STATISTICAL PROCESS MONITORING IN THE 21ST CENTURY

Michael Wood
University of Portsmouth Business School
Portsmouth
UK
michael.wood@port.ac.uk
19 October 2000

This is a draft of a paper published in Antony, J and Preece, D (eds), Understanding, managing and implementing quality: frameworks, techniques, cases, Routledge, London, 2002.

INTRODUCTION

The term "statistical process control" (SPC) refers to a loosely defined collection of techniques for monitoring a process so as to prevent deterioration and facilitate improvement. These techniques are used for monitoring a process (Box and Kramer, 1992), so the phrase "process monitoring" seems more appropriate than "process control" and I have used the former in the title. However, this is just a matter of terminology: SPC techniques always were about monitoring rather than control, and as the use of the word control causes problems - as we will see below - it is sensible to replace it with a more accurate term. I will use the term SPC/M in this paper.

The most prominent of these techniques, and the traditional focus of SPC/M, is the Shewhart control chart, named after its originator, Walter Shewhart. There are different types of Shewhart chart: the most widely used are charts for the mean (*X-bar*) and range (*R*), proportion defective (*p*), and number of defects (*c*). In each case, a graph of the appropriate statistic (mean, range, etc) of successive samples is plotted, and "control lines" superimposed to indicate points which are "out of control". These indicate "special" or "assignable" causes of variation which should be investigated and, if appropriate, action taken to adjust the process.

There are also a number of more sophisticated control charting methods (although, according to Gunter, 1998, these are "rarely used"). These include multivariate methods for monitoring several related variables simultaneously (Montgomery, 1996), methods for monitoring a single measurement (as opposed to one based on a sample) such as moving average charts and exponentially weighted moving average (EWMA) charts (see, for example, Montgomery, 1996), and cumulative sum (cusum) methods which are more sensitive than Shewhart charts for detecting small but consistent changes in the level of the measurement (Hawkins and Olwell, 1998).

In addition to control charts, the conventional SPC/M package incorporates ways of establishing the capability of a process (Rodriguez, 1992) - the most commonly used index here being c_{pk} - and a number of more elementary methods for solving problems and improving quality. For example, Montgomery (1996, p. 130) lists the "magnificent seven": histogram or stem-and-leaf-display, check sheet, Pareto chart, cause and effect diagram, defect concentration diagram, scatter diagram, as well as the control chart.

These methods, their implementation, and the concepts and philosophy underlying them, are covered in the many texts on SPC/M and related areas: eg Oakland (1999), Woodall and Adams (1998), Montgomery (1996), Bissell (1994), Mitra (1993). Woodall and Montgomery (1999) provide a helpful recent review of current issues and research in the area.

The purpose of this chapter is not to provide a summary of SPC/M and how to implement it. There are many excellent texts - such as those mentioned above - to which readers can refer for the technical and organizational details of the procedures, and an analysis of their potential benefits. Instead, this chapter aims to provide a critique of SPC/M, and some suggestions about how it needs to be adapted to the twenty first century. I am assuming that the reader has some familiarity with the main SPC/M techniques - although not necessarily with the details of formulae or the more advanced methods.

The value of SPC/M has been widely recognised over the last half century. According to Stoumbous et al (2000)

> control charts are among the most important and widely used tools in statistics. Their applications have now moved far beyond manufacturing into engineering, environmental science, biology, genetics, epidemiology, medicine, finance, and even law enforcement and athletics.

However, there is little recent, empirical evidence of widespread benefits from SPC/M in business, and, indeed, a few suggestions that all is not well. For example Gunter (1998, p. 117) suggests that "it is time to move beyond these now archaic and simplistic tools [control charts]" and complains that

> We have become a shockingly ingrown community of mathematical specialists with little interest in the practical applications that give real science and engineering their vitality.

Woodall and Montgomery (1999) suggest that the problem is that "in much of academia, the rewards are for publications, not usefulness" (p. 18). In the world of real applications, on the other hand:

> Many, if not most of the users of control charts have had little training in statistics. Thus, there is a reluctance to introduce more complex topics, such as the study of autocorrelation, into training materials. (p. 17)

The result of this situation is, as might be expected, disappointing. Dale and Shaw (1991, p. 40), on the basis of research in the late eighties, concluded that

> The findings of this piece of research must bring into question the effectiveness of the current methods of educating company managements on the use of SPC/M. The time, resources, and cost committed to SPC/M by organizations has been considerable and if a cost benefit analysis were to be performed it would be unfavourable.

Hoerl and Palm (1992) also comment on the "limited success" of many efforts to use Shewhart charts in industry, which, they say, is typically due to using the wrong formula, a poorly chosen sampling plan, or that the "improvement work demanded by the charts is so radical in the context of the organization's culture that the organization is unable to properly respond" (p.269).

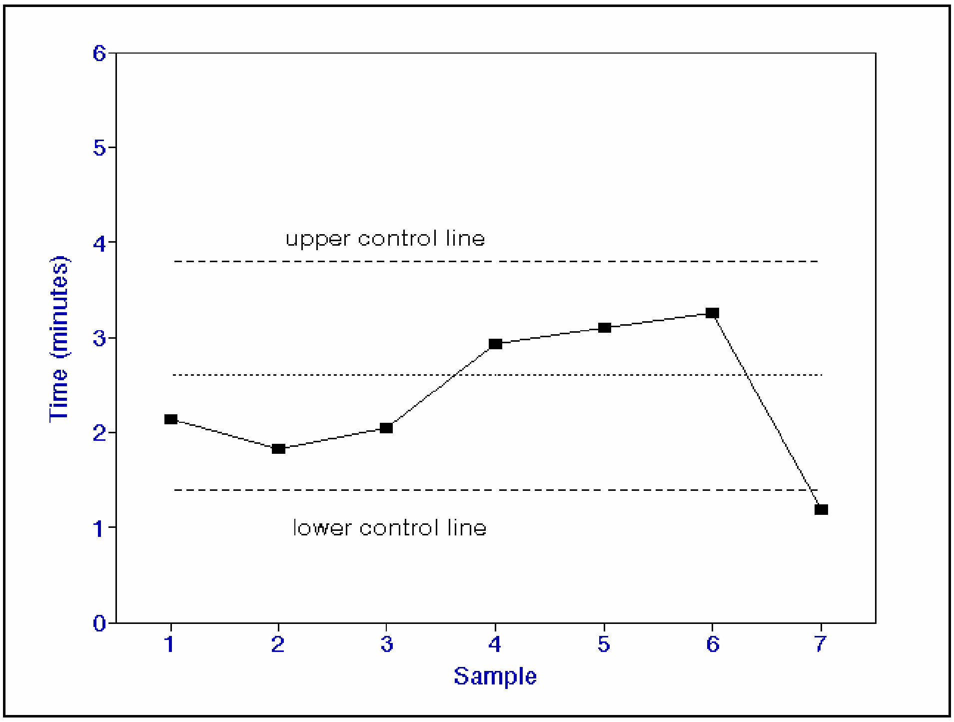
This chapter explores issues such as these. It starts with a discussion of the purpose and potential benefits of SPC/M. However, as we have seen, this is not the whole story; there are difficulties in practice. The section after gives a brief case study, based on a small manufacturing company, which illustrates many of the difficulties of trying to implement traditional SPC/M techniques, as well as some of the benefits. This leads on to a systematic discussion of these difficulties and how they might be resolved.

THE PHILOSOPHY, PURPOSE AND POTENTIAL BENEFITS OF SPC/M

The purpose of a control chart is to monitor a process so that patterns and trends over time are seen, and problems can be picked up in time to prevent defective output being produced. The control lines are designed to pick out those fluctuations which are too large to be normal "chance" fluctuations; the purpose of this is to avoid wasting time and money and disturbing the process unnecessarily by reacting to "normal" or "common cause" variations.

Figure 1, for example shows the mean times taken for patients to reach the admissions ward from casualty in a hospital. This is based loosely on the situation described by Jefferson and Humphrys (this volume, draft, p 9), although these authors are not advocating charts like this (I will return to this issue later in this chapter). Each point on the graph is the mean of the times taken by a random sample of 20 patients on the day in question. For a variety of obvious reasons, the times taken by individual patients will vary, with the result that the daily means will also vary; the control lines on the graph are calculated statistically to encompass 99.8% of these random fluctuations. Points within the control lines should be ignored - because the chart suggests that only "common causes" of variations are in play here[1]. On the other hand, points outside, such as Day 7, indicate that there is a "special" cause which should be investigated - in this case to see if anything can be learned from it to improve the process. (This chart is for monitoring the *mean* time; care obviously needs to be taken when comparing these means with the maximum allowable time for an *individual* patient.)

Figure 1: Mean chart of hospital journey time



1

The other techniques in the SPC/M toolbag support control charts in monitoring processes, and in finding ways to improve them. By detecting and dealing with special causes of variation at particular points in time, control charts stabilise a process, but other techniques may be useful to improve the process as a whole. SPC/M can be used for service and administrative processes as well as manufacturing processes, although care needs to be taken when adapting techniques developed in a manufacturing context to another context (Wood, 1994). This means that SPC/M is relevant across the whole spectrum of business processes, and is likely to be an important component of any quality management strategy.

There are a number of assumptions underlying SPC/M which are worth clarifying. The key problem is seen as *variation*: in any process there will always be differences from one hour (or day or month) to the next. In manufacturing contexts such variation is usually unwelcome, and so one aim of SPC/M is to reduce this variation. This is not necessarily the case in a service context (Wood, 1994) - a hospital is not, for example, in a position to reduce the variation between its patients even if it wanted to - but variation is still a problem in that it may hide underlying changes in the system.

This leads on to the second key assumption: the main task is to manage the underlying *process*, and to avoid reacting to chance events - which is likely to be counterproductive. If SPC/M is used for critical processes, this will help to ensure that the whole system will produce adequate and consistent quality levels. It is for this reason that the use of SPC/M is an important part of various quality standards, and TQM initiatives.

SPC/M is specifically concerned with monitoring processes *through time*. Typically, historical data is used to measure the performance of a process, and then a control chart established to monitor changes over time. This is a continuous improvement model, based on the assumption that the best way of improving is to make a detailed study of past performance. There are, needless to say, other possibilities: eg BPR (business process reengineering) is concerned with discontinuous change, and the benchmarking principle involves learning from other processes.

Statistics can also, of course, be used for other quality-related purposes besides SPC/M. Examples include the design of experiments (chapter ??? of this volume), and the SERVQUAL analysis of a police service (Capon and Mills, this volume): this is a useful snapshot at a particular time, but is too complex to be part of a routine monitoring process (although some aspects of it may be).

SPC/M is undoubtedly widely used. To take one example, the Ford Motor Company makes extensive use of SPC/M methods, and requires its vendors to do likewise. There are many other companies, large and small, in a similar position.

The benefits which should follow from SPC/M are similar to those from any other useful quality strategy: improved quality levels, reduced costs, enhanced reputation and improved market share. However, there is a disappointing lack of recent statistical evidence of the extent to which these benefits have been realised. One recent textbook on SPC/M (Oakland, 1999), for example, cites no references to empirical research in the section on *Successful users of SPC/M and the benefits derived*. This is perhaps understandable given the difficulties of research in this area, but it is unfortunate.

AN ILLUSTRATIVE CASE STUDY

This section is based on one of the cases outlined in Wood and Preece (1992). It illustrates some of the difficulties, as well as the benefits of SPC/M.

Company A makes small plastic components by an injection moulding process, typically at the rate of thousands per hour. Each component has a hole in it; the size and shape of this hole is crucial to performance. The quality problems which may occur are that the hole may be the wrong size - this is checked by measuring the rate at which air flows through the hole (the "flow rate"); the hole may have irregularities - checked by a visual inspection; or, most seriously, the hole may not exist at all - this can be picked up by the flow rate or the visual checks.

A number of changes have recently been made to the quality management system. Before these changes, the company operated an inspection system which worked as follows: after each hour's production a sample of, typically, twenty components was chosen at random and inspected visually and by means of a flow rate test, and the entire hour's production was scrapped if any component in the sample did not meet the specification. The inspections were not carried out by the operators, but by a separate inspection department.

There were a number of difficulties with this system. It was very expensive in labour - the number of inspectors was approximately the same as the number of operators; it led to conflicts between the operators and the inspectors; it tolerated a considerable amount of scrap; and the small sample sizes meant that problems with the manufacturing process were not always picked up.

Two changes have now been made. Firstly the quality department has been reorganised: the inspectors have been replaced by a smaller number of auditors whose role is to assist and train the operators so that these operators can carry out the inspections themselves. (Fortuitously, this change occurred at a time when orders were increasing so more operators were needed and no redundancies were necessary.) The operators then have the responsibility for the operation of the machines and the control of quality. This change was welcomed by all concerned.

The second change was the design and installation of equipment to measure flow rates automatically. This equipment could take a large sample of components (typically several hundred instead of twenty), measure the flow rates of each, and then analyse and display the data on a VDU. The software, written specifically for this purpose by a software house, was designed to produce control charts (mean and sigma), capability indices, and to display a histogram of the sample of data and thus indicate whether all components were within the specification.

Before the introduction of the automatic testing equipment there was a certain amount of apprehension that the operators/inspectors would be unable or unwilling to use the new automatic system. These fears proved quite unfounded: after a short induction period the new equipment is being used with few difficulties and considerable enthusiasm - as it replaces a particularly monotonous manual task. However the control charts and capability indices are not used; rather the histograms are used to check for any components outside the specification, and, just as before, each hour's production is only accepted if all components in the sample are within specification. However, the sample size used is now much larger (hundreds rather than tens) which clearly means that the sensitivity of the monitoring process has increased: problems are now more likely to be noticed. However, this was not an *intended* change, but just an accidental byproduct of the changed technology.

The quality manager had considered implementing a "proper" statistical control chart system. (He had experience of a statistical approach to quality in a previous post.) He gave the following reasons for not doing so:

(1) his pressure of work and therefore lack of time;

(2) the operators and setters of the machines would be wary of a system which appears to be tightening the controls;

(3) there was little point, anyway, in tightening the control further;

(4) the flow rate, which, in his view, was the obvious parameter to chart, was not really the critical one. (The system for monitoring flow rates automatically was set up before his arrival.)

(5) it would be difficult to apply standard techniques because different batches for different customers have different specifications, and because a number of other special features of the processes made them difficult to fit into standard techniques.

(6) he was particularly concerned when a defect arose in one of the cavities of a moulding machine which meant that components produced from that cavity had

no hole at all. If this fault was not spotted, and components with no holes were passed as acceptable, this would have very costly repercussions. The problem here is not the problem of monitoring a drift, but of preventing disaster.

All this (particularly the last two points) meant that the situation was seen as a "non-standard" one and so difficult to deal with using the standard textbook techniques for quality control. In fact, the standard techniques can be adapted to cope with the problem of changing specifications, and it is possible to use probability theory to calculate how large a random sample must be to be almost sure of detecting a fault leading to the absence of holes[2]: but the ability to work out things of this type seems most unlikely to result from a short training course in SPC/M techniques.

The impetus for the changes that were made appeared to be the experience of the quality manager who had seen various techniques working in a previous job in the motor industry, and his perception that the existing situation was unsatisfactory. The company experienced only a limited amount of pressure to use specified approaches, such as control charts. Most customers made only very general enquiries about the quality system - which were not sufficiently explicit to drive the company down one direction rather than another.

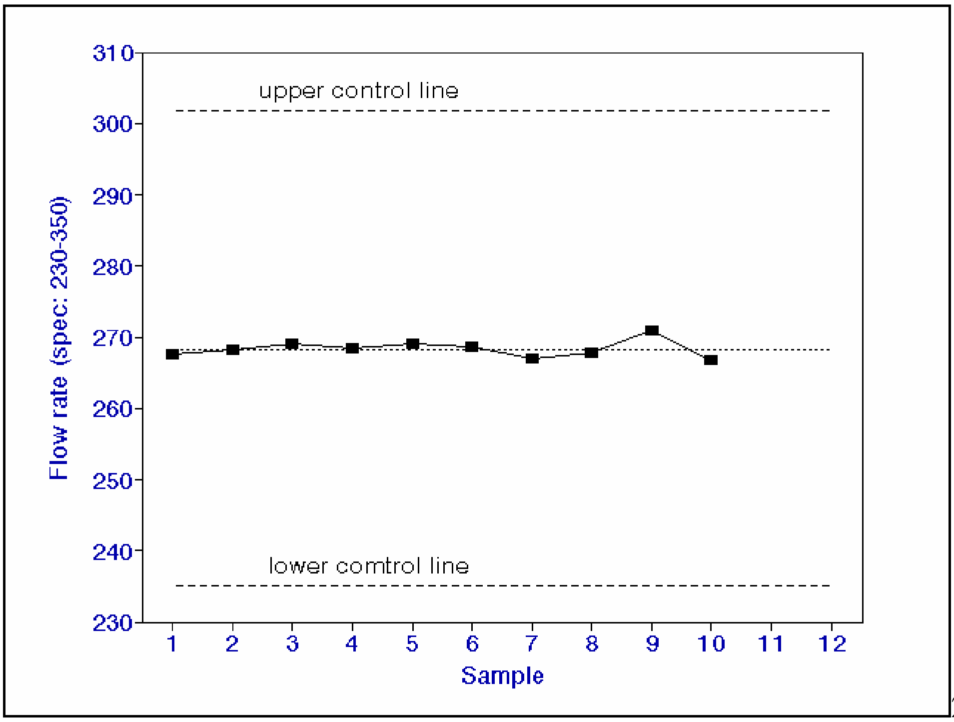
The company has managed to achieve considerable benefits from the new system. The reorganization of the quality department eliminated the conflicts between operators and inspectors by eliminating the inspectors, and, by giving the operator some of the responsibility for the whole process, increased the likelihood of inspection data being used constructively to monitor and improve the process. The monitoring process is undoubtedly cheaper and more efficient now, which has obvious implications for costs and quality.

However, the control charts are not used, and the checking process is still clearly oriented towards inspection and rejection rather than the future oriented attitude of proper process control. This is despite the fact that the process clearly did have room for improvement: one sample of data showed scrap rates ranging from 3 to 12%, with problems picked up by the flow rate test being the largest category.

An example of the type of control chart produced (but not used by the operators) by the automatic testing equipment appears in Figure 2 below. This chart is a simplified version of the one actually produced - which includes a standard deviation chart and capability indices as well. Figure 2 is a control chart monitoring the mean of a sample of 200 components, and clearly shows that the process is well "in control" - ie all points are well inside the control lines. However, the control lines are incorrectly calculated[3]: the correct picture (according to the standard textbook procedure) is as in Figure 3, which shows that the penultimate sample is *out of control* - ie just outside the control lines - and that corrective action should be taken. Figure 3 provides evidence that the process mean has *changed* very slightly, despite the fact that *it is still well within the specification*; this would provide the operator with the means to make the necessary adjustments *before* the process approaches the limits of the specification, thus preventing the production of future scrap. Figure 2, on the other hand, does not achieve this; the control lines here are much too far apart[4]. The erroneous calculation of the control lines was not noticed by the developers of the software, or by anyone at Company A.

Figure 3 is not however, ideal, either. This shows a monitoring system which can detect very small changes and is more sensitive than necessary; this is because the sample size (200) is determined by the capabilities of the monitoring equipment rather than the requirement of process monitoring. We will discuss the implications of this in the next section.

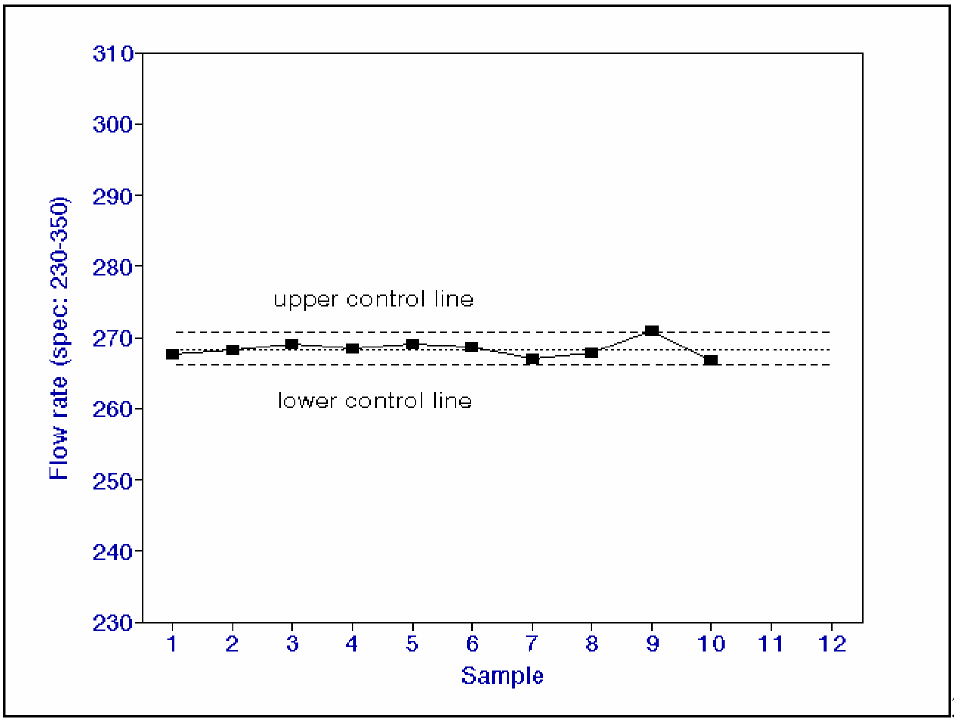
Figure 2: Mean chart similar to actual charts used



2

The estimate of the standard deviation of the individual measurements is 11.1, and the size of each sample is 200.

Figure 3: "Correct" version of the mean chart



3

The estimate of the standard deviation of the individual measurements is 11.1, and the size of each sample is 200.

As we have seen, the quality manager in this company was reluctant to use many aspects of SPC/M. The next section looks at some of the reasons behind this reluctance.

SPC/M IN PRACTICE - PROBLEMS AND SUGGESTED SOLUTIONS

There are a number of inter-related problems, which, for convenience, I will divide into four clusters of issues - each of which can be tackled in ways that are, in principle, straightforward.

Polarised views on the scope of the statistical approach

Statistics is a subject which tends to polarise people. There are those who think that all problems should be tackled statistically, and there are those who think that no problem should be tackled statistically. Needless to say, both of these attitudes are wrong; the best approach is likely to lie between these two extremes. This polarisation of attitudes may be a problem if it leads to lost opportunities.

Figure 1 above shows a mean chart based on a process in a hospital. Jefferson and Humphrys point out (p. 9) that relying on a numerical measure inevitably means that the detail of the particular situation is ignored; this detail is obviously important to any improvement strategy. On the other hand the statistical approach shows general patterns which may alert us to general trends and problems, and may suggest which particular incidents are worth looking at - this is, of course, the whole point of an control chart.

The quality manager in the case study company was reluctant to embrace SPC/M wholeheartedly because he felt - rightly - that SPC/M was useful for detecting a drift in the hole size, but it was unhelpful for the more urgent problem of dealing with components with no hole at all. This latter problem needs a suitable method of sampling[2], and then an investigation of *every* problem found - rather than a statistical analysis of the numbers of such problems. A sensible strategy would have been to have used this tactic, *and* SPC/M for the hole size.

Wood and Christy (1999) distinguish between statistical inferences, and illustrative inferences that can be drawn from specific cases. Both styles of investigation are useful and likely to complement each other: statistical analysis and the detailed investigation of possibilities - such as particular modes of failure (see FMEA chapter, this volume?).

The changing nature of business processes and technology

Control charts were originally designed for manufacturing processes in the 1920s and 1930s. The potential applications of SPC/M now encompass service processes, and manufacturing processes far more diverse than those for which SPC/M was originally developed. In addition, the methods of monitoring are now often automated, as are some aspects of the statistical analysis. This has a number of implications.

The greater diversity of processes means that the standard methods are now less likely to be appropriate: an approach may well have to be crafted for each particular set of circumstances. As an example, Company A's quality manager was correct in thinking that standard methods were not the answer to some of his problems: a customised solution was required[2]. This may be provided by a statistical specialist, although the danger here would be that the users of customised techniques may not have an adequate understanding to use them to their full effect.

The fact that the data for the monitoring process at Company A could be obtained automatically meant that samples were now much larger. This is a general trend (Gunter, 1998; Woodall and Montgomery, 1999): the costs of data may be reduced to almost zero in some cases, so SPC/M may have access to far more data and far bigger samples. With large samples control lines are simply not necessary because sampling error[5] is negligible. The width of the control lines in Figure 3 gives an indication of sampling error: this is obviously negligible in comparison to the specification interval. The sample is large enough for *any* noticeable fluctuation to be meaningful. Control lines were introduced to help distinguish random fluctuations from real change, but with large samples this is not an important problem. This means that control charts would be simple line graphs of a key measure plotted against time - provided that this measure is based on a large sample so that sampling error can be ignored.

Sometimes, there may be too much data. The problem now is one of monitoring so that important trends are noticed without being overwhelmed by the quantity of data. Gunter (1998) suggests that new techniques for data reduction and visualisation should become part of the new SPC/M toolkit for dealing with "data smog".

Another possibility raised by new technology is that of making automatic *adjustments* to the process, as well as *monitoring* it. This goes some way beyond traditional SPC/M which relies on human intervention to react to the signals produced by the monitoring system. The relationship between the two approaches is discussed by Box and Kramer (1992). There are technical problems in incorporating methods for making automatic process adjustments into SPC/M systems (discussed by Box and Kramer, 1992 - the details are too complex to summarise here), but in some contexts extending SPC/M in this direction may be very useful.

Difficulties of understanding statistical concepts and consequent misinterpretation, misuse or non-use

Some of the authors cited in the introduction draw attention to the difficulties of learning and understanding statistics, and the problems this causes. This would be confirmed by the comments of many students who have studied statistics: statistics is a discipline which is hard to master. In the case study above, the fact that nobody at the company picked up the fact that the control charts were incorrectly calculated (and the difference between Figures 2 and 3 is not trivial) illustrates this problem, which is exacerbated by the fact that the error was not recognised: the quality manager thought he understood correctly even when the problem was pointed out to him. In another of the case studies in Wood and Preece (1992), reference was made to the "horrendous" difficulties of understanding even seemingly elementary statistical concepts such as the standard deviation. Needless to say the "skewness and kurtosis checks of normality" created even more difficulty. And this was despite an extensive training programme incorporating sophisticated technology such as interactive video.

An obvious response to this problem is to try to make the statistics easier. This can be achieved by getting employees to learn and practise "the simple mechanics of the X-Bar-R chart without being exposed to derivations, exceptions ... [or] 'standard deviation' or 'rational subgroups'" (Evans, 1987, p. 37). Those who have difficulty with calculation can be tutored or can ask someone else to do their calculations.

The difficulty with this approach is that it is likely to lead to people carrying out computations by rote without any understanding of the underlying rationale. In essence, people are trained to do what a computer could do better. It does not address the problem of the quality manager at Company A who had insufficient appreciation of the rationale behind the formulae to see the error in the software.

This approach does not tackle misconceptions on the conceptual side. The most serious of these is the typical misconception surrounding the word "control". The phrase "in control" is often viewed as meaning that the process manager has *real* control: ie the necessary control to ensure that the process delivers what its customers want. This, after all, is what the term "control" implies in ordinary English. In fact, however, the word control, and the control lines, refer to a statistical fiction: namely the behaviour of the process when it is influenced only by "common causes of variation" and is behaving in a normal, but to some extent, random, manner.

The notion of statistical control as the state of a process under the influence of common or random causes of variation only is a subtle one. Levi and Mainstone (1987: 26) cite evidence that "people tend to perceive patterns and meaning in random events, and to impute more predictability to events than is warranted", which suggests that control charts are likely to be difficult to understand because they conflict with deep seated intuitions (although the fact that these intuitions need to be corrected does indicate the importance of the charts' message). What exactly are the common or random causes of variation which are incorporated in the notion of statistical control? In fact, the answer depends on the control chart model that is used (Wood et al, 1999), which means that the formulae and the interpretation are inextricably interwoven, and are far more subtle than SPC/M trainers tend to assume.

One way of tackling difficulties of this kind is to revise or rename the concepts. For example, the "in control" state of a process could be called "ordinary conditions", the corresponding zone between the control lines called the "expected zone", and the control lines themselves could be called "surprise limits" (Wood et al, 1998). Then, of course, statistical process *control* needs renaming: statistical process monitoring being an obvious alternative. Similarly, process capability indices[6] could be replaced by something along the lines of "predicted defectives per million" (Wood et al, 1998), and the (deliberately?) mystifying slogan "six sigma" replaced by the equivalent, but more transparent, "one defective in one thousand million"[7].

It is also possible to use methods which are more transparent to non-statisticians than the conventional ones. Some suggestions are made in Wood (1995), Wood et al (1998) and Wood et al (1999). Some of these suggestions are rough approximations to standard methods, but others, chiefly those based on resampling (Diaconis and Efron, 1983; Gunter, 1991; Gunter, February 1992; Gunter, April 1992; Simon, 1992; Wood et al, 1999) are often as rigorous and accurate as conventional methods, and sometimes more so. Resampling methods tend to be used by professional statisticians because they perform *better* than standard methods, or will provide answers to questions which standard methods cannot. It is against this background that Jones and Woodall (1998) conclude, on the basis of a simulation study of several resampling methods for control charting, that they "do not perform substantially better than the standard method ..." (p. 374), but, by implication, they might perform slightly better.

To illustrate this, resampling could be used to establish the control, or "surprise", limits for Figure 3. The procedure is to use the data on which Figure 3 is based to generate a large number of random samples of the appropriate size (called resamples for obvious reasons), work out the mean of each such sample, and then use the resulting distribution to assess the variability of sample means under "ordinary conditions", and to read off the appropriate "surprise" limits. This is easily done with a computer; the process is entirely transparent and free of all statistical concepts except the mean, percentiles and tally charts. In particular, no mention is made of standard deviations or the normal distribution. Essentially the same approach can be used to set up limits for a range chart, *p* chart, *c* chart (although this is not quite so straightforward), or for many other possibilities (see Wood et al, 1999, for more details). This approach has the advantage of flexibility: it can, for example, easily be adapted to the case study quality manager's problem of deciding how large a sample is required to be reasonably sure of picking up the problem of the absence of holes in the components[8].

Of course, if the samples are large and it is cheap to gather data, and control lines are deemed unnecessary for the reasons discussed above, then none of these techniques are called for. On the other hand, useful statistical analysis may benefit from going beyond the standard techniques to topics such as - to take the examples cited above - autocorrelation (Woodall and Montgomery, 1999), or data reduction and visualisation techniques (Gunter, 1998). Stoumbous et al (2000) point out two gaps: one between applications and developments published in applied research journals, and the other between these and theoretical statistics journals. Most practitioners are receiving little benefit from recent developments: if they are to do so the educational problem becomes even more urgent.

Conflicting purposes and interests

Another explanation for the relative lack of success of many SPC/M applications is the possibility that SPC/M may not sit comfortably with the predominant culture of an organization (Bushe, 1988; Hoerl and Palm, 1992; Preece and Wood, 1995): the time scale for SPC/M to produce benefits may be too long, or there may be a reluctance to use a technique which shows up problems, for example. These points may result is SPC/M being ignored or abandoned, or it may lead to SPC/M techniques being (mis)interpreted in ways which suit the interests of particular stakeholders. Such reinterpretation is more likely to occur if there is little real appreciation of the nature of the underlying concepts - as we suggested above was often the case.

The espoused philosphy of SPC/M emphasises its role in understanding a process, diagnosing potential problems in time to prevent defective output, and in recognising opportunities for improvement. In practice, however, it may have two other functions, which conflict with this role: its role in quality assurance, and its role for demonstrating that a process is "in control".

Given the potential benefits of SPC/M, its use is sometimes seen as ensuring satisfactory quality levels. This is the (perfectly sensible) rationale behind its incorporation into various quality standards and the insistence of some organizations that suppliers use SPC/M. However, the difficulty in practice is that this may result in the application of the letter of the law (of SPC/M) but not its spirit. Company A had no such pressure, and so did not bother to set up a formal SPC/M system. However, Company B, a supplier of the Ford Motor Company (see Wood and Preece, 1992) carried out various statistical calculations but made little use of them. Their motivation was to satisfy Ford that they were implementing SPC/M, not to gain any of the direct benefits of the application. As they also had problems understanding and interpreting the statistical techniques, this was perhaps inevitable. This corruption of the spirit of SPC/M is clearly less likely if the techniques and the ideas underlying them are properly understood by all parties concerned.

The second, related, potential "function" of SPC/M is as a way of demonstrating that a process is "in control". The suggestion that the operators at Company A would be wary of the control imposed by control charts is clearly based on this assumption. If the key goal of the SPC/M application is to demonstrating that the present and past process is in control, rather to monitor the process to diagnose problems and improve the future process, it is hardly surprising if the SPC/M methods are manipulated to achieve this end, possibly at the expense of their effectiveness at improving the future process. For example, Figure 2 was undoubtedly a more comfortable picture to the operators and managers at Company A, than was Figure 3, because "out-of-control" signals were clearly much less likely. This seems a plausible explanation for the refusal to accept that Figure 3 was the correct version. In general, smaller sample sizes produce charts with wide control lines like Figure 2. If Company A had decided to use control charts, they could have ensured charts with the "control" lines a comfortable distance apart by using small samples. This, however, clearly makes little sense from the diagnosis point of view, since the equipment measuring the flow rates will cope with large samples, and obviously larger samples provide more information. One means of resolving this conflict would be to perform two *separate* analyses: one for future oriented diagnosis and monitoring, and a second, separate analysis for present and past oriented performance measurement (Wood, 1994).

Similar issues arise in relation to Figure 1. Jefferson and Humphrys (this volume, p 10) point out the danger of numerical measures being manipulated to suit particular interests. This danger might be reduced by having separate measures that are used for diagnosis purposes via SPC/M charts, but which are not used for audit purposes.

CONCLUSIONS

Statistical monitoring of processes is as necessary now as it has ever been. However, techniques which made good sense seventy years may now be problematic from various points of view. Traditional methods may need adapting to the modern context. We can distinguish four main clusters of problematic issues for the implementation of SPC/M in the 21st century.

(i) The first is that of clarifying the role of the statistical approach. Statistics almost always has a role to play, but there is also almost always a role for non-statistical methods. The quality manager of the case study company could have used SPC/M to monitor gradual drifts in the hole size, but he also needed to have a system which was capable of immediate detection of the problem of components being made without any hole.

(ii) Difficulties in interpreting and using the concepts and methods are more important than is often supposed and need attention. This is partly a matter of education, but there is also scope for adjusting terminology to clarify the nature of the technique (eg avoiding the word "control" in control charts), and introducing more transparent methods. The goal should be that of ensuring that all parties appreciate how the statistical methods should be used and their results interpreted. There are many contexts in which benefit would be gained from the use of more advanced techniques than those traditionally used, which can only exacerbate this problem.

(iii) SPC/M is sometimes viewed as a means of demonstrating quality to third parties. This may lead to manipulation of methods and data to give a favourable impression. Similarly SPC/M might be seen as a means of tightening control by management (eg this was one reason given by the quality manager of the case study company for not using the control charts). The diagnostic and improvement functions of SPC/M are far more likely to be achieved if monitoring schemes for improving future performance are kept separate from those used for measuring past performance.

(iv) Finally, the changing business and technological environment means that SPC/M needs to be applied in an ever greater diversity of circumstances. The traditional toolkit of techniques is no longer likely to be adequate. Flexible, transparent methods are particularly valuable as they can be adapted to each particular context. Alternatively, new approaches must be crafted for each particular situation. The problem of determining a suitable sample size for the case study quality manager to detect the production of components without holes illustrates the two possibilities well. Note [2] is the standard mathematician's approach, whereas note [8] is a general purpose simulation approach adapted to this context.

On the other hand, some changes may make the statistical problems easier. When automation means that data samples are large, the traditional problem of computing control limits - the source of so many statistical problems, misconceptions and distortions - disappears because sampling errors are negligible and control lines are not worth plotting.

NOTES

1. Sometimes inner "warning" lines are included at the 2.5 and 97.5 percentiles, and the process investigated if two successive points fall outside the same warning line.

2. The problem of changing specifications in different batches can be catered for by taking the deviation from the specification as the variable; and the required sample size is given by

$$log(p)/log(1-1/c)$$

where p is the probability of *failing* to find the fault and c is the number of cavities on the moulding machine.

3. The formulae used for the control lines in figure 2 are

overall mean - 3s and *overall mean + 3s*

instead of

overall mean - $3s/\sqrt{n}$ and *overall mean +* $3s/\sqrt{n}$

where n is the sample size and s is the sample standard deviation.

4. It is true that the control lines are inside the specification limits, but they are close. If the mean were to drift down to the lower control limit, the variation between individual components would mean that some components would be outside the specification.

5. Sampling error refers to the error, when drawing conclusions from a sample, which is attributable to the normal variation between one sample and another. Obviously, sampling errors will usually be smaller with large samples: this is the whole point in taking a large sample.

6. These indices are distinctly odd. For example a value for c_{pk} of 1 indicates a defective rate of around 0.27%. The reason for this odd equivalence is bound up in the mathematics of the normal distribution, which is surely an irrelevance.

7. Six sigma means that the tolerance limit is six standard deviations (sigmas) away from the mean. This corresponds to this probability level if the process is centred, as the reader should be able to confirm using the normal distribution function in a spreadsheet. If the process is not centred the level of defectives will obviously be larger.

8. If there are c cavities, of which one is defective, the proportion of the output which is defective is $1/c$. This fact can be used to simulate a batch of output, from which a sample of, say, 50 components can be drawn. If this process is repeated, say 1000 times, the probability that samples of 50 will find the fault can easily be estimated. If this is unacceptably low, the simulation could be run again for a larger sample size. This crude, trial and error approach, is not as efficient as the formula given above, but its rationale is entirely transparent, and it could easily be adapted to different circumstances.

REFERENCES

Bissell, D. (1994). Statistical methods for SPC/M and TQM. London: Chapman and Hall.

Bushe, G. R. (1988). Cultural contradictions of statistical process control in American Manufacturing organizations. Journal of Management, 14(1), 19-31.

Box, G., & Kramer, T. (1992). Statistical process monitoring and feedback adjustment - a discussion. Technometrics, 34(3), 251- 267.

Dale, B. G., & Shaw, P. (1991). Statistical process control: an examination of some common queries. International Journal of Production Economics, 22, 33-41.

Diaconis, P., & Efron, B. (1983, May). Computer intensive methods in statistics. Scientific American, 248, 96-108.

Evans, W. D. (1987, March). Statistics-free SPC/M. Manufacturing Systems, 34-37.

Gunter, B. (1991, December). Bootstrapping: how to make something from almost nothing and get statistically valid answers - Part 1: Brave New World. Quality Progress, 97-103.

Gunter, B. (1992, February). Bootstrapping: how to make something from almost nothing and get statistically valid answers - Part 2: the Confidence Game. Quality Progress, 83-86.

Gunter, B. (1992, April). Bootstrapping: how to make something from almost nothing and get statistically valid answers - Part 3: examples and enhancements. Quality Progress, 119-122.

Gunter, B. (1998, April). Farewell Fusillade: an unvarnished opinion on the state of the quality profession. Quality Progress, 111-119.

Hawkins, D. M., & Olwell, D. H. (1998). Cumulative sum charts and charting for quality improvement. New York: Springer-Verlag.

Hoerl, R. W., & Palm, A. C. (1992). Discussion: integrating SPC and APC. Technometrics, 34(3), 268-272.

Jones, L. A., & Woodall, W. H. (1998, October). The performance of bootstrap control charts. Journal of Quality Technology, 30(4), 362-375.

Levi, A. S., & Mainstone, L. E. (1987). Obstacles to understanding and using statistical process control as a productivity improvement approach. Journal of Organizational Behavior Management, 9(1), 23-32.

Mitra, A. (1993). Fundamentals of quality control and improvement. New York: Macmillan.

Montgomery, D. C. (1996). Introduction to statistical quality control (3rd edition). New York: John Wiley & Sons.

Oakland, J. S. (1999). Statistical Process Control (4th edition). Oxford: Butterworth-Heinemann.

Preece, D., & Wood, M. (1995). Quality Measurements: who is using the sums and for what purpose? Human Resource Management Journal, 5(3), 41-55.

Rodriguez, R. N. (1992, October). Recent developments in process capability analysis. Journal of Quality Technology, 24(4), 176-187.

Simon, J. L. (1992). Resampling: the new statistics. Arlington, VA: Resampling Stats, Inc.

Stoumbous, Z. G., Reynolds Jr, M. R., Ryan, T. P., & Woodall, W. H. (2000, September). The state of statistical process control as we proceed into the 21st century. Journal of the American Statistical Association, 95(451), 992-998.

Wood, M. (1994). Statistical methods for monitoring service processes. International

Journal of Service Industry Management, 5(4), 53-68.

Wood, M. (1995). Three suggestions for improving control charting procedures. International Journal of Quality and Reliability Management, 12(5), 61-74.

Wood, M., Capon, N., & Kaye, M. (1998). User-friendly statistical concepts for process monitoring. Journal of the Operational Research Society, 49(9), 976-985.

Wood, M., & Christy, R. (1999). Sampling for possibilities. Quality & Quantity, 33, 185-202.

Wood, M., Kaye, M., & Capon, N. (1999). The use of resampling for estimating control chart limits. Journal of the Operational Research Society, 50, 651-659.

Wood, M., & Preece, D. (1992). Using quality measures: practice, problems and possibilities. International Journal of Quality and Reliability Management, 9(7), 42-53.

Woodall, W. H., & Adams, B. M. (1998). Statistical process control. in H. M. Wadsworth Jr. (ed), Handbook of statistical methods for engineers and scientists (2nd edition), (pp. 7.3-7.28). New York: McGraw-Hill.

Woodall, W. H., & Montgomery, D. C. (1999). Research issues and ideas in statistical process control. Journal of Quality Technology, 31, 376-386.