Variation from Measurement Systems

B. Abraham
University of Waterloo
M. Brajac
General Motors of Canada

RR-94-01 (NT)

January 1994

VARIATION FROM MEASUREMENT SYSTEMS

Bovas Abraham
Department of Statistics and Actuarial Science
University of Waterloo, Waterloo, Ontario, Canada

and

Mike Brajac
General Motors of Canada, Oshawa, Ontario, Canada

January 1994

ABSTRACT

Industries, government agencies and other organizations utilize large resources to collect data, sometimes using highly sophisticated equipment. These data may lead to very costly decisions. One should be very concerned about the actual data collection processes: methods of collection, the equipment used, personnel involved, etc. We look at some specific situations to illustrate the importance of good measurement systems.

Key words: Cost of Data, Measurement Process, Variation.

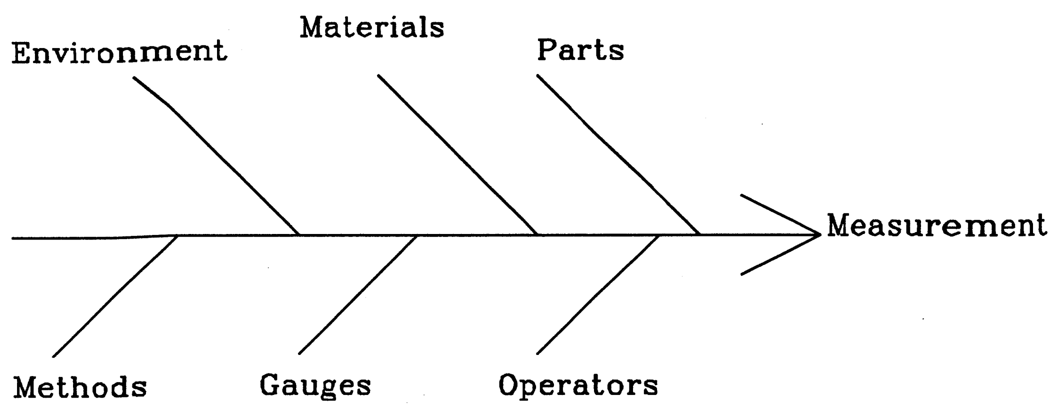
1. INTRODUCTION

Recently there is a greater awareness of Quality Improvement activities and this has brought, among other things, the concept of Data Based Decisions to the forefront of North American Industry. Deming (1986) urges that decisions must be based on facts supported by data. More managers are asking for data before decisions are taken and policies are implemented. This makes it extremely important to think seriously about data collection procedures (who, when, how, etc.) and the systems that produce measurements. Basing decisions or policies on inappropriately collected data can result in poor policies and inappropriate implementation schemes. Hunter (1980) makes the point that "Data of poor quality are a pollutant to clear thinking and rational decision making."

Today there are many sophisticated machines which automatically take measurements. Are we certain that they are supplying accurate and precise measurements? What percent of the total variation is really due to the measurement system itself? It is not uncommon to find a measurement system that is being used in production that accounts for a major portion of the process variability. There are occasions where the factors and levels for an experiment are well defined. However, the measurement system may be completely missed. In one case, one of us was working with two engineers from the "head office" who came to run an experiment on a robot that applied weatherstripping. These engineers had spent considerable time investigating various factors and levels and had come up with a suitable design. They had even started to run some trial runs in the morning prior to the actual experiment. The consultant asked the question: what is going to be measured? He was shocked to learn that this had been given very little thought and was told that "this was going to be figured out" prior to running the experiment later that day. To be fair, these engineers had just completed a design of experiments course and it was their first experiment. But this begs the question, why weren't measurement systems emphasized in the course?

Sources of Measurement Variation

The following diagram gives some ideas about the various sources that can contribute to the variation in measurements.



Variation in gauges, operators and methods is usually referred to as measurement system (MS) variation.

In assessing and/or improving an existing measurement system or in the development of a new system, the following points will be helpful:

(i) Selection of Units

The unit of measurement should be small enough to detect the variation in the process.

(ii) Repeatability and Reproducibility

Reproducibility refers to the variation from the measurement system (variation between operators and that between gauges). This variation should be small compared with the total variation; also it should be small compared with specification limits. Repeatability refers to the variation from repeat measurement on the same part by the same operator.

(iii) Accuracy

Average of the measurements on the same part repeatedly by the same operator and by the same gauge should be close to the true value.

(iv) Linearity

Variation or bias in measurements should not depend on the size of the part.

(v) Stability

Measurement system should be consistent (stable) over time. This requires measuring things over time.

2. DEVELOPMENT OF A NEW MEASUREMENT SYSTEM

Sometimes customer motivated quality improvement projects may face the reality that there is no measurement system. Soft trim in the automotive industry is a typical example. A trim plant typically does not work with blue prints. Customer requirements are often communicated by means of pictures. One engineer summarized this process of the communication as "Make more like this, but keep making them better."

The following points may help in the development of a measurement system for these kinds of situations.

(i) Define the characteristic that is to be measured. For improving a system, a proper characteristic is to be measured and optimized.

(ii) As a first step, try to devise a system that at least can ascertain if the characteristic is present.

(iii) Refine the system so that it can distinguish between various degrees of the characteristic. Insure that product that is available represents various degrees. For example, if a six point scale is developed, there should be six points of product available. Frequently a five or seven point scale will be applied to a product, but only two or three points are used.

(iv) There will often be several versions of the measurement system before one is finalized. Measurement system studies are useful to evaluate each iteration. In each of these iterations, the five points introduced in Section 1 will provide some guidance.

The statistician can play a role by ensuring that appropriate methods are used to evaluate the measurement system at each stage. This information can provide the feedback that is needed to measure the progress.

As an example of the amount of work that is required for such a situation, we will discuss a measurement study taken from General Motors of Canada Windsor Trim Plant, Liburdi, et al (1992). The most interesting part about this case is that it deals with the iterative nature of the scientific method to arrive at an acceptable solution.

The Windsor Trim Plant sews pieces of material together to form a seat cover. This is placed over a "bun", either in-house or at another company. One characteristic of seats, particularly for high-end leather seats is "wrinkles". The pattern of wrinkles adds an air of richness to the product. If the pattern does not appear to be natural, it is often perceived to be of poor quality. One characteristic critical to a rich looking appearance is shirring. Shirring, is the gathering of material along an edge. This gathering, imparts wrinkles in the piece of material. In order to increase the understanding of the effect of shirring on material, the plant decided that there must be an effective means of measuring shirring. There was no gauge on the market for measuring this.

A number of iterations was performed using a variety of home made gauges. This variety is a credit to the imagination of the individuals involved in the project. Some of the earlier iterations dealt with fitting the shirred piece of leather to a round object that had a ruler mounted to it. Some of the objects used were a round hollow drum core, a soft drink can, and a basketball. Repeatability and reproducibility was poor mainly due to the lack of solid fixturing for the part. During subsequent investigation on gauges that fixed the part securely, it was found that repeatability was particularly poor. This was due to the fact that the gauge was stretching the part, thereby destroying some of the resiliency of the leather. This work led to the development of a gauge that did not damage the part and also provided good repeatability and reproducibility.

The shirring case is typical of many measurement system studies that have been performed at the Windsor Trim Plant. Moreover, many of these studies were performed at various stages during the product life cycle, from early prototype development to mass manufacturing. It is now starting to be ingrained in the plant "culture" that measurement systems are a key issue. Before quality improvement projects are initiated in the plant, the question of measurement systems is raised. At first, this was somewhat frustrating for managers and engineers who were anxious to get the project started. The key to changing the thinking was to identify the measurement system study as an important step that is integral to the projects, not something that is done prior to the starting of the project.

Is everything done perfectly all of the time in the plant? The answer is no; but external recognition such as the Canadian Award for Business Excellence in Quality has proven that plant problem solving approaches have changed considerably. Measurement systems analysis has been a cornerstone of the plant's problem solving methodology.

3. ASSESSMENT OF MEASUREMENT SYSTEMS

There are well accepted statistical methods to assess various aspects of measurement system. "Repeatability and Reproducibility" studies which make use of analysis of variance techniques is a way of assessing the contribution of the measurement error variation to the total process variation. These can also lead to a comparison of measurement error variation to the specification limits. Similarly there are ways to check bias and other aspects of the measurement system. Since there are many papers and books on this topic, we will not discuss it in detail. However, we include one example here to illustrate some related issues.

A car manufacturer had complaints from the field about headlights being aimed too high which caused problems for on-coming vehicles. Some improvement in the headlight aiming operation, the last operation before a car goes out of the plant, was called for. A measurement system study was carried out to see, among other things, how big shift to shift (operator to operator) variation was. Management of the study was very important because of the personnel changes from Shift A and Shift B.

In this study several cars were used and three repeated measurements (forming a subgroup) from each headlight of a car were obtained during Shift A; their ranges are shown in the lower part of Figure 3.1. The same headlights were measured during shift B and the corresponding ranges are shown in the upper part of Figure 3.1

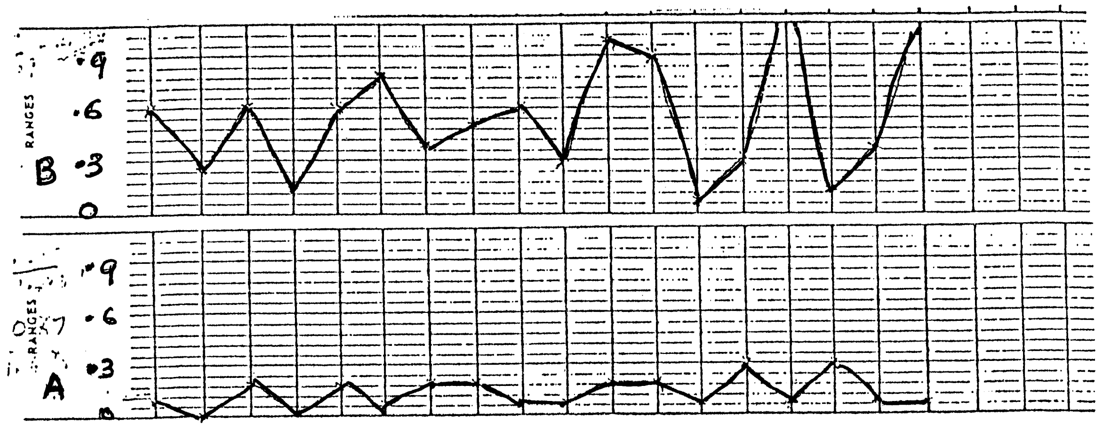


Figure 3.1 Headlight Aiming during Shifts A and B

Why are the ranges so large and more variable in Shift B? Should we have recommended that the operator in Shift B needed more training? How was that operator, who had been very cooperative during the study, going to react? It was necessary to improve the whole system of which the operator was only one component.

The whole headlight aiming process was reexamined and it was found that the aimers are quite sensitive in the way they sit on the headlight when the measurement is taken. There is some room for movement and this could change the measurement considerably. Hence, it was recommended that the aimers be modified so that they are insensitive (robust) to the operation and the possibility of movement is reduced. To implement the suggestion, the locators on the aimers were to be modified so that they would sit securely on the headlights. The supervisor of the department was receptive to the idea and there was a system in place, though not perfect, for implementation.

We suggested implementing the change on one aimer and conducting a study to evaluate the improvement; if it was substantial then the changes were to be implemented on the other aimers. We found out later that all aimers were modified; apparently people were happy with the modification on the first aimer! It should be noted that management of the study was very important. Improvement of the system as a whole, rather than particular components was given due importance. We learned that we should be sensitive to people issues.

4. IMPROVEMENT OF MEASUREMENT SYSTEMS

On-line gauges which are used to replace destructive testing equipment for inspection need to be calibrated adequately. In one of the projects that we were acquainted with there was little correlation between the gauge readings and those from the actual destructive test. A full factorial experiment was designed to optimize the gauge (high correlation between readings) by changing three of its controllable parameters. In this case the response for each of the runs is a set of on-line and off-line (destructive test) readings leading to a scatter plot. An additional complication was that two external factors were thought to effect the readings. Since these factors may vary during production it was important that the gauge was robust (insensitive) to this type of variation. Thus the experiment was enlarged so that each run (set up of the gauges) is tested under four conditions obtained from the two external (noise) factors at two levels each. Now the response for each of the runs consists of four sets of on-line and off-line readings (four graphs). The experimental set up is shown in Figure 4.1.

Control Factors			L	L	H	H	N_2
			L	H	L	H	N_1
A	B	C	1	2	3	4	Noise ↑
L	L	L	•	•	•	•	
H	L	L	•	•	•	•	
L	H	L	•	•	•	•	
H	H	L	•	•	•	•	
L	L	H	•	•	•	•	
H	L	H	•	•	•	•	
L	H	H	•	•	•	•	
H	H	H	•	•	•	•	

Figure 4.1. Experimental set up for improving correlation between on-line and off-line tests

A, B, C are the controllable factors and N1 and N2 are the uncontrollable external (or noise) factors; L and H indicate low and high levels of the factors. Each response is indicated by a "•" and this is a label for a scatter plot. There are thirty two scatter plots. Optimizing the gauge means here to find a gauge set up in which the correlation shown in the scatter plot is high and that this correlation is consistent (stays approximately the same) over the four noise conditions tested. In Taguchi's (1986) terminology this is a product array (inner and outer array) experiment to improve the gauge. In this particular study the results indicate that the gauge could not be optimized any further, leading to the conclusion that this gauge was not serving the purpose for which it was being used.

5. MEASUREMENT SYSTEMS IN NON-MANUFACTURING

We have, so far, concentrated on measurement systems in the manufacturing sector. However, the issues raised about measurement systems are also very important in non-manufacturing areas. For instance, inter-laboratory testing variation can be very large due to various factors such as procedures used, equipment available and personnel involved. The Centre for Disease Control in Atlanta sends out samples of blood to various laboratories for standardization of test procedures. In one of their studies of laboratory variation, they sent out blood containing lead to about 100 labs [see Hunter (1980)]. The best estimate of lead concentration was 41 micrograms/decilitre. However, the results from the labs varied from 33 to 55 micrograms/decilitre, quite a range. Apparently, the lead concentration in normal human blood is in the range of 15 to 20 micrograms/decilitre. With the amount of inter-laboratory variation found in the study, it is quite possible to report a value as large as 30 when the actual concentration is about 20. This situation is very serious as it can lead to unnecessary prescriptions, medications and even psychological consequences.

The performance appraisal system used by many companies is another example of a measurement system in non-manufacturing situations. It also is subject to the various sources of measurement error given in Section 1. Companies which apply measurement system concepts to the plant floor do not apply these same concepts to the performance appraisal system. From a practical point of view, it may not always be possible to conduct a measurement system study on the performance appraisal system. However, it is important for managers to think about the various aspects of a measurement system, listed in Section 1, in the context of performance appraisals.

6. CONCLUDING REMARKS

What are some of the factors that inhibit the development of good measurement systems? The lack of time and operation in a perpetual fire fighting mode are probably the key factors inhibiting development of good measurement systems.

Many managers do not grasp the concept of variation as it is applied to measurement data. Graphs or columns of numbers are accepted as gospel. Failing to recognize the concept, little support is given to engineers and technicians who wish to undertake measurement system studies. At the GM Windsor Trim Plant, management was supportive and allowed a quality technician to experiment with seven different iterations before an acceptable system was discovered. In many other organizations, the technician would have been forced to give up after two or three tries and told to use the "best gauge."

The statistician can play a very important role in consulting with a client. First of all, the statistician can clarify measurement systems issues and explain the various components of measurement error, such as repeatability, reproducibility, accuracy, linearity and stability over time. Often engineers, in particular, will confound all of these component errors in one term - accuracy. This is very confusing and does not lead to the kind of effort that is needed in developing good measurement systems. The magnitude of the constituent errors can be used as a guide in the development of a system. In the Windsor Trim Plant shirring example, poor repeatability led to the discovery that the gauge was altering the part. Finally, and most importantly the statistician can help others to ask the right questions about the integrity of data with respect to the measurement system.

7. REFERENCES

Deming, W.E. (1986). Out of the Crisis, MIT Centre for Advanced Engineering Study.

Hunter (1980). The National System of Scientific Measurements, Science Vol. 210, 21, p. 869-874.

Taguchi, G. (1986). Introduction to Quality Engineering, White Plains, NY: Unipub/Kraus International.

Liburdi, J. et al (1992). Measurement System for Seat Assembly, GM Symposium, p. 300-305.