 5 "1" and then 5 "0", repeating itself in a never ending series. We still have equal probability of "1" or "0" but the information in the signal is less with regard to the situation where there was no known pattern. This is due to the fact that we stand a better chance of guessing the next bit once we know the present, as we are aware of the pattern. This pattern is an other example of added structure.

Our discussion leads us to the statement :

* *Information content of a system is size and lack of structure.*

Interesting discussions on this may be found in the references (72),(86) and (100). The author has found the following visualization useful. The more effort that is needed to find some item in a body of material, the more information there is in the material.

Now let us bring up some aspects on information location. In reference (68) Wearn comments on information as a reduction of uncertainty. This is a very common pitfall. Reduction of uncertainty is in effect reduction of information. We can see that the comprehension of what is information is essential. What Wearn aims at is that if the information inherent in a system is moved to the mind of the observer then the observer becomes less uncertain about the system. This is a learning or cognitive process which is the speciality of Wearn. The location of information is obviously of importance. Issues like this have been discussed by Brillouin, (100).

Norman has a very interesting discussion on location of information in connection with user interfaces, in the book Psychology of everyday things. (70). A comparison of the driver environment of a modern car and the keypad for access of all features of a modern telephone is very illustrative. A common user exposed to a new car is capable of accessing most features after just a couple of minutes, without the aid of a manual. The same is not applicable for a modern telephone system. This show how information is shown to the user. Even though information theoretical terminology is not used the discussion is just about what Brillouin brings up. This is an interesting application of information theory in product development and quality engineering. We will come back to this in chapter 8.

Throughout this report we are looking into quality techniques resulting in a reduction of information in the system. The minimizing information level is one that allows the system to operate in just the desired way and nothing else.

In the discussion above we have assumed the observation space to be discrete. Hence we have worked with a discrete probability density function. Of course the same arguments can be made with a continuous observation space. We may illustrate the observation space with a continuous stochastic variable x. Then we have a probability density function $p(x)$. The information E is then written as $E = -\int p(x)ln(p(x))dx$. Integration is carried out over the entire observation space. From theoretical aspects the formulation over continuous observation space is of some interest. In application discrete observation space will however be created in some way or another. We shall see later that information on continuous observation space is somewhat different from information on discrete observation space. This will give certain consequences for the formulation of our quality metric.

In next section we will discuss quality in terms of information.

5.2. Poor quality – a surplus of information.

Consider a steering system of a truck. We may look at this system in terms of Figure 1.(i.e. the general dynamic system description from a robust engineering perspective). The turning angle of the steering wheel is the input signal. The turning radius of the truck is the function signal. The input signal contains all the information that the user wants to see in the function signal.

From the previous section we know that the information in the input signal may be found from the probability density function of the steering angle. Obviously we would like to see the same information in the probability density function for the turning radius of the truck. We know from experience that this will not be the case.

The reason for the difference in information is that one steering wheel angle gives different turning radii depending on the conditions. The conditions for the turning maneuvers may be weather conditions, road surface condition, tyre conditions etc. These conditions are of course nothing that a designer may put very much restraints on. From the probability density functions for the different factors representing these conditions we may calculate the information content of these factors. In relation to Figure 1 we may identify the condition factors above as noise factors.

The information of the noise factors is entering into the system. By influence on the system performance the noise information finds its way to the function signal of the system. We can see that the system function gets less determinate. In other words the structure of the system function signal is decreasing. Accordingly the system function signal contains more information than expected.

Assume that we have two different steering system designs with different susceptibility to noise factor influence on the system function signal. The difference may be differing values of the design factors, i.e. the control factors as identified in Figure 1. We would judge the one system with the lower susceptibility as representing better quality. This system would also show a closer agreement between the information of the input signal and the information of the system function signal.

Accordingly we conclude that poor quality is equivalent to an excess of information. Further this added information is entering the system as noise.

5.3. Information and quality evaluation.

In this section we will be engaged with some basic observations of properties of the information concept with respect to quality evaluation.

In reference (64) a paper feeding system of a copying machine, see Figure 4 was discussed. In this section we will discuss alternative quality characteristics for this system. This was done also in the given reference but we will use information arguments.

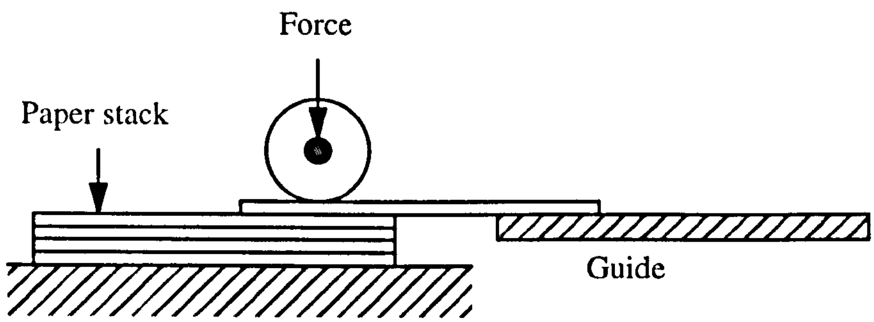


Figure 4 Simple sketch of a paper feeding system in a copying machine.

To be able to chose usable quality characteristics a function description of the system is needed. Figure 5 shows an example of a function description of a paper feeding system. We will concentrate on the function "feed one sheet". Failure to deliver this function either means feeding no sheet or feeding more than one sheet.

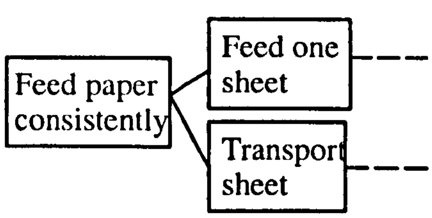


Figure 5 Function family tree of a paper feeding system.

Above we have seen that information content may be a useful quality measure. Let us evaluate two different quality characteristics. They will be described as they are analyzed.

The first quality characteristic to be considered is "fraction defective feedings", p, Figure 6. Two performance levels $p_I = 0.0001$ and $p_{II} = 0.00001$ are evaluated. Fraction defective represents a two class probability density function. The probabilities are p and $(1-p)$. We may calculate the information of each level of operation according to the formulae in section 5.1.

We will get $E_I = 1.021x10^{-3}$ and $E_{II} = 1.251x10^{-4}$ respectively. As performance level II obviously represents better quality we would expect lower information content in the signal from that case. This is also what we get. The difference is $E_I - E_{II} = 8.96x10^{-4}$. This observation agrees well with our previous discussion on poor quality as a surplus of information. Whether the dynamic in the signal is sufficient or not is an issue that we will come back to. Before we do that we will re–evaluate the system using another quality characteristic.

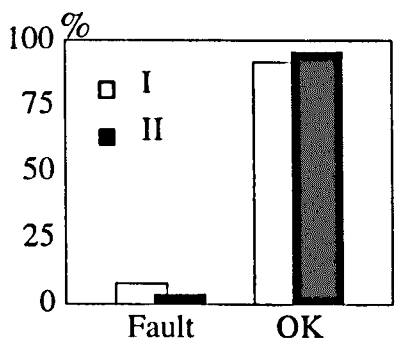


Figure 6 Diagram over two performance levels as shown by "fraction defective".

The second quality characteristic is constructed from the detailed knowledge of the design. As indicated above we have two failure modes for the function studied. That is, feeding no sheets and feeding two or more sheets. In relation to Figure 4 there are two threshold values for the feeding roller pressure force, F. The feeding roller pressure force is of course a subfunction of the main function feed one paper. There is one threshold for each failure mode, Figure 7.

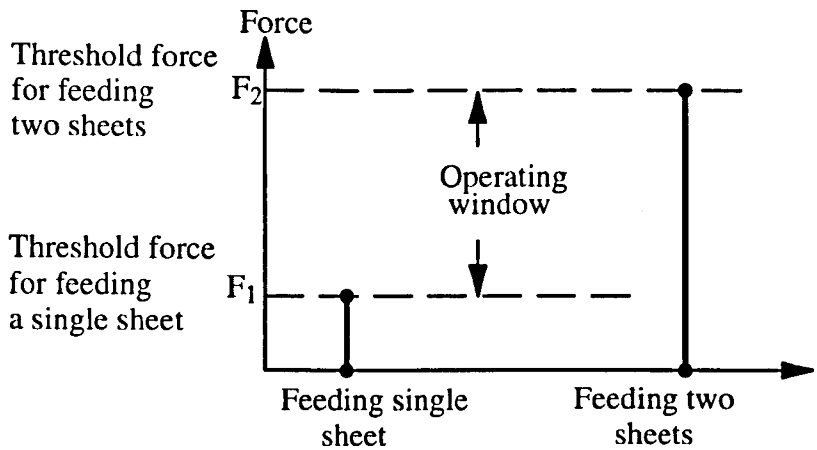


Figure 7 Operating window for roller pressure force.

The space between F_1 and F_2 makes up an operating window. Putting the operating force F_0 in the center of the operating window and keeping the variation of F_0, F_1 and F_2 as low as possible makes the design robust. We will get density functions for the thresholds and the operating force as shown in Figure 8.

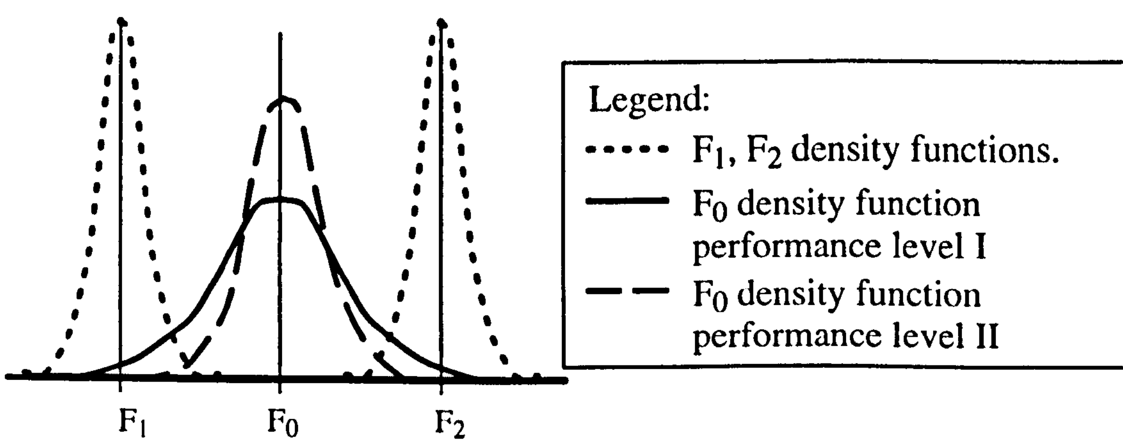


Figure 8 Two quality performance levels as shown by roller pressure force.

The overlap between the F_0 density function and the two threshold density functions represents the probability of failure. From Figure 8 we may see that three different robust engineering strategies are possible; 1, keep variations of thresholds down; 2, keep variation of operating force down; 3, a combination of 1 and 2. Figure 8 illustrates strategy 2. The two performance levels shown are the same as used for the "fraction defective characteristic, i.e. probability of failure being *0.0001* and *0.00001* respectively.

Normal probability density functions are used. Again using the formulae of section 5.1. we may calculate the information in the signal from each quality level. Continuous density functions are used. We get $E_{cI} = 4.452$ $E_{cII} = 4.143$. We observe once again that the performance level with the lower failure rate, i.e. higher quality, is producing a function signal with lower information content.

Further we observe the difference $E_{cI} - E_{cII} = 0.309$. Considering the quotient between $E_{cI} - E_{cII}$ and $E_I - E_{II}$ we get $Ir = \frac{E_{cI} - E_{cII}}{E_I - E_{II}} = 345$. The latter characteristic gives 345 times more information representing the same difference in quality level.

This difference in information between the different characteristics have two different reasons. The first is that a given design of quality characteristic can only represent a limited amount of information. In section 5.1. we saw examples of this. In our play of dice we had first a characteristic (number pairs) with a density function with 36 classes. The maximum information we get when all probabilities of the different classes are equal. In this case it corresponds to $E_{36max} = ln(36) = 3.584$. After our restriction we have only 11 classes left. Hence $E_{11max} = ln(11) = 2.400$. The actual information level that we got in the restrained case was $E_{sum} = 2.270$ which is lower than E_{11max}. Our restraint brings structure to the original set of outcomes in two ways. First it brings the number of classes down and then it brings structure to the eleven classes by a density function. We conclude that the maximum limit of information of a discrete valued characteristic is *ln(N)* where *N* is the number of classes.

The second reason for the information differences found above is very similar to the first. It is found as we try to derive the continuous probability function formulation of information as a limiting case of the discrete probability formulation. Trying to do that the continuous formulation will be found to be indeterminate by an arbitrary constant, see Shannon (102). The arbitrary constant is like the indeterminacy of the potential energy in mechanical engineering, i.e. an arbitrary choice of reference level gives you whatever potential energy number you want.

Before we come to any conclusions on the above discussion note that the information measure has the properties wanted, i.e. better quality implies lower information content. It is a monotonic relation. Further we know from robust engineering that a characteristic with more classes is better than one with fewer classes. The first statement is however true only within reasonable limits from an acceptable quality level. In our example we wanted to have all outcomes in one class, successful single sheet feeds. Thus the less information the better. In an example with pattern recognition on a photograph we may regard the color density function as a target probability function. That pattern represents the expected information. More information or less information in the photograph makes it harder to recognize the pattern.

We see that Figure 3, i.e. the Shannon information for a binary signal, is still valid for the two class characteristics discussed above in this section. Note that a system design consistently giving failures gives a system function signal with zero information. This observation is interesting as we try relate information content to quality level. A feeder system consistently giving good feedings also gives a system function signal with zero information. We need a mean to discriminate between the two system states.

In the next chapter we will see that a target distribution for an information distance evaluation does present such a mean.

6. A new quality metric

6.1. Information distance as a quality metric.

To design a quality measure based on information we need to have something that does not suffer from the indeterminacy of the continuous formulation. Further we need to direct the measure to what we want. Thus we wish to avoid getting an indication of good quality for something which is in just the opposite end of the spectrum. We recall the situation with the Shannon entropy for a binary signal in section 5.1., Figure 3.

On the basis of the above deficiencies, entropy or information has hitherto been disapproved of as a quality measure. Taguchi also argues this from practical reasons in terms of involved calculations, (80). The deficiencies could be handled however by using the differential measure of information using a fixed a priori probability density function, a target probability density function. The involved calculations are still valid as long as calculations are made by hand. The author believes this to be an irrelevant remark as the necessary calculations may easily be done on a pocket calculator. The author is in favor of an information theoretically based quality metric. Analysis of the results from a designed experiment contains two parts. Weighing the levels of a control factor against each other and evaluating the significance of the different factors against each other. The latter is done by analysis of variance, ANOVA. This part is not effected by the proposed new metric. The choice the best control factor levels is done using the new quality metric. The evaluations necessary for an information theoretically based quality metric is only a fraction of what it takes to do the ANOVA.

However over recent years researchers are directing attention to the possibility of using information metrics for quality. A very rigorous approach is made by Suh, (34). Suh proposes a very simple quality metric based on an analysis of the SN–ratio by Taguchi and the theory of information. This metric is given as a demonstration and it takes little account of anything else other than gross deviations. It does not have a clear connection to quality costs. However Suh is clearly advocating the relation between quality and information. An interesting study that tries to improve on the S/N–ratio concept is presented by de Boer *et. al.*, (56). After evaluation of several different adaptations of S/N–ratios originating from the quality loss function they propose a robustness coefficient. based on an evaluation of the overlap between the predicted output probability density distribution function and an expected or demanded tolerance interval around the target value. This value is based on pure mathematical statistics and does not draw on the physical process going on in the object under study. It is thus difficult to include a priori knowledge. Actually this measure comes quite close to the one proposed by Nam Suh, as stated above.

A very interesting investigation on quality metric or performance measures has been carried out by Leon *et. al.* (35). Their alternative metrics are evaluated on theoretical merits. SN–ratios as well as others are considered. Relations to quality costs are discussed. Leon *et. al.*, (88), also give a very interesting discussion on Performance Measures Independent of Adjustments, PerMIA. Box, (69), emphasis the need for PerMIA too The S/N–ratio due to Taguchi is not a PerMIA. Information metrics in terms of entropy are not discussed in those papers. Carroll *et. al.*, (57), give a very interesting investigation on PerMIA that is coming close to the metric proposed in this thesis.

In many areas information content has been successfully used to develop new technology and to advance the level of understanding. One such area is environmental monitoring as proposed by Wu *et. al.* (33). Another area is image recognition. In the terminology introduced above one would regard the searched pattern as the input signal. The actual picture where the pattern is to be found is the response signal. The difference between pattern and picture is the noise introduced via the picture reproduction system, (66),(86). This difference may also be regarded as added information. Special measures using differential information have been designed and used for this purpose, (66). In case the noise level is very high special adaptations of the information distance measure has to be made. At low noise levels the information distance as discussed below may be used as is. The a priori density function is then the searched pattern. The pattern is represented as a density distribution. This is compared with the density distribution for the picture. When the two density distributions are identical the differential information is zero.

We may regard the etching process in the printed circuit board production as an image reproduction process. Thus for this application the reasoning above is directly applicable. The differential information could be used as a quality measure. Quality is 100% when the differential information is zero.

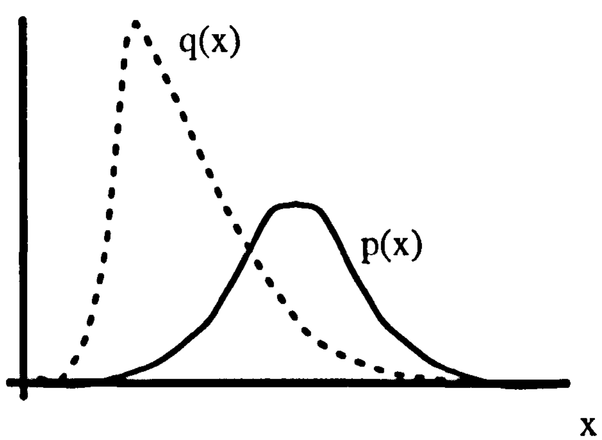


Figure 9 Information distance quantify difference between density functions q(x) and p(x).

Differential information could be written in different forms. One is the Kullback–Leibler distance, (86),(66),

$$D_{KL}(P:Q) = \Sigma_i\, p_i\, ln(p_i/q_i) \text{ or}$$

$$D_{KL}(P:Q) = \int p(x)\, ln(p(x)/q(x))\, dx$$

where Q, ($q(x)$), is the a priori density function and P, ($p(x)$), the actual density function, see Figure 9. This however has some drawbacks. Particularly this is true for continuous stochastic variables. One example is the values of x for which $p(x)$ or $q(x)$ are zero.

The a priori density function can be determined as the user has full control over that function. If the a priori density function is truly zero for some values x those values are impossible values and cannot be an outcome of the system function. The word impossible will be more understandable as we later talk about how to design the a priori or the target density function. The actual or observed density function, $p(x)$ is harder to handle. As it is based on a truncated population some x–values may be assigned zero probability even though it is a possible value. The problem in this case is that a computer implementation becomes more involved as limiting expressions for $p(x) \, ln(p(x))$have to be introduced as $p(x)$ approaches zero. In case that is not done overflow problems will appear.

Within information theory there are more general measures of differential information. These are often called information distances, (66). For our purposes one such class is particularly usable. This class stated as

$$D(P{:}Q) = \Sigma_i \, (1+p_i \,) f((1+q_i)/(1+p_i)) \text{ or}$$

$$D(P{:}Q) = \int (1+p(x)) f((1+q(x))/(1+p(x))) \, dx.$$

The function$f(.)$ is a general twice differential convex function. We propose to use $f(x)=xln(x)$. The added constant "1" may in fact be very arbitrary. However in combination with the particular choice of $f(x)$ the value "1" is very appropriate. That means the we get zero contribution for x–values where $p(x)$or $q(x)$ is zero.

Further we propose to make the measure symmetric in terms of density functions. A symmetric measure is introduced mainly for the sake of ease of use, i.e. minimizing the risk to misplace the two density functions in the formula. Hence our measure is

$D(P{:}Q) = \Sigma_i\ ((1+p_i\ )ln((1+p_i)/(1+q_i))+(1+q_i\ )\ ln((1+q_i)/(1+p_i)))$ or

$D(P{:}Q) = \int (((1+p(x))\ \ ln((1+p(x))/(1+q(x)))+(1+q(x))$
$f((1+q(x))/(1+p(x)))) dx$.

In these expressions p_i, q_i, $p(x)$ and $q(x)$ may be defined over a multidimensional stochastic variable x. For example the copper pattern on a, printed circuit board, PCB, can be defined as a density function of the position (x,y) on the PCB.

In the case of a discrete valued stochastic variable the above definition presents no problem. In the case of a continuous variable there are some problems in limiting cases. It may be noted that in real applications all variables will be handled as discrete variables. However the limiting cases give some good information about the nature of quality.

Limiting cases also represent strong arguments for the choice of distance definition. The present choice has some good properties in the limits as compared to the Kullback–Leibler distance. For example the former has a limiting value as the actual probability distribution goes to extremes with a given target distribution. Under the same conditions the Kullback–Leibler distance, D_{KL}, is increasing over any limit. See the following sections.

6.2. Information distance evaluated.

In this section we will evaluate the proposed information distance for some different cases. We will discuss the different properties disclosed in these exercises.

6.2.1. Smaller is better

6.2.1.1. a) Uniform probability density

The target probability density is assumed to be uniform in the interval (0,a). The response probability density function is assumed to be uniform in the interval (0,b). The information distance is plotted in Figure 10. Three different values for a, (0.95, 1.0, 1.05) is used. The distance comes to zero as b equals a. We note the fact that the distance is increasing as b becomes less than a. This is correct in terms of information distance but may be a little awkward in the light of quality. If we are at low cost, achieving a sharper probability density function than the target probability density function this should be judged to be high quality. Else the measure is contradictory. This situation puts emphasis on the choice of target probability density function to represent world class quality level. Properly handled this would not be a big problem in practice.

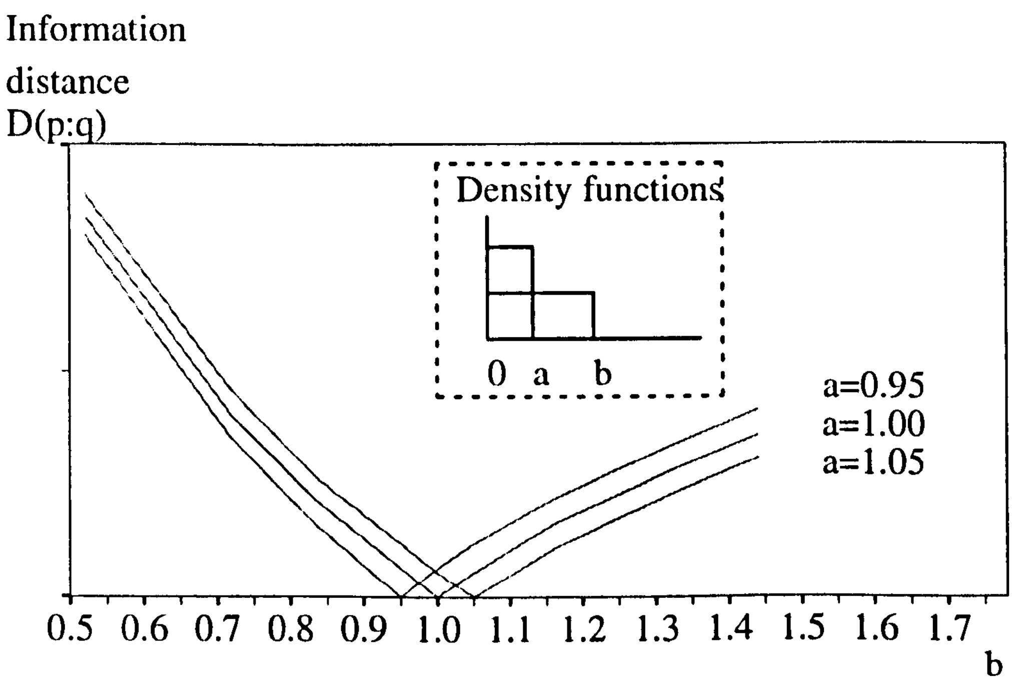


Figure 10 Information distance between Uniform probability density functions Uni(0,a) and Uni(0,b) as a function of the parameters a and b.

With the target distribution parameter, a, kept constant the proposed information distance approaches $D = \ln(1 + \frac{1}{a})$ as the actual distribution parameter b approaches ∞. As b approaches 0.0 D does not have a limiting value. In the light of the practical case this not particularly worrying. We often see a uncontrolled cost increase as we try to get a response to be on one single target value with 100% probability. In comparison the D_{KL}, *(Kullback–Leibler distance)* does not have any corresponding limiting values.

6.2.1.2. b) Exponential probability density function

The target probability density function in this case is $Q(x) = q\, e^{-qx}$. The probability density function of the actual response is $P(x) = p\, e^{-px}$. The information distance for q = 1.0 and varying values of p is plotted in Figure 11.

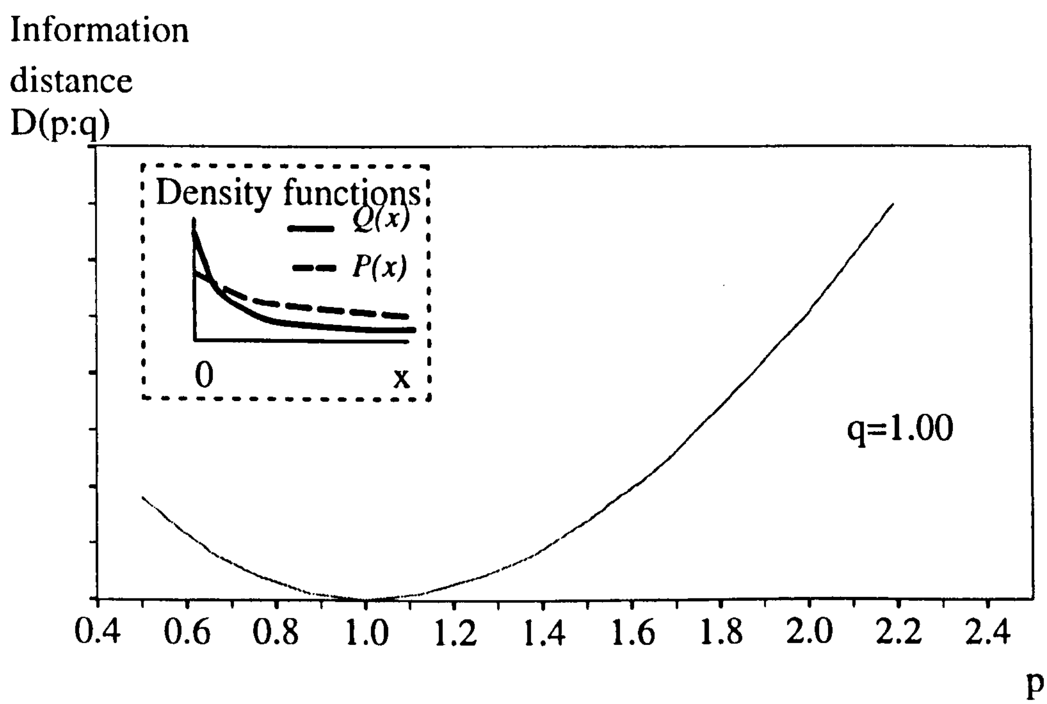


Figure 11 Information distance between exponential probability density functions $Q(x) = q\, e^{-qx}$ and $P(x) = p\, e^{-px}$.

Once again we see that the distance increases as the response comes closer to zero than the target. Still we have a minimum at p = q. With a good choice of target density function this will be satisfactory. In this context a good choice of target probability density function is one that reflects world class quality level. In that scenario an increasing information distance as the actual probability density function goes sharper than target reflects overspending. The quality level is not justified by the cost i takes to achieve it.

As the parameter p of the actual distribution approaches 0 or ∞ the distance approaches, $D = \frac{1}{q}\ln(1 + q) - 1 + \ln(1 + q)$. As a comparison the KL–distance is written $D_{KL} = \frac{p}{q} + \frac{q}{p} - 2$. This expression does not approach any limiting value.

6.2.2. <u>Nominal is best</u>

6.2.2.1. a) Uniform probability density function

In the case of nominal is best the target density function may be assumed to be a uniform density function in the interval (a_1,a_2). The response density function may be of different shape. Two different cases are represented in Figure 12 and Figure 13 respectively. The response is assumed to be uniform in the interval (b_1,b_2)

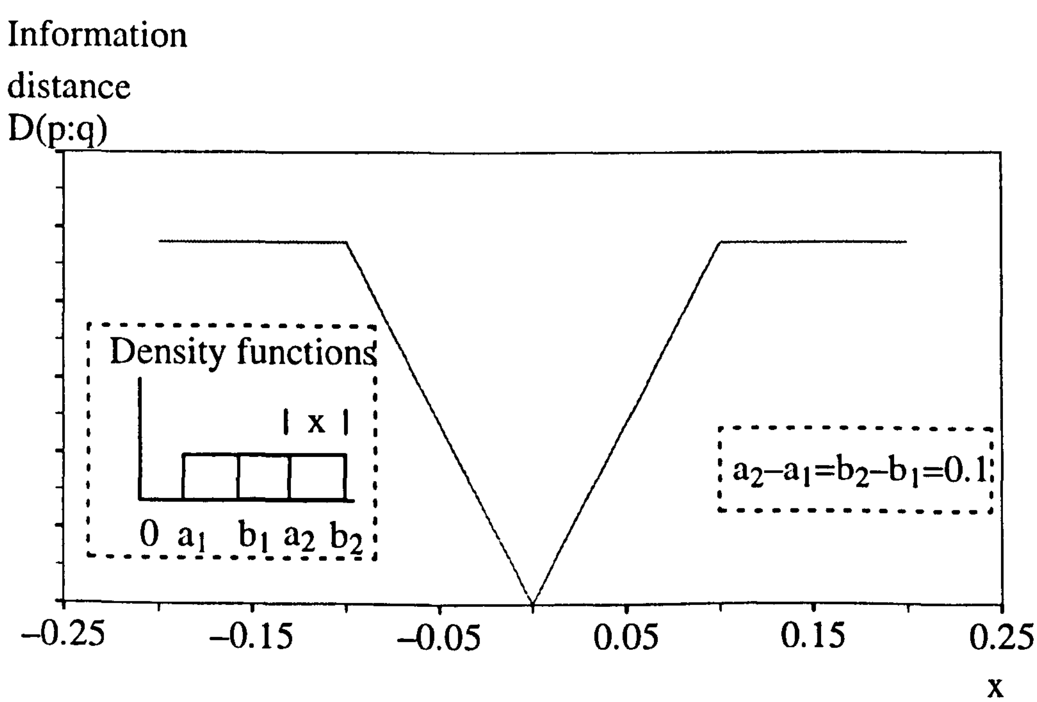


Figure 12 Information distance between two equal uniform probability density functions as a function of the offset x between them.

In the first case the width of the two intervals (a_1,a_2) and (b_1,b_2) is the same. The response is offset by the distance x from the target. The information distance as a function of the offset x is shown in Figure 12. We note that the information distance is zero when the offset is zero. When the offset becomes so big that there is no overlap between the two density functions the information distance remains constant no matter how big the offset x (assuming the width of (b_1,b_2) being constant). This limiting value is $D = 2\ln(1 + \frac{1}{a_2 - a_1})$.

This latter property is interesting. It puts some light on what is a target density function. In the present case, every response within the interval (a_1,a_2) is as good. Any response outside this interval is of no value. This is very well demonstrated with the information distance remaining constant when all responses fall outside (a_1,a_2), irrespective of the value of $(b_1+b_2)/2$.

The second case is one where the target and the response density functions are centered around the same point. The information distance as a function of x, half the difference in width between the density functions, is shown in Figure 13. The variable x is positive when (b_1,b_2) is wider than (a_1,a_2).

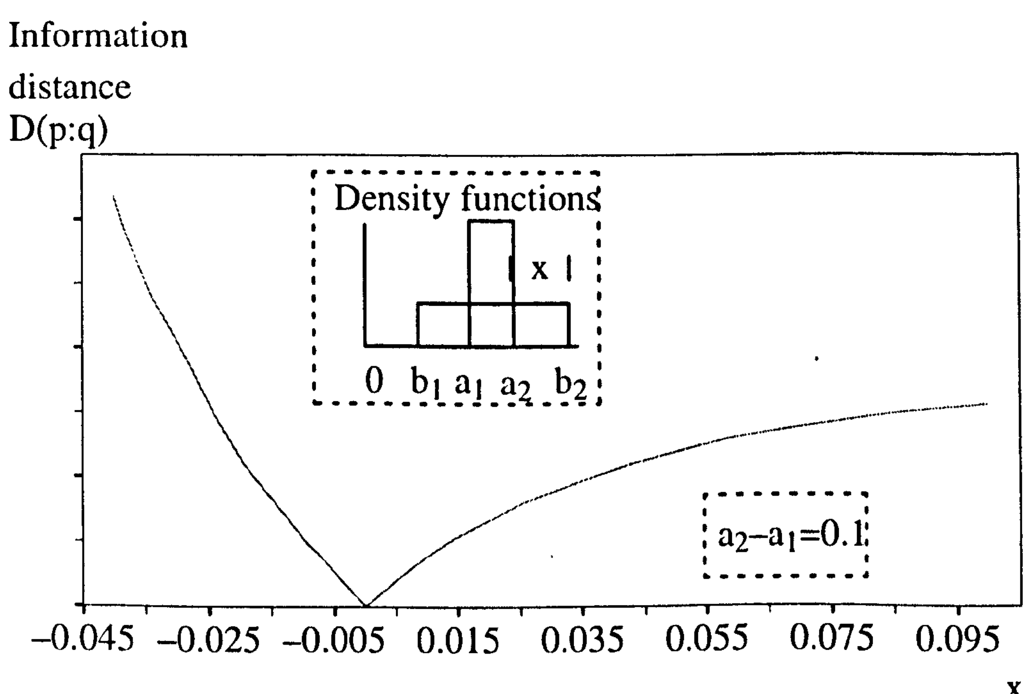


Figure 13 Information distance between two uniform probability density functions Uni(a_1,a_2) and Uni(b_1,b_2) with equal mean value.

We note that the information distance is minimum (zero) as the two density functions coincide. As the response density function becomes very sharp the information distance increases.

With the target distribution parameters a_1 and a_2 constant the proposed information distance approaches $D = \ln(1 + \frac{1}{a_2 - a_1})$ as the actual distribution parameter $(b_2 - b_1)$ approaches ∞. As $(b_2 - b_1)$ approaches 0.0 D does not have a limiting value. The D_{KL} does not have any corresponding limiting values.

In case of "nominal is best" and "larger is better" the exponential distribution can not be used. Evaluations for this case can then not be made.

6.3. Comments on the proposed information distances.

In the proposed formulae for information distance probability distributions, p(x), always appear as (1+p(x)). Because of this there is no contribution to the distance from a distribution p(x) or q(x) in an interval where the distribution has a zero value. In a process state where only waste is produced we will have no overlap between the two distributions p(x) and q(x). Thus the information distance has a limiting value built up of two independent terms.

$$D_{Limit} = \int q(x)\ln(1 + q(x))dx + \int p(x)\ln(1 + p(x))dx$$

Another extreme situation is where the actual distribution shows no concentration patterns whatsoever. This case could be illustrated with the exponential distribution limiting cases in section 6.2.1.2. above. There a limiting value exists.

As the actual probability density distribution gets sharper (i.e. shows more concentration) than the target probability density distribution we get another limiting case. It appears that at this limit a limiting value does not exist even for the proposed information distance. The quality losses at this extreme are often dominated with internal losses. Thus the company is overspending to achieve a performance better than expected, by the customer, i.e. a performance level far beyond the limit where the customer experiences an added value. For this there may appear losses at a magnitude that does not relate to the process performance at all. Accordingly the interpretation is that it is reasonable that the information distance may show very high values at this extreme. This will put the focus on unbalance of attention between different product processes. This is the same as saying that customer losses due to concurrent product processes are orders of magnitude larger.

6.4. Target probability density function.

The target density function is the best that would be expected from the output of the system. The response values corresponding to a very low probability intensity are values that are really not wanted even though they may appear under very special circumstances. From this point of view it may be seen that the target density function is related to the loss function. Response values that incur great losses to the company are also values that appear with a very low probability intensity.

The quality categories, smaller is better and nominal is best may be represented by a Kronecker $\delta(x-x_0)$–function as the target density function. When this is done it is also said that anything else but the value x_0 is of no value to the system user. This is seldom the case. The information distance introduced above tends to be infinite as the target density function goes towards $\delta(x-x_0)$. This then reflects that no values but x_0 can be accepted.

A similar situation appears for a target density function represented by a rectangular pulse in the x–interval (a,b). This target density function tells us that values for x outside this interval are of no value. That is to say that a value for $x = b + \Delta$ is as bad as $b + 10\Delta$. In terms of the above information distance, the distance is the same for either value of x.

None of the above cases are practical as they are almost never achievable. Instead a suitable target density function has to be chosen.

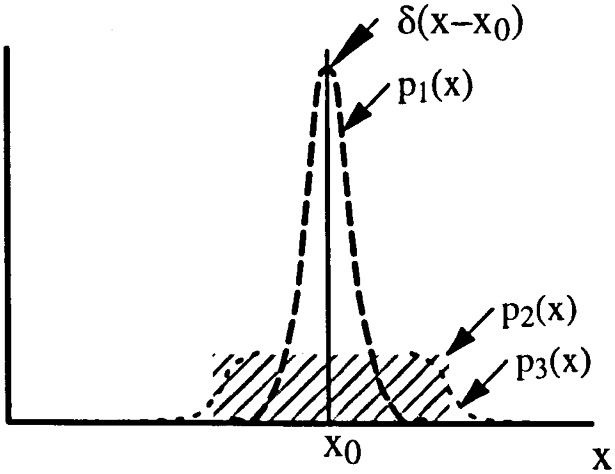


Figure 14 Special target probability density functions.

In cases where we are striving to achieve a $\delta(x-x_0)$ type of target it should be smoothed out like $p_1(x)$ in Figure 14. A similar procedure is valid for a rectangular box type target density function like $p_2(x)$ in Figure 14. That density function will get smooth edges as indicated by the density function $p_3(x)$. This procedure is very similar to a smoothing strategy used in some optimization algorithms. For example the penalty function technique may illustrate this, (95). An application in the area of quality engineering is demonstrated by Styblinski, (6). This shows how an optimization process may be tuned into the present stage of performance in terms of both positioning and dynamics.

6.4.1. Static characteristics.

A suitable target density function should be chosen such that performance becomes world class or such that quality losses come to acceptable levels in comparison with the present level of quality in the company.

The target density function is then designed with the aid of the quality loss function. To illustrate this we look at some examples.

For the categories smaller is better and larger is better, see Figure 15, a log-normal density function may be used for x and 1/x respectively as x is supposed to be non–negative. Some guidance for the choice of density function may be found in reference, (66). Using the loss function $L(x) = k\,x^2$ or $L(x) = k/x^2$, the average loss per piece becomes $k(\sigma^2+\mu^2)$ or $k(\sigma^2 \pm \mu^2)^3/\mu^8$, respectively. Here σ and μ is respectively representing sample standard deviation and sample mean

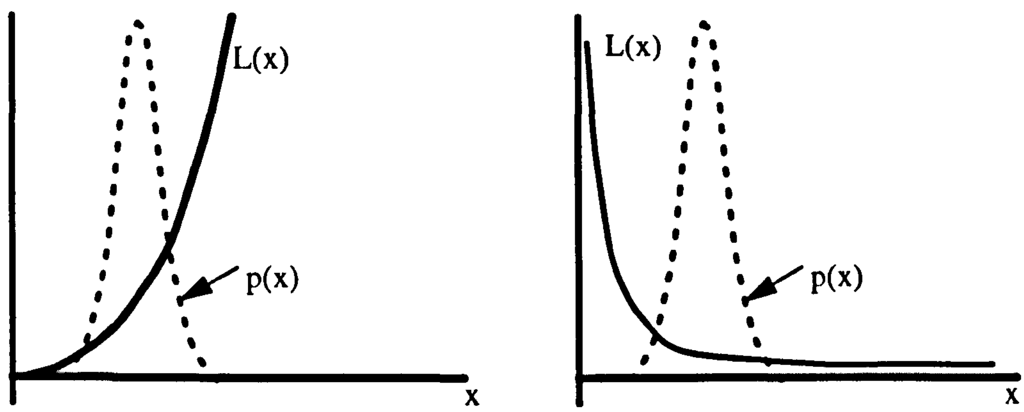


Figure 15 Loss function and target probability density function, "smaller is better" and "larger is better".

An exponential density function with parameter $1/\mu$ may be used as a target density function for smaller is better. The average loss is then $2k\mu^2$. A uniform density function in the interval (0,a) may also used as a target probability density function for smaller is better. This gives the average loss $k\,a^2/3$. In this case however the target probability density function properties as discussed in the previous section have to be considered.

Nominal is best has the loss function $L(x) = k\,(x-m)^2$, where m is the target value. If x is normally distributed the average loss is $k((\mu - m)^2 + \sigma^2)$, see Figure 16. With a uniform density function for x in the interval $(m-a/2, m+a/2)$ the average loss is $ka^2/12$. Again, the target density function properties need to be considered.

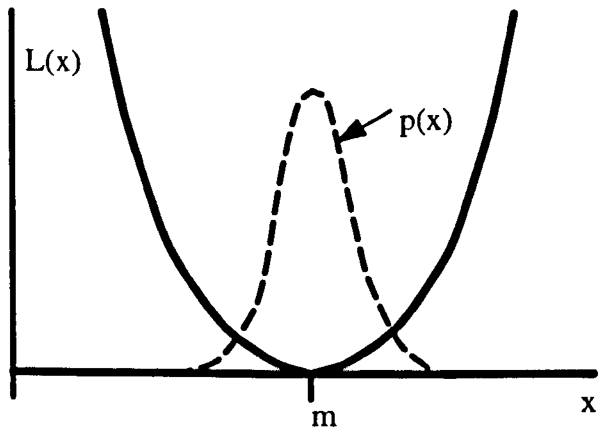


Figure 16 Loss function and target probability density function, "nominal is best".

By setting the target on average loss to world class, the parameters for the target distributions may be set according to the expressions given above. In this way the state of the art in the actual technology is directly tied into the development process. Assessment of the loss is very essential to this way of working. This puts emphasis on analysis of system function and system structure.

6.4.2. Dynamic characteristics.

The concepts from static quality characteristics may be readily transferred to dynamic quality characteristics. The loss function translates into a valley type function with its bottom following the ideal system function description in the s–F plane as shown in Figure 17. The cross section of the valley is the parabola that we have seen for "nominal is best" type characteristics.

The target function signal probability density function is a ridge type function. The locus of the maximum probability intensity is projected directly on to the ideal system function description in the s–F plane. The cross section of the target function signal density function may be a normal probability density function. The average of this normal probability density function is obviously for every s_i, coinciding with the ideal system function performance at s_i. The standard deviation $\sigma(s)$ is determined in an analogous way by evaluation of the overall loss according to the loss function.

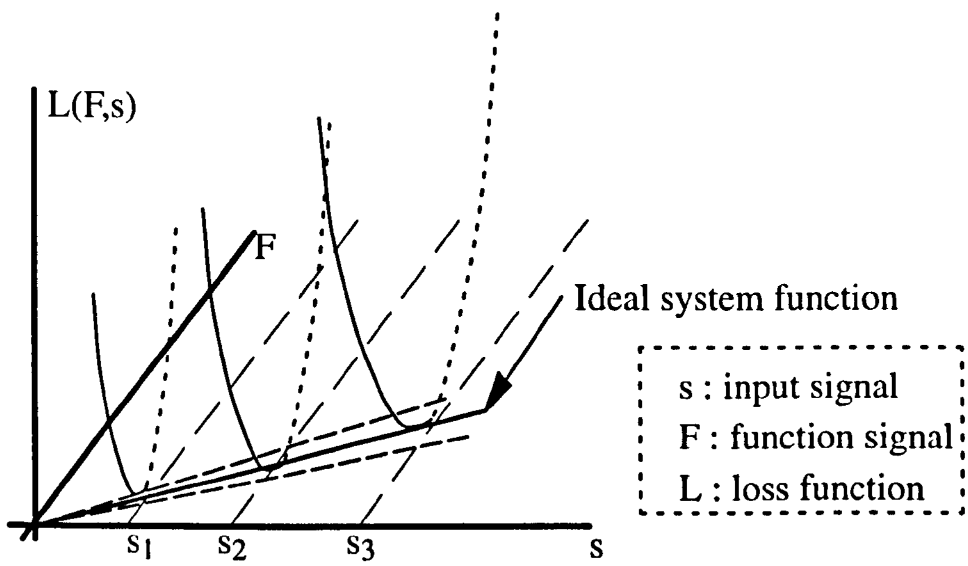


Figure 17 Loss function illustration for a dynamic characteristic.

In the general continuous valued case the procedure of determining $\sigma(s)$ is a matter of nonlinear regression. In most practical cases however there will be only a limited number of discrete values, s_1, s_2, s_3, ...

6.4.3. Information distance normalization

To create a quality measure applicable through the universe of processes a general normalization is needed. This is a task outside the scope of this thesis. However some fundamentals on the way towards that goal may be treated or discussed here. First of all we restrict ourselves to a sequence of improvement exercises to one and the same process observing the same function. Further we keep the target distribution function for this process function the same. The target function kept the same also tells us that the discretisation, i.e. the number of classes and the width of classes are kept constant.

Under this assumption the maximum assessable information is the same. As we alter our process under study, we address this same maximum amount of information. This means that comparison between process evaluations i.e. information distances from sequentially performed process improvement exercises are valid and relevant. Thus a process state that gives a smaller information distance than another is better.

Next we look into the possibility to quantify,quality absolutely. For most applications it is possible to identify some state where it is only producing waste. I.e. everything put into the process is turned into waste. That waste and the value of the adverse effects that the process results have on the environment represent an upper limit, L_{max}, for the quality loss connected to this process. In this state there is no overlap between the target probability distribution function and the actual process function probability distribution function. With the proposed information distance evaluation there exists a limiting value, D_{Limit}, for the information distance D corresponding to this case.

Thus we have D_0 corresponding to the acceptable (world class) loss level, L_{wc}, and D_{Limit} corresponding to maximum loss, L_{max}.In this way we have defined two fixed points for a scale linking information distance and the quality losses. Whether or not this scale is a linear scale or not is not proven in the general case. A special case is seen in section 6.2.2.1. From the first part of that section it may be judged that the scale for this special case may be linear. However this is dependent on the loss function being a box–type function as the target in that example. In the general case the loss function and the target probability distribution have a dual relationship. This suggests that to be able to put up an expression for the shape of the scale a valid expression for the loss function is needed. In most cases this expression is not known. Accordingly Taguchi has proposed a quadratic expression as a first approximation, (93). The argument for this expression goes along the lines of the actual performance not being unreasonably far from target.

For the limiting expression of information distance for exponential probability density function given in section 6.2.1.2. a linear relation may be derived between D and the distribution parameter q. This relation is valid for small values of q. In section 6.4.1. the parameter q, i.e. $1/\mu$, is related to the average loss corresponding to the target probability density distribution.

As said above the quadratic loss function is only valid in the neighborhood of the target. Further we note from above that a linearization does exist. Thus we propose a linearization of the information distance around the target distribution. We have set out a target probability density function to represent world class loss level. This corresponds to zero information distance. That is our first fixed point for our information distance scale. Next we use expressions like those in section 6.4.1. to calculate an actual performance probability density function p_{uf}(x), corresponding to $r+1$ times the losses at world class performance level. Then the information distance between that calculated performance probability density function, p_{uf}(x), and the target probability density function is calculated. This gives us the second fixed point, D_{uf}, for our linearized scale. At this point the excessive loss is r times the acceptable loss at world class performance level.

We thus define the normalized linear information distance as $D_{norm} = r\frac{D}{D_{uf}}$. D is the distance as evaluated using the formulae from section 6.1. D_{uf} is the calculated information at the upper fixed point as defined above. r is an integer defining the loss level at the upper fixed point as defined above.

With this scale the D_{norm} represents the size of the excessive loss in units of the acceptable loss at world class performance level. Thus $D_{norm} = 3.5$ says that the excessive loss is 3.5 times the acceptable loss.

6.5. Robust design using the new quality metric.

In chapter 4. the standard procedure for robust design according to Taguchi was described. The eight step procedure had the following headings:

- Quantifying system function
- System factor listing

- Categorizing system factors
- Design space and factor levels
- Experimentation planning
- Experimentation and data acquisition
- Data analysis, optimization and prediction
- Verification

This section will be engaged with the changes to the robust design procedure due to the new information distance metric. Looking at the metric layout above we note that the existence of a target probability density function is a prerequisite for the application. It is also obvious that the establishing of a target density function is a TQM activity. The target density function represents the long run quality target. Robust design is an activity to improve the system performance such that the target density function is approached.

6.5.1. System function evaluation.

In the previous sections the theory of the information metric was presented. It has been formulated for both discrete and continuous valued system function signals. The continuous valued cases are of interest in applications of theory development. In practice however the discrete case is almost always used. In fact the author does not know of any application with continuous formulation.

Above we have been talking about the choice of quality characteristics. Taguchi has argued against attribute data and in favour of continuous data. Consider an attribute characteristic using two classes, good and bad. From our discussion on information we may observe that the maximum information of a two class probability density function is $ln(2)$. As we increase the number of classes to, N, we get the maximum possible information $ln(N)$. From this we conclude that the Taguchi comment is sensible and the more classes the better.

In terms of optimization of the experimentation there is an optimum number of classes. On one hand we want the most information. On the other hand we want sensible sample density functions to an affordable experimentation cost. From the knowledge point of view we may adapt the class design to the present level of knowledge. I.e. if we know a lot we may have many and narrow classes. The experimentation cost aspect asks for few classes as that requires less observations in each sample.

Now as we know that the number of classes is related to the knowledge or information we may use that knowledge to chose amongst quality characteristics. The general rule is to chose quality characteristics that makes it easy to create many classes. That is an indication on the characteristic being rich in information. As a general rule we will search for characteristics that are rich in information.

The target density function given in general terms from the TQM activity may very well be presented as a continuous valued probability density function. The main properties of the target density function are that it represents the sharpness of the target and the value of improvement for the system function. For each single application, that continuous valued probability density function is adapted to the present design of classes.

6.5.1.1. Static characteristics.

In chapter 4. different groups of characteristics were discussed. They were: Simple discrete values (simple enumerative values), Simple continuous values, Fixed marginal enumerative values, Multi–fractional values, Multi–enumerative values, and Multi–variable values. Each group was assigned a preferred analysis method. Using the information distance metric as outlined above the same analysis method would be used through out all groups. This makes the group assignments obsolete. Instead we may concentrate on the design of classes. We observe that continuous valued characteristics will really be treated as classified data. This is achieved by dividing the value space into intervals or strata. Each interval represents one class. Irrespective of whether the classes are ordered or not the classes are given a sequence number. This sequence is then strictly adhered to. From that information cumulative classes from one end to the other are designed. Thus cumulative class I : sequence class 1, cumulative class II : sequence class 1 and 2, cumulative class III : sequence class 1, 2 and 3, etc are generated. The cumulative classes are needed to be able to make the predictions required. The predictions are made using the Logit–transform. In the accumulating analysis according to Taguchi the cumulative classes for the subjective judgements made for factor level settings are also needed.

Having chosen the appropriate quality characteristic and designed the relevant classes, we proceed with the target probability density function adaptation. It is easily shown that the information distance metric may discriminate the different test cases given by Box, (69), in his criticism of the S/N–Ratio. The discrimination could be to any level of accuracy.

As the target density function is set the previous robust design procedure is followed up to the analysis phase. The experimentation layout and the primary evaluation structure is shown in Figure 18.

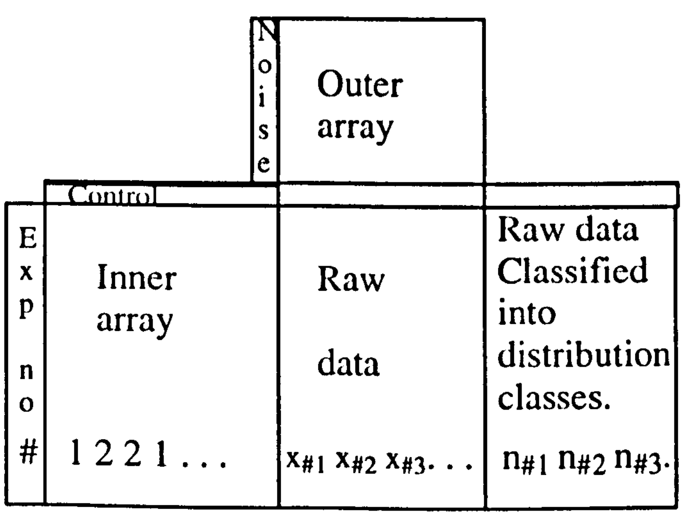


Figure 18 A standard experimentation layout for information metric application.

The raw data $x_{\#i}$ for each experiment # are classified into the same classes j as is used for the target probability densities, t_j. Thus the frequency numbers

$n_{\#j}$ are generated. On the basis of these frequency numbers the response table is calculated in accordance with the experimental plan.

Levels | Factors	...	f	...	
...				
k		$p_{fk1},\ldots, p_{fkj},\ldots$		
...				

Table 6 Response table with relative frequencies for control factors, f and levels k.

By dividing the frequency numbers in the response table with the total number of observations for all control factor set–ups #, that assemble factor level k, the relative frequencies, p_j are created. (See Table 6). Using the formulae proposed in section 6.1. the information distances $D_{fk}(p_{fkj}, t_j)$ are calculated.

The objective of the robust design activity is to minimize the information distance. Hence the results in the response table are utilized to choose the optimum levels of the control factors to give a minimum information distance. For each control factor, f, the level k which has the lowest information distance is chosen.

If D_{norm} is used in this analysis the actual level of improvement in excessive loss is assessed.

6.5.1.2. Dynamic characteristics.

Dynamic characteristics have not yet been studied in real applications, using the procedure outlined in this report. Hence there will not be any application discussed below. In principle however there is no difference between static and dynamic type quality characteristics using the proposed procedure.

In our discussion we will refer to section 6.4.2. and Figure 17 where the target density distribution is discussed. When handling a dynamic characteristic experimentally as many signal factor levels as judged necessary is chosen. Practically this will document itself as a repetition of the two rightmost columns of "rooms" in Figure 18, one for each level of the signal factor. The target density function will be a multidimensional function.

In an application with one signal factor and one system function, there is one set of system function value intervals, defining classes to the density function, for each signal factor level. The number of classes may vary for each signal factor level. The common approach will however be to use the same number of classes for each signal factor level. The major effort in the application to dynamic type characteristics will be in the design of the target density function and the class definition intervals. Apart from the level of quality this will also influence the sensitivity of the experimentation.

The evaluation will be performed in just the same way as for static characteristics the only difference being that the total number of classes is the sum of classes at each signal factor level. The frequency numbers for each class is assembled in the same way as before and the relative frequency is calculated based on the total number of observations. The information distance is calculated using the total number of classes.

In this way there is a response table with information distances generated as in the case of static characteristics.

6.5.2. Prediction.

With respect to prediction the traditional robust design procedure relies on an additivity property in the quality characteristic itself. One of the differences between Taguchi's version and the Fisher–Box version of experimentation is to be found here. Whereas Taguchi puts emphasis on the work before the actual experimentation Fisher–Box put emphasis on the analysis after the experimentation. The former tries to achieve additivity through a good choice of control factors and characteristics. The latter tries to validate a valid model, additive or otherwise, through statistical rigor in the analysis.

This is of course very appealing from a mathematician's point of view. However most practical engineers have a stronger ability in the domain of system performance for the studied product.

In the proposed procedure we always get a response table in terms of information distance in addition to the frequency response table. The prediction activity may be run either way, according to Taguchi or Fisher–Box. The author has a preference to the Taguchi way as the domain knowledge is more important than knowledge of statistics.

The prediction activity is done using the Logit–transform on cumulative frequencies corresponding to the control factor levels chosen from the response table. The prediction will give the cumulative frequencies $p_{COj,}$ for the optimum conditions. Through analysis of variance it will also be possible to get the confidence limits for the cumulative frequencies. See reference (80). The ANOVA is performed using the frequency data just as it has always been done for classified data.

It is important to emphasize what has been pointed out above. The information distance is a metric to evaluate the performance level of the product process. The ANOVA on the other hand is is a tool to evaluate the efficiency and quality of the experiment. It is not a tool to choose the factor levels giving the best product process performance level.

The level of improvement is evaluated using the information distance $D_O(p_{Oj}, t_j)$. Here p_{Oj} is the relative frequencies corresponding to the cumulative frequencies p_{COj}.

For dynamic characteristics the procedure is the same. I.e. create a sequence of all classes and work with the cumulative frequencies.

We observe that what we have achieved by the introduction of information distance is a generalization and a formalization of the accumulation analysis according to Taguchi, (80). The formalization is that judgement for the best control factor level is made from the information distance rather than a subjective judgement as proposed by Taguchi. The generalization is explained in the following way.

In accumulating analysis the cumulative frequencies corresponding to each control factor level are calculated. The cumulative frequencies are an estimate of the probability distribution function. The judgement used to chose the best control factor level is to chose that level that corresponds to the cumulative frequency that approaches 1.0 fastest. In this way it is possible to handle, smaller is better and larger is better, type characteristics in a reasonable way. Nominal is best is not as easily handled.

Looking at the probability distribution function what is wanted is a function rising from 0.0 to 1.0 as fast as possible to represent world class, as discussed above. The target density function is in the continuous case the derivative of the target probability distribution function. Thus the target density function represents the steepness and the location of the rise in the probability distribution function.

In this way we may consider the proposed procedure a generalization of accumulating analysis that is able to handle, smaller is better, lager is better, nominal is best and dynamic type characteristics in just the same way throughout.

It has been said above that the Logit–transform is used for the prediction of the probability distribution function. It is always applied to the distribution function (i.e. the cumulative frequencies) to get a stable performance of that prediction procedure. If non–cumulative classes had been used the prediction in different classes would be uncoupled and the resulting probability density function could add up to an accumulated probability larger than 1.0.

Looking at the prediction in one class the Logit–transform may be analyzed by applying the information concept.

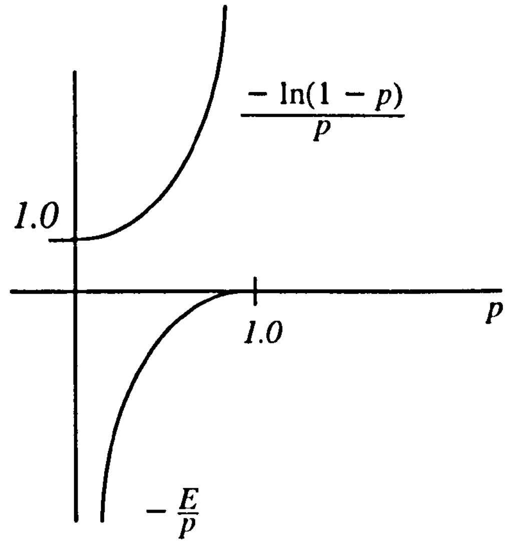


Figure 19 A information breakdown of the Logit–transform.

First we introduce a slight redefinition of the Logit–transform.

We call it Ω_{ln}. We write $\Omega_{\ln} = -\ln\frac{1-p}{p}$.

This may be rewritten as $\Omega_{\ln} = -\frac{E + \ln(1-p)}{p}$.

Here E is the Shannon information for a two class, i.e. binary, signal, $E = -p\ln(p) - (1-p)\ln(1-p)$.

The two terms of the Ω_{ln} expression are illustrated in Figure 19. Going back to our discussion in section 5.1. we recognize that E represent structure of the data represented in the presently analyzed class. A high value of E is representing a low level of structure. Next we take a look at the term $-ln(1-p)$. Let p represent the probability of the outcome being what we want to get. Then if we want to learn anything more about the outcomes represented by the probability $(1-p)$ we need to resolve the information to a corresponding degree of accuracy. What is the corresponding degree of accuracy? The number of observations needed to get any relevant number of observations of "$(1-p)$"–outcomes are of the order $\frac{1}{1-p}$. We remember again from section 5.1. that the information in such a number of observation is proportional to $\ln\frac{1}{1-p} = -\ln(1-p)$. This is exactly the second term. Thus we may say the amount of information we need to interrogate in order to learn anything new is given by the second term.

The factor $\frac{1}{p}$ may be interpreted as an uncertainty over what we know about the outcome that we want to get, i.e. the data represented by the probability p. Looking at E we want it to be low. This is the case both for p close to 0.0 and 1.0. However as p is close to 0.0 we know very little about the outcome we want to get. As p is close to 1.0 the situation is the opposite. This is represented by the factor $\frac{1}{p}$. The term $-\frac{E}{p}$ is negative as p is close to 0.0. This represents a lack of information to say anything with any certainty for the outcome we want to get. Any new observation will at this point, with a high probability be of "*(1–p))*" group. Thus we will learn more about that group of data to say whether it may be anything of value or not.

Now we may consider the Ω_{ln} to be a measure of the amount of information to be interrogated to improve our knowledge of the out come we want. Of course we want this number to be as high as possible. I.e. if our knowledge is low very little is needed to get us more knowledged and vice versa.

The prediction procedure using Logit–transform may in this light be described in the following way:

The grand average of the analyzed cumulative class represents the present level of knowledge in terms of additional information to be interrogated for improvement. For each design factor, that factor level that will increase the information needed to be interrogated the most is chosen. Compare the actual calculations performed at the end of section 13.1. in the appendix A.

The author has also noted that Ω_{ln} is the negative of the derivative with respect to p of the Shannon information in the present class: $\Omega_{\ln} = -\frac{dE}{dp}$. Analyzing this novel formulae that has not been observed before we see: As we want to increase p we have to decrease E. That is the same as striving for big values of Ω_{ln}. We can say that the E–price for change in p should be high. *That is the principle of prediction using Logit transform.* This also agrees with the principles of maximum entropy discussed by Kapur, (66).

6.5.3. Verification.

The result from the prediction activity is a probability distribution function. In the long run that distribution function will be verified by production results. The verification activity in the robust design procedure is an additional experimentation run. The setup of that experimentation run is the optimum control factor levels generating the minimum information distance according to our analysis of the designed experiment plan.

In the verification we use the distribution function from the prediction. An ANOVA presents a possibility to calculate a confidence interval for the probability distribution function for the characteristic studied. This will be demonstrated in the following applications.

6.6. Summary of the new metric contributions.

In the present chapter a new quality metric has been introduced. It has been demonstrated that it is unlike any other metrics tieing TQM activities like benchmarking into the metric through the target distribution. It has been shown that while traditional quality engineering is designing unique metrics for different applications the new quality metric is applicable throughout all applications. The target probability density function is designed from product process focus. This gives a natural focus on product process performance rather than statistical rigor of experimentation by design of the target function. The new metric relies on frequency table results. Thus the development activity of the experimentation efficiency could be focused on one form of data production.

Further the new metric is designed from the basis of information theory. This has lead the way to the tight connection between quality engineering and information theory. A novel interpretation of the Logit–transform from this perspective has been presented. Further application of the information theoretical perspective that give new interpretations will be presented in later chapters.

7. Robust design application of new quality metric

In the following sections the application of the proposed procedure to some examples will be illustrated and discussed. The first is the example from Phadke, (64). This will be followed by some applications from manufacturing of major plastic appliances, such as sewage pipes.

7.1. Etching process for silicone chips.

In section 5.5 of the book Quality engineering using robust design, (64), Phadke demonstrates a special technique, accumulating analysis, to analyze ordered categorical data. The same set of raw data has also been analyzed using the traditional analysis in the same reference. These analyses are reproduced in appendix A.

The example of appendix A is an application of robust engineering in production process development. The example is used to illustrate the application of the new quality metric.

The example is a process improvement in the manufacturing of VLSI chips. This example includes several different characteristics such as the number of surface defects and the thickness deposition rate. The part of the example that we consider is that one that deals with the number of surface defects.

The control factors for the manufacturing process used in the experiment are listed in Table 7 below.

The control factors are assigned to the inner array. The noise factors are treated as compound factors in this example. The different noise factor levels are position of an individual wafer in the furnace and position of a cut chip on an individual wafer. This compounds the variation in gas flow, gas temperature and gas concentration. Thus the outer array is more or less collapsed into a row of observations.

Table 7 Control factors.

Factor
A. Deposition temperature (C^o)
B. Deposition pressure (mtorr)
C. Nitrogen flow (sccm)
D. Silane flow (sccm)
E. Settling time (min)
F. Cleaning method

This set of raw data has been used to show the properties of the new procedure. Two cases have been analyzed using two different target probability density functions. The class definition used in the present analysis is the same as in reference (64), Table 8.

Table 8 Categories used in ordered categorical data analysis of defects.

Category number	Observation category (defects)	Cumulative category (defects)
I	0 – 3	0 – 3
II	4 – 30	0 – 30
II	31 – 300	0 – 300
IV	301 – 1000	0 – 1000
V	> 1000	$0 - \infty$

The first probability density function is, probability 1.0 for class I and zero for class II through V. The second target density function is binomial Bin(5,0.05). The response tables in terms of information distance are shown in Table 11 and Table 12 respectively.

The analysis by Phadke gives from a visual inspection of the charts in figure 5.5 page 125, of reference (64), the following optimum choice A_1, B_1, C_3, D_2 (D_1), E_1 / E_2, F_2. The analysis with S/N–ratio in chapter 4, of reference (64), gives A_1, B_1, C_1, D_1, E_2, F_2. In a trade off between different system function the optimum choice becomes $A_1, B_2, C_1, D_3, E_2, F_2$. (See appendix A .)

The ANOVA is carried out using the accumulated frequency number data. See reference (80).

Table 9 ANOVA–table for 5 class data

ANOVA for 5 class data				
Factor	dof	Square Sum	Variance	F–ratio
A	8	133.6	16.69	716.
B	8	104.3	13.04	559.
C	8	24.6	3.08	114.
D	8	15.4	1.93	83.
E	8	10.0	1.25	54.
F	8	33.5	4.19	180.
Error	110	2.6	0.02	
Total	158	324		

Table 10 Predictions for optimum combination, on 5 class data.

Predictions for 5 class data				
Total class	probabilities			
0.303	0.469	0.642	0.778	1.000
Logit–transform prediction		(A1,B1,C1,	D1,E1,F2)	
3.66	12.02	24.54	35.88	
Predicted class probabilities				
0.6992	0.9402	0.9965	0.99997	1.0000
Upper class limits				
3	30	300	1000	∞

We note from Table 10, that with a probability of 0.778 we will get fewer than 1000 surface defects judged from the overall data. The predicted opti-

PAGE MISSING IN ORIGINAL

1 to factor level 2 as the binomial distribution parameter goes from 0.0001 to 0.1. This compares well with the graphs for the cumulative probabilities found in Appendix, A. The cumulative graphs corresponding to level 1 and 2 respectively are quite different. Accordingly we are not surprised that they come close to different target distributions. We further note that the significance of the factor effects are good (see Table 9). Still there may be an ambiguity left in the choice between two factor levels out of three.

From the above discussion we may see that the procedure of factor level choice based on information distance is more stringent than the visual inspection proposed earlier for accumulating analysis. On the other hand we see that the choice of target distribution really makes a difference. Accordingly with the information distance procedure we get a method to consistently control where we are heading with our improvement efforts.

Table 13 Response table. Target probability density function Bin(5, 0.1)

Factor	Level 1	Level 2	Level 3	Optimum
A	0.05	0.29	0.40	A_1
B	0.04	0.13	0.50	B_1
C	0.15	0.27	0.17	C_3 (C_1)
D	0.17	0.18	0.24	D_1 (D_2)
E	0.14	0.15	0.25	E_1 (E_2)
F	0.24	0.11	0.29	F_2

To investigate the importance of the target density function further we made another study of the surface defects problem using 11 classes for our density function.

We note that the 11 classes data agrees well with the 5 classes data. Further we see that the statistics are very good. We get enough observations in all classes. Obviously we get more information with 11 classes that do we with 5 classes.

Table 14 ANOVA–table for 11 class data

ANOVA for	11 classes data			
Factor	dof	Square Sum	Variance	F–ratio
A	20	309.4	15.47	25.70
B	20	225.8	11.29	18.75
C	20	57.2	2.86	4.75
D	20	33.5	1.67	2.78
E	20	18.7	0.93	1.55
F	20	78.5	3.92	6.52
Error	1490	897.0	0.60	
Total	1610	1620		

Table 15 Predictions for optimum combination, on 11 class data.

Predictions on 11 class data										
Total class probabilities										
0.099	0.247	0.358	0.414	0.494	0.537	0.617	0.710	0.796	0.957	1.000
Logit transform predicted										
4.47	9.75	10.65	10.97	12.69	13.68	17.00	21.40	25.74	—	—
Predicted class probabilities (A1,B1,C1,D1,F2)										
0.732	0.904	0.921	0.926	0.949	0.959	0.980	0.993	0.997	1.000	1.000
Upper class limits										
0	2.5	6.3	15.8	39.8	100.0	251.2	630.9	1585	3981	9999

Phadke makes a prediction for the optimum choice given in Table 16. A prediction of the probability distribution function was made and verified by Phadke. That verified distribution function was used to calculate the information distance with respect to the two target density functions given in this section, i.e. Bin(5,0.0001) and Bin(5,0.05). The information distances for the optimum choice of factor levels is 0.05 and 0.03 respectively.

Table 16 Response table. Target probability density function Bin(11, 0.0001)

Factor	Level 1	Level 2	Level 3	Optimum
A	0.45	0.79	0.80	A_1
B	0.49	0.65	0.84	B_1
C	0.58	0.74	0.62	C_1
D	0.57	0.74	0.62	D_2
E	0.65	0.61	0.65	E_1
F	0.57	0.59	0.81	F_1 (F_2)

From this we see that the accumulating analysis procedure ends up in a choice which is closer to the binomial density function, Bin(5,0.05) than is it to the density function (1.0, 0.0, 0.0, 0.0, 0.0), (Bin(5,0.0001)). In this exercise the starting combination was A_2, B_2, C_1, D_3, E_1, F_1. The response for this set up has the information distance to our two target density functions 0.55 and 0.33 respectively.

We observe the procedure of prediction used. It is a two step approach. First we use the response table as above to make the optimum factor level choice. This is the application of the new metric. As would be expected the metric to evaluate the quality level is used. Then we take the probability data response as analyzed by Phadke. With that data using the Logit–transform a prediction is made for the probability distribution function. From the predicted probability distribution function the predicted information distance may be calculated. Evaluating the predicted quality level.

7.2. Rotation casting of plastic well.

This application was carried out at the company UPONOR AB in Fristad, Sweden. The application is one where large appliances such as sewage wells and other plastic containers are produced. The process used is called rotation casting. It is run in such a way that the plastic powder compound is charged in an appropriate amount into a hot mould. The mould is then rotated in several directions and around several rotation axes to get the compound powder to penetrate into every vicinity of the mould.

As the powder hits the walls of the mould it melts and a plastic skin is formed. Depending on different parameters the wall thickness around the well varies and the finished product deviates from the design intent. A special study was made to investigate the wall thickness distributions, (19).

In short the process at UPONOR is as follows:
- The plastic compound is charged into the mould.
- The mould is placed in an apparatus with two rotation arms.
- The apparatus with the moulds is placed in an oven.
- The mould is spun around the two arms.
- The apparatus is removed from the oven.
- The mould is cooled down using two cooling systems, air and water mist.
- The plastic well is removed from the mould and openings are cut in the well.

In this process the following parameters were identified and used:

A	Temperature of oven,	3 levels.
B	Time in oven,	3 levels.
C	Rotation speed of arm 1,	2 levels.
D	Rotation speed of arm 2,	2 levels.
E	Time of rotation,	2 levels.
F	Time of air cooling,	2 levels.
G	Time of water mist cooling,	2 levels.
H	Internal air pressure during cooling	2 levels.

The following interactions were studied:

AxF	Oven temperature x time of air cooling.
AxG	Oven temperature x time of water mist cooling.

ExC Rotation time x rotation speed of arm 1.

ExD Rotation time x rotation speed of arm 2.

This experiment was planned into a L_{32} orthogonal array. Factors A and B were handled by column merging and dummy treatment.

The wall thickness of the well was measured at 20 points. Thus there are 20 data items for each control factor set up. To evaluate the 32 different runs with an information distance approach a target distribution was designed, Table 17. This distribution was compiled of two binomial distributions over 5 classes. The binomial distribution parameter was set to 0.0001. To account for the target value 8.5 mm the classes were centered around that value. The nearest classes are 1 mm wide.

Table 17 Target density function and class definition for 8.5 mm target thickness.

Class probability	Upper class limit (mm)
5.E–17	0.
2.E–12	6.3
3.E–08	7.3
0.0002	7.5
0.4998	8.5
0.4998	9.5
0.0002	9.7
3.E–08	10.7
2.E–12	14.4
5.E–17	36.0

The evaluation was performed as described in section 6.5. In addition to that an Anova was carried out. The Anova is summarized in Table 18. There the significance level of the different factors may also be seen.

Information distance per factor level and interaction levels are summarized in Table 19 and Table 20. All factor levels but for factor B and H were decided on the basis of interactions. In Table 21 predictions and verifications are given. The original exercise was evaluated using SN–ratio metric. It turned out that the exercise did not give any improvement to the process as compared to the running conditions. This judgement is based on the information distance evaluated, Table 21, column "Info dist.".

Table 18 ANOVA–table for 10 class data, with 8.5 mm target thickness.

ANOVA for 8.5 mm target thickness					
Factor	dof	Square Sum	Variance	F–ratio	Significance level
A	18	37.2	2.06	2.15	0.995
B	18	42.7	2.37	2.46	0.995
C	9	25.3	2.81	2.92	0.995
D	9	59.9	6.66	6.92	0.995
E *	9	8.8	0.97		
F	9	14.5	1.62	1.68	0.90
G *	9	12.0	1.34		
H *	9	10.7	1.19		
ExC	9	15.4	1.72	1.78	0.90
ExD *	9	12.1	1.34		
AxF *	27	29.5	1.09		
AxG	27	54.7	2.03	2.11	0.995
Error1	99	187.8	1.90		
Error2	5490	5249.4	0.96		
Pooled	5526	5312.9	0.96		
Total	5571	5760.0			

* Pooled effects

However the evaluation using information distance gave optimum factor levels resulting in some improvement. Primarily this is believed to be due to the significant factors B and F being off target from the SN–ratio evaluation.

In the original report (19) a discussion was made over a change of average thickness from 8.5 mm to 9.5 mm. As the process was running at 8.5 mm more material has to be added to the moulds to achieve that. This kind of information has to be taken into account when designing the target distributions. However it may also be seen that some shrinkage effects that could give some variation to the average thickness.

To be able to study this a new target density function was set up. It was made in the same way as the 8.5 mm target thickness. Class limits were changed to have a centering around 9.5 mm. The class probabilities were the same.

Table 19 Response table. Information distance per factor level

Factors \ Levels	1	2	3
A	0.155	0.115	0.145
B	0.137	**0.129**	0.166
C	0.147	0.140	
D	0.177	0.112	
E	0.147	0.132	
F	0.131	0.132	
G	0.136	0.143	
H	**0.130**	0.148	

Table 20 Response table. Information distance per interaction levels.

A1F1	0.142	A1G1	0.136	E1C1	0.164	E1D1	0.184
A1F2	0.182	A1G2	0.182	E2C1	**0.133**	E2D1	0.171
A2F1	0.123	A2G1	0.117	E1C2	0.139	E1D2	0.120
A2F2	**0.115**	A2G2	**0.115**	E2C2	0.142	E2D2	**0.106**
A3F1	0.137	A3G1	0.147				
A3F2	0.154	A3G2	0.145				

Bold characters in Table 19 and Table 20 indicate optimum choices of factor levels. Accordingly we get optimum combination $A_2B_2C_1D_2E_2F_2G_2H_1$.

The information distances calculated for 9.5 mm target thickness are given in Table 22 and Table 23. We note that the optimum factor levels come out quite differently. In the prediction activities we also get some difficulties as the main data will be concentrated in very few classes. In this way we do not get a lot of information out of the data. The density function has to be designed to extract as much information as possible from the data. This was not done in this case. We also note that the physics behind is giving difficulties as there really needs to be more raw material in the mould.

Table 21 Predictions for optimum combination, on 8.5 mm target thickness.

Logit transform predicted $A_2B_2C_1D_2E_2F_2G_2H_1$ as evaluated with information distance.										Info dist.
–	–7.64	–5.43	–3.57	2.41	10.99	11.81	24.95	—	—	—
Predicted class probabilities (accumulated)										
0.000	0.147	0.223	0.305	0.635	0.926	0.938	0.997	1.000	1.000	
Predicted class probabilities										
0.000	0.147	0.076	0.082	0.330	0.291	0.012	0.059	0.003	0.000	0.087
Process probabilities before process improvement exercise $A_3B_{2-3}C_2D_2E_2F_1G_{1-2}H_2$										
0.000	0.150	0.100	0.000	0.200	0.450	0.050	0.050	0.000	0.000	0.104
Verification1 of SN evaluation $A_2B_1C_2D_2E_2F_1G_1H_1$										
0.000	0.000	0.250	0.000	0.100	0.250	0.150	0.250	0.000	0.000	0.302
Verification2 of SN evaluation $A_2B_1C_2D_2E_2F_1G_1H_2$										
0.000	0.100	0.150	0.000	0.100	0.500	0.050	0.100	0.000	0.000	0.166

Finally we note that the significance level needs to be taken into consideration as the prediction activity is performed. Non significant factors entered into the predictions may give rise to impossible results.

In the next example we will study the effect of applying the factor discrimination procedure proposed by Taguchi, (80). Some comments over this issue will also be given in chapter 10.

Table 22 Response table. Information distance per factor level, 9.5 mm target thickness.

Factors \ Levels	1	2	3
A	0.201	0.234	0.229
B	**0.197**	0.247	0.260
C	0.275	0.182	
D	0.221	0.236	
E	0.220	0.226	
F	0.212	0.233	
G	0.211	0.234	
H	**0.217**	0.229	

Table 23 Response table. Information distance per interaction levels, 9.5 mm target thickness.

A1F1	**0.157**	A1G1	**0.179**	E1C1	0.271	E1D1	0.229
A1F2	0.259	A1G2	0.234	E2C1	0.283	E2D1	**0.217**
A2F1	0.200	A2G1	0.238	E1C2	**0.182**	E1D2	0.221
A2F2	0.272	A2G2	0.232	E2C2	0.183	E2D2	0.253
A3F1	0.251	A3G1	0.216				
A3F2	0.209	A3G2	0.243				

Bold characters in Table 22 and Table 23 indicate optimum choices of factor levels. Accordingly we get optimum combination $A_1B_1C_2D_1E_2F_1G_1H_1$.

7.3. Hot forming of joining sleeve of plastic sewage pipe.

UPONOR AB manufactures sewage pipes from PVC. The joining system of such pipes comprises an expanded sleeve in one end of each pipe. By entering the straight end of a pipe into the expanded end of another pipe a joint is made. The joints are sealed with rubber gaskets located in a special groove formed in the sleeve. The pipes themselves are manufactured by extrusion. The sleeves are manufactured in a separate process after that the extruded pipe has been cut into standard lengths.

This sleeve forming process has been studied with the following factors included in the study.

• Cooling delay	0. seconds, 10. seconds
• Expanding kernel temperature	55. °C, 70. °C, 85. °C.
• Heating duration	14. min. 18. min. 22. min.
• Heating temperature	230. °C, 210. °C, 190. °C.
• Cooling water temperature	10. °C, 20. °C, 30. °C.
• Forming duration	4. min. 8. min. 12. min.
• Distance between pipe end and furnace bottom. Measured along the longitudinal axis of the pipe.	10 mm, 30 mm, 50 mm.
• Cooling water application grove length.	50 mm, 100 mm, 200 mm.

Those factors are referenced as A through H respectively in the text below. The factors were assigned to an L_{18} array. Factor A was assigned to column 1, factor B was assigned to column 2 and so on.

Several different properties of the manufactured sleeves were evaluated. Crack volume and groove diameter (gasket location), ware studied in detail.

7.3.1. Crack volume evaluation.

When the sleeves are formed the pipe material may crack if the process is not properly run. This tendency was evaluated by measuring the crack volume as the sleeves were cut apart. Three cuts were cut in each sleeve. Three sleeves were evaluated for each control factor setting. The volume in each cut was noted. Thus we get 9 observations for each control factor setting.

This property is of the type smaller is better. Accordingly the target density function was designed to handle that, see Table 24.

Table 24 Target density function and class definition.

Class probability	Upper class limit (mm)
0.9994	1.1
0.006	2.5
1.5E–07	18.5
2.E–11	179.7
1.5E–15	1015.2
6.E–20	2964.2
1.E–24	∞

Table 25 Response table. Information distance per factor level

Factors \ Levels	1	2	3
A	**0.061**	0.102	—
B	0.212	**0.018**	0.067
C	**0.005**	0.182	0.120
D	0.098	0.168	**0.016**
E	0.110	**0.060**	0.075
F	**0.046**	0.102	0.096
G	**0.027**	0.056	0.205
H	**0.055**	0.076	0.109

Table 26 Response table. Information distance per interaction levels.

A1B1	0.233	A2B1	0.237
A1B2	0.073	A2B2	**5.4E–07**
A1B3	**5.4E–07**	A2B3	0.278

Bold characters in Table 25 and Table 26 indicate optimum choices of factor levels. Accordingly we get optimum combination $A_2B_2C_1D_3E_2F_1G_1H_1$.

The information distance was evaluated as given in Table 25 and Table 26. The optimum level of factor A and B was set by the interaction data. Note that the interaction data gives two equally valid choices. From trade–off reasons A_2B_2 was chosen.

Table 27 summarizes the ANOVA of this case. It is seen that the interaction effect between A and B is dominating the variation. When trying to make a prediction it was found that either A2B2 or A1B3 100% zero crack volume. For this situation the Logit–transform does not work and we have to accept the findings from the factor effect analysis. I.e. the factor effect analysis is our prediction.

Table 27 ANOVA–table for 7 classes data

ANOVA for crack volume					
Factor	dof	Square Sum	Variance	F–ratio	Significance level
A	6	32.7	5.46	9.98	0.999
B	12	86.4	7.20	13.2	0.999
C	12	92.1	7.68	14.0	0.999
D	12	68.7	5.73	10.5	0.999
E	12	20.7	1.73	3.16	0.999
F	12	11.5	0.96	1.76	0.95
G	12	71.5	5.96	10.9	0.999
H	12	10.5	0.87	1.59	0.95
AxB	12	105.6	8.80	16.1	0.999
Error	864	472.2	0.55		
Total	966	972			

The traditional SN–ratio analysis, (13), gave $A_2B_2C_1D_3E_3F_2G_2H_1$ as the optimum combination. It is seen that the information distance analysis gives a discrepancy for factors E, F and G. From Table 25 it may be seen that differences for factors E and G are not dramatic. For factor F however the information distance analysis give a strong indication of a optimum factor level other than that from the S/N–analysis. In the SN–ratio analysis the factor F was less significant and the choice was made from previous practice. We note that the main factors of importance is factors A and B. It makes the choice of factor levels of the other factors less important.

7.3.2. Groove diameter evaluation.

For each of the three sleeves being cut the groove diameter was measured. Thus we get three data items for each control factor setting. This is a very small number and it is interesting to see how this corresponds to the number of classes in the density function. Some discussion about this kind of problem is found later in this report as the application to SPC is being analyzed.

Accordingly this property has been evaluated for two different density functions. The first has 7 classes and the second has 11 classes. The density functions were designed on classes around the nominal diameter as shown in Table 28 and Table 33. For the seven class density function two four class binomial functions were used. The different class probabilities were calculated using binomial probabilities with the parameter $p_1 = 0.0001$ and $p_2 = 0.9999$ respectively. These two binomial density functions overlapped in class number 4 in Table 28, where they accumulated. The density function in Table 33 was created in a similar way.

Table 28 7 classes target density function and class definition.

Upper class limit	Class probability
$\Phi_0-0.3$	5.E–13
$\Phi_0-0.1$	1.5E–8
$\Phi_0-0.02$	0.00015
$\Phi_0+0.02$	0.9997
$\Phi_0+0.1$	0.00015
$\Phi_0+0.3$	1.5E–8
∞	5.E–13

The information distance factor response tables for the 7 class density function are given in Table 29 and Table 30.

Table 29 Response table. Information distance per interaction levels.

A1B1	1.270	A2B1	**1.0896**
A1B2	1.386	A2B2	1.0896
A1B3	1.270	A2B3	1.129

Table 30 Response table. Information distance per factor level, 7 classes density function.

Factors \ Levels	1	2	3
A	1.305	**1.078**	—
B	**1.164**	1.213	1.169
C	**1.169**	1.169	1.213
D	1.169	**1.074**	1.324
E	1.324	1.164	**1.074**
F	**0.960**	1.386	1.386
G	**1.164**	1.224	1.169
H	1.270	1.213	**1.104**

Bold characters in Table 30 and Table 29 indicate optimum choices of factor levels. Accordingly we get optimum combination $A_2B_1C_1D_2E_3F_1G_1H_3$.

Comparing Table 31 and Table 36 it may be seen that the significance level is decreasing somewhat when the number of classes is increased. One may speculate about the relevance of performing the Anova on accumulated frequency numbers as proposed by Taguchi in reference (80). A very significant factor in a downstream class will contaminate the information in upstream classes.

Table 31 ANOVA–table for 7 class data

ANOVA for groove diameter					
Factor	dof	Square Sum	Variance	F–ratio	Significance level
A	6	36.9	6.1	20.5	0.999
B	12	9.8	0.8	2.7	0.995
C	12	9.8	0.8	2.7	0.995
D	12	20.6	1.7	5.8	0.999
E	12	24.0	2.0	6.7	0.999
F	12	90.2	7.5	25.1	0.999
G	12	9.8	0.8	2.7	0.995
H	12	48.6	4.1	13.6	0.999
AxB	12	9.8	0.8	2.7	0.995
Error	216	64.6	0.3		
Total	318	324.			

Table 32 Predictions for optimum combination, 7 classes density function.

Logit transform predicted $A_2B_1C_1D_2E_3F_1G_1H_3$, A_2xB_1 as evaluated with information distance.							Info dist.
—	—	7.82	7.82	7.82	14.56	—	
Predicted class probabilities (accumulated)							
0.000	0.000	0.858	0.858	0.858	0.966	1.000	
Predicted class probabilities							
0.000	0.000	0.858	0.000	0.000	0.108	0.034	
Process probabilities before process improvement exercise $A_3B_{2-3}C_2D_2E_2F_1G_{1-2}H_2$							
0.000	0.111	0.000	0.111	0.278	0.222	0.278	
Verification1 of SN evaluation $A_2B_2C_1D_3E_3F_2G_2H_1$ (crack volume optimum)							
0.000	0.000	0.000	0.000	0.000	0.143	0.857	
Total class probabilities							
0.093	0.000	0.019	0.000	0.000	0.093	0.796	

In Table 37 there are also some results given over predictions using only those factors contributing more than 5% to the total sum of squares. The relative contribution is calculated using factor effects with the error sources subtracted, proposed by Taguchi, (80). We can see from these results that the predictions are consistent. The spurious effects created when using all factors in the prediction are due to the fact that the prediction is to a large extent based on error.

Table 33 11 classes target density function and class definition.

Upper class limit	Class probability
Φ_0–0.3	5.E–13
Φ_0–0.1	1.4E–8
Φ_0–0.02	0.00015
Φ_0+0.02	0.9995
Φ_0+0.1	0.00035
Φ_0+0.2	1.E–7
Φ_0+0.3	1.7E–11
Φ_0+0.5	1.7E–15
Φ_0+0.7	1.E–19
Φ_0+0.9	3.5E–24
∞	5.E–29

The information distance factor response tables for the 11 classes density function are given in Table 34 and Table 35.

Table 34 Response table. Information distance per factor level, 11 classes density function.

Factors \ Levels	1	2	3
A	0.922	**0.879**	—
B	**0.914**	0.945	0.925
C	**0.892**	0.911	0.895
D	0.900	**0.897**	0.948
E	0.992	0.875	**0.844**
F	**0.881**	0.961	1.083
G	0.920	0.927	**0.918**
H	0.931	**0.866**	1.021

Table 35 Response table. Information distance per interaction levels.

A1B1	0.957	A2B1	**0.901**
A1B2	0.996	A2B2	0.936
A1B3	0.957	A2B3	0.975

Bold characters in Table 34 and Table 35 indicate optimum choices of factor levels. Accordingly we get optimum combination $A_2B_1C_1D_2E_3F_1G_3H_2$.

When comparing 7 classes results with 11 classes results we note that the factor response for factors G and H change. This is because we get more information about changes appearing far from the target when we are using the 11 classes density function.

As we get poor statistics we ought to redesign the classes. The majority of observations are in the higher classes. We actually note the same problem as with the wall thickness of the well in the previous example. There we got problems as we got more classes with no observations as we moved the target to 9.5 mm instead of 8.5 mm. Wall thickness 9.5 mm (target) was exceeding the overall average thickness, which was between 8.5 mm and 9.0 mm. accordingly we got severe interactions as the material supplied was less than needed to achieve the target. It may be a worthwhile exercise to investigate possibility that the reason for the spurious effects is that there are interactions between the design factors.

Table 36 ANOVA-table for 11 class data

ANOVA for grove diameter					
Factor	dof	Square Sum	Variance	F-ratio	Significance level
A	10	43.5	4.4	12.0	0.999
B	20	26.0	1.3	3.6	0.999
C	20	14.3	0.7	2.0	0.99
D	20	25.6	1.3	3.5	0.999
E	20	34.9	1.7	4.8	0.999
F	20	165.3	8.3	22.7	0.999
G	20	23.2	1.2	3.2	0.999
H	20	61.9	3.1	8.5	0.999
AxB	20	14.1	0.7	1.9	0.99
Error	360	131.1	0.4		
Total	530	540.			

Both 7 and 11 classes analysis show however a good agreement with the SN-ratio analysis over factor effects, (13). The verification run was according to crack volume optimization and other production preferences. Accordingly we get little information about the groove diameter data. We note however that the diameter is varying less but is off target Φ_0. The prediction of the optimum factor setting for groove diameter is indicating a result which is more on target. This is only an indication as the prediction as said above needs to be developed further.

Table 37 Predictions for optimum combination, 11 classes density function.

Logit transform predicted $A_2B_2C_1D_2E_2F_2G_2H_1$ as evaluated with information distance.										
—	—	7.82	7.82	7.82	7.36	14.6	13.9	—	—	
Predicted class probabilities (accumulated)										
.000	.000	.858	.858	.858	.845*	.966	.961*	1.00	1.00	1.00
Predicted class probabilities										
.000	.000	.858	.000	.000	–.013*	.121	–.005*	.000	.000	.000
Logit transform predicted $A_2E_3F_1H_2$ as evaluated with information distance. Only factors contributing more than 5% to the Sum of Squares.										
—	—	.897	.897	.897	4.08	6.71	11.4	—	—	
Predicted class probabilities (accumulated)										
.000	.000	.551	.551	.551	.719	.824	.932	1.00	1.00	1.00
Predicted class probabilities										
.000	.000	.551	.000	.000	.168	.105	.108	.068	.000	.000
Process probabilities before process improvement exercise $A_3B_{2-3}C_2D_2E_2F_1G_{1-2}H_2$										
.000	.111	.000	.111	.278	.167	.056	.000	.000	.000	.278
Verification1 of SN evaluation $A_2B_2C_1D_3E_3F_2G_2H_1$ (crack volume optim.)										
.000	.000	.000	.000	.000	.000	.144	.285	.286	.285	.000
Total class probabilities										
.093	.000	.019	.000	.000	.037	.056	.148	.241	.352	.056

* These predicted results are obviously faulty values (negative probabilities). This is an indication that our statistics is poor. Compare the discussion above.

8. Information approach to quality engineering.

Quality engineering has over time had a strong emphasis on variation. This is true at least in the later stages in the product delivery process. It has also been recognized that the major causes of this variation are to be found in the earlier stages of the same process. In these stages the variation concept is not as easily identified. Information and information flow in that part of the process has for long been the subject of many studies. Those studies have however not been targeted directly towards quality improvements. This section is aimed at demonstrating how the information concept is suited to describe the quality engineering process in a unified way. Ways to measure quality of the product delivery process as well as the product and production processes designed will be discussed.

8.1. Added information – noise.

In this section we will discuss how the noise influences are documented as added information in the processes. Consider a general system as illustrated in Figure 1. Ideally the information found in the signal factor should be the same as the information found in the function of the system. The control factors contribute no information as per definition. Then if we have a difference between the information in the signal factor and the function this information has to come from the noise factors.

Assume the signal factor to be normally distributed, $N(\mu_1,\sigma_1)$. This signal represents the information $E_1=(\ln(2\pi e\sigma_1^2))/2$. Let us further assume that the noise is normally distributed, $N(\mu_2,\sigma_2)$. Let us further assume that $\mu_2 = 0$. Now if the system is simple the distribution of the function may be $N(\mu_1,\sqrt{\sigma_1^2+\sigma_2^2})$. Accordingly the function information is $E_f=(\ln(2\pi e(\sigma_1^2+\sigma_2^2)))/2$. Whether the resulting function distribution is exactly as above depends on the characteristics of the system. We note however the important point that the influence of noise is documented in the system function as an increase of information. Another point of interest is that information content itself is not a useful measure of quality level. Quality is, as we have seen above, a function of variation around a target. The information content of a signal does not take any account of average distance from the target. The information distance as proposed in this thesis does allow for both a shift of mean and a variation around mean.

Noise entering into the system represents an increase of information in the system function. It also represents an increased information distance from the target. Thus it represents as expected a degradation of quality.

The viewpoint of information flow is supported by other researchers. Singpurwalla, (32), investigates the Taguchi philosophy for the product delivery process, (93), in the framework of decision making under uncertainty. In this paper Singpurwalla is touching upon the issues discussed in this chapter. In general the approach is sound. The author has however another perception of the contributions of Taguchi. Of course the analysis of Singpurwalla is correct from the perspective that he sets. The author believes that decision making is just a part of the contribution of Taguchi. Singpurwalla discusses decision making in general terms with no specific metric recommended. SN–ratios are said to be one metric acceptable in limited cases. The uncertainty that Singpurwalla introduces is obviously noise influencing the process. The amount of uncertainty left after each decision is noise that managed to get into the information flow as the filtering, i.e. decision making process, was applied.

Along the corse of the product delivery process noise is an influence all the way. It may be either from external sources or noises present internally, i.e. not yet removed from the information being processed. This was realized by Pugh, (63), as he designed the Total Design model. The concept selection process is one filtering process in that model. Another is the product delivery specification. That acts as a filter to external noises. Accordingly it could be used as a means to measure quality. This will be discussed later.

8.2. Quality engineering.

Having introduced some thoughts about noise, i.e. excessive information flowing in the structures of the designs and the product delivery process we will look into some interpretations and consequences of this concept.

In Figure 20 a schematic of a complex system is shown. Noise factors affect the system by the addition of their information content. In Figure 20 this information entry is shown as entering into different subsystems. This information is unwanted as opposed to the information entering the system through the signal factors. As has been said above the information content in the function should be the same as that found in the signal factors. The difference that may be found is a measure of poor quality. In robust engineering we call a system, with all factors shown in Figure 1 present, a dynamic system. A system with no signal factor is called a static system, (64), (93).

We may at this instance again note the difference from a dynamic system in the context of control engineering. A dynamic system in a quality engineering context may or may not be a dynamic system in a control engineering context; it could well be a static system in control engineering context. From a control engineering point of view a system is dynamic if it contains energy accumulating devices, (50). In this thesis the word dynamic is used with a robust design interpretation.

For dynamic systems the target distribution is assembled from the signal distributions. In case of a static system the target distribution is designed from other sources, i.e. through benchmarking, etc.. In effect this designed distribution becomes the signal factor of the product delivery process, i.e. the design intent. In the vocabulary of Pugh, (63), this is the product delivery specification. Also for dynamic systems you should take benchmarking information is taken into account when designing the target distribution.

From the above it may be concluded that quality engineering is a matter of stopping the noise informations flowing from their entry points to the system function.

This way of defining quality engineering gives an interesting background to different methods used. Consider the system illustrated in Figure 20. Information entering into the system at C1 may take three different paths to the function. Thus the complexity number introduced by Pugh, (63), is a relevant quality tool. As the complexity number is reduced the number of transmission paths for noise are also reduced. That way the task of stopping the flow of noise information, is made easier. The complexity number is a good way of analyzing the information content in the system design. This also ties directly back to the design axioms by Suh, (62).

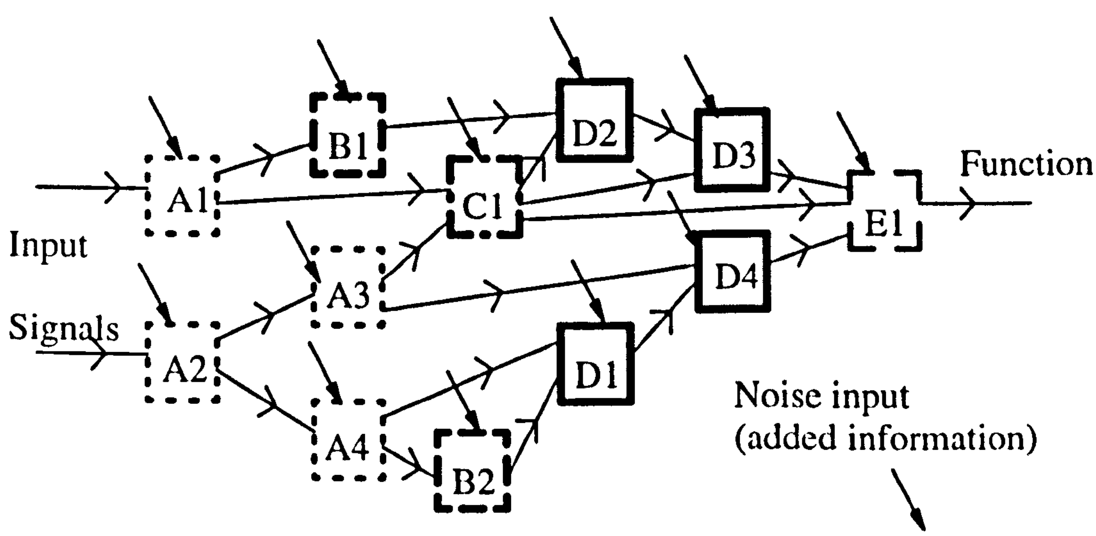


Figure 20 Information flow in a complex system.

The complexity number is defined as $C_f = (\sqrt[3]{N_t N_p N_i})/f$. N_t is the number of types of subsystems or parts. N_p is the total number of subsystems or parts. N_i is the number of interfaces between subsystems or parts. Finally f is the number of functions that the system is performing. It is very important that the complexity number is evaluated on a consistent system level. Consider a car. At a high system level there are the types of subsystems car body, powertrain and chassis. If in this context the number of fasteners between car body and chassis are brought up the analysis is failing. This is because the number of fasteners refers to a lower system level. It is in the complexity of the interface between chassis and car body. Thus there is a mix of information which is not consistent. Below we will show that the complexity number is a convenient measure of information content in the design. The system in Figure 20 has $C_f = 10.1$, derived from $N_p = 12$, $N_t = 5$ and $N_i = 17$. Thus it is rather complicated and the information content is rather high, compared with the systems in Figure 21.

One of the design axioms of Suh concerns the minimum amount of information in a design. In that perspective it is interesting to investigate the meaning of the complexity number. From statistical physics, (92), we may recall that the entropy for a gas may be considered a function of probable states.

Let us now assume that we have N different components to build a particular system. We may then chose different ways to produce the system by distributing the components over different categories. In our case we have the categories, number of types of subsystems, total number of subsystems and the number of interfaces between subsystems. A particular design may then be characterized by the numbers N_t, N_p and N_i. The probability of a component being type, part or interface may then be calculated as N_t/N, N_p/N and N_i/N respectively. The entropy for this design may then be derived as :

$$E_d = -\left[\frac{N_t}{N}\ln\left(\frac{N_t}{N}\right) + \frac{N_p}{N}\ln\left(\frac{N_p}{N}\right) + \frac{N_i}{N}\ln\left(\frac{N_i}{N}\right)\right]$$

This information is bounded upwards by:

$$E_B = -\left[\frac{1}{3}\ln\left(\frac{N_t}{N}\right) + \frac{1}{3}\ln\left(\frac{N_p}{N}\right) + \frac{1}{3}\ln\left(\frac{N_i}{N}\right)\right]$$, see appendix A.

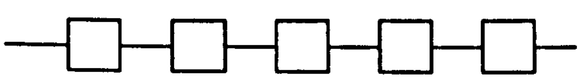
E_B may be rewritten as

$$E_B = -\left(\frac{1}{3}\ln(N_t) + \frac{1}{3}\ln(N_p) + \frac{1}{3}\ln(N_i)\right) + \ln N \qquad \text{or}$$

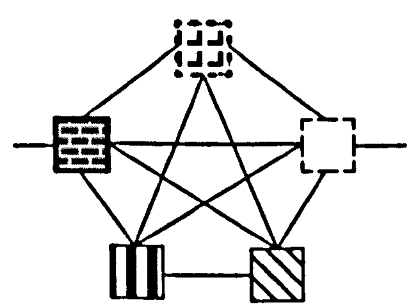
$$E_B = -\ln C + \ln N \text{ where } C=\sqrt[3]{N_i N_t N_p}.$$

From this observation we may conclude that the complexity number is related to the information of the system. In fact from practical calculations it may be seen that $\ln C \approx \ln\frac{N}{9} + E_d$, see Appendix C . Judging from the results in that Appendix this expression is valid to within 10% for practically achievable systems. Thus we see that a minimization of the complexity number is the same as a minimization of the system size and the structural information, E_d.

Consider a system containing I parts. The simplest system within that category has $N_i=I-1$, $N_t=1$ and $N_p=I$. The most complex system $N_i = \sum_{j=2}^{I}(j-1)$, $N_t=I$ and $N_p=I$. Those systems are illustrated for $I=5$ in Figure 21. Further information on limiting cases for the numbers N_i, N_t and N_p may be found in Table 55 of Appendix C .



a. C=2.71



b. C=6.30

Figure 21 The simplest and the most complex system containing five subsystems.

We can see that the complexity number is very tightly related to information content of the system at the system level under study. It is important to keep track of the system level. In case system levels are mixed, results will still be generated but they do not reflect the information. The complexity number should be applied hierarchically throughout the system levels. Details of an interface may also be considered as a system level. Thus it is possible to use the complexity number as an analysis tool for interface design. In this case new categories instead of parts, types and interfaces probably need to be introduced. From the information theoretical basis another complexity number valid for this application may be deduced. We will look at that next.

In parallel with the above discussion a complexity number for a situation with four categories instead of three may be designed. This definition is $C_4 = \sqrt[4]{N_w N_x N_y N_z}$. N_w, N_x, N_y and N_z represents the number of components in each of the four categories, whichever they may be. $N = N_w + N_x + N_y + N_z$ represents the total number of components. The information analysis of the C_4 measure gives the following relation $\ln C_4 \approx \ln\frac{N}{16} + E_{d4}$. Where

$$E_{d4} = -\left[\frac{N_x}{N}\ln\left(\frac{N_x}{N}\right) + \frac{N_y}{N}\ln\left(\frac{N_y}{N}\right) + \frac{N_z}{N}\ln\left(\frac{N_z}{N}\right) + \frac{N_w}{N}\ln\left(\frac{N_w}{N}\right)\right]$$

The denominator 16 is not clearly defined theoretically. It seems though that the square of the number of categories is a good choice, judging from the evaluations in Appendix C . Accordingly denominator 9 is valid for three categories and 25 would have been appropriate for five categories. These relations may be documented as is done for three categories in appendix C . The beauty of the complexity number in relation to a strict information analysis is the ease of use.

The information theoretical approach also shows that a modularization strategy within product design, (9), is a quality technique. As we minimize the complexity number we minimize the information in the system in accordance with the design axiom by Nam Suh (62),(1),(2). The same kind of reasoning based on entropy is also introduced by Hitchins, (29). We have also seen that we minimize the number of possible flow paths for noises.

Thus modularization implies careful design of interfaces between modules. Those interfaces are parts of the flow paths for the noises. They actually act as noise filters and have to be designed as such.

8.3. The RFCA procedure.

So far we have discussed how the concept of information may be used to develop (the more general complexity number given above) quality engineering tools and to analyze the nature of quality (poor quality added information). Now we will address the issue how the product delivery process itself is an information filtering process.

Sontow and Clausing, (15), have revised the enhanced QFD procedure, (49), to integrate further quality engineering tools. They propose the name Requirements and Failure Cause Analysis, RFCA, for the new procedure.

The original QFD procedure encompassed four phases. It did not distinguish between product and process development. This was achieved in the enhanced QFD procedure, EQFD. EQFD relies heavily on the system structure for the product to be developed. Concurrent engineering was a prerequisite for EQFD.

The main improvement in EQFD was the integration of system structure into the planning matrices. The resulting procedure was improved upon further as function analysis was introduced.

Function analysis was not explicitly pointed out as a component in the original QFD. It was however usually applied in the needs or voice of the customer section of the first matrix. A structured function analysis was elaborated such that a tight bond with the system structure was developed. The user function structure is broken down into subsystem function structure. In this process a concept selection exercise is conducted at each system level to generate the next lower system level. The prioritizing process from QFD is performed at each system level, first amongst subsystems and then for individual subsystems. The latter activity is performed for high priority subsystems only. In this way one of the major drawbacks with the QFD is eliminated. Matrices of a size 100*100 need never be seen.

As the system structure gets into EQFD it is natural to get the production process into the procedure. Subsystem structure is of course tightly connected with assembly processes. Accordingly you get an assembly process planning is tied into the total system planning matrix. As the system structure is refined the system planning matrices are accompanied by production planning matrices. In this way concurrent engineering is supported and facilitated by EQFD.

The next step in the evolution of QFD–like processes was to integrate Failure Mode and Effect Analysis, FMEA, into the procedure. Through this step another important drawback was eliminated. This was to support noise analysis according to robust design in the procedure. The FMEA table is modified to a fault matrix. In this matrix failure modes are characterized as a combination of function and noise. The significance of each function in relation to system function is rated. Further the significance of each fault as characterized above is graded. This gives a possibility to grade different noises afflicting the system.

The produced noise grading is fed into the concept selection matrix as evaluation criteria. With this last addition the RFCA procedure is produced. The layout of planning matrices in RFCA, even though somewhat complex, shows the way noises may flow through a design.

QFD which is the base from which RFCA is derived is actually a way of putting structure to information. The information handled is all information influencing the product delivery process. Of course RFCA is dealing with this information in an even more structured way. We now recall the example with the play of dice in section 5.1. When we put the restriction to the evaluation of the number pairs, i.e. to only evaluate the sum of the two numbers, we took information out of the data. In just the same way the structure that RFCA is giving the information influencing the product delivery process is taking information out of that original lot of information. The RFCA procedure is nicely mapping onto the different steps of the product delivery process. In every step structure is added. Accordingly it is shown that the product delivery process is a process where information is consistently taken out of the information being processed. Eventually the design preferred is arrived at. According to Suh, (62), that design is the one with the least information content. This is consistent with the effect of the RFCA as described above.

8.4. TQM and information flow.

Figure 20 was presented to illustrate a product system. However it may as well be a product development system, i.e. the Total design process, (63). The signal factors in this system are the company visions. In the system most noise entry points are found in the product specification. Leaving one of the headings in the specification checklist, (63), unattended is the same as leaving a door open to noise entry. The noise may in this case be a variety of opinions among designers about what is the design intent. This leads to poor designs and an inefficient process. Noise may also be diffuse organizational rules, etc. Needless to say the function signal of this process is the set of products developed and the resources actually consumed to do it.

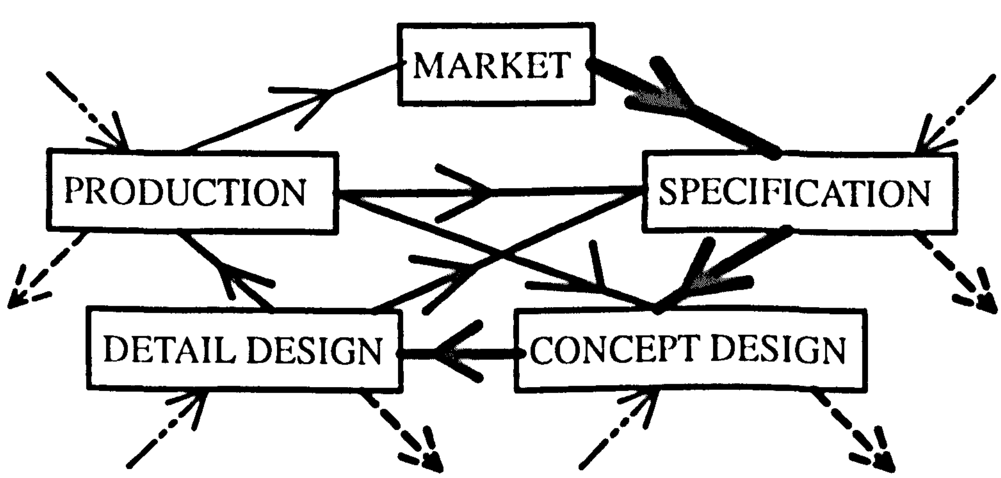


Figure 22 Information flow in the product development process.

Figure 22 illustrates the main information flow in the product development process in a somewhat different way. The different activities given in the boxes of the diagram is due to Pugh, (63). The process starts with a market activity, and enters into a specification activity. From there the conceptual design is started to establish the foundation for the detailed design and production activity. Then the market activity of selling and assembling information for the next project is begun. Meanwhile there is information exchange between the activities. Figure 22 illustrates the cyclic nature of the product development process.

The weight of the arrows of the internal flows is supposed to illustrate the amount of information flowing. Arrows drawn with dash and double dot represent information entering into the process from exterior sources other than the market. That may for example be experiences from the different persons involved as pointed out above. The dashed arrows illustrate the information extracted and found excessive in the process. This is in agreement with the design axioms put up by Suh, (62).

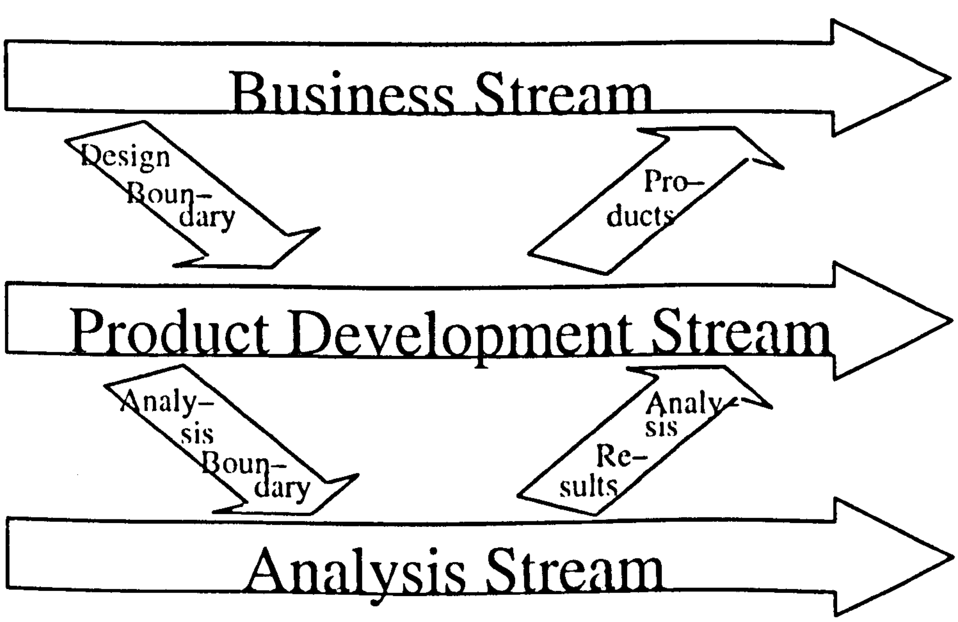


Figure 23 Business process breakdown

We may see all companies as information processing systems. The information processing has several levels, Figure 23, that are hidden in the total design model. Three main information streams can be identified: the business stream, the product development stream and the analysis stream. The business stream is actually the everyday activity of selling, producing and delivering. A sub–process of the business stream is the product development stream. This process is engaged with improvement of the product and production processes active in the business stream. The interface between the business stream and the product development stream is the product design specification as outlined in total design. In the same way the analysis stream is a sub–process of the product development stream. The interface between the product development stream and the analysis stream is an analysis specification. This specification may be outlined very much in the same way as the product development specification.

These specifications serve very well as a basis to develop a target distribution to apply the quality metric proposed in this thesis. The distribution of the quantification of the specification and the timeliness in delivery is measurable. Comparing this target distribution to the actual outcome in the process gives a possibility to quantify the quality of the process. Staffing along the process may also be regarded as a density function for which a target and an actual outcome will exist. This is another possibility to measure the quality of the process.

The total design model mainly engages itself with the product development stream. The RFCA covers also the analysis stream. It does take into account the noises affecting the product being developed. It does not take into account the noises affecting the product development process itself. Those latter noises are as said above addressed by the total design model, even though it is not spelled out explicitly.The RFCA model really works with information flow. Thus it is a good basis for further refinements of information flow analysis. TQM procedures evolving over the last decades address the business stream. The TQM procedure is usually qualitative and metrics for quality measurements are nonexistent. The new metric proposed in this thesis presents a possibility to quantify process quality, as demonstrated above.

What is important is that a function description of the system to be assessed for quality level has to exist. Once the function description exists, an information measure to use may be found. The modern quality systems like ISO 9000 call for function descriptions at all organizational levels. The European excellence model, (4), is even more process oriented. This makes the potential for information metrics very good.

As we have the function description of our system broken down to different subsystems we can identify target distributions for different performance measures at all subsystem interfaces. Measuring information distances at each of these interfaces gives us a way to trace the flow of added information. In this way we can see the contribution of different noise sources to the system function deterioration.

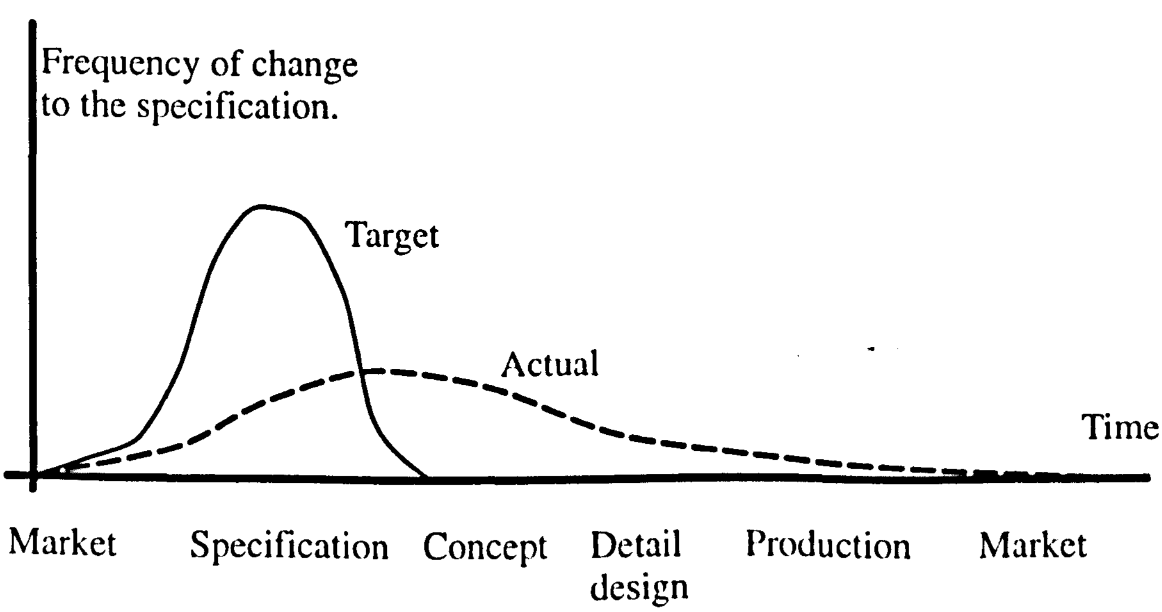


Figure 24 Frequency of change distribution functions demonstrating process efficiency.

For product and process development process performance we may use the product specification as a measuring tool. The number of changes in the specification is a measure of this performance. The distribution of product specification changes may be targeted and used to measure information distance, see Figure 24.

In Figure 25 some examples of evaluations of are given. We may regard the sample distributions given there as approximations to the distributions referred to by Fox, (14), as he describes differences between Japanese and Western industrial processes. His description gives a good qualitative picture. With the proposed metric we are able to work with quantifications for the process performance. The TQM activities are directed towards estimation of process losses. That gives us the possibility to quantify the information metric in the same way as has been discussed for product quality metrics earlier in this thesis, (section 6.4.3.). As regards the losses in the product development process a very nice discussion is given by Clausing in a recent book, (5).

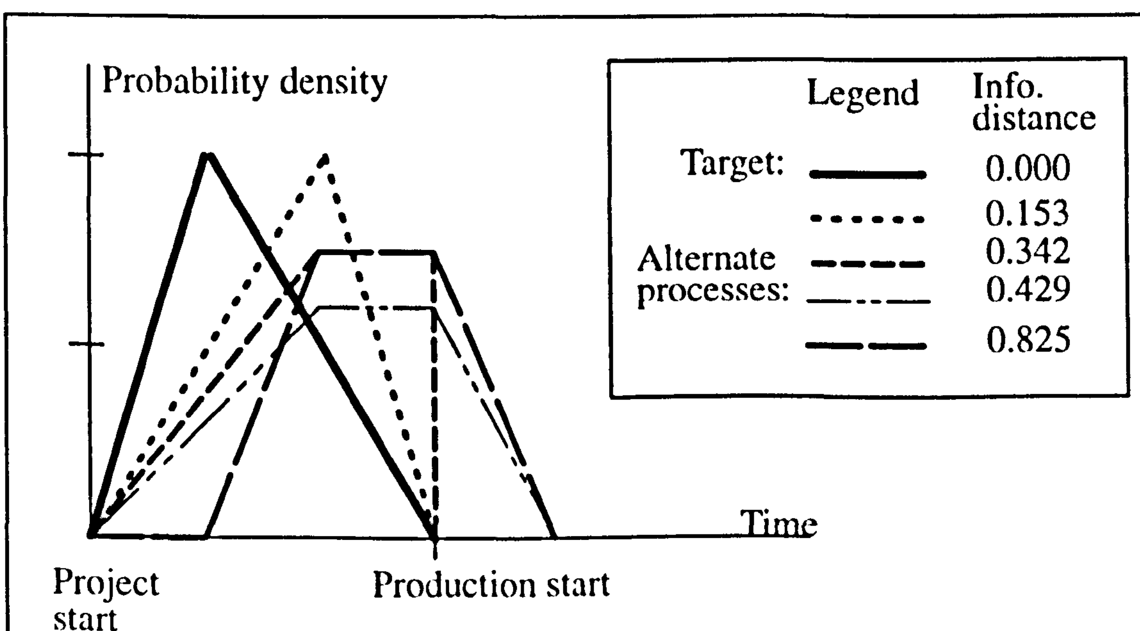


Figure 25 Product development process performance evaluated by application of information distance metric.

The density of functions exemplified in Figure 24 is valid for a posteriori estimation only. To be able to make estimates during the running process these density functions may be truncated to the present time. The information distance is then calculated on the basis of the truncated target and actual density functions of changes to the specification.

8.5. Robust design planning.

When we talk about robust design planning we talk about robust design using designed experiments. The normal procedures outlined for this kind of exercise focuses only on the individual subsystem under study, (See chapter 4.). From the above discussion we see that this may be too restricted. We know that quality engineering is a matter of blocking noise flow paths. The outline shown in the RFCA–procedure includes a good example of a robust design planning activity for systems with a higher level of complexity.

The robust design exercises should take into account noises in the product development process as well as in the product system. Product system noises should include the noises entering directly into the subsystem as well as the noises entering via the interfaces. Noises of the product development process document themselves as variations in subsystem layout and performance, (60).

From this discussion we conclude that down to the system level under consideration the information flow has to be analyzed in detail. Further factors at the same system level have to be used. This is also a consequence of the function design axiom by Suh, (34).

The robust design exercise has to be preceded by system analysis. This analysis has to give as a result the product noise factors and the development process noise factors. The latter will be in the shape of interface specifications including targets and possible limits of variation. The process, from this point on, is normal as given in any of the textbooks on robust design, (73). However the system level has to be recognized. Discussions and recommendations along the same lines from other view points have been given by Ramberg *et. al.*, (30).

In the light of the discussion over noise flow paths in the last section we can see that the robust design exercises will contribute to the mapping of the flow of added information. From this perspective the importance of the system level becomes obvious. Figure 20 shows that the usage of lower level control factors increases the complexity of the robust design exercise. This is due to the fact that each of the boxes in that figure may itself be such an involved structure of sub–subsystems. Hence the likelihood of interactions and other complications appearing is of course increasing. We can see that noises from the same source may enter into the overall system function at different points and through different paths. The complications arising from this is easily understood.

The system analysis from an information theoretical point of view as discussed above gives a solid base for steps 1 through 3 and step 6 in the procedure outlined in Chapter 4. In the same chapter the issue of quality measure was discussed. With reference to Chapter 5., the quality measure should be the one with the largest information content. That quality measure should be one that can be analyzed with the highest number of classes as was discussed in connection with the information distance analysis of Chapter 5.

When assessing the relations between system function and control factors the information flow should be used as guidance.

The analysis of the results of the application of designed experiments is of course carried out using the new quality metric, information distance. This analysis is the same irrespective of what quality measure was used in the experiment data acquisition. The prediction is an optimization exercise, as discussed in Chapter 5.

In this way the system function and the information flow is always the prime objective of the robust exercises. Through the loss function, the target distributions and benchmarking the results are tied directly back into the business stream.

8.6. Information approach in summary

In this chapter it has been shown that different quality tools introduced in practice earlier, on a common sense basis can be motivated from an information theoretical point of view. The progression from QFD to RFCA has been analyzed in a perspective of information processing. System analysis is widened to include product and process system. In that way the information distance metric has been shown to be applicable and useful even in TQM activities.

The complexity number as introduced from empirical reasoning has been shown to reflect information content in a system. A concept of information flow has been introduced to analyze the quality aspects of a system. Through that methodology it has been possible to argue that modular design is a quality improving strategy. Finally the information flow approach has been found to give a basis for the preparatory steps in the robust design procedure.

With the addition of on-line quality control methods to be discussed in the next chapter, information theory has been shown to be a foundation for quality engineering.

9. SPC as an application of information theory.

In this chapter we will discuss the possible application of the information distance metric in the area of process control. We do that to show how the new information theoretical approach can be used to integrate the different areas of quality engineering.

First some general discussion about the noise flow in production systems is made. Then a summary of the traditional SPC is given. Finally some demonstrations of the implementation of information distance based SPC are made.

9.1. Noise flow in production systems.

From an information theoretical point of view a production system is no different from any other process. We have earlier discussed the product development process. The observations made there are relevant also for production processes. There is a flow of information in the process.

The main flow is the production intent. That is, what to do, how much to do, and when to do it. This information comes to the process as sale orders. This information is broken down into the bill of materials with the assistance of the design documentation and the production system planning to give a production plan. In this way we can envision a system layout as in Figure 26 for the production process.

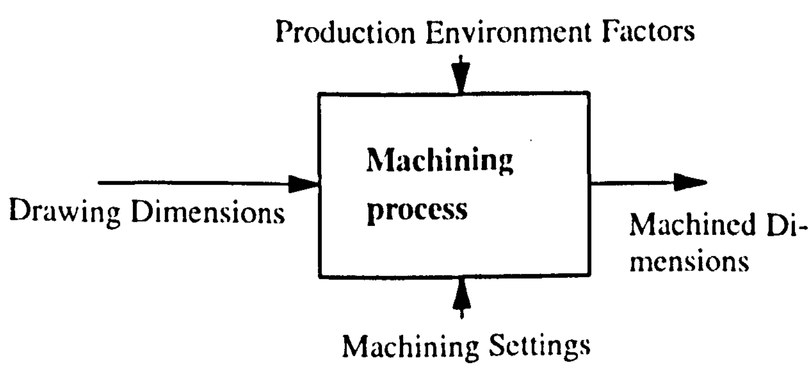


Figure 26 A machining process shown as a general dynamic system

The input signal is the production intent as discussed above. The function signal is the actual products being produced. The control factors are obviously the different parameters controllable in the production process itself. It may the cutting speed of a milling machine or curing time of an elastomer injection molding machine, etc. The noise factors are environmental influences, control factor variations and system wear, i.e. tooling wear.

Figure 26 is of course an over-simplified view. The illustration in Figure 20 is a better representation. In a breakdown like that we may recognize that the function signal of one subsystem is the input signal of another subsystem. Process control is a matter of monitoring the performance of these intermediate signals. When planning the monitoring system a good knowledge of the performance at the relevant system level is required, (47), (17). The basic idea of the SPC is to observe that the system is in statistical control.

The variations monitored in the different signals represents the amount of noise penetrating through the subsystems to the function. Under normal circumstances the variation will be of stable magnitude. A systematic change will indicate either a change in the noise filtering performance of the subsystem or an excessive increase in the noise level exposed to the system.

We observe from the above discussion that the noise may enter into a subsystem either as external noise as indicated in Figure 20 or via the input signal as contamination. The planning of a production control system totally relies on the existence of a good process description. A good process description is as we know a consequence of a correct implementation of ISO 9000.

9.2. SPC the traditional way.

W. A. Shewhart, (105), was the first to introduce the concept of statistical control charts into quality control. This was done before World War II. A widespread application did however not come into place until the 1960 and thereafter. It has one main objective; to monitor the stability of a process in statistical control.

The observed statistic is plotted in a diagram, Figure 27, with a time axis. On the basis of the anticipated distribution function for the control statistic control limits, upper (UCL) and lower (LCL) are calculated. It is very common that these limits are calculated such that the probability of an observation in the interval between the limits are 997 ‰. This corresponds to $\pm 3\sigma$ for a normally distributed stochastic variable. The calculations are usually performed assuming normal distribution. When choosing the control statistic certain care is exercised to assure the normality of the distribution function.

Sometimes a set of warning limits are also calculated. On the basis of the set control and warning limits, different analysis functions are constructed, (71), (67). The evaluation of these analysis functions generates a number of different alarms. The alarms may or may not be tied to different failure mode characteristics, (67). The original Shewhart chart was built on statistically independent observations. Decision rules used individual observations. As decision rules were enhanced several observations were taken into account. Thus the different decisions became statistically dependent as they were based partly on the same data. Some special control charts have been developed that are based on statistically dependent data. Two such charts are briefly discussed in section 9.2.2. .

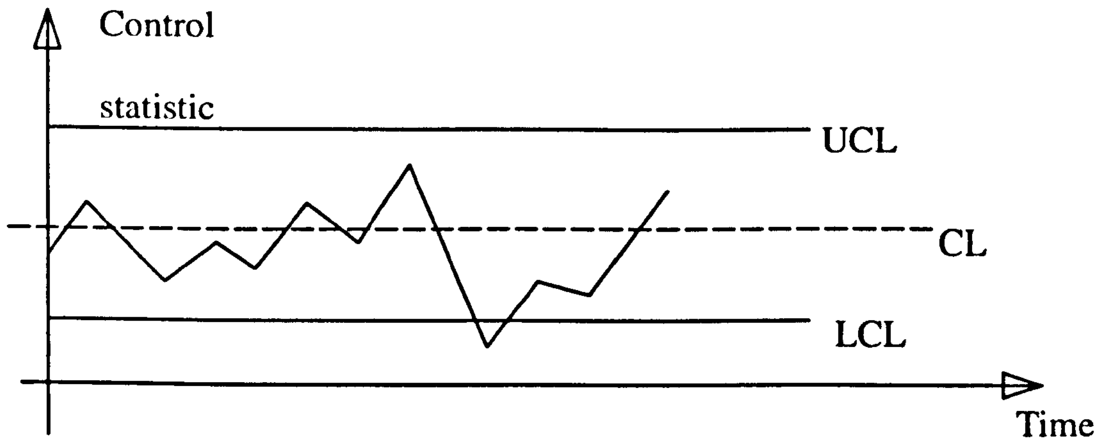


Figure 27 Control chart as used in SPC.

9.2.1. $\overline{X}$– and R–Charts.

It is common to use at least two control charts together; one for local mean and one for spread. The chart used for local mean is usually called $\overline{x}$–chart. This chart is used to trace position. As the distribution of the observed variable is usually not known, the average of several observations rather than the individual observations is plotted. Hence the name $\overline{x}$–chart. By using the average the applicability of a normal distribution may be claimed to be more relevant. This is of course motivated by the central limit theorem, (71). The correctness of this claim is very much dependent on the size of the group of observations used for averaging. A common compromise is to use 5 observations in each group.

The spread of the observations is usually monitored by the range, R, of each group of observation. Here again we see the motivation for the name of this chart. However the sample standard deviation, s, may also be used as the plotted variable. The range is however much easier to calculate. The predictive quality of R and s are the same. The group size to some extent influences the absolute quality of prediction. Accordingly range predictivity is a parameter in the group size decision process.

The start up of an SPC monitoring process is essentially the determination of control limits and central line (CL in Figure 27). The data necessary to do this is the grand average of the first plotted instances (something like 20 instances) in the $\bar{x}$– and the R–charts. In addition to this we obviously need an assumption about the distribution applicable as discussed above.

In this way we get two charts like the one shown in Figure 27. The control statistic in either one of them being $\bar{x}$ and R respectively. We further note that the lower control limit, LCL of the R–chart is of limited value as R is non negative. Hence it is usually set to zero.

9.2.2. MOSUM and CUSUM.

Some applications exist where a traditional $\bar{x}$–chart is too sensitive. Such a situation may be a control chart for a chemical component charged into a chemical process plant. One outlying observation may not give a drastic effect on the process result as a mixing between batches is present in the process. In this case an average over the last batches charged is a better predictor of what will happen with the process result.

A special chart sometimes called MOSUM has been designed to handle this situation. In the simplest form it involves plotting the average over the n_o last groups rather than each individual group. This obviously will smooth out the rapid variations while a persistent shift will accumulate. There are other more complicated forms applying an exponential weighting function, (47). Control limits are calculated and applied in very much the same way for this type of chart as for traditional charts.

In other situations the traditional charts are insensitive to small drift in the mean value of the process. This results in a value of ARL which is too high. Another special chart called CUSUM (Cumulative sum) has been designed for this case.

In a CUSUM chart a cumulative sum of the deviation from a process target is plotted for each observation. The process target must be very carefully chosen. It may be the center line applicable to a traditional $\bar{x}$–chart. In this kind of chart a consistent shift in mean will document itself as an average slope of the plotted graph.

The control limits of this kind of chart is then a so called V–mask. As the name indicates it has the shape of a V. The V is laying down with the vertex in the last plotted point. The legs pointing backwards up and backwards down respectively. The angle of the legs towards the horizontal are chosen such as to give appropriate alarms. Alarms are given as the plotted graph cuts the legs of the V–mask too few instances, i.e. plotting intervals, away from the last plotted point.

Several modifications to the shape of the V–mask are common. The motivations for the modifications are of course different alarming properties desired.

9.3. SPC using information metric.

As we have seen in the previous sections there has been a trend to tailor special charts for special requirements. In this section we will introduce a special class of charts based on the information distance metric. This class of charts has the property of adjustability to a desired decision model. As has been demonstrated above, each different type of traditional chart requires a non negligible amount of special expertise. The proposed class of charts will be require and the same knowledge for different applications. Further the tailoring or adaptation to the current case is the benchmarking and process loss identifications that is used in the procedure to set up the target density distribution function. That is done any how in a properly organized TQM oriented business. Hopefully this will put focus on the application rather than on the tool.

9.3.1. Conceptualizing information distance SPC

In chapter 6. an information metric for quality was discussed in detail. The metric proposed was information distance. In this section we will demonstrate and discuss how this metric can be applied to an SPC monitoring process. In principle we will have a control chart as illustrated in Figure 28 .

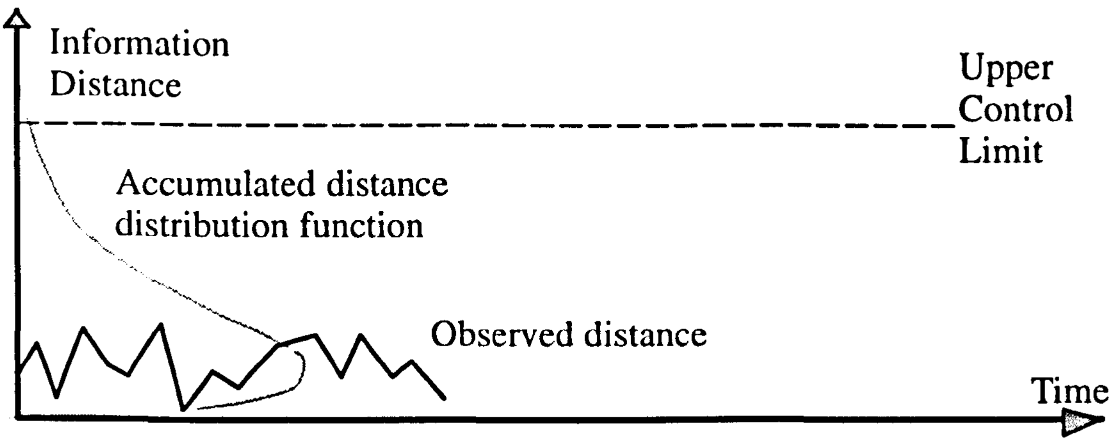


Figure 28 Information distance control chart.

The procedure to plot this chart is very similar to the one used for $\bar{x}$–chart. A number of observations are made and grouped. The probability distribution function for the group of observations is calculated. This probability distribution function will be over classes. Using the same terminology as in chapter 6. this calculated probability distribution function is compared to a target probability distribution function. The comparison is made as an information distance. To be able to do this the two distribution functions have to have identical class definitions.

The target distribution function may be chosen in the same way as discussed in chapter 6. i.e. on the basis of benchmarking activities. There is however also another possibility; that is to make a choice that is very similar to the discussion in section 9.2.1. above. This means to accumulate a distribution function for the present process as the monitoring is started. Which way the choice is made depends on what is to be achieved by the monitoring process.

A third basis of choice of target distribution function is obvious from Figure 26. The SPC activity is an activity to monitor that the performance of the product verified in the development process is really achieved and maintained in production. Thus the results of robust design activities are really forming the target distribution function. With this choice we get a very concrete integrating effect of the proposed quality metric. At the same time as we monitor the statistical stability we keep track of the process being on target in relation to results produced during the process development.

Information distance is taking into account position and spread at the same time. This was discussed in chapter 6. According to that discussion the target distribution function can be used to put emphasis more on position than on spread or vice verse. We note that by using a control chart based on information distance we may manage with one chart what we need two charts to do using traditional charts. Further we note that a careful choice of target distribution handles the trade off between position and spread. With the traditional control charts this trade off, whenever it has to be made, has to be left up to the end user.

With the target distribution generated partly from a benchmarking activity the monitoring process will also give an indication of the absolute quality level. This is not handled at all with the traditional approach.When the target distribution is chosen to be the accumulated one (as is sometimes recommended in traditional SPC) you just get what you get in terms of trade off between position and spread.

As discussed earlier the information distance will grow large as the process probability distribution function becomes narrower than the target distribution function. This is not a problem for two reasons. Firstly most changing processes that are not deliberately planned changes in production systems are degrading processes. In the rare cases that we get converging changing processes which are not planned changes we still benefit from getting an alarm to learn what has happened out of our control, and to improve the production system.

Information distance is not as easy to calculate as the local mean or the group range. Still, it is not prohibitively difficult. In any case this is not a big problem today as measurement systems are more and more being computerized. Evaluations will then be transparent for the user. The biggest problem in the implementation is to evaluate the control limit. Depending on the knowledge of the distribution function of the property being observed Monte Carlo simulations may be used to generate the control limit.

There are further parameters generating the alarming properties of an information distance SPC. These are the number of classes used in the target distribution function and the number of observations in each group. The influence of these parameters will be demonstrated in the simulations documented in the next section.

9.3.2. Monte Carlo simulated properties of Information distance SPC

As a demonstration, a Monte Carlo simulation is made assuming the observed property to be exponentially distributed with the distribution parameter being equal to 1.0 . This is then the base for the target distribution. Hence the simulation may be said to be using the accumulated target distribution function as discussed above.

The simulation is performed in two steps. First the control limit is generated as the value of information distance above which the probability to get an estimated information distance is 0.003. This corresponds to the alarm level of three standard deviations used for traditional SPC.

This control limit simulation is made using the process exponential distribution transformed to classes as a target distribution. The number of classes is varied as is the number of observations in each group. The observations in each group are generated and the corresponding distribution function is calculated. From this data the information distance between the target distribution function and the observation group distribution function is calculated. The observation generation is made using the continuous exponential distribution function. In each simulation 100,000 groups have been generated and evaluated.

The control limit generated will be a function of the number of classes and the number of observations in each group.

For each set of combination of number of classes and number of observations in the groups Average Run Length curves have been calculated in the second step of the simulation. The ARL–curves are generated by shifting the mean of the exponential process distribution. With these shifted means the information distances between the observations group distribution function and original target distribution function are calculated. The probability of these informations distances being greater than the 0.003 probability limit is observed. Again 100,000 groups are generated for each point on the ARL–curve. The results of the simulations are summarized in appendix D .

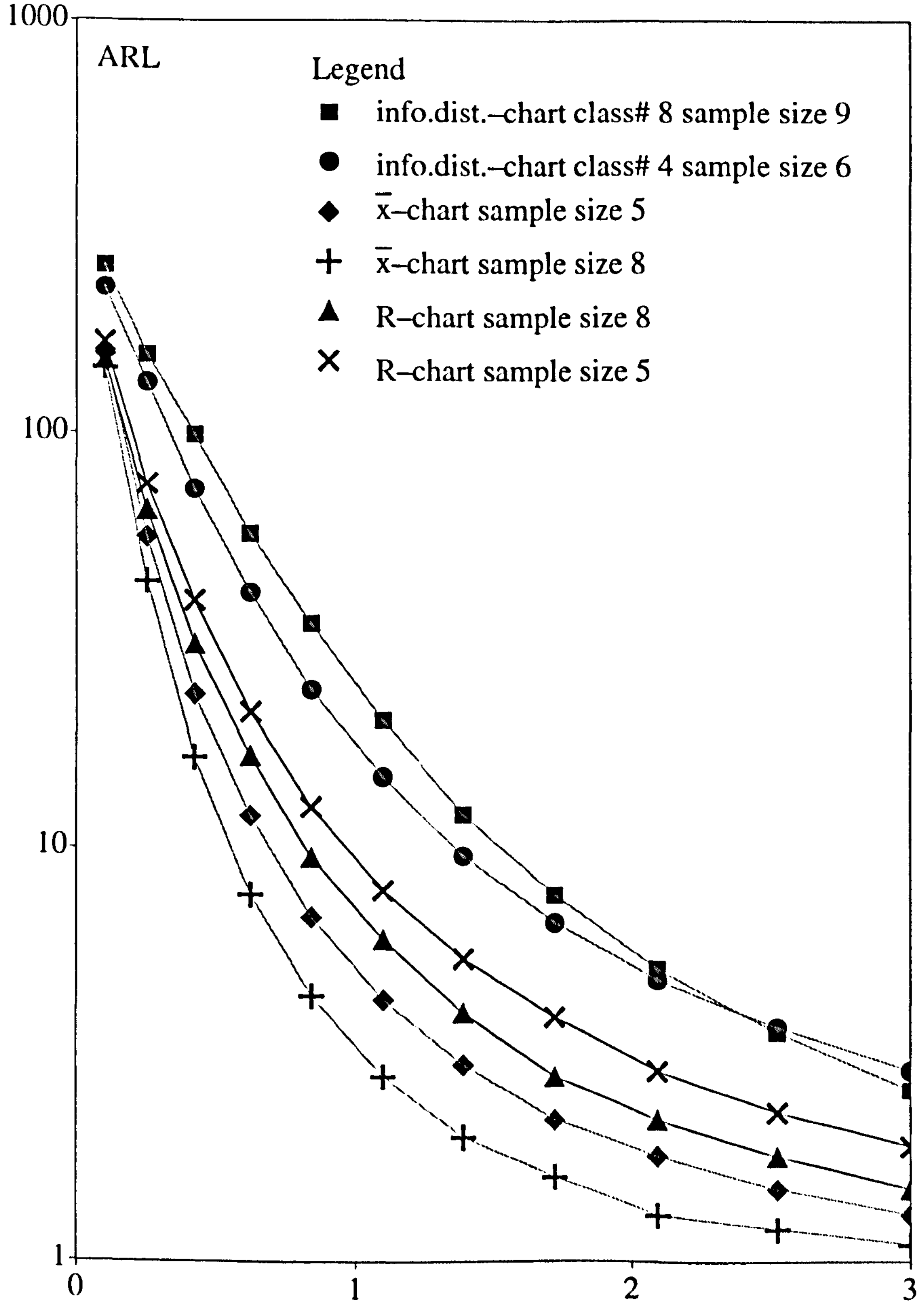


Figure 29 Comparisons of ARL number for traditional control charts simple information distance charts.

A similar simulation has also been made of the traditional control charts $\bar{x}$ and R. The results of these simulations are also given in appendix D . The ARL–curves for the traditional SPC charts and two of the best Information distance SPC charts are shown in Figure 29.

From Figure 29 we conclude that we get the expected alarm property of the information distance charts. The average run lengths are a little longer than those from the traditional charts. As the information distance chart ARL curves are consistently higher they can be shifted downwards simply by altering the alarm limit.

The characteristic shape of the curves may be altered in several different ways. One way is demonstrated in Figure 29, i.e. by the choice of number of classes and the sample size. In the above simulations we have used equidistant class width plus a tail class. Other class width distributions may of course also used to yield other decision properties. Another way to manipulate the shape is the choice of target probability distribution function. Indications of this may be seen in the discussions in chapter 6. It is concluded that with the information distance–SPC it is possible to design virtually the kind of ARL–curve wanted.

9.4. Information distance SPC summary

Above it has been shown that the information distance metric may be used to design a control chart applicable to different processes. In previous chapters we have shown the relevance of the information distance metric to most processes in industrial or pure administrative businesses. Accordingly the presented new chart may be widely applied. It will use the same administration and knowledge of the chart required wherever it is applied.

When applied it will be tying in the upstream benchmarking and other TQM activity results into the evaluations with the chart. As the new chart will be possible to design along different requirements on trade–off between location and spread, trade–offs are made centrally and consistently.

The traditional charts monitor statistical stability. The new chart does the same but in addition monitors quality level with respect to the set target density distribution function.

10. Discussion.

In this section we will discuss some of the observation reported above. Some merits and shortcomings of different concepts proposed will be brought up. Finally some recommendations for further research will be made.

10.1. Information distance metric.

In previous sections a quality metric has been proposed and has been used in different applications. All issues connected with quality have a probablistic nature. The proposed metric emphasizes the statistical approach. That is a strength but it may at first be considered a weakness. To get a good judgement more observations are needed. In connection to the discussion about SPC and information distance it was however shown that the predictivity of information distance was fair even with a rather small amount of statistical material. It is obvious that some care has to be exercised in choosing the sample size. That is of course true whichever method used, but with the proposed method it is rather more important.

In section 4.1. we made a listing of different quality characteristics. There was also an indication of different analysis methods, i.e. metrics, suitable for each one. With the proposed quality characteristic the analysis will be the same in all cases. That is valid even for dynamic characteristics. In the latter area we get a particular benefit as it is not necessary to go into details about statistical regression and other aspects of mathematical rigor. However to understand or develop a formulae for the SN–ratio of a dynamic characteristic requires considerable skills in regression.

In the reported applications we have seen that the absolute magnitude of the metric varies with the number of classes used in the density function. This is of course a weakness. However it reflects the amount of information that is being assessed. With other quality metrics there are also different magnitudes depending on what is chosen to be observed. This was clearly shown in the discussion about the paper feeding mechanism in section, 5.3. For other metrics there is only the expectation to measure relative levels of quality. With information distance you have the potential to measure absolute level of quality. The absolute level of quality needs a globally valid normalization strategy. This problem is not yet resolved. In section 6.4.3. some ideas on normalization was discussed. Below some further suggestions for normalization is given.

The actual choice of formulation of information distance belongs to a general class of distance definitions:

$D(P:Q) = \Sigma_i \, ((a+bp_i)ln((a+bp_i)/(a+bq_i))+(a+bq_i) \, ln((a+bq_i)/(a+bp_i)));$
$a+bq_i>0; \; a+bp_i>0,$ (See reference (66))

In chapter 5. $a=b=1.0$ was justified by simplicity in computer implementations. However by choosing those values there is minimum contribution to the distance as either p_i or q_i is zero. The distance is also limited from above and from below. It is always bigger than 0.0 and less than $4ln(2)$. The upper value is valid for extreme density functions. Thus it is not appropriate for realistic density functions q_i. The use of $4ln(2)$ is possible as normalizing factor. However that will not take the actual density function q_i into account. Thus coupling of the metric to the actual quality loss is lost. This way of normalization would not be a first choice according to the point of view of the author.

From the discussion about target density function, a smooth density function taking into account the quality losses was argued for. Losses are set to reflect world class performance. The argument for a smooth density function was to get the good dynamic in the quality system. Now let w be the information distance between the target density function q_i and the ideal performance. The latter is a density function with probability 1.0 in one class and 0.0 in all other classes. The unit probability class should be centered around the discrete target value. A proposal for normalization may be to divide the information distance between the performance density function, p_i, and the target density function, q_i, by w. That way distance will always be in fractions of w which is informative. The same class definitions for all three probability density functions discussed in this example is necessary to tie the normalization back to the loss level.

The choice of normalization is not straightforward. The property discussed above about the metric giving an indication of the amount of information assessed is a valuable property. It gives the user an awareness of the efficiency of the characteristic used. It is of interest to consider even the non–normalized information distance to some extent. In particular this is true as different characteristics are evaluated for use.

In our discussion we have compared this approach with alternative metrics from other researchers. None of them has had universal applicability. Very few have had a strong connection to the observable effects of poor quality, i.e. loss to society. Through the target density function and benchmarking, the proposed metric has such a connection.

10.2. Robust design and experimentation

In the reported applications we have seen some problems with the prediction procedure. In some places negative class probability was predicted. The prediction procedure adopted is the one given by Taguchi in reference (80). In the actual case, Table 37, the number of classes was high and several classes had missing observations. Accordingly the statistics of the data is poor. We also observe that with fewer classes the analysis of the same example works alright.

The analysis of variance for the two data analyses referred to above does not give good information as to what may be the problem. We note that significance levels are marginally lower from Table 36 than from Table 31. The data in those tables are calculated using accumulated results. Further it is an overall assessment. We may very well experience that one factor is very dominant and significant in a few classes but not in other classes. When using accumulated data these kinds of effects are easily hidden. When making a prediction only significant factors should be considered. Following the above reasoning predictions of some classes are made with non significant factors. Some researchers, (53), have also raised some questions over the relevance for the ANOVA of accumulated classes. This area needs to be considered further.

A prediction exercise based on the relative contribution to total variance gives other conditions for predictions. Disregarding factors contributing less than 5% gives that with 7 classes data factors A, D, E, F H should be considered. For 11 classes data factors A, E, F and H should be considered.

The reason for the effect of one single factor being significant at different levels through the different classes should also be considered. It may be that the prediction difficulties experienced due to these circumstances is an indication that interactions outside the experimentation model is present. The prediction part of the data analysis needs to be investigated in further depth. It may be possible to find an analysis tool to disclose model errors.

As regards the prediction method used a more rigorous analysis was made in section 6.5.2. That investigation gave as a novel result the prediction principle governing the Logit transform. In particular the Logit transform was found to be the derivative of the Shannon entropy with respect to the class probability p. We also note that the situation where the Logit–transform is normally used is one where the goal is to get p as close to unity as possible. That is also inherent to our investigation. However the prediction principle is interesting. In the general case with a target distribution the prediction principle is not really valid. Instead of striving to unity, there is a specific value for each class probability. The entropy pricing for fractions of p is no longer relevant. Rather, the entropy pricing for fractions of information distance may be valid. That may be a useful track to follow in order to develop a suitable prediction procedure.

In the applications, we have seen that the design of the target distribution is of vital importance for several reasons. Firstly, there is the statistical validity for prediction as discussed above. Secondly, there is the ability to extract information of the available observations. Thirdly, there is the relevance to world class performance.

The basic target density function may preferably be continuous and parameterized. This is for reasons that it is then conveniently tied to the benchmarking results via the loss function. In that way it is nicely controlled to reflect world class performance. In real applications the target density function is usually discretized. This discretisation may be manipulated by the parametrization of the basic target density function.

A manipulation making the basic target density function a little sharper may be incorporated as a standard evaluation procedure. When introducing the new metric we saw that an actual density function sharper than the target density function gives a positive distance. A proper application of the target density function should with a high probability indicate an overspending for sharper than target. A good evaluation package should of course spell out that this is happening. This may be accomplished by evaluation of the incremental change of information distance as the target density function is made slightly sharper. If the information distance is decreasing as you make the target density function becomes sharper the actual density function is sharper than target. At that stage whether there is overspending or the result is being achieved at low cost needs to be assessed. In the latter case new standards of performance are being set.

The ability of the target density function to extract information in the available material is essentially a matter of class definitions. Two different situations may be identified. Firstly there is a situation where the actual performance is far off from target. In that case classes need to be assigned to get good dynamic for gross changes towards compliance with the target. Secondly, there is the situation where the actual performance is on average close to the target. Then of course the class definition should be more refined around the center of the target density function.

The basic target density function is the same all the time. Depending on the actual density function for performance characteristics under consideration the discretization is varied as discussed above.

10.3. TQM

TQM may be regarded as a toolbox including all quality engineering tools. From a traditional viewpoint there is certainly nothing else but the toolbox and the objective, quality, to keep the tools together. However in the present report starting with the new metric information theory is introduced. It has been shown that the quality tools in the light of information theory may be considered as a means to filter out unwanted information from a process.

In Chapter 9. the application of information distance in the scope of statistical process control was demonstrated. This concept was based on the conceptual idea of the manufacturing system as a information processing system as shown by Figure 1.

In Chapter 8. the product development process was illustrated in the same way. The reasoning in this context gave a more solid foundation (towards information theory) to the empirically derived complexity number tool. Next we went on to the Total Design formalism and argued that this process was nothing else but to systematically refine the minimum information needed to make a product. For example Pugh called the product development specification, a design boundary. This is very true as the more the specification is quantified the more the variations during the design process are decreased.

A tentative quality measure for the product development process was proposed. With the information theoretical viewpoint it is conceptually easy to design quality measures for the product development process. The difficulty in pursuing the application is to define the target distribution. Here again the benchmarking tool is relevant. Over the last years it has become common to benchmark the product development process between companies.

Through the demonstrations discussed above we have seen that there is a common property of all quality tools. That property is an information refinement. The basic science concerning quality engineering is information theory. For years, since at least the pioneering work by Shannon, this has been realized by communication engineers. As this report puts the general quality engineering in the same framework it is possible for quality engineering to draw on the achievements within communication engineering.

The new metric requires a function or process description of the system. This is a big strength of the proposed procedure. It puts the focus on organization of information in early stages of development. This organization does not of course come about by itself. The information structuring does call for resource allocation in those stages. Many researchers have for long emphasized the importance of this kind resource allocation, (14). In real life this has however rarely happened. The information theoretical approach may give further support to a reallocation of resources within the development processes.

10.4. Further research

From the discussion above it is clear that this report is just the start of a new approach to quality engineering. Accordingly more research is needed. Below we will discuss some of the more relevant areas of research.

The first area is one concerned with target density functions. The dynamic of the quality metric in relation to the discretization of the target density function is of prime importance to the application. Further the benchmarking activity needs to be adapted to give a relevant target density function.

The RFCA procedure relies on a recognition of the system levels in the product design. In effect the base of the RFCA is a function description of the product. The different functions identified are actually being used as characteristics in the robust design activities. An area of research is the procedure to systematically break down the quality losses through the system levels to allocate appropriate loss levels to each function. This is a prerequisite for a proper design of the target density function.

The application of quality assessment using information distance in organizational processes is a large area where research is needed. The modern quality standards, i.e. ISO 9000, (81), or TQM models, (4), ask as said above for function descriptions of the company processes. Through these identified functions, quality characteristics may be designed. The efficiency of quality evaluation using information distance for these characteristics is a very interesting area of further investigation.

In Chapter 8. the information theory approach has been applied to system complexity. It is a strong belief of the author that a continued application of the information theoretical concept to DFX (Design For Xxxx) activities will be very rewarding. One entry point into this area may be the costing strategies as outlined by Pugh, (63). This will then draw on a widening of the complexity number. Another entry point is the modularization of design as outlined by Erixon *et. al.*, (12). Of particular interest is the impact of module interface design on quality, as discussed briefly in chapter 8.

Within the area of robust design an investigation into prediction strategies would be worthwhile. At present the Logit–transform is used (up till now) out of empirical reasons from a long application within robust design. It is working reasonably well but we have seen some spurious things in the applications above. With an information theoretical approach a revised procedure may be designed. Some suggestions have already been made.

Finally it should be of interest to investigate the quality metric, information distance, itself. What properties does it have in different aspects. One of the more important is the issue of normalization. In what way can the amount of information indication be conserved at the same time as you make a global quality standard is made?

11. Conclusions.

The present research set out to show the relevance of information theory as a basic theory for quality engineering. In the preceding chapters this has been done.

First a new metric was designed. A particular formulation of information distance was used. This was shown to have good quality evaluation properties. This has been done in both robust design applications and SPC simulations.

Further the applicability of the new quality metric has been demonstrated in both product and process quality evaluation. Processes considered include, development and manufacturing. The basis for applicability is a function based description. This assures applicability throughout all targeted organized activities.

Secondly the information theory has been shown to be a valid common base for quality engineering. Thus it has been shown that the information theory is an integrating factor for all quality related activities.

Poor quality has been identified to be a flow of surplus information in the system under consideration. Thus quality engineering tools have been interpreted as filters letting through only the information wanted in the system function. This is also completely in line with the modern system engineering theories.

Accordingly we may conclude that the task of a designer is information processing. A statement that is not that much appreciated amongst designers but nevertheless it is true.

12. References.

(1)N.,P., Suh,"Axiomatic Design of Mechanical Systems.",Trans. of The ASME, **117**, June 1995, pp 2 – 10.

(2)N.,P., Suh,"Designing-in of Quality Through Axiomatic Design..",IEEE Trans. on Reliability, **44**, No 2, June 1995, pp 256 – 264.

(3)B. Svensson, "Poor Quality –a Surplus of information.", Procedings of Productmodeller –95, pp 349 – 362, Linköping, Sweden, April, 1995.

(4)E., C., Foley,"Winning European Quality.", European Foundation for Quality Management, 1994.

(5)D. Clausing, "Total quality development: a step–by–step guide to world–class concurrent engineering.", ASME Press, 506p, 1994.

(6)M. A. Styblinski, " Generalized Formulation of Yield, Variability, Min-Max and Taguchi Circuit Optimization Problems.", Microelectronic Reliability, **34**, 1994:1, pp 31 – 37.

(7)J. L. McGovern, " A Critique of the Taguchi Approach – Part I: A Presentation of Some Deficiencies and How these Limits Its Efficiency and Validity.", Journals of Coatings Technology, **66**, March 1994:830, pp 65 – 70.

(8)J. L. McGovern, " A Critique of the Taguchi Approach – Part II: An Alternative that is more Efficient. ", Journals of Coatings Technology, **66**, April 1994:831, pp 55 – 61.

(9) G. Erixon, A. Erlandsson, A. v. Yxkull, B. M. Östergren, "Modulindela produkten.", Förlags AB Industrilitteratur, Stockholm, 1994. (In Swedish)

(10) P. D. T. O'Connor, "The practice of engineering management.", John Wiley & Sons Ltd, Chichester, England, 1994.

(11) P. Karlsson, "Assessment of customer value in the truck market.", Internal report to Volvo Truck Corp., Gotheburg, Sweden, 1994.

(12) G. Erixon, A. Erlandsson, A.v. Yxkull, B. M. Östergren, "Modularize the Product.", Swedish Manufacturing Industry Publishing Company AB, Stockholm, Sweden, 1994.

(13)M. Petzäll, "Process improvement for PVC–pipe processing.", Internal repport to UPONOR AB, Fristad, Sweden, 1994, (In Swedish).

(14) J. Fox, "Quality Through Design: The key to successful product delivery.", McGraw–Hill, Maidenhed, UK, 1993

(15) K. Sontow, D. P. Clausing, "Integration of Quality Function Deployment with Further Methods of Quality Planning.", LMP–93–005 (Working paper), Massachussetts Institute of Technology, Cambridge, Massachussetts, USA, 1993.

(16) J. Krottmaier, "Optimizing Engineering designs.", McGraw–Hill Publishing Company, London, 1993

(17)L. E. Shirland, "Statistical Quality Control with Microcomputer Applications.", John Wiley & Sons, Inc., New York, 1993

(18)F. Mistree, U. Lautenschlager, S. O. Erikstad, J. K. Allen, "Simulation Reduction Using the Taguchi Method.", NASA Contract Report 4542, University of Houston, Houston, Texas, Oct, 1993.

(19)A. Vedin, J. Bäckman, "Robust–design av rotationsgjutningsprocess.", Diploma work at The University of Borås, Borås, Sweden, 1993, (In Swedish).

(20)T. N. Goh, "Taguchi Methods: Some Technical, Cultural and Pedagogical Perspectives.", Quality and Reliability Engineering International, **9**, 1993, pp 185 – 202.

(21)G. K. Robinson, "Improving Taguchis Packaging of Fractional Factorial Designs.", Journal of quality Technology, **25**, January 1993:1, pp 1 – 11.

(22)G. Abdul–Nour, "On Some Factors Affecting The Just–In–Time Production System Output Variability: A Simulation Study Using Taguchi Techniques.", Computers and Industrial Engineering, **25**, 1993:1–4, pp 461 – 464.

(23)G. Abdul–Nour, "The Customer Service Level as Affected by Machine Unreliability in a JIT Production System Environment: A Simulation Study Using Taguchi Techniques.", Institute of Industrial Engineers 2nd Industrial Engineering Research Conference Proceedings, 1993, pp 415 – 419

(24)Y. H. A. Liou, P. P. Lin, R. R. Lindeke, H. D. Chiang, "Tolerance Specification of Robot Kinematic Parameters Using an Experimental Design Technique – The Taguchi Method.", Robotics and Computer Integrated Manufacturing, **10**, 1993:3, pp 199 – 207.

(25)M. Hamada, "Reliability Improvement Via Taguchi's Robust Design.", Quality and Reliability Engineering International, **9**, 1993, pp 7 – 13.

(26)K. N. Otto, E. K. Antonsson, "Extension to the Taguchi Method of Product Design.", Journal of Mechanical Design, **115**, March 1993:1, pp 5 – 13.

(27)J. P. Bentley,"An Introduction to Reliability and Quality Engineering.", Longman Scientific & Technical, Longman Group UK Ltd, Harlow, 1993.

(28)D. J. Wilde,"Monotonicity Analysis of Taguchi's Robust Circuit Design Problem.",Transaction of the ASME, Vol 114, Dec 1992, pp 616 – 619.

(29)D. K. Hitchins, "Putting Systems to Work.", John Wiley & Sons, Ltd., Chichester, UK, 1992

(30)J. S. Ramberg, J. J. Pignatiello Jr., S. M. Sanchez, "A Critique and Enhancement of The Taguchi Method.", ASQC Quality Congress Transactions – Nashville, 1992, pp 491 – 498,

(31)A. Kraslawski, T. Koiranen, L. Nyström, C. Gourdon, "Concurrent Engineering: Fuzzy Simulation and Similarity in Quality in Quality Loss Function Development and Applications.", Computers & Chemical Engineering, **16**, 1992, pp 361 – 368.

(32)N. D. Singpurwalla, "A Bayesian Perspective on Taguchi's Approach to Quality Engineering and Tolerance Design.", IIE Transactions, **24**, November 1992:5, pp 18 – 32.

(33)S. Wu, J. V. Zidek, # An Entropy–based Analysis of data from selected NADP/NTN Network Sites For 1983 – 1986.", Atmospheric Environment, **26A**, 1992:11, pp 2089 – 2103.

(34)N. Suh, "Design Axioms and Quality Control.", Robotics & Computer–Integrated Manufacturing, **9**, 1992:4/5, pp 367 – 376.

(35)R. V. LeÓn, C. F. J. Wu, "A Theory of Performance measures In Parameter Design.", Statistica Sinica, **2**, 1992, pp 335 – 358.

(36)Y.–S. Chen, K. Tang, "A Pictorial Approach to Poor–Quality Cost Management.", IEE Transaction on Engineering Management, **39**, May 1992:2, pp 149 – 157.

(37)S. Ashley, "Applying Taguchi's Quality Engineering to Technology Development.", Mechanical Engineering, July 1992, pp 58 – 60.

(38)T. N. Goh, "An Organizational Approach to Product Quality via Statistical Experimantal Design.",International Journal of Production Economics, **27**, 1992, pp 167 – 173.

(39)A.E. Freeny, V. N. Nair, "Robust parameter design with uncontrolled noise variables.", Statistica Sinica, **2**, (1992), pp 313 – 334.

(40)R. H. Myers, A. I. Khuri, G. Vining, "Respons surface alternatives to Taguchi robust parameter design approach." The American Statistician, **46,** May 1992:2, pp 131 – 139.

(41)J. L. Burati Jr,J. J. Farrington, W. B. Ledbetter, "Causes of Quality Deviations in design and Construction.", Journal of Construction Engineering and Management., **118**, 1992:1, pp 34 – 49.

(42)P. Sandvik–Wiklund, "Some contributions to industrial design of experiments.", LiU–Tek–Lic–1992:07, Linköping University, Sweden, 1992.

(43) I. E. Morely, D.–M. Hosking, "Total Design – Teamworking.", Design division, University of Strathclyde, Glasgow, 1992.

(44) B. Pease, "What's All This Taguchi Stuff, Anyhow?", Electronic Design, June 25, 1992, pp 83 – 84

(45) B. Pease, "What's All This Taguchi Stuff, Anyhow? (part II)", Electronic Design, June 10, 1993, pp 85 – 92

(46) R. Andersson, "QFD A system for efficient product development.", Studentlitteratur, Lund, Sweden, 1991, (In Swedish)

(47) D. C. Montgomery, "Introduction to Statistical Quality Control.", John Wiley & Sons Inc., New York, 1991.

(48) K. Akiyama, "Function analysis: Systematic improvement of quality and performance.", Productivity press Inc., 1991

(49) D. P. Clausing, S. Pugh, "Enhanced Quality Function Deployment.", pp 15 – 25, Proceedings Vol. 1. Design Productivity International Conference, Honolulu, Hawaii, 1991.

(50) L. Ljung, T. Glad, "Modellbygge och simulering.", Studentlitteratur, Lund, Sweden, 1991. (In Swedish)

(51)G. Bateson, "Further steps to an ecology of mind.", Cornelia & Bessie Book, New York, 1991.

(52)A. Rosenblatt, G.F Watson, "Concurrent Engineering – Special report.", IEEE Spectrum, July 1991, pp 22 – 37.

(53)M. Hamada, C. F. J. Wu, "Analysis of Censored data from Highly Fractionated Experiments.", Technometrics, **33**, Ferbruary 1991:1, pp 25 – 38.

(54)A. C. Shoemaker, K.–L. Tsui, C. F. J. Wu, "Economical Experimentation Methods for Robust Design.", Thechnometrics, **33**, November 1991:4, pp 415 – 427.

(55)J. A. Nedler, Y. Lee, "Generalized Linear Models for the Analysis of Taguchi–type Experiments.", Applied Stochastic Models and Data Analysis, **7**, 1991, pp 107 – 120.

(56)J. H. de Boer, A. K. Smilde, D. A. Doornbos, "Introduction of a Robustness coefficient in Optimization Procedures: Implementation in Mixture Design Problems. Part III: Validation and Comparison With Competing Criteria.", Chemometrics and Intelligent Laboratory Systems, **15**, 1991, pp 13 – 28.

(57)R. J. Carroll, P. Hall, "Nonparametric Estimation of Optimal Performance Criteria in Quality Engineering.", The Annals of Statistics, **18**, 1990:1, pp 181 – 302.

(58)W. J. Welch, T.–K. Yu, S. M. Kang, J. Sacks, "Computer Experiments for Quality Control by Parameter Design.", Journal of Quality Technology, **22**, January 1990:1, pp 15 – 22.

(59)L. Liu, W. A. Nazaret, R. G. Beale, "Computer–Aided Design for Quality (CADQ).", AT&T Technical Journal, May/June 1990, pp 46 – 60.

(60) B. Svensson, "Integrating analysis tools through the use of Taguchi methods.", Proceedings of the 1st international conference on integrated technology management, pp 42 – 53, IFS–conferences, London, 1990.

(61) Y, Akao (editor), "Quality Function Deployment, Integrating Customer Requirements into Product Design.", Productivity Press, Portland, Oregon, 1990.

(62) N. P. Suh, "The Principles of Design", Oxford University Press, New York, 1990

(63) S. Pugh, "Total design.", Addison–Wessley, Wokingham,England, 1990.

(64) M. S. Phadke, "Quality engineering using robust design.", Prentice–Hall, 1989.

(65) G. Taguchi, E. A. Elsayed, T. Hsiang, "Quality Engineering in Production Systems.", McGraw–Hill Publishing Company, New York, 1989.

(66)J. N. Kapur, "Maximum entropy models in science and engineering.", John Wiley & Son Inc, New York, 1989.

(67)M. Owen, "SPC and Continuous Improvement.", IFS Publications/ Springer–Verlag, Berlin, 1989.

(68)Y. Wearn, "Cognitive aspects of computer supported tasks.", John Wiley & Sons Ltd, Chichester, England, 1989.

(69)G. Box, "Signal–to–Noise Ratios, Performance Criteria, and Transformations.", Technimetrics, **30**, February 1988:1, pp 1 – 40.

(70)D. A. Norman, "Psychology of everyday things.", Basic Books Inc., New York, 1988.

(71)E. L. Grant, R. S. Leavenworth, "Statistical Quality Control, 6th ed.", McGraw–Hill Book Company Inc., San Franscisco, 1988.

(72)K. Lindgren, "Physics and information theory.", Dissertation, Division of Physics Chalmers University of Technology, Gothenburg, Sweden, 1988.

(73) P. J. Ross, "Taguchi Techniques for Quality Engineering.", McGraw-Hill Book Company Inc., San Franscisco, 1988.

(74) G. Pahl, W. Beitz, "Engineering design, a systematic approach.", Springer Verlag, 1988.

(75) L. A. Ealey, "Quality by Design, Taguchi Methods and U.S. Industry.", ASI Press, Dearborn, Michigan, 1988.

(76) S. Pugh, I. E. Morely, "Total Design – Towards a Theory of Total Design.", Design division, University of Strathclyde, Glasgow, 1988.

(77) L. Sandblom, "Quality Control.", Studentlitteratur, Lund, Sweden, 1988. (In Swedish)

(78) J.M. Juran, "Quality Control Handbook, Fourth Edition" McGraw-Hill, New York, 1988.

(79) J.J. Pignatiello Jr, "An Overview of the Strategy and Tactics of Taguchi.", IIE Transaction, Vol 20, Nr 3, pp 247 – 254. Sept 1988.

(80) G. Taguchi, "System of experimental design", Vol 1 &2, Unipub Kraus International Publications, New York, 1987

(81) ISO–9000, Quality management and quality assurance standards – Guidelines for selection and use., 1987.

(82) ISO–9001, Quality systems – Model for quality assurance in design/development, production, installation and servicing., 1987

(83) ISO–9002, Quality systems – Model for quality assurance in production and installation., 1987

(84) ISO–9003, Quality systems – Model for quality assurance in final inspection and test., 1987

(85) ISO–9003, Quality management and quality system elements – Guidelines., 1987

(86)K.–E. Eriksson, K. Lindgren, B. Å. Månsson, "Structure, context, complexity, organization.", Singapore World Science Corp, 1987.

(87)G.Bateson, "Steps to an ecology of mind.", Aronson Coperation, Northvale, New Jersey, 1987. (Chandler publications for health sciences)

(88)R. V. LeÓn, A. C. Shoemaker, R. N. Kacker, "Performance Measures Independent of Adjustment.", Technometrics, **29**, August 1987:3, pp 253 – 285.

(89)G. E. P. Box, R. D. Meyer, " Dispersion Effects From Fractional Designs.", Technometrics, **28**, February 1986:1, pp 19 – 27.

(90)V.N. Nair, D. Pregbon, "A data analysis strategy for quality engineering experiments.", AT&T Thechnical Journal, **65**, May/June 1986:3, pp 73 – 84.

(91)P. W. Atkins, "Physical Chemistry.", Oxford University Press, Oxford, 1986.

(92) P. W. Atkins, "Physical chemistry.", 3rd–edition, Oxford University Press, Oxford, 1986.

(93) G. Taguchi, "Introduction to quality engineering.", Asian Productivity Organization, 1986.

(94) K. Ishikawa, "What is Total Quality Control? The Japanese Way.", Prentice Hall, Eaglewood Cliffs, N.J., 1985.

(95)P. E. Gill, W. Murray, M. Wright, "Practical Optimization.", Academic Press, London, 1982.

(96)D. Pregibon, "Godeness of Link Tests for Generalized Linear models.", Applied Statistics, **29**, 1980:1, pp 15 – 24.

(97) The Asahi, Japanese language newspaper, April 15 1979. (As referenced by Phadke in (64).)

(98) G. E. P. Box, W. G. Hunter, J. S. Hunter, "Statistics for experimenters.", John Wiley & Sons, Inc. , New York, 1978.

(99) N. Lundqvist, R. Fridlund, "Value Analysis.", 2nd edition, Swedish Manufacturing Industry Publishing Company AB, Stockholm, Sweden, 1972. (In Swedish)

(100)L. Brillouin, "Science and information theory.", Academic Press NY, 1962.

(101) J.M. Juran, "Quality Control Handbook." McGraw-Hill, New York, 1951.

(102)C. E. Shannon, "A mathematical theory of communication.Part I -IV.", The Bell system technical journal., Vol. XXVII, No. 3, pp 379 - 423; 623 - 656, 1948.

(103) A. Wald, "Sequential analysis.", John Wiley &Sons, New York, 1947

(104) H.F. Dodge, H.G. Roming, "Single sampling and double sampling inspection tables.", The Bell System Technical Journal, No 20, pp. 1 - 61, 1941.

(105) W.A. Shewhart, "Statistical Method. From the Viewpoint of Quality Control.", Graduate School of the Department of Argiculture, Washington D.C., 1939.

(106) R. A. Fisher, "Statistical methods for research workers.", Oliver and Boyd, 1935.

(107) R. A. Fisher, The design of experiment., Oliver and Boyd, 1935.

(108) R. Becker, H. Plaut, I. Runge, "Anwendungen der Mathematischen Statistik auf Problem der Massfabrikation.", Springer-Verlag, 1931. (In German)

(109) W.A. Shewhart, "Economic Control of Quality of Manufactured product.", Van Norstrand, New York, 1931.

(110) H.F. Dodge, H.G. Roming, "A method of sampling inspection.",The Bell System Technical Journal, No 8, pp. 613 – 631, 1929.

(111) K.H. Daeves, "The utilization of statistics. A New and valuable aid in industrial research and in the evaluation of test data.", Testing, March, pp 173 – 189, 1924.

(112)F. X. Brown, R. W. Kane, "Quality cost and Profit Performance.", ASQC Technical Conference Transactions, Milwaukee, pp 505 – 514.

13. Appendices.

13.1. Appendix A . A Robust Engineering Example.

For reference purpose we give a short summary of a robust engineering example presented by Phadke, (64).

By manufacturing of very large scale integrated (VLSI) circuits there are many process steps. One example of these are the polysilicon deposition process. This step has been subject of a robust design study. The process is performed in a reduced pressure reactor. The reactor consists of a quartz tube which is heated by a 3–zone furnace. Silane and nitrogen gases are introduced at one end and pumped out at the other. The silane pyrolizes, and a polysilicon layer is deposited on top of the oxide layer on the silicone wafers. The wafers are mounted in quartz carriers. Two carriers, each carrying 25 wafers, are placed inside the reactor at a time so that polysilicon is deposited simultaneously on 50 wafers.

At the start of the study there were two main problems with this process: 1) too many surface defects were encountered, and 2) too large a thickness variation within wafers and among wafers. These two deficiencies caused a lot of scrap down stream in the process. It was decided to use the number of surface defects, polysilicon layer thickness and deposition rate as function characteristics.

Next the noise factors influencing the process have to be considered. As discussed above it is not always necessary to use an orthogonal array for the noise factors. The important point is to assure that there is a large amount of noise influencing the observed results of the experiments. The nonuniform thickness and the surface defects are caused by the variations in the parameters involved in the chemical reactions associated with the deposition process. As the silane gas decomposes during its way through the reactor we have a concentration gradient along the length of the reactor. As the wafers themselves are obstacles to the gas flow there is a nonuniform flow pattern for the gases at different wafers. The gas flow also causes a temperature variation in the reactor. Further there may be other noises such as, topography of the wafer surface before polysilicon deposition, variation in pumping speed and variation in gas supply. In the experimentation it was decided to evaluate wafer number 3, 23 and 48 (out of the fifty) along the reactor. In addition each wafer was evaluated in three different positions, top, middle and bottom. These steps were judged to capture the significant noises active in the process.

To finalize the planning activity we look into the choice of control factors. The general practice is to adapt the number of factor and factor levels to the level of knowledge of the process under study. In the present case the knowledge level was reasonable. Accordingly fewer factors and more levels for each factor were chosen. Six factors and three levels as shown in Table 38 were used. The two last factors may need further comment. It is important to establish thermal and pressure equilibrium inside the reactor before the reaction is allowed to start. This equilibrium is establish during a time allowance (settling time) between reactor charging and gas flow start. Before charging in the reactor, the wafers may be cleaned to reduce the number of surface defects. Three different cleaning methods were tested.

Table 38 Design factors and their levels.

Factor	Levels*		
	1	2	3
A. Deposition temperature (C^o)	T_o–25	$\underline{T_o}$	T_o+25
B. Deposition pressure (mtorr)	P_o–200	$\underline{P_o}$	P_o+200
C. Nitrogen flow (sccm)	$\underline{N_o}$	N_o–150	N_o–75
D. Silane flow (sccm)	S_o–100	S_o–50	$\underline{S_o}$
E. Settling time (min)	$\underline{t_o}$	t_o+8	t_o+16
F. Cleaning method	$\underline{\text{None}}$	CM_2	CM_3

*Operating levels before improvements are identified by underscore

The interactions among the control factors were not very well understood. Accordingly no interaction was accounted for in the experimentation plan. The six control factors were allocated to the columns of a L_{18} matrix as is shown in Table 39 .

Table 39 L_{18} orthogonal array indicating factors assigned to different columns.

Expt. no.	Column numbers and factor assignments*							
	1 e	2 A	3 B	4 C	5 D	6 E	7 e	8 F
1	1	1	1	1	1	1	1	1
2	1	1	2	2	2	2	2	2
3	1	1	3	3	3	3	3	3
4	1	2	1	1	2	2	3	3
5	1	2	2	2	3	3	1	1
6	1	2	3	3	1	1	2	2
7	1	3	1	2	1	3	2	3
8	1	3	2	3	2	1	3	1
9	1	3	3	1	3	2	1	2
10	2	1	1	3	3	2	2	1
11	2	1	2	1	1	3	3	2
12	2	1	3	2	2	1	1	3
13	2	2	1	2	3	1	3	2
14	2	2	2	3	1	2	1	3
15	2	2	3	1	2	3	2	1
16	2	3	1	3	2	3	1	2
17	2	3	2	1	3	1	2	3
18	2	3	3	2	1	2	3	1

* Empty columns are identified by e. Effects analyzed in those columns are indicating model errors.

The detailed specifications of the 18 different experiments to be carried out are given in Table 40.

Table 40 Experimenters log

Expt. no.	Tempera-ture	Pressure	Nitrogen	Silane	Settling time	Cleaning Method
1	T_o–25	P_o–200	N_o	S_o–100	t_o	*None*
2	T_o–25	P_o	N_o–150	S_o–50	t_o+8	CM_2
3	T_o–25	P_o+200	N_o–75	S_o	t_o+16	CM_3
4	T_o	P_o–200	N_o	S_o–50	t_o+8	CM_3
5	T_o	P_o	N_o–150	S_o	t_o+16	*None*
6	T_o	P_o+200	N_o–75	S_o–100	t_o	CM_2
7	T_o+25	P_o–200	N_o–150	S_o–100	t_o+16	CM_3
8	T_o+25	P_o	N_o–75	S_o–50	t_o	*None*
9	T_o+25	P_o+200	N_o	S_o	t_o+8	CM_2
10	T_o–25	P_o–200	N_o–75	S_o	t_o+8	*None*
11	T_o–25	P_o	N_o	S_o–100	t_o+16	CM_2
12	T_o–25	P_o+200	N_o–150	S_o–50	t_o	CM_3
13	T_o	P_o–200	N_o–150	S_o	t_o	CM_2
14	T_o	P_o	N_o–75	S_o–100	t_o+8	CM_3
15	T_o	P_o+200	N_o	S_o–50	t_o+16	*None*
16	T_o+25	P_o–200	N_o–75	S_o–50	t_o+16	CM_2
17	T_o+25	P_o	N_o	S_o	t_o	CM_3
18	T_o+25	P_o+200	N_o–150	S_o–100	t_o+8	*None*

As the experiments were performed data over the three different characteristics were collected and assembled as shown in Table 41 and Table 42.

Table 41 Surface defect data (Defects/unit area)

Expt. No.	Test wafer 1			Test wafer 2			Test wafer 3		
	Top	Center	Bottom	Top	Center	Bottom	Top	Center	Bottom
1	1	0	1	2	0	0	1	1	0
2	1	2	8	180	5	0	126	3	1
3	3	35	106	360	38	135	315	50	180
4	6	15	6	17	20	16	15	40	18
5	1720	1980	2000	487	810	400	2020	360	13
6	135	360	1620	2430	207	2	2500	270	35
7	360	810	1215	1620	117	30	1800	720	315
8	270	2730	5000	360	1	2	9999	225	1
9	5000	1000	1000	3000	1000	1000	3000	2800	2000
10	3	0	0	3	0	0	1	0	1
11	1	0	1	5	0	0	1	0	1
12	3	1620	90	216	5	4	270	8	3
13	1	25	270	810	16	1	225	3	0
14	3	21	162	90	6	1	63	15	39
15	450	1200	1800	2530	2080	2080	1890	180	25
16	5	6	40	54	0	8	14	1	1
17	1200	3500	3500	1000	3	1	9999	600	8
18	8000	2500	3500	5000	1000	1000	5000	2000	2000

Table 42 Thickness and deposition rate data.

Ex. No.	Thickness (Å)									De-pos. Rate*
	Test wafer 1			Test wafer 2			Test wafer 3			
	Top	Center	Bottom	Top	Center	Bottom	Top	Center	Bottom	
1	2029	1975	1961	1975	1934	1907	1952	1941	1949	14.5
2	5375	5191	5242	5201	5254	5309	5323	5307	5091	36.6
3	5989	5894	5874	6152	5910	5886	6077	5943	5962	41.4
4	2118	2109	2099	2140	2125	2108	2149	2130	2111	36.1
5	4102	4152	4174	4556	4504	4560	5031	5040	5032	73.0
6	3022	2932	2913	2833	2837	2828	2934	2875	2841	49.5
7	3030	3042	3028	3486	3333	3389	3709	3671	3687	76.6
8	4707	4472	4336	4407	4156	4094	5073	4898	4599	105.4
9	3859	3822	3850	3871	3922	3904	4110	4067	4110	115.0
10	3227	3205	3242	3468	3450	3420	3599	3591	3535	24.8
11	2521	2499	2499	2576	2537	2512	2551	2552	2570	20.0
12	5921	5766	5844	5780	5695	5814	5691	5777	5743	39.0
13	2792	2752	2716	2684	2635	2606	2765	2786	2773	53.1
14	2863	2835	2859	2829	2864	2839	2891	2844	2841	45.7
15	3218	3149	3124	3261	3205	3223	3241	3189	3197	54.8
16	3020	3008	3016	3072	3151	3139	3235	3162	3140	76.8
17	4277	4150	3992	3888	3681	3572	4593	4298	4219	105.3
18	3125	3119	3127	3567	3563	3520	4120	4088	4138	91.4

First an ordinary signal to noise ratio, (SN–ratio) analysis was performed. The intermediate results are summarized in Table 43 . The first line of that table is derived as follows:

Number of surface defects a "Smaller the better" type of characteristic,

$$\eta = -10\log_{10}\left[\frac{1}{9}\sum_{i=1}^{3}\sum_{j=1}^{3} y_{ij}^2\right] =$$

$$-10\log_{10}\left(\frac{(1^2+0^2+1^2)+(2^2+0^2+0^2)+(1^2+1^2+0^2)}{9}\right) =$$

$$-10\log_{10}\left(\frac{8}{9}\right) = 0.51$$

Polysislicon layer thickness, a "Nominal is best" type of characteristic,

$$\mu = \frac{1}{9}\sum_{i=1}^{3}\sum_{j=1}^{3}\tau_{ij} =$$

$$\frac{1}{9}((2029+1975+1961)+(1975+1934+1907)+(1952+1941+1949)) =$$

$$1958.1\,\AA$$

$$\sigma^2 = \frac{1}{8}\sum_{i=1}^{3}\sum_{j=1}^{3}(\tau_{ij}-\mu)^2 =$$

$$\frac{1}{8}\left((2029-1958.1)^2+\ldots+(1949-1958.1)^2\right) = 1151.36\,(\AA)^2$$

$$\eta' = 10\log_{10}\frac{\mu^2}{\sigma^2} = 10\log_{10}\frac{1958.1^2}{1151.36} = 35.22\,dB$$

Deposition rate, a "Larger the better" type of characteristic. Not properly treated as you have only one data observation per control factor setting,

$$\eta'' = 10\log_{10}r^2 = 20\log_{10}r = 20\log_{10}(14.5) = 23.23\,dBam$$

here r is the deposition rate. The remaining 17 lines were calculated in the same way. The different characteristics were then analyzed for factor effects. The analysis comprised of both average factor level responses and the analysis of variance (ANOVA).

Table 43 Data summary by experiment.

	Experiment Condition	Surface Defects	Thickness		Deposition Rate
Expt. No.	Matrix* e A B C D E e F	η (dB)	μ (Å)	η' (dB)	η'' (dB)
1	1 1 1 1 1 1 1 1	0.51	1958	35.22	23.23
2	1 1 2 2 2 2 2 2	–37.30	5255	35.74	31.27
3	1 1 3 3 3 3 3 3	–45.17	5965	36.02	32.34
4	1 2 1 1 2 2 3 3	–25.76	2121	42.25	31.15
5	1 2 2 2 3 3 1 1	–62.54	4572	21.43	37.27
6	1 2 3 3 1 1 2 2	–62.23	2891	32.91	33.89
7	1 3 1 2 1 3 2 3	–59.88	3375	21.39	37.68
8	1 3 2 3 2 1 3 1	–71.69	4527	22.84	40.46
9	1 3 3 1 3 2 1 2	–68.15	3946	30.60	41.21
10	2 1 1 3 3 2 2 1	–3.47	3415	26.85	27.89
11	2 1 2 1 1 3 3 2	–5.08	2535	38.80	26.02
12	2 1 3 2 2 1 1 3	–54.85	5781	38.06	31.82
13	2 2 1 2 3 1 3 2	–49.38	2723	32.07	34.50
14	2 2 2 3 1 2 1 3	–36.54	2852	43.34	33.20
15	2 2 3 1 2 3 2 1	–64.18	3201	37.44	34.76
16	2 3 1 3 2 3 1 2	–27.31	3105	31.86	37.71
17	2 3 2 1 3 1 2 3	–71.51	4074	22.01	40.45
18	2 3 3 2 1 2 3 1	–72.00	3596	18.42	39.22

The analysis results are summarized in Table 44, Table 45 and Table 46. The task of selecting optimal factor level settings is now a little difficult as there needs to be a trade–off between the different characteristics. The main analysis results for all characteristics are therefore summarized in Table 47.

Table 44 Analysis of surface defects data*

Factor	Average η by factor level (bB)			Degree of freedom	Sum of squares	Mean square	F
	1	2	3				
A. Temperature	–24.23	–50.10	–61.76	2	4427	4414	27
B. Pressure	–27.55	–47.44	–61.10	2	3416	1708	21
C. Nitrogen	–39.03	–55.99	–61.10	2	1030	515	6.4
D. Silane	–39.20	–46.85	–50.04	2	372	186	2.3
E. Settling time	–51.52	–40.54	–44.03	2	378	189	2.3
F. Cleaning method	–45.56	–41.58	–48.95	2	164 °	82	
Error				5	405 °	81	
Total				17	10192		
(Error)				(7)	(569)	(81)	

* Overall mean h = –45.36 dB. Underscore indicates starting level.
° Indicates the sum of squares added together to form the pooled error sum of squares shown in parentheses. Pooling is a procedure where you treat, factors with a variance less than the average error variance, as errors.

Table 45 Analysis of thickness data*

Factor	Average η' by factor level (bB)			Degree of freedom	Sum of squares	Mean square	F
	1	2	3				
A. Temperature	35.12	34.91	24.52	2	440	220	16
B. Pressure	31.61	30.70	32.24	2	7 °	3.5	
C. Nitrogen	34.39	27.86	32.30	2	134	67	5.0
D. Silane	31.68	34.70	28.17	2	128	64	4.8
E. Settling time	30.52	32.87	31.16	2	18 °	9	
F. Cleaning method	27.04	33.67	33.85	2	181	90.5	6.8
Error				5	96 °	19.2	
Total				17	1004	59.1	
(Error)				(9)	(121)	(13.4)	

* Overall mean η' = 31.52 dB. Underscore indicates starting level.
° Indicates the sum of squares added together to form the pooled error sum of squares shown in parentheses.

Table 46 Analysis of deposition rate data*

Factor	Average η" by factor level (bBam)			Degree of freedom	Sum of squares	Mean square	F
	1	2	3				
A. Temperature	28.76	34.13	39.46	2	343.1	171.5	553
B. Pressure	32.03	34.78	35.54	2	41.0	20.5	66
C. Nitrogen	32.81	35.29	34.25	2	18.7	9.4	30
D. Silane	32.21	34.53	35.61	2	36.3	18.1	58
E. Settling time	34.06	33.99	34.30	2	0.3 °	0.2	
F. Cleaning method	33.81	34.10	34.44	2	1.2 °	0.6	
Error				5	1.3 °	0.26	
Total				17	441.9	25.9	
(Error)				(9)	(2.8)	(0.31)	

* Overall mean η" = 34.12 dBam. Underscore indicates starting level.

° Indicates the sum of squares added together to form the pooled error sum of squares shown in parentheses.

The deposition temperature has the largest effect on all three characteristics. A 25 C° reduction of temperature, compared to the starting level, results in a 26 dB improvement of, the number of defects, characteristics. The same level change does give a negligible effect on thickness and a 5.4 dB reduction of deposition rate. Thus we get a 20–fold reduction in number of defects and a 2–fold reduction of the deposition rate.

The deposition pressure has the next largest effect on surface defects and deposition rate. A 200 mtorr reduction of the pressure results in a 10–fold reduction of surface defects and 37 percent reduction of deposition rate. The effect on thickness variation is very small.

Nitrogen flow rate has a moderate effect on all three characteristics. The starting level for this factor gives the highest SN–ratios for surface defects and thickness variation. An indication of further improvement by increase of the flow rate of this dilutant gas is noted for future experiments. A reduction of the flow rate may increase the deposition rate slightly.

Table 47 Summary of factor effects.

Factor	Level	Surface defects η (dB)	F	Thickness η' (dB)	F	Deposition rate η'' (dBam)	F
A. Temperature	A_1: T_o–25	–24.23		35.12		28.76	553
	A_2: T_o	–50.10	27	34.91	16	34.13	
	A_3: T_o+25	–61.76		24.52		39.46	
B. Pressure	B_1: P_o–200	–27.55		31.61		32.03	
	B_2: P_o	–47.44	21	30.70	–	34.78	66
	B_3: P_o+200	–61.10		32.24		35.54	
C. Nitrogen	C_1: N_o	–39.03		34.39		32.81	
	C_2: N_o–150	–55.99	6.4	27.86	5.0	35.29	30
	C_3: N_o–75	–41.07		32.30		34.25	
D. Silane	D_1: S_o–100	–39.20		31.68		32.21	
	D_2: S_o–50	–46.85	2.3	34.70	4.8	34.53	58
	D_3: S_o	–50.04		28.17		35.61	
E. Settling time	E_1: t_o	–51.52		30.52		34.06	
	E_2: t_o+8	–40.54	2.3	32.87	–	33.99	–
	E_3: t_o+16	–44.03		31.16		34.30	
F. Cleaning method	F_1: *None*	–45.56		27.04		33.81	
	F_2: CM_2	–41.58	–	33.67	6.8	34.10	–
	F_3: CM_3	–48.95		33.85		34.44	
		–45.36		31.52		34.12	

Silane flow rate has a moderate effect on all characteristics. The best response for thickness variation is when the flow rate is reduced by 50 sccm. This may also give a small reduction in the number of defects. Deposition rate is slightly reduced by a reduction in silane flow rate.

By an 8 minutes increase of the settling time, surface defects may be improved by 10 dB. A further increase of settling time gives an increased number of defects as compared with the 8 minutes level. The thickness variation performance is also at its best as the settling time is increased by 8 minutes. However the performance change is within the standard deviation of the error. Deposition rate is not affected by settling time.

Cleaning method has no effect on deposition rate. The effect on surface defects is within the standard deviation of the error. Thickness variation may give 6 dB improvement by application of a cleaning method irrespective of which one.

Table 48 Predictions using the additive model.

Factor	Starting condition				Optimum condition			
	Setting	Contribution ° (dB)			Setting	Contribution ° (dB)		
		Surface defects	Thick-ness	Depo-sition rate.		Surface defects	Thick-ness	Depo-sition rate.
A*	A_2	–4.74	3.39	0.01	A_1	21.13	3.60	–5.36
B	B_2	–2.08	0.00	0.66	B_2	–2.08	0.00	0.66
C	C_1	6.33	2.87	–1.31	C_1	6.33	2.87	–1.31
D	D_3	–4.68	–3.35	1.49	D_3	–4.68	–3.37	1.49
E*	E_1	–6.16	0.00	0.00	E_2	4.82	0.00	0.00
F*	F_1	0.00	–4.48	0.00	F_2	0.00	2.15	0.00
Overall mean		–45.36	31.52	34.12		–45.36	31.52	34.12
Total		–56.69	29.95	34.97		–19.84	36.79	29.60

*Indicates the factors whose levels are changed from the starting to optimum conditions.

° By *contribution* we mean the deviation from the overall mean caused by the particular factor level.

The optimum setting of settling time and cleaning method is obvious from these observations. Level 2 is chosen. By the choice of appropriate levels for the remaining factors a trade off between productivity and quality loss has to be made. It was decided to take the quality loss benefit of the temperature factor. Level 1 was then chosen for temperature. The remaining factors were kept at their starting level. The final choice was then $A_1B_2C_1D_3E_2F_2$.

On the basis of these factor levels a performance prediction was made for the optimum setting. The prediction is shown in Table 48. The last line of that table gives the actual predictions while the other lines summarizes the contributions of the individual factors. In accordance with the Taguchi procedure a verification run was made. The results from this verification run are given in Table 49. The agreement achieved is good. We can thus conclude that the additive model assumed is justified. One remark has to be added at this point. The optimum settings as discussed above do not consider the target value of the thickness characteristics.

Table 49 Results of verification experiment

		Starting Conditions	Optimum condition	Improvement
Surface defects	rms	600/cm^2	7/cm^2	
	η	–55.6 dB	–16.9 dB	38.7 dB
Thickness	std. dev.*	0.028	0.013	
	η'	31.1 dB	37.7 dB	6.6 dB
Deposition rate	rate	60 Å/min	35 Å/min	
	η''	35.6 dBam	30.9 dBam	–4.7 dBam

*Standard deviation of thickness is expressed as a fraction of the mean thickness.

For instance the average response of thickness for A_1, A_2 and A_3 is 4151, 3060 and 3770 respectively. Considering that the target value is 3600 and that the difference in deposition rate effect between temperature and pressure is substantial the trade off may come out completely different. Taking this into account the optimum choice may well be A_2 and B_1.

The data for the surface defect characteristics were also analyzed as ordered categorical data. The data were assessed in classes as shown in Table 50.

Table 50 Categories used in ordered categorical data analysis of defects.

Category number	Observation category (defects)	Cumulative category (defects)
I	0 – 3	0 – 3
II	4 – 30	0 – 30
II	31 – 300	0 – 300
IV	301 – 1000	0 – 1000
V	> 1000	$0 - \infty$

The analysis used was accumulating data. For that reason cumulative classes were also assigned. Smaller is better or larger is better type characteristics may be treated by accumulation towards either end of the range of the ordered classes, (80). The last cumulative class will contain all observations. Thus this cumulative class is of no further interest.

Table 51 Categorized data for surface defects.

Expt. No.	Number of observations by categories					Number of observations by cumulative categories				
	I	II	III	IV	V	I	II	III	IV	V
1	9	0	0	0	0	9	9	9	9	9
2	5	2	2	0	0	5	7	9	9	9
3	1	0	6	2	0	1	1	7	9	9
4	0	8	1	0	0	0	8	9	9	9
5	0	1	0	4	4	0	1	1	4	9
6	1	0	4	1	3	1	1	5	6	9
7	0	1	1	4	3	0	1	2	6	9
8	3	0	2	1	3	3	3	5	6	9
9	0	0	0	4	5	0	0	0	4	9
10	9	0	0	0	0	9	9	9	9	9
11	8	1	0	0	0	8	9	9	9	9
12	2	3	3	0	1	2	5	8	8	9
13	4	2	2	1	0	4	6	8	9	9
14	2	3	4	0	0	2	5	9	9	9
15	0	1	1	1	6	0	1	2	3	9
16	3	4	2	0	0	3	7	9	9	9
17	2	1	0	2	4	2	3	3	5	9
18	0	0	0	2	7	0	0	0	2	9
Total	49	27	28	22	36	49	76	104	126	162

The observations for surface defects presented in terms of the above classes may be seen in Table 51. When analyzing data like this, work is with the cumulative classes only. In the present case classes I through IV will be looked at. Each of the classes is analyzed as if it had been a single characteristic. Hence we get a response table for each class. Notice that the response table for the last class (V) is trivial.

The response table calculated in this way is found in Table 52. There is also a response table where the responses have been normalized to the trivial fifth class, giving probabilities. The normalized response tables have been plotted in Figure 30.

Table 52 Categorized data for surface defects.

Factor	Level	Number of observations by cumulative categories.					Probabilities for the cumulative categories.				
		I	II	III	IV	V	I	II	III	IV	V
A. Temperature	A_1: T_o–25	34	40	51	53	54	0.63	0.74	0.94	0.98	1.0
	A_2: T_o	7	22	34	41	54	0.13	0.41	0.63	0.76	1.0
	A_3: T_o+25	8	14	19	32	54	0.15	0.26	0.35	0.59	1.0
B. Pressure	B_1: P_o–200	25	40	46	51	54	0.46	0.74	0.85	0.94	1.0
	B_2: P_o	20	28	36	43	54	0.37	0.52	0.67	0.80	1.0
	B_3: P_o+200	4	8	22	32	54	0.07	0.15	0.41	0.59	1.0
C. Nitrogen	C_1: N_o	19	30	32	39	54	0.35	0.56	0.59	0.72	1.0
	C_2: N_o–150	11	20	28	39	54	0.20	0.37	0.52	0.72	1.0
	C_3: N_o–75	19	26	44	48	54	0.35	0.48	0.81	0.89	1.0
D. Silane	D_1: S_o–100	20	25	34	41	54	0.37	0.46	0.63	0.76	1.0
	D_2: S_o–50	13	31	42	44	54	0.24	0.57	0.78	0.81	1.0
	D_3: S_o	16	20	28	41	54	0.30	0.37	0.52	0.76	1.0
E. Settling time	E_1: t_o	21	27	38	43	54	0.39	0.50	0.70	0.80	1.0
	E_2: t_o+8	16	29	36	42	54	0.30	0.54	0.67	0.78	1.0
	E_3: t_o+16	12	20	30	41	54	0.22	0.37	0.56	0.76	1.0
F. Cleaning method	F_1: *None*	21	23	26	34	54	0.39	0.43	0.48	0.63	1.0
	F_2: CM_2	21	30	40	46	54	0.39	0.56	0.74	0.85	1.0
	F_3: CM_2	7	23	38	46	54	0.13	0.43	0.70	0.85	1.0

The optimum choice of factor levels from this analysis is made on the basis of these plots. The choices A_1,B_1 and F_2 are obvious. There is no strong preference for choices of levels for factors C and D. C_3 and D_2 may be argued for. For the factor E the levels E_1 and E_2 are equally good.

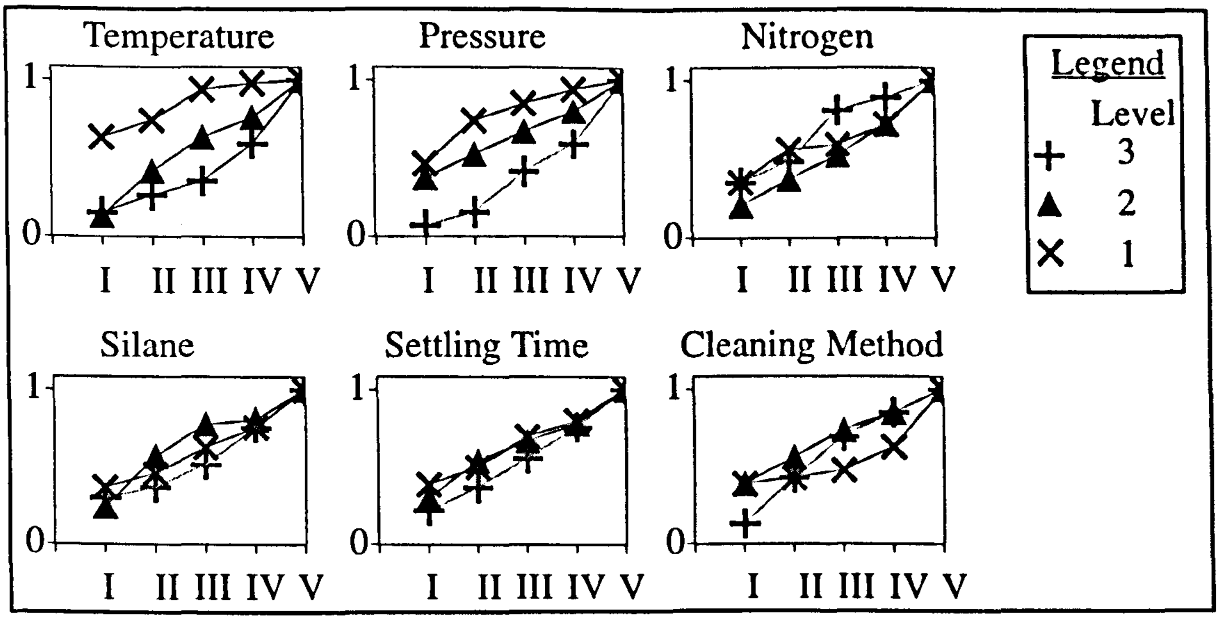


Figure 30 Line plots of the factor effects for the categorized surface defect data.

When making predictions for 0. – 1. data, like the probability data above, Taguchi recommends the use of the omega–transform, (80). The omega–transform is expressed as: $\omega(p) = 10\log_{10}\frac{p}{1-p}$ where p is a probability value. The prediction is made for each column of cumulative probability categories. For the optimum setting $A_1B_2C_1D_3E_2F_2$, using the values in Table 52 we get:

$$
\begin{aligned}
\omega_{A_1B_2C_1D_3E_2F_2} &= \omega_{\mu(I)} + \left(\omega_{A_1(I)} - \omega_{\mu(I)}\right) + \left(\omega_{B_2(I)} - \omega_{\mu(I)}\right) + \left(\omega_{F_2(I)} - \right. \\
&= \omega(0.30) + (\omega(0.63) - \omega(0.30)) + (\omega(0.37) - \omega(0.30)) + (\omega(0.39) - \\
&= -3,68 + (2.31 + 3.68) + (-2.31 + 3.68) + (-1.94 + 3.68) \\
&= 5.42\ dB
\end{aligned}
$$

Table 53 Predicted probabilities for the cumulative classes.

Control Factor Setting	Number of observations by cumulative categories.					Probabilities for the cumulative categories.				
	I	II	III	IV	V	I	II	III	IV	V
Optimum $A_1B_2C_1D_3E_2F_2$	5.42	6.98	14.53	19.45	∞	0.78	0.83	0.97	0.99	1.0
Starting $A_2B_2C_1D_3E_1F_1$	−3.68	−1.41	0.04	2.34	∞	0.23	0.42	0.50	0.63	1.0

The transform value 5.42 dB corresponds to a probability $p = 0.78$. The remaining columns are predicted in the same way. The prediction says that with a probability of 0.78 we will get 3 surface defects or less, with a probability of 0.83 we will get 30 surface defects or less. In the same way we find that in the starting condition we will get 1000 or less surface defects with a probability of 0.63. These predicted results are plotted in Figure 31. The outcome of the verification experiment as shown in Table 49 agrees well with these results.

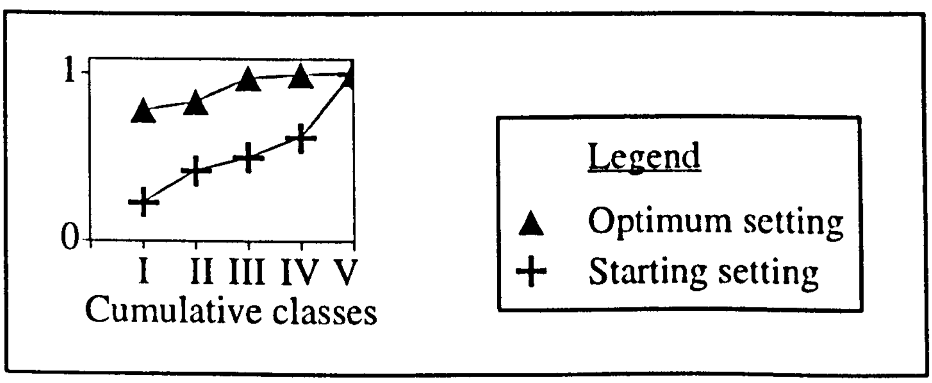


Figure 31 Predicted probabilities for the cumulative classes.

13.2. Appendix B . Information bounds.

Consider a discrete valued stochastic variable X with the value space $\{x_1, x_2, \ldots, x_k, \ldots, x_n\}$. A discrete distribution $\{p_1, p_2, \ldots, p_k, \ldots, p_n\}$ is connected to the value space. Assume that x_i is ordered such that the p_i an increasing series. If all values, x_i, were equally probable all p_i would be equal to 1/n.

The information content in the variable X may be calculated as :

$$E_d = -\sum_{i=1}^{n}(p_i \ln p_i)$$

We would like to show that this information is bounded upwards by:

$$E_B = -\sum_{i=1}^{n}\left(\frac{1}{n}\ln p_i\right)$$

Consider the difference:

$$E_B - E_d = -\sum_{i=1}^{n}\left(\left(\frac{1}{n} - p_i\right)\ln p_i\right)$$

Let k be the highest index i for which $p_i \leq 1/n$, then :

$$E_B - E_d = -\sum_{i=1}^{k}\left(\left(\frac{1}{n} - p_i\right)\ln p_i\right) - \sum_{i=k+1}^{n}\left(\left(\frac{1}{n} - p_i\right)\ln p_i\right)$$

Now find two small numbers δ_1 and δ_2 such that

$$E_B - E_d \geq -\sum_{i=1}^{k}\left(\left(\frac{1}{n} - p_i\right)\ln\left(\frac{1}{n} - \delta_1\right)\right) - \sum_{i=k+1}^{n}\left(\left(\frac{1}{n} - p_i\right)\ln\left(\frac{1}{n} + \delta_2\right)\right)$$

If p_k is truly less than 1/n we have

$$\sum_{i=1}^{k}\left(\frac{1}{n} - p_i\right) = m \quad \text{and} \quad \sum_{i=k+1}^{n}\left(\frac{1}{n} - p_i\right) = -m$$

Hence

$$E_B - E_d \geq -m\ln\left(\frac{1}{n} - \delta_1\right) + m\ln\left(\frac{1}{n} + \delta_2\right)$$

The two small numbers δ_1 and δ_2 may be chosen arbitrarily small and δ_1 may even be zero. As ln(z) is monotonically increasing and negative in the interval $0.0 < z < 1.0$ we can conclude that $E_B - E_d$ is greater than zero. In case p_k is identically equal to 1/n the summations above would still be of the same magnitude and of opposite signs. Thus the inequality is still true.

Finally we note that E_d is maximum as all p_i is equal to 1/n. See reference (66). In that case $E_B - E_d$ is equal to zero. Accordingly it has been shown that E_B is an upper bound for E_d.

13.3. Appendix C . Complexity and information.

Table 54 Raw data for complexity versus information relations

N_i,N_t,N_p	N_i,N_t,N_p	N_i,N_t,N_p	C	ln(C)	N	E_d	ln(9C/N)	E_d/ln(9C/N)
1	1	5	1.71	0.54	7	0.80	0.79	1.01*
1	2	4	2.00	0.69	7	0.96	0.94	1.01*
1	3	3	2.08	0.73	7	1.00	0.98	1.02
2	2	3	2.28	0.83	7	1.08	1.08	1.00
1	1	6	1.82	0.60	8	0.74	0.72	1.03*
1	2	5	2.15	0.77	8	0.90	0.89	1.02*
1	3	4	2.29	0.83	8	0.97	0.95	1.03
2	2	4	2.52	0.92	8	1.04	1.04	1.00*
2	3	3	2.62	0.96	8	1.08	1.08	1.00
1	1	7	1.91	0.65	9	0.68		1.05*
1	2	6	2.29	0.83	9	0.85		1.02*
1	3	5	2.47	0.90	9	0.94		1.04*
1	4	4	2.52	0.92	9	0.96		1..04*
2	2	5	2.71	1.00	9	1.00		1.00*
2	3	4	2.88	1.06	9	1.06		1.00
3	3	3	3.00	1.10	9	1.10		1.00
1	1	8	2.00	0.69	10	0.64	0.59	1.09*
1	2	7	2.41	0.88	10	0.80	0.77	1.04*
1	3	6	2.62	0.96	10	0.90	0.86	1.05*
1	4	5	2.71	1.00	10	0.94	0.89	1.06
2	2	6	2.88	1.06	10	0.95	0.95	1.00*
2	3	5	2.92	1.13	10	1.03	1.03	1.00*
2	4	4	3.04	1.16	10	1.05	1.05	1.00
3	3	4	3.30	1.19	10	1.09	1.09	1.00
1	1	9	2.08	0.73	11	0.60	0.53	1.13*
1	2	8	2.52	0.92	11	0.76	0.72	1.05*
1	3	7	2.76	1.01	11	0.86	0..81	1.06*
1	4	6	2.88	1.06	11	0.92	0.86	1.07
1	5	5	2.92	1.07	11	0.93	0.87	1.07
2	2	7	3.04	1.11	11	0.91	0.91	1.00*
2	3	6	3.30	1.19	11	0.99	0.99	1.00*
2	4	5	3.42	1.23	11	1.04	1.04	1.01
3	3	5	3.56	1.27	11	1.07	1.07	1.00
3	4	4	3.63	1.29	11	1.09	1.09	1.00
1	1	10	2.15	0.77	12	0.57	0.48	1.18
1	2	9	2.62	0.96	12	0.72	0.68	1.07
1	3	8	2.88	1.06	12	0.82	0.77	1.07

N_i,N_t,N_p	N_i,N_t,N_p	N_i,N_t,N_p	C	ln(C)	N	E_d	ln(9C/N)	E_d/ln(9C/N)
1	4	7	3.04	1.11	12	0.89	0.82	1.08
1	5	6	3.11	1.13	12	0.92	0.85	1.09
2	2	8	3.17	1.16	12	0.87	0.87	1.00
2	3	7	3.48	1.25	12	0.96	0.96	1.00
2	4	6	3.63	1.29	12	1.01	1.00	1.01
2	5	5	3.68	1.30	12	1.03	1.02	1.01
3	3	6	3.78	1.33	12	1.04	1.04	1.00
3	4	5	3.91	1.36	12	1.08	1.08	1.00
4	4	4	4.00	1.39	12	1.10	1.10	1.00
1	1	13	2.35	0.85	15	0.49	0.34	1.41*
1	2	12	2.88	1.06	15	0.63	0.55	1.14*
1	3	11	3.21	1.17	15	0.73	0.65	1.11*
1	4	10	3.42	1.23	15	0.80	0.72	1.12*
2	2	11	3.53	1.26	15	0.76	0.75	1.02*
1	5	9	3.46	1.27	15	0.85	0.76	1.13**
1	6	8	3.63	1.29	15	0.88	0.78	1.13**
1	7	7	3.66	1.30	15	0.89	0.79	1.13
2	3	10	3.91	1.36	15	0.86	0.85	1.01*
2	4	9	4.16	1.43	15	0.93	0.91	1.01*
2	5	8	4.31	1.46	15	0.97	0.95	1.02
3	3	9	4.33	1.46	15	0.95	0.95	1.00*
2	6	7	4.38	1.48	15	0.99	0.97	1.03
3	4	8	4.57	1.52	15	1.01	1.01	1.00*
3	5	7	4.72	1.55	15	1.04	1.04	1.00
3	6	6	4.76	1.56	15	1.05	1.05	1.00*
4	4	7	4.82	1.57	15	1.06	1.06	1.00*
4	5	6	4.93	1.60	15	1.09	1.09	1.00
5	5	5	5.00	1.61	15	1.10	1.10	1.00
1	1	16	2.52	0.92	18	0.43	0.23	1.84*
1	2	15	3.11	1.13	18	0.56	0.44	1.26*
1	3	14	3.48	1.25	18	0.65	0.55	1.18*
1	4	13	3.73	1.32	18	0.73	0.62	1.17*
2	2	14	3.83	1.34	18	0.68	0.65	1.05*
1	5	12	3.91	1.36	18	0.79	0.67	1.17*
1	6	11	4.04	1.40	18	0.83	0.70	1.18**
1	7	10	4.12	1.42	18	0.85	0.72	1.18**
1	8	9	4.16	1.43	18	0.87	0.73	1.18**
2	3	13	4.27	1.45	18	0.78	0.76	1.02*
2	4	12	4.58	1.52	18	0.85	0.83	1.02*
3	3	12	4.76	1.56	18	0.87	0.87	1.00*

N_i,N_t,N_p	N_i,N_t,N_p	N_i,N_t,N_p	C	ln(C)	N	E_d	ln(9C/N)	E_d/ln(9C/N)
2	5	11	4.79	1.57	18	0.90	0.87	1.03*
2	6	10	4.93	1.60	18	0.94	0.90	1.04
2	7	9	5.01	1.61	18	0.96	0.92	1.04
2	8	8	5.04	1.62	18	0.96	0.92	1.04
3	4	11	5.09	1.63	18	0.93	0.93	1.00*
3	5	10	5.31	1.67	18	0.98	0.98	1.00
4	4	10	5.43	1.69	18	1.00	1.00	1.00*
3	6	9	5.45	1.70	18	1.01	1.00	1.01
3	7	8	5.52	1.71	18	1.03	1.01	1.01
4	5	9	5.65	1.73	18	1.04	1.04	1.00
4	6	8	5.77	1.75	18	1.06	1.06	1.00
4	7	7	5.81	1.76	18	1.07	1.07	1.00
5	5	8	5.85	1.77	18	1.07	1.07	1.00
5	6	7	5.94	1.78	18	1.09	1.09	1.00
6	6	6	6	1.79	18	1.10	1.10	1.00

* signifies combinations of N_i,N_t and N_p that are not possible to appear i the creation of a real system.

** signifies systems that are very highly interconnected and hardly an example of a realistic system.

The heading of column 1, 2 and 3 in Table 54 is meant to illustrate that the figures indicate either of the numbers N_i,N_t and N_p whichever may be appropriate.

Table 55 Limiting specifications of systems N_i,N_t and N_p.

N_p	N_{tn}	N_{tM}	N_{im}	N_{iM}	N_m	N_M
2	1	2	1	1	4	5
3	1	3	2	3	6	9
4	1	4	3	6	8	14
5	1	5	4	10	10	20
6	1	6	5	15	12	27
7	1	7	6	21	14	35
8	1	8	7	28	16	44
9	1	9	8	36	18	54
10	1	10	9	45	20	65
11	1	11	10	55	22	77
12	1	12	11	66	14	90
13	1	13	12	78	26	104
14	1	14	13	91	28	119
15	1	15	14	105	30	135
16	1	16	15	120	32	152
17	1	17	16	136	34	170
18	1	18	17	153	36	189
19	1	19	18	171	38	209
20	1	20	19	190	40	230

N_p number of parts

N_{im} minimum number of interfaces in a realistic system

N_{iM} maximum number of interfaces in a realistic system

N_{tm} minimum number of types of parts

N_{tM} maximum number of types of parts

N_m minimum number of components, $N_p + N_{im} + N_{tm}$

N_M maximum number of components, $N_p + N_{iM} + N_{tM}$

13.4. Appendix D . Simulation data for Information distance SPC

Table 56 ARL–curves for information distance SPC charts.

Class count	Sample size	Tail border	Process mean offset	ARL	Info distance mean
3	3	1.232	0.0	333.3	0.337
			0.10	114.8	0.368
			0.25	66.0	0.412
			0.42	39.2	0.458
			0.62	25.3	0.503
			0.84	16.9	0.550
			1.10	12.1	0.600
			1.39	8.9	0.642
			1.72	6.9	0.687
			2.09	5.4	0.731
			2.52	4.4	0.773
			3.00	3.7	0.813
3	4	1.035	0.0	333.3	0.293
			0.10	196.1	0.324
			0.25	139.9	0.367
			0.42	94.5	0.412
			0.62	64.2	0.458
			0.84	39.7	0.505
			1.10	26.7	0.553
			1.39	18.0	0.600
			1.72	12.9	0.647
			2.09	9.5	0.693
			2.52	7.2	0.738
			3.00	5.7	781

Class count	Sample size	Tail border	Process mean offset	ARL	Info distance mean
3	5	1.001	0.0	333.3	0.267
			0.10	199.2	0.297
			0.25	107.8	0.339
			0.42	58.5	0.384
			0.62	33.4	0.431
			0.84	20.5	0.478
			1.10	13.3	0.527
			1.39	9.2	0.575
			1.72	6.7	0.624
			2.09	5.1	0.671
			2.52	4.0	0.717
			3.00	3.3	0.761
3	6	0.902	0.0	333.3	0.249
			0.10	202.8	0.279
			0.25	104.0	0.321
			0.42	59.1	0.365
			0.62	34.8	0.412
			0.84	21.0	0.460
			1.10	13.9	0.509
			1.39	9.5	0.558
			1.72	7.0	0.608
			2.09	5.3	0.656
			2.52	4.2	0.703
			3.00	3.4	0.748

Class count	Sample size	Tail border	Process mean offset	ARL	Info distance mean
3	7	0.843	0.0	333.3	0.236
			0.10	191.6	0.265
			0.25	95.01	0.308
			0.42	50.5	0.352
			0.62	28.7	0.398
			0.84	16.8	0.447
			1.10	10.9	0.496
			1.39	7.4	0.546
			1.72	5.4	0.596
			2.09	4.0	0.645
			2.52	3.2	0.692
			3.00	2.6	0.739
3	8	0.821	0.0	333.3	0.236
			0.10	206.2	0.256
			0.25	121.5	0.297
			0.42	64.5	0.342
			0.62	33.4	0.388
			0.84	19.9	0.437
			1.10	13.2	0.487
			1.39	8.7	0.537
			1.72	6.2	0.587
			2.09	4.6	0.637
			2.52	3.6	0.685
			3.00	2.9	0.732

Class count	Sample size	Tail border	Process mean offset	ARL	Info distance mean
3	9	0.766	0.0	333.3	0.219
			0.10	177.3	0.248
			0.25	92.1	0.289
			0.42	46.9	0.334
			0.62	26.9	0.381
			0.84	15.8	0.429
			1.10	10.2	0.479
			1.39	6.8	0.530
			1.72	4.9	0.580
			2.09	3.6	0.630
			2.52	2.9	0.679
			3.00	2.3	0.726
4	3	1.090	0.0	333.3	0.276
			0.10	190.1	0.302
			0.25	136.4	0.338
			0.42	85.5	0.376
			0.62	56.6	0.414
			0.84	37.3	0.455
			1.10	25.0	0.494
			1.39	17.3	0.536
			1.72	12.5	0.577
			2.09	9.3	0.617
			2.52	7.1	0.657
			3.00	5.7	0.696

Class count	Sample size	Tail border	Process mean offset	ARL	Info distance mean
4	4	0.884	0.0	333.3	0.231
			0.10	200.0	0.256
			0.25	177.1	0.291
			0.42	66.9	0.327
			0.62	41.1	0.366
			0.84	27.4	0.406
			1.10	18.6	0.447
			1.39	12.7	0.489
			1.72	9.2	0.531
			2.09	6.9	0.574
			2.52	5.3	0.616
			3.00	4.2	0.658
4	5	0.794	0.0	333.3	0.204
			0.10	229.4	0.228
			0.25	136.1	0.261
			0.42	76.2	0.298
			0.62	44.4	0.336
			0.84	27.3	0.375
			1.10	17.1	0.418
			1.39	11.4	0.460
			1.72	8.1	0.504
			2.09	5.9	0.548
			2.52	4.6	0.548
			3.00	3.7	0.592

Class count	Sample size	Tail border	Process mean offset	ARL	Info distance mean
4	6	0.748	0.0	333.3	0.185
			0.10	226.8	0.209
			0.25	133.9	0.242
			0.42	73.0	0.277
			0.62	41.2	0.315
			0.84	24.1	0.355
			1.10	14.9	0.397
			1.39	9.6	0.440
			1.72	6.6	0.485
			2.09	4.8	0.530
			2.52	3.7	0.575
			3.00	2.9	0.619
4	7	0.679	0.0	333.3	0.172
			0.10	193.4	0.195
			0.25	116.7	0.228
			0.42	64.4	0.263
			0.62	35.7	0.301
			0.84	20.5	0.341
			1.10	12.4	0.383
			1.39	8.1	0.427
			1.72	5.6	0.472
			2.09	4,1	0.517
			2.52	3.2	0.562
			3.00	2.5	0.608

Class count	Sample size	Tail border	Process mean offset	ARL	Info distance mean
4	8	0.647	0.0	333.3	0.162
			0.10	211.4	0.185
			0.25	125.3	0.217
			0.42	75.4	0.252
			0.62	37.6	0.290
			0.84	20.2	0.330
			1.10	12.5	0.372
			1.39	8.0	0.416
			1.72	5.5	0.461
			2.09	4.1	0.507
			2.52	3.1	0.553
			3.00	2.5	0.599
4	9	0.607	0.0	333.3	0.154
			0.10	194.2	0.177
			0.25	99.6	0.209
			0.42	54.7	0.243
			0.62	29.5	0.281
			0.84	17.0	0.321
			1.10	10.6	0.364
			1.39	6.9	0.408
			1.72	4.7	0.453
			2.09	3.5	0.500
			2.52	2.7	0.546
			3.00	2.2	0.592

Class count	Sample size	Tail border	Process mean offset	ARL	Info distance mean
5	3	0.987	0.0	333.3	0.238
			0.10	243.3	0.259
			0.25	205.8	0.290
			0.42	150.6	0.322
			0.62	105.7	0.355
			0.84	73.6	0.389
			1.10	50.3	0.423
			1.39	33.3	0.460
			1.72	22.8	0.496
			2.09	15.9	0.533
			2.52	11.5	0.570
			3.00	8.7	0.606
5	4	0.781	0.0	333.3	0.193
			0.10	199.6	0.213
			0.25	147.1	0.242
			0.42	74.0	0.272
			0.62	46.6	0.304
			0.84	29.6	0.337
			1.10	19.6	0.372
			1.39	13.0	0.409
			1.72	9.1	0.447
			2.09	6.6	0.485
			2.52	5.0	0.524
			3.00	4.0	0.563

Class count	Sample size	Tail border	Process mean offset	ARL	Info distance mean
5	5	0.680	0.0	333.3	0.165
			0.10	243.3	0.184
			0.25	166,7	0.212
			0.42	97.8	0.241
			0.62	60.3	0.273
			0.84	37.9	0.306
			1.10	23.9	0.341
			1.39	15.7	0.378
			1.72	10.7	0.416
			2.09	7.6	0.456
			2.52	5.7	0.496
			3.00	4.4	0.537
5	6	0.638	0.0	333.3	0.147
			0.10	263.9	0.165
			0.25	161.8	0.192
			0.42	98.4	0.220
			0.62	59.2	0.251
			0.84	35.2	0.284
			1.10	21.4	0.319
			1.39	13.3	0.357
			1.72	8.9	0.396
			2.09	6.3	0.436
			2.52	4.6	0.477
			3.00	3.6	0.519

Class count	Sample size	Tail border	Process mean offset	ARL	Info distance mean
5	7	0.572	0.0	333.3	0.133
			0.10	208.8	0.151
			0.25	140.1	0.177
			0.42	81.5	0.205
			0.62	46.4	0.236
			0.84	26.2	0.269
			1.10	15.9	0.304
			1.39	9.8	0.341
			1.72	6.5	0.381
			2.09	4.7	0.421
			2.52	3.5	0.463
			3.00	2.7	0.506
5	8	0.542	0.0	333.3	0.123
			0.10	316.5	0.141
			0.25	170.6	0.166
			0.42	104.8	0.194
			0.62	49.6	0.224
			0.84	26.7	0.257
			1.10	16.4	0.292
			1.39	9.6	0.330
			1.72	6.2	0.369
			2.09	4.4	0.411
			2.52	3.3	0.453
			3.00	2.5	0.496

Class count	Sample size	Tail border	Process mean offset	ARL	Info distance mean
5	9	0.502	0.0	333.3	0.115
			0.10	226.2	0.132
			0.25	120.8	0.157
			0.42	64.1	0.185
			0.62	36.9	0.215
			0.84	20.1	0.248
			1.10	12.2	0.283
			1.39	7.6	0.321
			1.72	5.1	0.361
			2.09	3.6	0.402
			2.52	2.7	0.445
			3.00	2.2	0.488
6	3	0.904	0.0	333.3	0.214
			0.10	261.8	0.232
			0.25	235.3	0.258
			0.42	188.0	0.286
			0.62	158.2	0.313
			0.84	126.3	0.343
			1.10	89.4	0.373
			1.39	59.0	0.405
			1.72	39.8	0.437
			2.09	27.2	0.470
			2.52	18.4	0.504
			3.00	13.2	0.538

Class count	Sample size	Tail border	Process mean offset	ARL	Info distance mean
6	4	0.684	0.0	333.3	0.169
			0.10	196.1	0.185
			0.25	150.4	0.209
			0.42	96.1	0.235
			0.62	64.3	0.262
			0.84	45.3	0.291
			1.10	30.7	0.321
			1.39	20.5	0.352
			1.72	13.9	0.385
			2.09	9,8	0.419
			2.52	7.1	0.455
			3.00	5.3	0.491
6	5	0.602	0.0	333.3	0.141
			0.10	248.8	0.157
			0.25	166.9	0.179
			0.42	100.9	0.204
			0.62	59.2	0.230
			0.84	36.3	0.258
			1.10	22.8	0.288
			1.39	14.5	0.319
			1.72	9.6	0.352
			2.09	6.7	0.388
			2.52	4.9	0.424
			3.00	3.7	0.462

Class count	Sample size	Tail border	Process mean offset	ARL	Info distance mean
6	6	0.553	0.0	333.3	0.123
			0.10	273.2	0.137
			0.25	174.8	0.159
			0.42	118.2	0.183
			0.62	73.2	0.208
			0.84	44.1	0.236
			1.10	26.2	0.265
			1.39	15.8	0.297
			1.72	10.2	0.331
			2.09	7.0	0.366
			2.52	4.9	0.403
			3.00	3.7	0.442
6	7	0.490	0.0	333.3	0.109
			0.10	219.8	0.123
			0.25	142.0	0.145
			0.42	90.2	0.168
			0.62	54.5	0.192
			0.84	30.8	0.220
			1.10	18.2	0.249
			1.39	11.3	0.281
			1.72	7.3	0.315
			2.09	5.1	0.351
			2.52	3.7	0.388
			3.00	2.8	0.428

Class count	Sample size	Tail border	Process mean offset	ARL	Info distance mean
6	8	0.461	0.0	333.3	0.099
			0.10	316.5	0.113
			0.25	160.3	0.133
			0.42	109.5	0.156
			0.62	57.7	0.181
			0.84	29.6	0.208
			1.10	18.3	0.237
			1.39	10.9	0.269
			1.72	7.0	0.303
			2.09	4.8	0.339
			2.52	3.4	0.377
			3.00	2.6	0.417
6	9	0.422	0.0	333.3	0.091
			0.10	216.0	0.105
			0.25	124.4	0.125
			0.42	66.8	0.147
			0.62	39.3	0.171
			0.84	22.6	0.198
			1.10	13.4	0.228
			1.39	8.3	0.259
			1.72	5.5	0.293
			2.09	3.9	0.330
			2.52	2.9	0.368
			3.00	2.2	0.408

Class count	Sample size	Tail border	Process mean offset	ARL	Info distance mean
7	5	0.538	0.0	333.3	0.127
			0.10	238.1	0.140
			0.25	159.7	0.158
			0.42	96.8	0.179
			0.62	65.4	0.201
			0.84	41.7	0.225
			1.10	27.1	0.250
			1.39	17.2	0.277
			1.72	11.3	0.306
			2.09	7.8	0.337
			2.52	5.6	0.369
			3.00	4.2	0.404
7	6	0.490	0.0	333.3	0.108
			0.10	264.6	0.120
			0.25	175.4	0.137
			0.42	118.7	0.158
			0.62	80.5	0.179
			0.84	47.4	0.202
			1.10	29.0	0.227
			1.39	17.6	0.254
			1.72	11.3	0.283
			2.09	7.6	0.314
			2.52	5.3	0.348
			3.00	3.9	0.382

Class count	Sample size	Tail border	Process mean offset	ARL	Info distance mean
7	7	0.431	0.0	333.3	0.095
			0.10	226.2	0.106
			0.25	150.2	0.123
			0.42	94.1	0.142
			0.62	60.4	0.163
			0.84	34.5	0.186
			1.10	20.9	0.211
			1.39	12.9	0.238
			1.72	8.3	0.267
			2.09	5.7	0.298
			2.52	4.1	0.331
			3.00	3.1	0.367
7	8	0.414	0.0	333.3	0.085
			0.10	383.1	0.096
			0.25	196.5	0.112
			0.42	146.4	0.131
			0.62	83.5	0.151
			0.84	42.8	0.174
			1.10	24.8	0.198
			1.39	14.2	0.225
			1.72	8.8	0.254
			2.09	5.8	0.286
			2.52	4.1	0.320
			3.00	3.0	0.355

Class count	Sample size	Tail border	Process mean offset	ARL	Info distance mean
7	9	0.366	0.0	333.3	0.077
			0.10	229.9	0.087
			0.25	135.7	0.103
			0.42	78.2	0.122
			0.62	46.5	0.142
			0.84	27.2	0.164
			1.10	16.2	0.188
			1.39	9.9	0.215
			1.72	6.4	0.244
			2.09	4.4	0.276
			2.52	3.2	0.310
			3.00	2.4	0.346
7	10	0.333	0.0	333.3	0.071
			0.10	207.9	0.081
			0.25	108.9	0.096
			0.42	61.2	0.114
			0.62	35.2	0.134
			0.84	20.5	0.156
			1.10	12.6	0.181
			1.39	7.8	0.207
			1.72	5.2	0.237
			2.09	3.6	0.268
			2.52	2.6	0.303
			3.00	2.0	0.339

Class count	Sample size	Tail border	Process mean offset	ARL	Info distance mean
7	11	0.296	0.0	333.3	0.066
			0.10	197.6	0.075
			0.25	93.2	0.091
			0.42	50.9	0.108
			0.62	28.2	0.128
			0.84	16.1	0.150
			1.10	9.6	0.174
			1.39	5.9	0.201
			1.72	4.0	0.230
			2.09	2.8	0.262
			2.52	2.1	0.296
			3.00	1.7	0.333
8	5	0.494	0.0	333.3	0.119
			0.10	267.4	0.129
			0.25	183.5	0.144
			0.42	113.0	0.162
			0.62	74.2	0.181
			0.84	51.0	0.201
			1.10	33.3	0.223
			1.39	21.5	0.246
			1.72	14.2	0.271
			2.09	9.6	0.299
			2.52	6.7	0.327
			3.00	4.9	0.358

Class count	Sample size	Tail border	Process mean offset	ARL	Info distance mean
8	6	0.445	0.0	333.3	0.100
			0.10	284.9	0.110
			0.25	198.0	0.124
			0.42	126.1	0.141
			0.62	88.6	0.159
			0.84	53.8	0.178
			1.10	33.7	0.200
			1.39	20.5	0.223
			1.72	13.1	0.248
			2.09	8.6	0.275
			2.52	5.9	0.304
			3.00	5.3	0.336
8	7	0.388	0.0	333.3	0.087
			0.10	235.3	0.096
			0.25	153.4	0.109
			0.42	99.3	0.125
			0.62	63.6	0.143
			0.84	39.6	0.162
			1.10	23.7	0.183
			1.39	14.9	0.206
			1.72	9.4	0.231
			2.09	6.3	0.258
			2.52	4.5	0.288
			3.00	3.3	0.320

Class count	Sample size	Tail border	Process mean offset	ARL	Info distance mean
8	8	0.370	0.0	333.3	0.077
			0.10	375.9	0.85
			0.25	187.6	0.098
			0.42	167.2	0.114
			0.62	90.4	0.131
			0.84	48.3	0.150
			1.10	30.9	0.170
			1.39	16.8	0.193
			1.72	10.4	0.218
			2.09	6.7	0.246
			2.52	4.6	0.275
			3.00	3.3	0.308
8	9	0.328	0.0	333.3	0.069
			0.10	256.4	0.077
			0.25	156.5	0.090
			0.42	99.4	0.105
			0.62	56.9	0.121
			0.84	34.7	0.140
			1.10	20.3	0.160
			1.39	12.1	0.183
			1.72	7.7	0.208
			2.09	5.1	0.236
			2.52	3.6	0.266
			3.00	2.6	0.298

Class count	Sample size	Tail border	Process mean offset	ARL	Info distance mean
8	10	0.293	0.0	333.3	0.063
			0.10	215.5	0.070
			0.25	116.7	0.083
			0.42	67.5	0.097
			0.62	39.4	0.114
			0.84	23.2	0.132
			1.10	14.3	0.152
			1.39	8.9	0.175
			1.72	5.8	0.200
			2.09	3.9	0.227
			2.52	2.8	0.258
			3.00	2.1	0.290
8	11	0.262	0.0	333.3	0.058
			0.10	204.9	0.065
			0.25	97.3	0.077
			0.42	55.5	0.091
			0.62	31.0	0.107
			0.84	18.1	0.126
			1.10	11.2	0.146
			1.39	7.0	0.168
			1.72	4.6	0.193
			2.09	3.2	0.221
			2.52	2.3	0.251
			3.00	1.8	0.284

Class count	Sample size	Tail border	Process mean offset	ARL	Info distance mean
9	5	0.444	0.0	333.3	0.115
			0.10	230.4	0.123
			0.25	153.6	0.135
			0.42	95.4	0.150
			0.62	65.4	0.166
			0.84	44.2	0.184
			1.10	30.1	0.203
			1.39	20.3	0.223
			1.72	13.7	0.246
			2.09	9.3	0.269
			2.52	6.7	0.295
			3.00	4.9	0.322
9	6	0.416	0.0	333.3	0.096
			0.10	260.4	0.103
			0.25	176.1	0.115
			0.42	144.1	0.129
			0.62	100.1	0.144
			0.84	66.8	0.161
			1.10	43.8	0.180
			1.39	27.2	0.200
			1.72	17.5	0.222
			2.09	11.2	0.245
			2.52	7.5	0.271
			3.00	5.2	0.299

Class count	Sample size	Tail border	Process mean offset	ARL	Info distance mean
9	7	0.355	0.0	333.3	0.083
			0.10	238.7	0.089
			0.25	155.5	0.100
			0.42	105.9	0.114
			0.62	68.8	0.128
			0.84	43.1	0.145
			1.10	27.9	0.163
			1.39	17.4	0.183
			1.72	11.0	0.204
			2.09	7.3	0.228
			2.52	5.0	0.254
			3.00	3.6	0.283
9	8	0.336	0.0	333.3	0.073
			0.10	350.9	0.079
			0.25	179.2	0.089
			0.42	157.7	0.102
			0.62	108.1	0.116
			0.84	56.1	0.132
			1.10	36.4	0.150
			1.39	19.9	0.170
			1.72	12.2	0.191
			2.09	8.1	0.215
			2.52	5.3	0.241
			3.00	3.6	0.270

Class count	Sample size	Tail border	Process mean offset	ARL	Info distance mean
9	9	0.296	0.0	333.3	0.065
			0.10	248.8	0.071
			0.25	157.5	0.081
			0.42	107.0	0.093
			0.62	62.0	0.107
			0.84	40.0	0.123
			1.10	23.4	0.140
			1.39	14.1	0.160
			1.72	8.8	0.181
			2.09	5.8	0.205
			2.52	4.0	0.231
			3.00	2.8	0.260
9	10	0.265	0.0	333.3	0.059
			0.10	237.5	0.074
			0.25	130.5	0.074
			0.42	78.1	0.886
			0.62	46.7	0.099
			0.84	27.5	–.115
			1.10	16.9	0.132
			1.39	10.5	0.151
			1.72	6.7	0.173
			2.09	4.5	0.197
			2.52	3.2	0.223
			3.00	2.4	0.252

Class count	Sample size	Tail border	Process mean offset	ARL	Info distance mean
9	11	0.239	0.0	333.3	0.054
			0.10	223.2	0.059
			0.25	108.6	0.068
			0.42	63.9	0.080
			0.62	36.9	0.093
			0.84	21.9	0.108
			1.10	13.5	0.125
			1.39	8.4	0.145
			1.72	5.5	0.166
			2.09	3.7	0.190
			2.52	2.7	0.216
			3.00	2.0	0.245
10	5	0.425	0.0	333.3	0.113
			0.10	268.1	0.119
			0.25	188.3	0.130
			0.42	119.0	0.142
			0.62	83.2	0.156
			0.84	58.0	0.172
			1.10	41.4	0.188
			1.39	27.7	0.206
			1.72	18.8	0.226
			2.09	12.5	0.247
			2.52	8.5	0.269
			3.00	6.1	0.294

Class count	Sample size	Tail border	Process mean offset	ARL	Info distance mean
10	6	0.396	0.0	333.3	0.095
			0.10	289.0	0.100
			0.25	217.4	0.109
			0.42	167.8	0.121
			0.62	125.2	0.134
			0.84	84.9	0.149
			1.10	58.9	0.165
			1.39	37.1	0.183
			1.72	24.9	0.2o2
			2.09	15.6	0.222
			2.52	10.2	0.245
			3.00	6.7	0.270
10	7	0.338	0.0	333.3	0.082
			0.10	304.9	0.086
			0.25	188.7	0.095
			0.42	133.5	0.106
			0.62	91.7	0.118
			0.84	59.0	0.132
			1.10	38.6	0.148
			1.39	24.1	0.165
			1.72	15.1	0.184
			2.09	9,7	0.205
			2.52	6.5	0.228
			3.00	4.5	0.253

Class count	Sample size	Tail border	Process mean offset	ARL	Info distance mean
10	8	0.306	0.0	333.3	0.072
			0.10	375.9	0.076
			0.25	183.5	0.084
			0.42	158.0	0.094
			0.62	102.4	0.106
			0.84	54.8	0.120
			1.10	38.6	0.135
			1.39	20.6	0.152
			1.72	13.2	0.171
			2.09	8.7	0.191
			2.52	5.7	0.215
			3.00	3.9	0.240
10	9	0.274	0.0	333.3	0.064
			0.10	263.9	0.068
			0.25	163.1	0.075
			0.42	118.9	0.085
			0.62	71.2	0.097
			0.84	47.8	0.110
			1.10	28.6	0.125
			1.39	17.2	0.142
			1.72	10.7	0.160
			2.09	6.9	0.181
			2.52	4.7	0.204
			3.00	3.3	0.230

Class count	Sample size	Tail border	Process mean offset	ARL	Info distance mean
10	10	0.249	0.0	333.3	0.058
			0.10	273.2	0.061
			0.25	161.0	0.068
			0.42	102.8	0.078
			0.62	62.4	0.089
			0.84	38.0	0.102
			1.10	22.7	0.117
			1.39	13.7	0.134
			1.72	8.6	0.152
			2.09	5.6	0.173
			2.52	3.9	0.196
			3.00	2.8	0.222
10	11	0.222	0.0	333.3	0.052
			0.10	259.7	0.056
			0.25	131.9	0.063
			0.42	77.6	0.072
			0.62	46.3	0.083
			0.84	27.8	0.096
			1.10	17.2	0.110
			1.39	10.5	0.127
			1.72	6.7	0.145
			2.09	4.5	0.166
			2.52	3.1	0.189
			3.00	2.3	0.215

Class count	Sample size	Tail border	Process mean offset	ARL	Info distance mean
11	5	0.408	0.0	333.3	0.114
			0.10	251.3	0.118
			0.25	184.2	0.126
			0.42	129.2	0.137
			0.62	89.9	0.149
			0.84	66.4	0.163
			1.10	48.4	0.177
			1.39	33.7	0.193
			1.72	23.3	0.210
			2.09	15.7	0.229
			2.52	10.6	0.249
			3.00	7.4	0.271
11	6	0.374	0.0	333.3	0.095
			0.10	300.3	0.099
			0.25	223.7	0.106
			0.42	180.8	0.116
			0.62	133.0	0.127
			0.84	94.0	0.140
			1.10	68.4	0.154
			1.39	44.3	0.169
			1.72	29.6	0.186
			2.09	19.1	0.205
			2.52	12.4	0.225
			3.00	8.0	0.247

Class count	Sample size	Tail border	Process mean offset	ARL	Info distance mean
11	7	0.321	0.0	333.3	0.0823
			0.10	286.5	0.0852
			0.25	187.6	0.092
			0.42	146.8	0.100
			0.62	102.0	0.111
			0.84	70.2	0.123
			1.10	47.3	0.137
			1.39	30.2	0.152
			1.72	19.2	0.169
			2.09	12.1	0.87
			2.52	8.0	0.207
			3.00	5.4	0.229
11	8	0.288	0.0	333.3	0.072
			0.10	325.7	0.075
			0.25	194.9	0.081
			0.42	165.6	0.089
			0.62	113.8	0.099
			0.84	62.8	0.111
			1.10	47.1	0.124
			1.39	25.5	0.139
			1.72	16.6	0.155
			2.09	10.5	0.173
			2.52	6.8	0.194
			3.00	4.6	0.216

Class count	Sample size	Tail border	Process mean offset	ARL	Info distance mean
11	9	0.269	0.0	333.3	0.064
			0.10	327.9	0.067
			0.25	239.2	0.072
			0.42	173.9	0.080
			0.62	118.3	0.089
			0.84	74.3	0.101
			1.10	47.9	0.114
			1.39	28.2	0.129
			1.72	16.8	0.145
			2.09	10.2	0.163
			2.52	6.6	0.183
			3.00	4.4	0.206
11	10	0.234	0.0	333.3	0.058
			0.10	267.4	0.060
			0.25	170.6	0.065
			0.42	117.1	0.073
			0.62	74.3	0.082
			0.84	46.2	0.093
			1.10	28	0.106
			1.39	17.8	0.120
			1.72	11.0	0.136
			2.09	6.9	0.154
			2.52	4.7	0.174
			3.00	3.2	0.197

Class count	Sample size	Tail border	Process mean offset	ARL	Info distance mean
11	11	0.216	0.0	333.3	0.053
			0.10	264.6	0.055
			0.25	168.9	0.059
			0.42	116.7	0.067
			0.62	73.4	0.076
			0.84	43.9	0.087
			1.10	26.7	0.099
			1.39	16.1	0.113
			1.72	9.8	0.129
			2.09	6.1	0.147
			2.52	4.1	0.167
			3.00	2.9	0.190

Table 57 $\bar{x}$–chart ARL numbers as a function of sample size for exponentially distributed process with distribution parameter $\lambda = 1.0$.

Sample size and tail limit / Mean offset	3 2.47	4 2.31	5 2.24	6 2.13	7 2.04	8 2.02	9 1.94	10 1.89	11 1.85
0.0	333.3	333.3	333.3	333.3	333.3	333.3	333.3	333.3	333.3
0.1	157.7	128.9	156.5	133.9	120.5	144.9	120.0	103.2	104.1
0.25	67.0	58.7	56.0	45.7	41.0	43.9	35.4	31.0	28.1
0.42	32.2	27.1	23.5	19.0	16.4	16.6	13.6	11.9	10.7
0.62	17.2	13.0	12.0	9.5	8.0	7.7	6.4	5.7	5.1
0.84	10.2	7.7	6.8	5.5	4.6	4.4	3.6	3.3	2.9
1.10	6.6	5.1	4.3	3.5	2.9	2.8	2.4	2.1	2.0
1.39	4.6	3.4	3.0	2.4	2.1	2.0	1.7	1.6	1.5
1.72	3.4	2.6	2.2	1.9	1.7	1.6	1.4	1.3	1.3
2.09	2.6	2.1	1.8	1.5	1.4	1.3	1.2	1.2	1.1
2.52	2.1	1.7	1.5	1.3	1.2	1.2	1.1	1.1	1.1
3.0	1.8	1.5	1.3	1.2	1.1	1.1	1.1	1.0	1.0

Table 58 R–chart ARL numbers as a function of sample size for exponentially distributed process with distribution parameter $\lambda = 1.0$.

Sample size and tail limit / Mean offset	3 6.49	4 6.95	5 7.16	6 7.37	7 7.51	8 7.65	9 7.85	10 8.02	11 8.07
0.6	333.3	333.3	333.3	333.3	333.3	333.3	333.3	333.3	333.3
0.1	184.2	179.9	167.5	163.9	155.8	151.5	162.3	161.8	154.8
0.25	92.0	88.0	75.1	73.3	67.7	65.1	65.1	65.7	65.0
0.42	48	44.7	39.4	36.0	33.4	31.0	31.6	32.0	29.6
0.62	27.6	24.7	21.3	19.6	17.9	16.7	16.6	16.3	15.2
0.84	17.1	14.8	12.6	11.2	10.2	9.5	9.3	9.1	8.4
1.10	11.3	9.4	7.9	7.1	6.4	6.0	5.8	5.5	5.2
1.39	7.8	6.5	5.4	4.8	4.3	4.0	3.8	3.7	3.4
1.72	5.7	4.7	3.9	3.4	3.1	2.8	2.7	2.6	2.4
2.09	4.4	3.5	2.9	2.6	2.4	2.2	2.1	2.0	1.9
2.52	3.4	2.8	2.3	2.1	1.9	1.8	1.7	1.6	1.5
3.0	2.8	2.3	1.9	1.7	1.6	1.5	1.4	1.4	1.3