

Application

# *Implementing Evaluation of the Measurement Process in an Automotive Manufacturer: a Case Study*

Graeme Knowles<sup>\*,†</sup>, Gordon Vickers and Jiju Anthony

*Warwick Manufacturing Group, International Manufacturing Centre, University of Warwick, Coventry CV4 7AL, U.K.*

*Reducing process variability is presently an area of much interest in manufacturing organizations. Programmes such as Six Sigma robustly link the financial performance of the organization to the degree of variability present in the processes and products of the organization. Data, and hence measurement processes, play an important part in driving such programmes and in making key manufacturing decisions. In many organizations, however, little thought is given to the quality of the data generated by such measurement processes. By using potentially flawed data in making fundamental manufacturing decisions, the quality of the decision-making process is undermined and, potentially, significant costs are incurred. Research in this area is sparse and has concentrated on the technicalities of the methodologies available to assess measurement process capability. Little work has been done on how to operationalize such activities to give maximum benefit. From the perspective of one automotive company, this paper briefly reviews the approaches presently available to assess the quality of data and develops a practical approach, which is based on an existing technical methodology and incorporates simple continuous improvement tools within a framework which facilitates appropriate improvement actions for each process assessed. A case study demonstrates the framework and shows it to be sound, generalizable and highly supportive of continuous improvement goals. Copyright © 2003 John Wiley & Sons, Ltd.*

KEY WORDS: measurement systems; variability reduction; capability; repeatability; reproducibility

## **INTRODUCTION**

**M** easurement processes are endemic within the manufacturing environment. Every process has some form of measurement of key characteristics associated with it. The data generated is used to draw a variety of conclusions about products and processes. Reviewing Deming<sup>1</sup> and Wheeler and Lyday<sup>2</sup>,

\*Correspondence to: Graeme Knowles, Warwick Manufacturing Group, International Manufacturing Centre, University of Warwick, Coventry CV4 7AL, U.K.

†E-mail: g.knowles@warwick.ac.uk

we can suggest that this may take the form of: enumerative data, where measurements are used to characterize products that have not been measured; analytical data, where measurements are used to characterize the underlying production process; or descriptive data, where measurements are used to characterize the products measured.

All of the potential modes of using measurement information are associated with some form of decision-making. The one thing that all these decisions have in common is that they are predicated on the belief that the data supporting them is reliable. The decisions made are only as reliable as the data that underpins them. If the data is variable, then so is the quality of the decision-making. As Shewhart<sup>3</sup> noted:

*'In any program of control we must start with observed data; yet data may be either good, bad or indifferent. Of what value is the theory of control if the observed data going in to that theory are bad?'*

## **THE PURPOSE OF MEASUREMENT IN MANUFACTURING**

As noted in the introduction, all meaningful decisions within the control and evaluation of any manufacturing process require measurements to support them. No accurate decisions about the acceptability of products, stability of processes or distribution of outputs can be made without recourse to a measurement system, the consequences of poor quality data in these decisions will usually lead to increased overall cost (Knowles *et al.*<sup>4</sup>). Clearly, the quality of decisions can only be as good as the data which goes into them and, hence, the measurement systems providing the data.

Measurement processes have essentially one of two purposes when deployed in a manufacturing environment:

- compare the process performance with its natural limits (establish control);
- compare the output of a process with the acceptable limits of performance (establish capability).

Both of these are crucial to any notion of improvement.

It is recognized that measurement data are used in non-manufacturing applications and, in particular, during product development and research activities. The conclusions of this paper could also apply in such circumstances, but this was not the focus of this particular study.

## **MEASUREMENT VARIATION**

There is uncertainty in every measurement that is taken and this can be attributed to a number of key inputs, such as gauges, operators, methods and environment. These have been identified in, amongst others, Hoyle<sup>5</sup> and Weckenmann and Rinnagl<sup>6</sup>.

Clearly, there will be consequences and costs associated with any incorrect decisions generated by such uncertainty. Should the measurement system show that a process is out of control when it is actually in control it will trigger improvement efforts which are not necessary and which cannot have any benefit. In the opposite case, a process will perform unpredictably for longer than necessary with consequences for the quality of the output. Should a product be incorrectly defined as inside specification, an unacceptable product will be passed from the process to the next process or final customer, with likely consequences of late rejection, consequential production costs or customer dissatisfaction. Should a good product be falsely rejected the company will lose the cost of its manufacture. Raouf *et al.*<sup>7</sup>, amongst others, have tried to model the cost of the unnecessary rejection of good parts for various levels of measurement variability, but many of the other costs are difficult to quantify and little effort has been made to define them.

Whilst for an individual decision the measurement error may generate either an  $\alpha$  or a  $\beta$  error, the long-term tendency will be for the measurement error to inflate the observed process variability over time. As noted by Weckenmann and Rinnagl<sup>6</sup>, the observed standard deviation ( $\sigma_o$ ) is related to the true process standard

deviation ( $\sigma_p$ ) and measurement system standard deviation ( $\sigma_m$ ) by the following equation:

$$\sigma_o = \sqrt{\sigma_p^2 + \sigma_m^2} \quad (1)$$

From the above discussion, it is clear that the measurement capability is an integral and important part of any effective process control and improvement strategy. This has been recognized by the inclusion of some form of measurement system evaluation as part of several frameworks for implementing statistical process control (for example, Does *et al.*<sup>8</sup> and Antony *et al.*<sup>9</sup>), and by its inclusion as a key element of the Six Sigma approach to improving process capability (e.g. Pande *et al.*<sup>10</sup>) and in the influential QS 9000 Automotive Standard<sup>11</sup>.

When seeking to improve the observed variation in a process, an important question is how to effectively discriminate between the condition where most of the variation is due to the process (i.e. it is reasonable to work on the process) and where a significant proportion of the observed variation is due to the measurement system (i.e. it is reasonable to work on the measurement system).

### **KEY CHARACTERISTICS OF A MEASUREMENT PROCESS**

The following criteria describe a 'good' measurement process. It should be:

1. *stable*, as defined by a control chart (i.e. no special causes present);
2. *linear* (i.e. it responds proportionately over the range of equipment);
3. *accurate*, so as to correspond closely to a known reference standard;
4. *consistent*, having little scatter of individual readings of the same product;
5. able to support *analysis*, by generating variable rather than attribute data;
6. *robust*, in that it is not sensitive to changes of operator, environmental conditions or other noise factors.

Points 2, 3 and 5 are fundamental aspects of the design and maintenance of a measurement system and will not be discussed further in this paper. Point 1 relates to ongoing control. Points 4 and 6 are the elements traditionally associated with a measurement system evaluation approach. They are analogous to repeatability (4) and reproducibility (6). These are defined below<sup>12</sup>.

- *Repeatability*. A measure of variation when the same operator, using the same gauge and the same method, measures the same item.
- *Reproducibility*. A measure of variation when two or more operators measure the same item, using the same gauge and the same method.

The combination of these measures will allow one to establish the answer to the question posed in the previous section.

### **METHODS OF EVALUATING MEASUREMENT VARIATION**

There are a number of different methods for analysing measurement variation available to companies today. These have been continuously evolving, particularly over the last 10 years. They are listed below as identified by Ermer and E-Hok<sup>13</sup>.

- Original Gauge R&R studies.
- Modified R&R studies.
- ANOVA.
- Evaluating The Measurement Process (EMP) methodology.

Ermer and E-Hok<sup>13</sup> reviewed the strengths and weaknesses of the available methods, noting that the first two methods have fundamental flaws undermining their accuracy and that ANOVA is the purest form of analysis,

but does not lend itself to easy identification of the causes of variability. From a practical viewpoint, however, the EMP approach proposed by Wheeler and Lyday<sup>2</sup> was shown to give stronger indications of how to go about improving the process, whilst still giving accurate results. Central to this technique is the graphical display of the average-range, bias and inconsistency control charts. The whole philosophy is based on the measurement process stability, predictability, centring and spread, which is aligned with comments of both Wheeler and Lyday<sup>2</sup> and Weckenmann and Rinnagl<sup>6</sup>. This technique was selected by the company as appropriate and aligned to the ongoing variability reduction drive within the organization. The detail of the EMP technique is shown in the case study later in the paper.

### ***SUMMARY OF THE STRENGTHS OF EMP***

The EMP approach was selected by the company due to the following strengths.

- *Accurate reflection of gauge capability*: obviously crucial to the usefulness of any approach.
- *Graphical representation of data*: ease of interpretation is a significant element of the utility of any technique. Graphical data conveys signals much more concisely and clearly than purely numerical data. This also aligned well with the cultural preferences of the organization, which was well versed in the use of charting techniques in variability reduction.
- *Clear indication of areas for improvement*: as Juran<sup>14</sup> states, it is important that any measure indicates what should be done differently. Merely designating measurement systems as acceptable or in need of improvement is less helpful than indicating the elements of the process that should be worked on.
- *Minimal statistical knowledge required to complete the analysis*: since the reliability of measurement data is extremely important, it is desirable that studies can be carried out, interpreted and worked on by as many people in the company as possible. The more accessible the technique, the wider its possible and likely use.

### ***APPLICATION OF EMP WITHIN AN AUTOMOTIVE COMPANY***

#### *Company context*

This organization became interested in the application of measurement variability reduction as the variability reduction drive in the organization began to reach maturity. A data driven culture had been established and the organization was becoming comfortable with the concepts of control and capability in the manufacturing processes. A growing recognition was developing in the company that the improvement effort, whilst largely successful, was being hampered by poor measurement systems in some areas.

Whilst EMP was selected as the best approach, it was recognized as requiring a framework in order to ensure consistent application and further enhancement of its usefulness to the company. These are explained below and a practical example taken from the organization included, exemplifying the approach and its associated benefits.

#### *The modified procedure*

The model below was developed to facilitate the consistent implementation of the technique by a wide variety of practitioners within the organization.

Without altering the substantive elements of the EMP procedure, an enhanced practical framework has been created. Observing and documenting the process in sufficient detail to create a flowchart clarifies the focus for variability reduction. Interpretation guidelines are also provided in order to facilitate appropriate decision-making and action at the various stages of analysis. By revisiting and standardizing the process flowchart the model ensures that learning is captured and the process 'refrozen'<sup>15</sup> in its improved state. This simplifies the approach and leads practitioners through the phases creating a more accessible and consistent methodology in practice (see Figure 1).

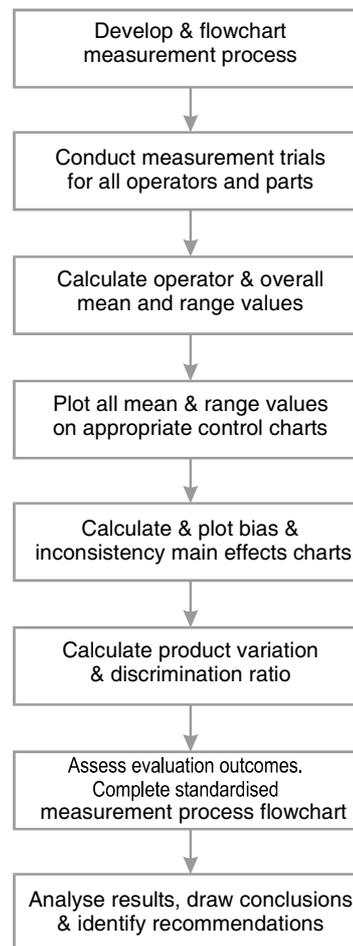


Figure 1. Company procedure for assessing measurement system variability

An example of this methodology is worked through below to demonstrate the achievable benefits when using this approach.

#### *Case study*

The case study involves the measurement of a door gap on an automobile. The steps of the flowchart were followed rigorously and are summarized below.

#### *Develop and flowchart the measurement process*

The initial process was already in existence, using a 'carrot' gauge and followed the process shown in Figure 2.

#### *Conduct measurement trials for all operators and parts*

Four door gap conditions were prepared to represent the variability in conditions that might be encountered during normal production. Five operators who were normally involved in the inspection operation were selected, to ensure that the range of operator skill levels, methodologies etc. were included. Each operator measured each part twice. Trials were conducted in a random order and without knowing which part was which, to avoid biasing the results. The raw data for the initial process (carrot gauge) is shown in Table I.

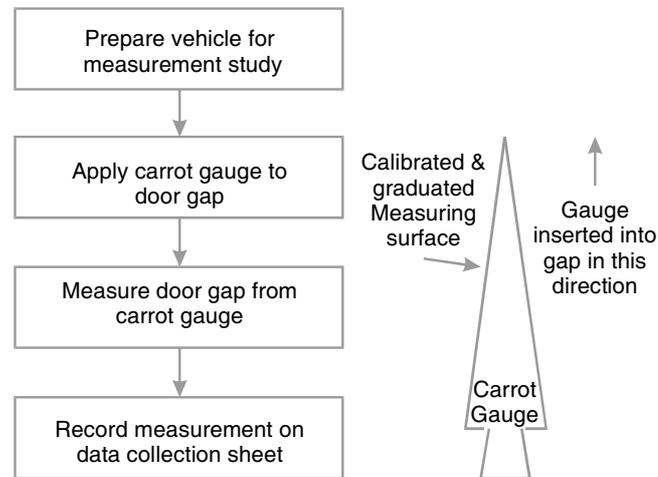


Figure 2. Carrot gauge and measurement system

Table I. Raw data for the carrot gauge. Overall mean of data: 5.150

Part No.	Op. A		Op. B		Op. C		Op. D		Op. E	
	T1	T2								
1	4.3	4.2	4.8	4.7	5.7	5.5	5.7	5.7	6.2	5.8
2	4.2	4.2	4.7	4.7	5.1	5.5	5.4	5.4	5.5	6.0
3	4.1	4.1	4.7	4.7	5.7	5.8	5.3	5.4	5.9	5.9
4	4.1	4.1	4.7	4.7	5.6	5.6	5.3	5.3	5.8	5.9

Table II. Statistics by operator for the carrot gauge: mean of operator means = 5.150; mean of operator ranges = 0.100; standard deviation of part means = 0.0821; number of operators (NO) = 5; number of parts (NP) = 4; tests per part per operator = 2. Standard deviation estimates: test-retest error = 0.0874; parts = 0.0773; combined = 0.117

Part No.	Op. A		Op. B		Op. C		Op. D		Op. E		Part means
	Mean	Range									
1	4.25	0.10	4.75	0.10	5.60	0.20	5.70	0.00	6.00	0.40	5.26
2	4.20	0.00	4.70	0.00	5.30	0.40	5.40	0.00	5.75	0.50	5.07
3	4.10	0.00	4.70	0.00	5.75	0.10	5.35	0.10	5.90	0.00	5.16
4	4.10	0.00	4.70	0.00	5.60	0.00	5.30	0.00	5.85	0.10	5.11
Means	4.163	0.02	4.713	0.02	5.563	0.18	5.438	0.02	5.875	0.25	5.15

Calculate operator and overall mean and range values

The mean range of statistics by operator are shown in Table II.

Plot all mean and range values on appropriate control charts

*Average chart.* This technique utilizes the graphical display and prediction benefits of the average-range control chart. However, as it is constructed by isolating the measurement error within the same subgroup, it compares the variability of the process with the measurement error. A high number of out-of-control points is desirable, indicating a low measurement error in comparison with the process variability. The chart limit calculations for the carrot gauge are shown in Table III.

Table III. Standard deviation estimates and chart limits

Control charts		LCL	CL	UCL
3-sigma	Operator-part means	4.965	5.150	5.335
	Operator-part ranges	n/a	0.100	0.327
Main effects charts		LDL	CL	UDL
5%	Operator means (bias)	5.073	5.150	5.214
	Operator ranges (inconsistency)	0.0375	0.100	0.177
1%	Operator means (bias)	5.067	5.150	5.227
	Operator ranges (inconsistency)	0.0282	0.100	0.194

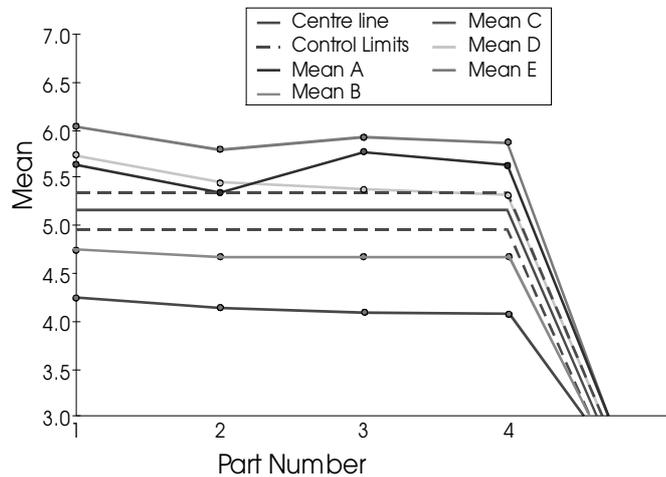


Figure 3. Part means by operator

The part means chart (Figure 3) shows low variability in the operator measurements compared with the control limits (based on measurement error), indicating that the measurement variability is excessive compared with the dimensional variation. Since this is true for all individual operators, it reflects a repeatability problem rather than a reproducibility problem.

*Range chart.* Although the range chart is constructed from the measurement error ranges, this chart is interpreted in exactly the same way a normal range control chart would be. There should be no out-of-control points detectable from the chart. Any out-of-control points would require further investigation and their subsequent removal, as these would indicate a level of inconsistency in the way the overall process is being operated.

Calculations were carried out as in Tables II and III.

The test-retest chart (Figure 4) indicates that the process is not being performed consistently. Operators C and E exhibit points outside the control limit and can thus be concluded to have significantly more variability in their approach than would be expected. This is an indication that standardization would improve the measurement error significantly. Whatever other conclusions we draw from the further charts, this means that fundamental action is required.

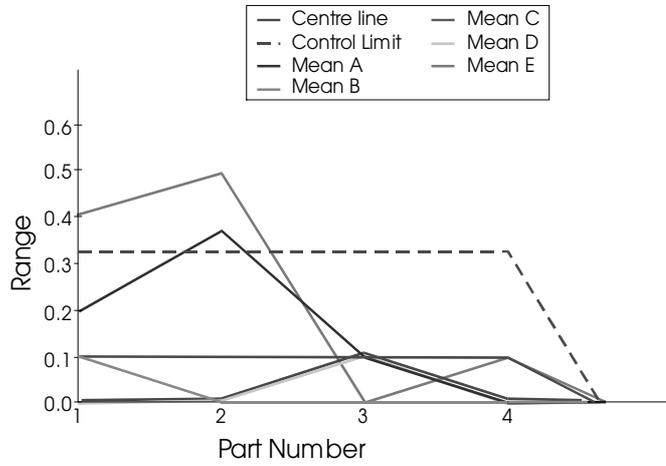


Figure 4. Test-retest ranges by operator

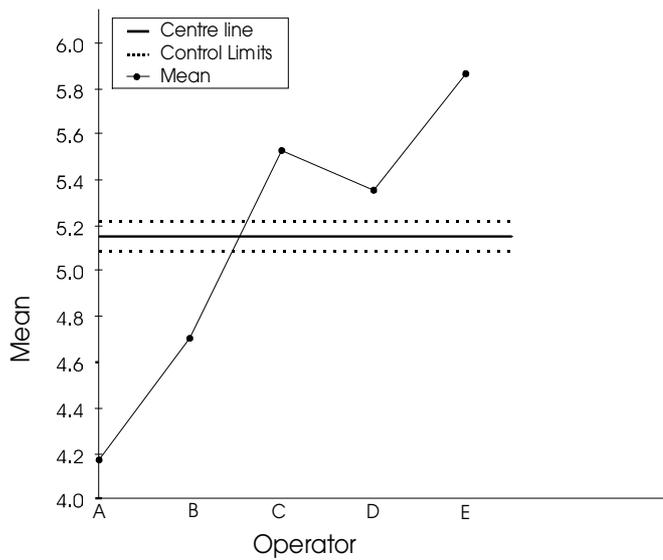


Figure 5. Bias plot

*Calculate and plot bias and inconsistency main effects charts*

There are two additional control charts that are introduced when using this technique. These are the bias and inconsistency main effect charts. Each chart focuses on a particular ‘nuisance’ aspect of measurement variation.

*Bias main effects.* The bias main effect chart plots the overall average value for each operator against the overall process mean and modified control limits. If any of the operator averages falls outside a control limit, then this identifies that there is a detectable difference between the operator averages, i.e. a bias exists. Furthermore, the chart will identify which operator causes the bias.

The highly out-of-control bias plot (Figure 5) indicates that there are very significant differences between the operators in terms of their bias (or central tendency).

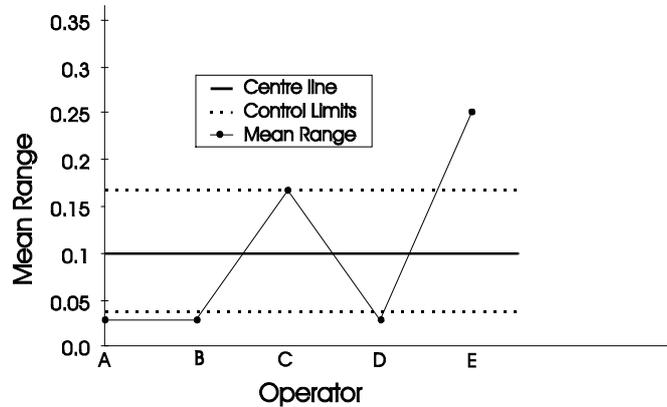


Figure 6. Inconsistency plot

*Inconsistency main effects.* The inconsistency chart works in a similar manner to the bias chart, except the operator average range value is plotted. Again, modified control limits are calculated and the chart interpreted. No operator average range value should fall outside a control limit, since this would indicate that there is a detectable difference between the operator test–retest values (because this chart uses the measurement error ranges). Again, the chart will indicate which operator is inconsistent with the rest.

The highly out-of-control inconsistency plot (Figure 6) indicates significant differences in the variability of the measurement process between operators. Operator E is the most inconsistent and can be seen as the main contributor to this problem.

*Intraclass correlation coefficient (ICC)*

This is a ratio of the true product variation and the measured value variance. It measures the proportion of the variation in the measured values which is due to the true product variation and thus forms a useful ratio for the evaluation of a measurement process. The calculation for ICC is:

$$ICC = \frac{\text{est. variance (true product)}}{\text{est. variance (measured value)}} = \frac{\hat{\sigma}_p^2}{\hat{\sigma}_m^2} = \frac{\hat{\sigma}_m^2 - \hat{\sigma}_e^2}{\hat{\sigma}_m^2} = 1 - \frac{\hat{\sigma}_e^2}{\hat{\sigma}_m^2} \tag{2}$$

where  $\hat{\sigma}_e^2$  is the estimated variance of measurement error.

*Calculate product variation and discrimination ratio*

As the ICC involves variances whose units are the square of the measurement units for the data collected, the magnitude of the ratio can be difficult to interpret<sup>16</sup>. In order to overcome this limitation Wheeler and Lyday<sup>2</sup> offer a transformation to the ICC called the discrimination ratio, which is more intuitive. The discrimination ratio characterizes the relative usefulness of a given measurement for the specific product being analysed. This ratio is the relationship between the estimated variance of measurement error and the estimated variance of the product measured values and thus it is an intuitively sensible measure. The calculation is

$$\text{estimated discrimination ratio } (D_R) = \sqrt{\frac{2\hat{\sigma}_m^2}{\hat{\sigma}_e^2} - 1} = \sqrt{\frac{1 + ICC}{1 - ICC}} \tag{3}$$

The discrimination ratio provides an indication of whether the measurement process is able to detect product variation. If the measurement error is excessive, it will obscure this product variation. Thus this relationship will provide an insight into whether the measurement process requires improvement or is sufficient to distinguish product variation. However, at what ratio does the measurement process require subsequent improvement before it can reliably measure a given product?

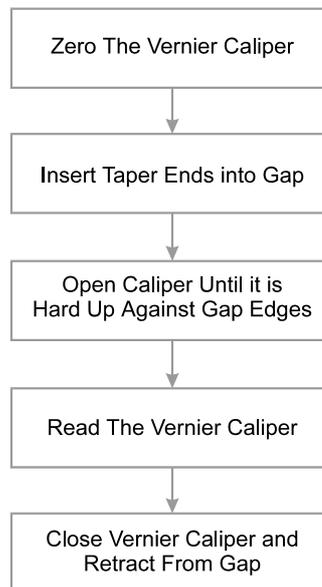


Figure 7. Modified measurement process using Vernier callipers

Wheeler and Lyday<sup>2</sup> offer an empirical recommendation, which has been adopted by the company in question:

*'That for simple measurement processes it might well be worth working on the measurement process if the Discrimination Ratio is down in the region of 4.0.'*

This ratio can be visually represented by an intraclass correlation plot<sup>2</sup>. This graphically compares the variances as the discrimination ratio and may, therefore, be preferable for some operatives, but will not be dealt with in this paper.

The discrimination ratio is calculated as in Equation (3) at 1.60, indicating that the measurement process currently in use is not adequate to discriminate within the levels of variability present.

#### *Assess the outcomes of the evaluation*

The team then assessed the measurement process for potential causes of the observed patterns of variation, coming to the following conclusions.

- The 'carrot' gauge approach was highly susceptible to operators using different pressures, gauge deterioration, parallax errors and reading errors.
- It was unlikely that such fundamental weaknesses could be addressed by minor changes to operation or hardware. This was supported by the very low discrimination ratio.
- A completely new approach to the measurement of this feature was required.

A new approach was designed using a Vernier calliper modified by the addition of extended measurement faces to allow parallel access to narrow gaps.

#### *Complete standardized measurement process flow*

The new process flow is shown in Figure 7.

#### *Analyse results, draw conclusions and identify recommendations*

The new measurement process was analysed following the steps above. A summary of the raw data (Table IV) and calculations (Tables V and VI) is presented below, and associated charts are shown in Figures 7–10.

Table IV. Raw data for Vernier callipers

Part No.	Op. A		Op. B		Op. C		Op. D		Op. E	
	T1	T2								
1	4.22	4.25	4.21	4.23	4.30	4.31	4.28	4.30	4.20	4.27
2	5.18	5.19	5.19	5.19	5.18	5.19	5.21	5.19	5.19	5.19
3	4.81	4.81	4.82	4.77	4.73	4.75	4.79	4.77	4.81	4.82
4	4.92	4.93	4.84	4.85	4.84	4.90	4.90	4.89	4.97	4.91

Table V. Statistics by operator for Vernier callipers: mean of operator means = 4.783; mean of operator ranges = 0.0220; standard deviation of part means = 0.38; number of operators (NO) = 5; number of parts (NP) = 4; tests per part per operator = 2. Standard deviation estimates: test-retest error = 0.0192; parts = 0.389; combined = 0.390

Part No.	Op. A		Op. B		Op. C		Op. D		Op. E		Part means
	Mean	Range									
1	4.24	0.03	4.22	0.02	4.31	0.01	4.29	0.02	4.24	0.07	4.26
2	5.19	0.01	5.19	0.00	5.19	0.01	5.20	0.02	5.19	0.00	5.19
3	4.81	0.00	4.80	0.05	4.74	0.02	4.78	0.02	4.82	0.01	4.79
4	4.93	0.01	4.85	0.01	4.87	0.06	4.90	0.01	4.94	0.06	4.90
Means	4.789	0.01	4.763	0.02	4.775	0.03	4.791	0.02	4.795	0.03	4.78

Table VI. Standard deviation estimates and chart limits for Vernier callipers

Control charts		LCL	CL	UCL
3-sigma	Operator-part means	4.742	4.783	4.823
	Operator-part ranges	n/a	0.0220	0.0719
Main effects charts		LDL	CL	UDL
5%	Operator means (bias)	4.766	4.783	4.797
	Operator ranges (inconsistency)	0.0083	0.0220	0.0389
1%	Operator means (bias)	4.764	4.783	4.799
	Operator ranges (inconsistency)	0.0062	0.0220	0.0427

All range points for test-retest errors (Figure 8) are in control, indicating that the new measurement process is performed consistently across all operators.

The means chart (Figure 9) shows a wide variation of measurements against the control limits (based on measurement error), indicating that the measurement error is small compared with the part variation. This is again a good indication that the measurement process is an improvement on the previous system.

The inconsistency chart (Figure 10(a)) indicates that there are no significant operator inconsistency effects.

The bias chart (Figure 10(b)), however, shows that there may be a problem with operator B in terms of bias.

The discrimination ratio is calculated to be 28.65, which establishes the new measurement process as highly capable and a vast improvement on the previous one.

The recommendations were to adopt the new process and retrain operators fully; particular attention should be paid to operator B in order to minimize the potential bias effect.

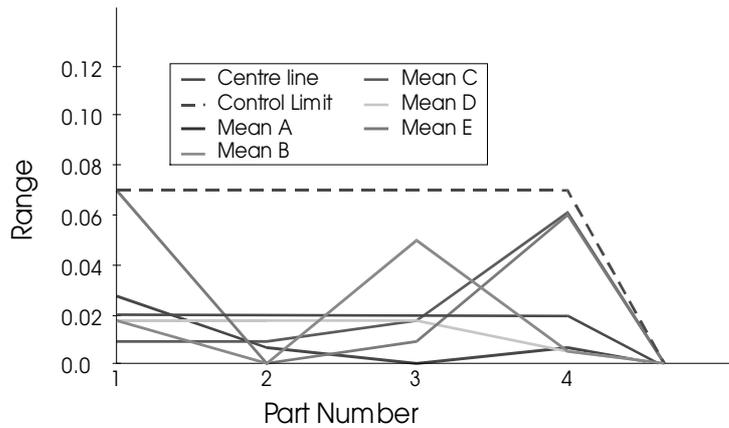


Figure 8. Test-retest ranges for Vernier callipers

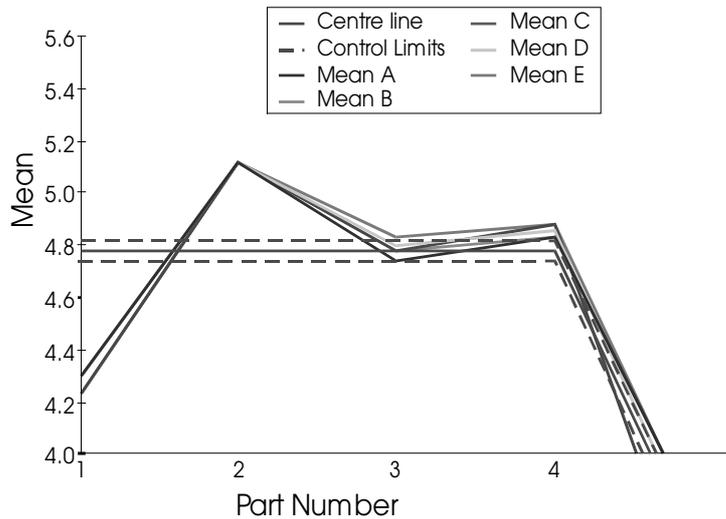


Figure 9. Part means by operator for Vernier callipers

After retraining a restudy was done, which indicated that operator B was no longer a problem with regard to bias. At this point, the new process methodology and training for any new operators were formally integrated into the departmental operating procedures.

## DISCUSSION

This successful case study indicates that the proposed measurement evaluation process and associated framework are effective. By adding the additional improvement-focused steps, the group were facilitated in establishing appropriate responses to the measurement process failings identified by the application of EMP. They were also better able to embed the new process quickly. The documentation generated by the process will also support further action on the process, should this become necessary due to future improvements in process variability. In short, the company now has a process that is under better control, because the measurement system

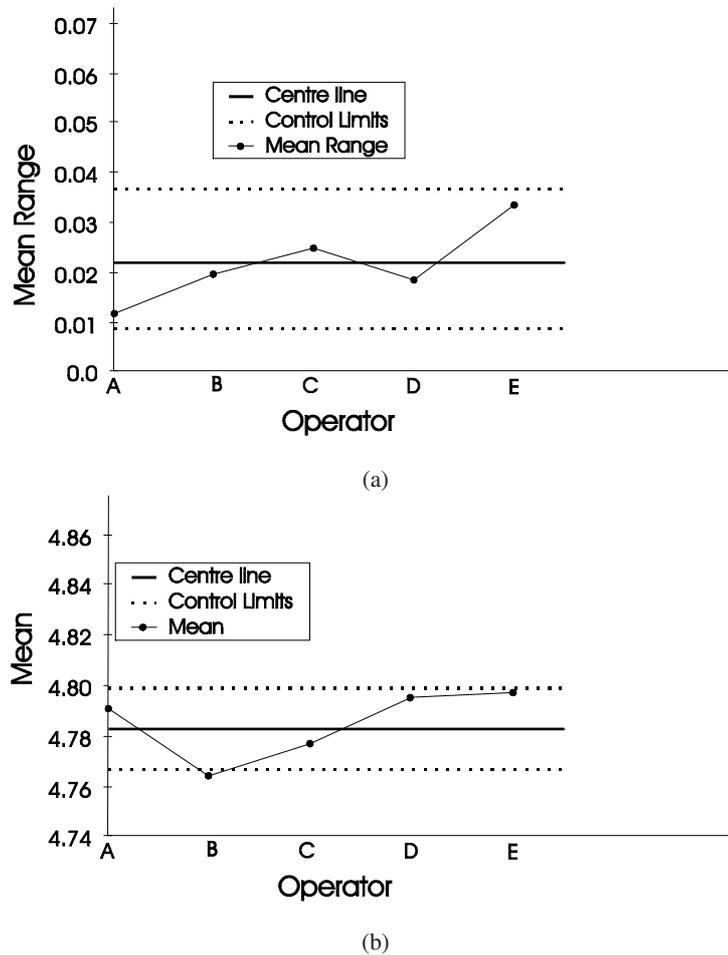


Figure 10. (a) Inconsistency plot for Vernier callipers; (b) bias plot for Vernier callipers

now more accurately reflects what is happening in the process. Attempts to apply SPC had previously failed, but are now successful as the data going into the control chart are now reliable.

The company has applied the approach to a large number of processes since its initial introduction and have made a significant impact on the quality of data in the organization. For example, of the first 30 studies carried to completion, only five met all the criteria for acceptability on the first study, all did so afterwards. It is also interesting to note that of the 30 studies, 20 results were different to those obtained using the less systematic approach in place prior to the introduction of EMP to the company. In all cases, the quality of decisions taken on the level of control of the process and the acceptability of products has been improved, saving money previously wasted by incorrect data. This has taken the form of:

- fewer parts rejected incorrectly;
- fewer defective parts passed to later operations and customers;
- fewer false alarms and missed special causes on control charts.

It is concluded from this that the process is not only effective in its application, but makes a significant business contribution to ensuring that improvement effort is targeted on the appropriate element of the process/measurement combination.

It is worth noting that the methodology is a good logical fit with approaches like the Six Sigma Define, Measure, Analyse, Improve, Control approach<sup>9</sup>. This again supports the wide applicability of the methodology.

Perhaps the only significant weakness of the process is the failure to address changes over time in a measurement process (as measurement system analysis in QS 9000<sup>11</sup> does), it is very much a 'snapshot' of performance. This was dealt with by incorporating an EMP study into recurring calibration schedules and recognizing that, as an element of the manufacturing process, any significant (i.e. special cause) changes would show up on the SPC chart associated with the process.

## CONCLUSIONS

This paper has demonstrated the importance of evaluating measurement systems in order to generate reliable data on which management decisions can be sensibly based. A brief comparison of the most widely used techniques is reported, and a judgement that the EMP methodology proposed by Wheeler and Lyday represents the best combination of an accurate assessment of gauge capability, usability and clarity of improvement actions has been given.

An extended EMP process has been proposed which puts an organizational framework around the mechanics of the process in order to ensure that improvements are fully captured and embedded. The case study demonstrates a significant improvement in the quality of data available after the application of the methodology and the further applications confirm the more general applicability of the approach, supported by its affinity with established variability reduction approaches.

The methodology has delivered significant business benefit through a reduction in costs associated with measurement error and improved process control through the provision of reliable data.

The methodology proposed in this paper builds upon the sound basis of EMP practice to produce a robust, consistent and transferable procedure for evaluating measurement systems and driving improvements where appropriate.

## REFERENCES

1. Deming WE. *Out of the Crisis*. Massachusetts Institute of Technology: Cambridge, MA, 1982.
2. Wheeler DJ, Lyday RW. *Evaluating the Measurement Process* (2nd edn). SPC Press: Knoxville, TN, 1989.
3. Shewhart WA. *Economic Control of Quality of Manufactured Product*. Van Nostrand: New York, 1931. (Re-published in 1980 by the American Society of Quality Control.)
4. Knowles G, Antony J, Vickers G. A practical methodology for analysing and improving the measurement system. *Quality Assurance* 2000; **8**(2):59–75.
5. Hoyle D. *QS-9000. Quality Systems Handbook*. Society of Automotive Engineers Inc: Warrendale, U.S.A., 1995.
6. Weckenmann A, Rinnagl M. Acceptance of processes: Do we need decision rules? *Precision Engineering* 2000; **24**:264–269.
7. Raouf A, Duffuaa SO, Shuaib AN. On process capability and product tolerancing as affected by measuring device. *International Journal of Quality & Reliability Management* 1995; **12**(8):74–81.
8. Does RJMM, Schippers WAJ, Trip A. A framework for implementation of statistical process control. *International Journal of Quality Science* 1997; **2**(3):181–198.
9. Antony J, Balbontin A, Taner T. Key ingredients for the effective implementation of statistical process control. *Work Study* 2000; **49**(6):242–247.
10. Pande PS, Neuman RP, Cavanagh RR. *The Six Sigma Way: How GE, Motorola, and Other Top Companies Are Honing Their Performance*. McGraw-Hill: New York, NY, 2000.
11. Chrysler Corporation, Ford Motor Company and General Motors Corporation. *Quality System Requirements, QS 9000* (3rd edn). Automotive Industry Action Group: Southfield, MI, 1998.
12. Antony J, Knowles G, Roberts P. Gauge capability analysis: Classical versus ANOVA. *Journal of Quality Assurance* 1999; **6**(3):173–182.
13. Ermer DS, E-Hok RY. Reliable data is an important commodity. *The Standard* 1997; **Winter**:15–30.
14. Juran J. *Juran on Planning For Quality*. The Free Press: U.S.A., 1988.
15. Siegal W, Church AH, Javitch M, Waclawski J, Burd S, Bazigos M, Yang T-F, Anderson-Rudolph K, Burke WW. Understanding the management of change: An overview of managers' perspectives and assumptions in the 1990s. *Journal of Organisational Change Management* 1996; **9**(6):54–80.
16. Quesenberry CP. *SPC Methods for Quality Improvement*. John Wiley & Sons Inc: New York, U.S.A., 1997.