Journal of Intelligent Manufacturing (1993) **4**, 245–257

Quality control expert systems: a review of pertinent literature

TSUANG KUO and ANIL MITAL

Industrial Engineering, University of Cincinnati, Cincinnati, OH 45221-0116, USA

Received April 1992 and accepted July 1992

Statistical quality control (SQC) is an effective tool that ensures quality products and services by means of control charts, the essence of SQC, and sampling plans. While the computation of sample statistics and the development of control charts are routine exercises, the interpretation of chart patterns, trends and the associated diagnosis of assignable causes requires expert knowledge. The present trend is to develop a quality control system and apply it throughout the company (company-wide quality control – CWQC or total quality control – TQC). This frequently means involvement of non-quality personnel in QC teams. Additionally, many companies are faced with a shortage of experienced quality controllers and individuals who can train and educate others on statistical quality control techniques. Quality control expert systems (QCESs) are considered as one way to alleviate these difficulties. In recent years, quality control expert systems have attracted the attention of both quality researchers and practitioners. This paper reviews existing quality control expert systems and recommends a set of quality engineering techniques that should be used to form a knowledge base, the heart of an expert system.

Keywords: Knowledge-based expert systems, statistical process control, statistical quality control

1. Introduction

Minute variations or differences always exist in any production process, regardless of how well designed or carefully it is maintained. Most variations are essentially small and uncontrollable and are usually referred to as natural variability or 'background noise'. When only background noises exist in a process, it is considered to be acceptable, or in statistical control. Other kinds of variability may also be occasionally present in the output of a process. The sources of such variability, which are generally large when compared to the background noise, are referred as 'assignable causes'. A process that is operating in the presence of assignable causes is said to be out of control (Western Electric, 1956; Montgomery, 1985; Wadsworth *et al.*, 1986; Banks, 1989).

A typical control chart contains a center line, called the center line (CL), and two other horizontal lines, called the upper control limit (UCL) and the lower control limit (LCL). If all the sample values of a characteristic appear within the control limits, the process is considered to be statistically in control. On the other hand, if a point falls outside the control limits or certain unnatural patterns exist, the process is interpreted as out of control. Investigative and corrective actions are required to identify and eliminate the assignable cause(s) responsible for this behavior.

Montgomery (1985) cites five reasons for the long history of use of control charts:

(1) A successful control chart program reduces scrap and rework, which increases productivity, decreases costs and increases production capacity;

(2) Control charts are effective in preventing nonconformity;

(3) Control charts prevent unnecessary process adjustments by distinguishing between background noise and abnormal variations;

(4) Control charts provide diagnostic information through the pattern of points plotted on the control charts;

(5) Control charts provide information about process capability – an index of the process stability over time.

These advantages place an increased emphasis on statistical quality control. In addition, CWQC, or TQC, has become an essential requirement at the present. A shortage of experienced quality control experts (Evans and Lindsay, 1988), the need to train and educate non-quality experts, and the inclusion of non-quality personnel in the TQC team are resulting in the extensive use of computers in many areas. The applications of computers lie primarily in automated test facilities, data collection, storage and analysis, and QCES. Adding expert systems technology has become a necessity to enhance the computers' capabilities, to provide more user friendliness, and to ensure that these ever more automated and integrated systems can be managed by responsible experts and users.

Traditionally, control charts are selected by quality engineers based on their experience, knowledge and standard procedural documents. Alternative charts might be available if certain system parameters change. For instance, the availability and capability of test equipment, software and operators will affect the quality plans differently. Owing to the availability of information on SQC and the process itself and/or the reluctance of people to change, the same charts are commonly used repeatedly (Alexander and Jagannathan, 1986). When the appropriate charts have been selected, the control chart parameters need to be set, i.e. the center line, upper limit, lower limit, subgroup size, and sampling interval. These parameters should be carefully evaluated with respect to statistical and economic aspects (Montgomery, 1985). The statistical approach is most often used in professional work. Empirical solutions for determining control chart parameters are widely used and available from published literature, i.e. ANSI Z1.3 and SQC textbooks. These solutions usually do not result in an optimal setting of the process. Unfortunately, the statistical approach is sufficiently flexible that explicit statistical and/or economic criteria are not usually considered. Successful control chart implementation requires the persistent attention of quality control personnel to 'read' control charts and monitor the patterns of plotted points regularly. Although many unnatural patterns can be recognized easily, some need experience and expertise (Nelson, 1984).

In order to be competitive in the global market, all steps in the use of SQC should be carried out effectively. Therefore, more attention should be paid to extracting and formalizing global quality engineering knowledge in a systematic manner, i.e. information used by the quality engineer should be stored in computers and not in the engineer's mind. Rules of thumb, thus, will be preserved for future usage.

Expert systems are computer programs which mimic the behavior of human experts. From the viewpoint of users, availability, consistency and unlimited testability ahead of deployment are the major advantages of expert systems (Braun, 1990). From the viewpoint of programming, expert systems have two advantages over conventional computer programs – modularity and symbolic manipulation. Modularity means that the domain knowledge can easily be updated without having to modify other portions of the system. Symbolic manipulation enables expert systems to handle knowledge which cannot be cast easily into mathematical formulae, or, with all its details, into conventional algorithmic programs (Barr and Feigenbaum, 1981). All of these characteristics and requirements ensure that quality control applications will benefit from the use of expert systems.

Many existing industrial QCESs are one-of-a-kind systems that are not very likely to apply to other fields. QCESs developed in the academic environment suffer from insufficient knowledge and are experimental in nature; they lack practicality. This paper focuses on the components of a practical general-purpose knowledge base for QCESs, and reviews existing QCESs in that light. Functionally, the knowledge base of a QCES should at least include rules for selection of control charts, construction of control limits, economic design of control charts, interpretation of control charts and diagnosis of assignable causes. Useful quality engineering techniques for these purposes are briefly discussed.

2. Quality control expert systems (QCES) knowledge base

Structurally, expert system programs consist of a knowledge base, a working memory and an inference mechanism. The knowledge base contains facts and heuristics associated with the problem. The working memory is used for keeping track of input data for the particular problem, the problem status and what has been done. The inference mechanism controls how knowledge can be used to reach a solution (Alexander and Jagannathan, 1986).

2.1. *Building-blocks of QCES knowledge base*

One of the essential components of a QCES is the knowledge base which accommodates expertise. The knowledge base plays a critical role in evaluating and comparing different expert systems. Based on the functions of the QCES, the knowledge base may be divided into several parts: (1) selection of control charts; (2) construction of control limits; (3) economic design of control charts; (4) interpretation of control charts; and (5) diagnosis of assignable causes.

2.1.1. *Selection of control charts*

Control chart selection is based on several factors (Alexander and Jagannathan, 1986; Hosni and Elshennawy, 1988). These include:

(1) Nature of quality characteristics – attributes versus variables;
(2) Underlying distribution of quality characteristics;
(3) Type of inspection – destructive versus non-destructive;
(4) Cost of inspection;
(5) Inspection time;
(6) Bulk versus discrete production;
(7) Type of defects;
(8) Production rate;
(9) Sensitivity on shift detection.

The expert system should ask a sequence of questions regarding these factors from the user. The answers are then used to select at least one of the following control charts: x-bar and σ-charts, x-bar and R-charts, c-charts, u-chart, median chart, individual chart, p-chart, np-chart, modified control chart, CUSUM chart, and moving average chart.

2.1.2. *Construction of control limits*

Many iterative steps are involved before a process can be concluded to be in statistical control. First, data are collected in subgroups of 25 or more (ANSI/ASQC Z1.3, 1985). It is a standard practice in the USA to use 3σ control limits regardless of the distribution of the quality characteristic (Montgomery, 1985). Based on the sample statistics, trial control limits are computed. The quality characteristic of interest is plotted against the trial limits. If all the subgroup points fall within the control limits and no apparent trends exist, the process is determined to be in statistical control for that period of 25 subgroups. Otherwise, processes are out of control. Investigative and corrective actions are taken to remove assignable causes. Once these remedies are made, out-of-control sample data are discarded and new data are collected. New control limits are again computed, and the points are plotted. This process often iterates many times before statistical control is achieved.

Precontrol is a simplified procedure for monitoring a process. To be correctly used, it requires three assumptions: first, that the quality characteristic of interest is normally distributed, and that the natural tolerance limits ($\mu \pm 3\sigma$) and the specification limits exactly coincide; secondly, that 1–3% non-conforming is acceptable; lastly, that the process capability ratio is at least 1.15 (Montgomery, 1985). Precontrol divides the specification limits (SL) into quartiles. Each quartile is coded as green, yellow or red from the center line to the specification limits (Fig. 1). The process is considered to be in control if five successive data points fall within the green region. If two consecutive yellows or a red appear, the process is out of control. In this case, actions for identifying and removing assignable causes are taken, the process is reset and the procedure is restarted (MacKertich, 1990).

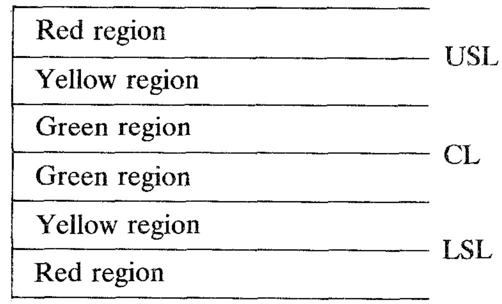

Fig. 1. Precontrol chart.

2.1.3. *Design of control charts*

The design of control charts includes selection of the sample size, the sample frequency and the control limits. Statistical designs usually consider the power of the test and the probability of the type I error which leads to the optimal settings. In many cases, general guidelines are available and widely used (i.e. ANSI Z1.3). Economical designs often utilize operations research methods to minimize expected net cost or maximize expected net profit. The single assignable cause model developed by Duncan (1956) is widely used in optimization of control chart parameters. The assumption is that assignable causes occur according to a Poisson process with an intensity of λ occurrences per hour. The expression $E(L)$, representing the expected loss per hour incurred by the process, is minimized by determining optimal values of the control chart parameters (Montgomery, 1985).

$$E(L) = \frac{a_1 + a_2}{h} + \frac{a_4\,[(h/1-\beta) - \tau + gn + D] + a_3 + a_3'\,\alpha\varepsilon^{-\lambda h}/(1-\varepsilon^{-\lambda h})}{1/\lambda + (h/1-\beta) - \tau + gn + D} \quad (1)$$

where:

α	= probability of false alarm, $\alpha = 2\int_k^\infty \phi(z)\,dz$
$\phi(z)$	= standard normal density
k	= specified values of the upper and lower control limits
h	= time interval for sampling
τ	= expected time of occurrence of single assignable cause between jth and $j+1$th samples
$1/\lambda$	= expected length of in-control period
β	= probability of type II error
δ	= magnitude of single assignable cause
n	= sample size
g	= time required to take a sample
D	= time required to find the assignable cause

a_1 = fixed cost of sampling and testing
a_2 = variable cost of sampling and testing
a_3 = cost of finding an assignable cause
a_3' = cost of investigating a false alarm
a_4 = hourly penalty cost associated with production in the out-of-control state.

2.1.4. *Interpretation of control charts*

Measurements of particular quality characteristics are recorded; sample statistics, such as the mean and range, are computed and plotted on a control chart. Control charts are then examined by an analyst for patterns and trends. A conclusion is reached as to whether or not the process is in statistical control.

2.1.4.1 *Patterns on control charts:* The most common signals on a control chart that indicate a change in the process are (Western Electric, 1956):

(1) Cycles;
(2) Freaks;
(3) Plotted points falling outside control limits;
(4) Gradual changes in level;
(5) Systematic variations;
(6) Trends;
(7) Mixtures;
(8) Abnormal fluctuations.

Some of the unnatural patterns may be recognized easily. Others may need experience and expertise. Nelson (1984) provides eight tests to interpret control charts on a uniform and scientific basis (Fig. 2):

Test 1 – one point beyond zone A;
Test 2 – nine points in a row in zone C or beyond;
Test 3 – six points in a row steadily increasing or decreasing;
Test 4 – 14 points in a row alternating up and down;
Test 5 – two out of three points in a row in zone A or beyond;
Test 6 – four out of five points in zone B or beyond;
Test 7 – 15 points in a row in zone C, above and below the center line;
Test 8 – eight points in a row on both sides of the center line with none in zone C.

Champ and Woodall (1987) provide another set of supplementary runs rules which use probability limits ($\alpha = 0.002$). The runs rule, which signals whether k of the last m standardized sample means fall in the interval (a, b), $a < b$, is denoted by $[t(k, m, a, b)]$:

Rule 1 – C1 = $[T(1, 1, -\infty, -3), T(1, 1, 3, \infty)]$;
Rule 2 – C2 = $[T(2, 3, -3, -2), T(2, 3, 2, 3)]$;
Rule 3 – C3 = $[T(4, 5, -3, -1), T(4, 5, 1, 3)]$;
Rule 4 – C4 = $[T(8, 8, -3, 0), T(8, 8, 0, 3)]$;
Rule 5 – C5 = $[T(2, 2, -3, -2), T(2, 2, 2, 3)]$;
Rule 6 – C6 = $[T(5, 5, -3, -1), T(5, 5, 1, 3)]$;
Rule 7 – C7 = $[T(1, 1, -\infty, -3.09), T(1, 1, 3.09, \infty)]$;
Rule 8 – C8 = $[T(2, 3, -3.09, -1.96), T(2, 3, 1.96, 3.09)]$;
Rule 9 – C9 = $[T(8, 8, -3.09, 0), T(8, 8, 0, 3.09)]$.

These rules increase the type I error. However, the type II error decreases. Montgomery (1985) provides the following expression for computing the overall probability of a false alarm when t separate rules are used to indicate an out-of-control situation:

$$\alpha = 1 - \prod_{i=1}^{t} (1 - \alpha_i) \tag{2}$$

assuming all rules are independent.

Supplementary runs rules cause the Shewhart charts to be more sensitive to small shifts in the mean, but not as sensitive as the CUSUM chart (Champ and Woodall, 1987).

2.1.4.2. *Process capability:* The process capability index is used to show how control charts and other statistical techniques can be used to estimate the natural capability of a process, and determine how it will perform relative to specifications on the product. Two widely used indices are defined as following:

$$C_p = \frac{\mathrm{UCL} - \mathrm{LCL}}{6\sigma} \tag{3}$$

where LCL = lower control limit, UCL = upper control limit, σ = standard deviation of the process population. Another measure of process capability is C_{pk}, defined as:

$$C_{pk} = \min\left(\frac{\mathrm{UCL} - \mu}{3\sigma}, \frac{\mu - \mathrm{LCL}}{3\sigma}\right) \tag{4}$$

where μ = mean of the process.

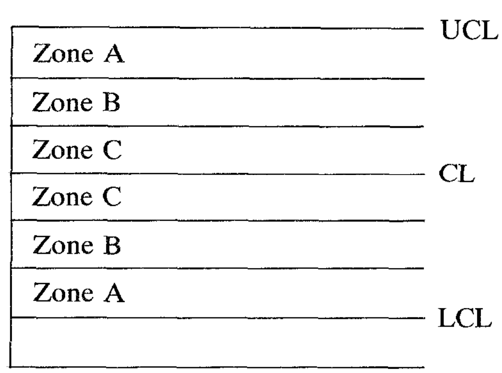

Fig. 2. Control chart.

2.1.4.3. *Regression:* Control charts are used to monitor the process output which detects changes in inputs and brings the process back to the in-control state. The regression model that relates the influence of inputs to process output is helpful in determining the required magnitude of the adjustments (Brillhart and Wible, 1989). Regression analysis provides a line of best fit, $\hat{y}_i = a + b\ (x_i - \bar{x})$, to describe the relationship between inputs (x_i) and outputs (y_i). The unknown parameters, a and b, can be derived by minimizing the sum of the squares of the differences between computed $\hat{y}_i$s and corresponding observed y_is (Ott, 1975).

$$a = \frac{\Sigma y_i}{n} = \bar{y} \tag{5}$$

$$b = \frac{\Sigma\ (x_i - \bar{x})\ (y_i - \bar{y})}{\Sigma\ (x_i - \bar{x})^2} \tag{6}$$

2.1.4.4. *Time-series analysis:* Most methods used for interpreting control charts are primarily based on the assumption that the data are statistically independent or uncorrelated. Such methods are not directly applicable to data which are dependent. The nature of the dependence, or dynamics, distinguishes one model from another. Mathematical models – the autoregressive moving average (ARMA) models – can be used to characterize the dynamics, or the 'memory', of this dependence.

The ARMA model expresses the dependence between x_t and x_{t-1}, x_{t-1} and x_{t-2}, and so on. Term x_t denotes the observation at time t. The model may be written as:

$$x_t = \Phi_1 x_{t-1} + \Phi_2 x_{t-2} + a_t - \Theta_1 a_{t-1} \tag{7}$$

or

$$x_t - \Phi_1 x_{t-1} - \Phi_2 x_{t-2} = a_t - \Theta_1 a_{t-1} \tag{8}$$

where Φ_1, Φ_2 and Θ_1 are constants, $a_t \sim \mathrm{NID}(0, \delta_a^2)$.

This model expresses the dependence of x_t on its two preceding values, x_{t-1} and x_{t-2}, and has an 'autoregressive dependence' of order two. It also includes the dependence on preceding a_t values of order one. This model is called the autoregressive moving average model of order two and one respectively. It is denoted as ARMA (2, 1).

Every ARMA (2, 1) model has its own characteristic polynomial:

$$\Phi(\mathrm{x}) = 1 - \Phi_1 X - \Phi_2 X^2 \tag{9}$$

and a corresponding characteristic equation

$$1 - \Phi_1 x - \Phi_2 x^2 = 0. \tag{10}$$

Since a quadratic equation always has two roots (possibly complex), the roots of the equation can be easily found to be:

$$\lambda_1, \lambda_2 = \frac{\Phi_1 \pm (\Phi_1^2 + 4\Phi_2)^{1/2}}{-2\Phi_2} \tag{11}$$

Terms λ_1, λ_2 are then used to gain quantitative information about the process. The real roots represent constancy, growth or decay trends, and the complex roots represent seasonality with the period given by the imaginary part (Pandit and Wu, 1983).

2.1.5. *Diganosis of assignable causes*

If the process is concluded to be out of control, a determination must be made to identify the location and type of assignable cause for the quality deviation. Because SQC techniques are generic and interpretation may be product specific, no universal knowledge exists for the diagnosis purpose. When a lack of control is detected, the usual recommendation is to 'search for an assignable cause', or 'take remedial action', etc., without any specific advice on how to proceed. However, some general guidelines available in the published literature help to start the diagnostic process (Ott, 1975; Montgomery, 1985; Nelson, 1985):

Test 1 (one point beyond zone A) –
- a shift in the position of the mean;
- an increase in the standard deviation of the process;
  - If a range chart is used and remains in control, an increase in variation can be ruled out;
- a mistake in calculation;
- an error in measurement;
- bad raw material;
- a breakdown of equipment;

Test 2 (nine points in a row in zone C or beyond) –
- a shift in the process average;

Test 3 (six points in a row steadily increasing or decreasing) –
- tool wear;
- depletion of chemical baths;
- deteriorating maintenance;
- improvement of skill;

Test 4 (14 points in a row alternating up and down) –
- two machines, spindles, operators or vendors used alternately;

Test 5 (two out of three points in a row in zone A or beyond) –
- a shift in the process average;
- an increase in variation.

Test 6 (four out of five points in zone B or beyond)
- a shift in the process mean.

Test 7 (15 points in a row in zone C, above and below the center line) –
- stratification problem;

control limits are too wide;
an arithmetic mistake;
Test 8 (eight points in a row on both sides of the center line with none in zone C) –
stratification problem;
control limits are too wide;
an arithmetic mistake.

Process capability indices may be used to monitor the process mean. Useful decision rules, combinations of C_p and C_{pk}, are listed below (Hosni and Elshennawy, 1988):

$- C_p < 1$,	The process is not capable of meeting the specification limits;
$- C_p = 1$, $- C_{pk} < 1$	Shift the process mean;
$- C_{pk} = 1$	The process is running very close to the specification limits;
$- C_p > 1$, $- C_{pk} < 1$	The process is capable of meeting the specification limits, but the process mean needs to be shifted as its setting is incorrect;
$- C_{pk} > 1$	The process is meeting the specification limits and there is some room for the natural variability of the in-control process.

2.1.6. *Concluding remarks*

Although SQC has proven to be a very effective tool to increase productivity and quality, most shop-floor applications are still practicing the $\pm 3\sigma$ control limits with points falling outside the control limits as an action rule. The main reason for using this approach is simplicity, and techniques, such as expert systems, unfortunately, have remained mainly of academic interest.

With advances in computer technology, use of expert systems to monitor and enhance quality should receive a boost. Already data collection and analysis procedures, necessary to develop control charts, have witnessed a move toward automation. Furthermore, a real-time and on-line decision and corrective action can be accomplished by expert system technology. This will lead to more effective, consistent and reliable control of quality in manufacturing and service industries.

3. Features and capabilities of existing QCESs

3.1. *Alexander and Jagannathan (1986)*

The authors point out that the performance of different control charts can be varied in several ways, e.g.: type I and type II errors, sensitivity to detect changes in the process mean and/or variance, and the average run length. These factors have to be considered thoroughly in order to select appropriate control chart(s) for each potential application. An expert advisory system is developed for appropriate control chart selection.

Control chart selection rules are derived from the published literature. Rules for control chart selection are based on the nature of the quality characteristic, such as whether measurements are possible (visual inspection), whether measurements are practical (too expensive), or whether a 100% check is required. Other factors include type of defects, sensitivity requirement of shift, and type I error. The rules are divided into two categories: rules for attribute control charts and rules for variables charts.

The rules to select an attribute chart are as follows:

(1) If an attribute can be classified as defective or non-defective, then a p-chart is selected;
(2) If (1) does not hold, then attribute data are defects-based;
(3) If data are defects-based and the probability of defects occurring in all defect types is considered equal, a c-chart is recommended;
(4) If a c-chart is recommended and the inspection unit is not the same as the production unit, then a u-chart is preferred;
(5) If data are defects-based and defect types are not equally important, then a D-chart is preferred;
(6) If a D-chart is recommended and the inspection unit is not the same as the production unit, then a u-chart is preferred.

The rules to select a variables control chart are as follows:

(1) If a variable chart is selected and the time required to obtain measurements is small and the capability to detect shift is not critical, then x-bar and R-charts are recommended;
(2) If a variable chart is selected and the time required to obtain measurements is small and the capability to detect shift is critical, then a modified control chart with a warning limits chart is recommended;
(3) If a variable chart is selected and the time required to obtain measurements and the capability to detect shift are critical, then a moving average chart is recommended;
(4) If a moving average is recommended and the sensitivity to detect small shift is desired, then a geometric moving average chart is preferred;
(5) If a variable chart is selected and the sensitivity to detect small shifts in average and variance is desired and allowed type I error is small, a CUSUM chart is recommended.

Guidelines for the construction and interpretation of control charts can be obtained through a database query which is not performed by the expert system. Optimization of control chart parameters is suggested as a possible future enhancement.

Knowledge representation: production rule system.
Shell used: GENIE (IQ LISP).
Hardware: IBM PC/XT/AT or compatible.

3.2. *Gipe and Jasinski (1986)*

This Bell Communications Research Laboratory project proposes a three-level evolutionary approach to building expert systems. The first stage is to help the quality engineer in picking out the quality problems and determining assignable cause(s). Stage two is to perform detailed analysis on the vast amounts of data looking for quality problems hidden from the standard analysis routines. The final stage is an integrated system that uses all available data to report on the overall quality effort. Based on a product's previous performance and its component parts, this system is able to enhance standard analysis routines and even predict the field quality of products.

The authors list the following goals:

(1) Analyses of audit data – interpretation of control charts and location of out-of-control samples;

(2) Definition of field performance study (FPS) action items – use of the collected data to identify significant quality problems;

(3) Recommendations for QA activities – suggestions for quality programs for new products;

(4) Discovery of hidden FPS test problems – examination of large quantities of data to search for patterns;

(5) An integrated system – study of products from manufacture through field performance;

(6) An enhanced reliability prediction method – prediction of the performance of equipment;

(7) Enhanced auditing procedures – suggestions for improvements to current auditing procedures.

Hardware: Pyramid 90x super-minicomputer.

Software: Operating system: OSx, a combination of AT&T system V Unix and BSD 4.2 Unix.

Expert systems languages: OPS5 developed at Carnegie-Mellon University, based on LISP. B-EXPERT developed at Bellcore based on C.

Database: Ingres database management system.

Statistical software: 'S' statistical analysis package.

3.3. *Scott and Elgomayel (1987)*

This knowledge-based diagnosis system has two major goals: identification and interpretation of the random and assignable patterns on x-bar and R/s-charts. The source of knowledge is primarily Western Electric (1956). The knowledge base consists of five basic components:

(1) x-Bar and R/s control limits determination – control limits are obtained by either previously established or calculation-based sample data statistics;

(2) Basic control charts testing – control charts are first tested using the 3σ limits rule which locates out-of-control sample(s). Further investigations of runs in the data are implemented for process instability.

(3) Basic control charts interpretation – the R/s chart should be analyzed first and once the state of control is reached, then the x-bar may be analyzed (Western Electric, 1956). The presence of out-of-limits points, unnatural patterns or trends on the R/s-chart is evidence of increased process variability which is usually caused by inaccurate inspection procedures. Another important factor is the level of automation of the process: automated or manual operation. For an automated operation, process variability is a function of machine capability. An increased process variability indicates a fundamental breakdown in the machine. For a manual operation, it usually means an inconsistency in work method, or a lack of operator care or concentration. An out-of-control signal in the x-bar chart reflects a fluctuation of the process average. Frequently, a shift indicates a tool wear problem which is an unavoidable part of the process. In such a case, the process should be allowed to operate until a maximum degree is reached. The interpretation system provides two measurements which further evaluate the state of control. With processes in statistical control, histograms are produced by a FORTRAN program in order to understand the underlying data distributions. Process performance is compared with engineering specifications to estimate the percentage of units which exceed specification.

(4) Advanced control chart testing – the process is now further tested for out-of-control signals by the complete set of 15 patterns, identified by Western Electric (1956). These patterns are cycles, freaks, gradual change in level, grouping or bunching, instability, interaction, mixtures, natural pattern, stable forms of mixture, stratification, sudden shift in level, systematic variables, tendency of one chart to follow another, trends, and unstable forms of mixture.

(5) Comparison of process capability against engineering specifications – The process is determined to be capable of meeting specifications if the following two conditions hold true:

(a) $2 * (M_1 + 1)\ \sigma \leq$ upper specification – lower specification where M_1 is the x-bar control limits multiplier, usually M_1 is 3.

(b) the process is centered within $\sigma/3$ of the nominal specification values.

Knowledge representation: LISP.

3.4. *Dagli and Stacey (1988)*

This expert system assists the quality engineer in the selection and design of the best control chart for a given application. During the consultation, the knowledge-based system asks questions to determine the type of attribute to be tracked, the resources available for chart calculation and the information needed from the control chart. The expert system then determines the best control chart for the situation and gives recommendations as to the best application of the suggested chart.

The reasoning process for control chart selection depends on chart type (variable or attribute). For the variable chart, the expert system determines the appropriate sample size depending on the nature of the quality control test (destructive or non-destructive) and the homogencity of thc sample population. It then recommends a control chart with a confidence factor based on the resources available for calculating the various values used in the control charts. Attribute charts are determined by the type of data to be tracked, either defects per unit or number of proportion defective items, and sample variability. Duncan's formula (1956) is used to optimize control chart parameters: sample size, sample interval, and upper and lower control limits values.

The expert system is currently structured to select at least one of the following control charts: p-chart, np-chart, x-bar and σ-charts, x-bar and R-charts, x-bar and moving R-charts, c-chart, u-chart, and median chart.

Knowledge representation: production rule system.

Shell used: M1.

Hardware: IBM PC/XT/AT or compatible.

3.5. *Evans and Lindsay (1988)*

This expert system is designed for interpretation of x-bar and R-charts. The knowledge base is partitioned into three sets: analysis rules, interpretative rules and diagnostic rules.

The analysis rule base consists of 36 rules for automatic determination of out-of-control conditions. These rules are derived from supplemental runs rules. The conclusion reached from the analysis rule base is one of the following: in control, out of control or suspected to be out of control. Lack of statistical control is concluded from the following rules:

(1) A sample point lies outside the control limits;
(2) Two or three consecutive points lie outside the 2σ warning limits;
(3) Four or five consecutive points lie outside the 1σ warning limits;
(4) Six or seven consecutive points lie on one side of the center line;
(5) A run of six or seven consecutive points up or down exists;
(6) Cyclical variation is observable.

Suspect instability in the process is detected by the following rules:

(1) Three consecutive points lie outside the 1σ warning limits;
(2) Six or seven consecutive points lie on one side of the center line;
(3) A run of five or six consecutive points up or down exists.

If the process is lacking control, the interpretative rule base uses these conclusions along with a decision tree to determine the type of pattern in the chart.

The diagnostic rule base uses the pattern found with specific process information to conclude assignable cause(s). Assignable causes include: points at or near control limits, patterns and trends, and cyclic variation. For x-bar charts, diagnosis could be one of these: change in process setting, change in material, minor part failure, tool wear, environmental factors, and rotation of operators. For R-charts, diagnosis includes operator error, poor material, operator fatigue, operator skill improvement, and maintenance cycles.

Knowledge representation: production rule system.

Shell used: EXSYS.

Hardware: IBM PC/XT/AT or compatible.

3.6. *Hosni and Elshennawy (1988)*

This expert system is marketed by the Institute of Industrial Engineers as an IIE Microsoftware Statistical Quality Control package. The scope of application includes control charts selection, interpretation, and diagnosis. Through dialogue with the user and utilizing a decision tree, this system directs users to the appropriate chart(s).

The decision to select a chart is made based on the following factors: qualitative versus quantitative, underlying distribution, type of inspection, destructive versus non-destructive, cost of inspection, inspection time, availability of sampling during production, bulk versus discrete production, in-process versus preprocess inspection, and lot size. Control charts are then selected from individual charts: x-bar and s-charts, x-bar and σ-charts, x-bar and R-charts, p-chart, np-chart, u-chart, and c-chart.

A decision matrix is provided which relates out-of-control signals and possible assignable causes. The signals are:

(1) One point lies beyond the control limits;
(2) Two out of three consecutive points lie close to a control limit;

(3) Approximately two-thirds of the data lie in the middle third (mean $\pm\sigma$) of the control chart;
(4) Five consecutive points lie on the same side of the center line;
(5) A run of at least five consecutive points up or down;
(6) An erratic, non-random pattern;
(7) A series of five or more points lie close to the center line;
(8) One point on the R- or σ-chart lies above the upper control limit;
(9) One point on the R- or σ-chart lies below the lower control limit.

Possible assignable causes are:

(1) A shift in the process parameter of the chart (u, R, s, p or c);
(2) An increase in the process variability (R, s);
(3) A decrease in the process variability (R, s);
(4) Unstable variability, lower precision (R, s);
(5) A drift or trend in the process parameter of the chart;
(6) Increased precision in the process parameter of the chart;
(7) Cyclic or erratic shift;
(8) Increased accuracy in the process parameter of the chart.

The process capability indices c_p and c_{pk} are used to interpret the process performance. Sample size is determined from MIL-STD 414, Normal Inspection, Level IV.

3.7. *Love and Simaan (1988, 1989)*

This knowledge-based system is developed for the detection and diagnosis of out-of-control events on an aluminum strip rolling mill operation. Automatic detection and diagnosis is a comminatory application of statistical process control principles and process knowledge. Data from the manufacturing process are collected into a signal database. A signal of interest is chosen from the database and the system is initiated to produce reports of diagnosed problems in the process.

Based on nonlinear filtering techniques, a two-level procedure for automatic process diagnosis is provided. At the first level, interpretation is carried out in the following steps. Using numerical-signal processing techniques, the raw signals are preprocessed to eliminate noise. A combination of nonlinear filters is then used to isolate particular primitive variations in the signal. These nonlinear filters are a median filter, a slope filter, a horizontal threshold filter, an integrator and an amplitude thresholder. Finally, input data are classified into three features: peaks (impulses), steps (mean-shift), ramps (linear trend).

At the second level a rule-based program diagnoses special cause(s) of variation. The signal interpreter process is written in LISP. The filtered signals are segmented into pieces. A combination of the three features forms a data structure to represent the primitive variations. Combinations (events) are used to identify particular signal objects and transform the signals into their symbolic descriptions. The symbolic information is then used for interpretation.

Each event has an associated rule set which contains the specific cause of variation of that event. Separate rule sets help both conceptual and software modularity which makes it easier to add knowledge when it becomes available. Diagnosis is reached by backward chaining on the appropriate rule set to a given event. Undefined events are automatically reported and new rules may be developed and added to rule sets later.

Knowledge representation: production rule system.
Shell used: FLAVORS.
Hardware: SYMBOLICS (LISP machine).

3.8. *Brillhart and Wible (1989)*

Harris Semiconductor developed a real-time expert system (PREXPERT) for monitoring, characterizing and controlling a front-end photolithographic process for optimal product throughput, quality and yield. PREXPERT is not only capable of detecting out-of-control products but also continuously improving the process aim and tightening the process variation. This application combines the concepts of real-time data acquisition, recency-weighted process characterization and automatic system tuning.

Lot history and product information are stored in a COMETS (Consilium's on-line manufacturing and engineering tracking system) database. The knowledge base of PREXPERT comprises the statistical process control techniques, the customer requirements and the engineering expertise of the photolithographic process. Process capabilities indices (c_p and c_{pk}) are utilized to aim the process mean and to reduce the variability. With the linear regression from exposure energy to the resist measurement and the regression model from the resist measurement to the final etched measurement, PREXPERT is able to predict the final etched dimension with a 95% confidence interval which eliminates the production pilot run. The adjusted R-square of the regression model is 0.95 indicating that 95% of the variability has been accounted for by this model.

This plant is capable of producing over 220 different product types and over 2600 masks. At any given time, there are more than 50 active products and over 700 masks in the line. Most of the products have their own unique process flow. To combine these products into a single control chart, the normalization z-control chart is employed. Currently, run tests about the center line are the only non-random pattern detection; such as, eight points that fall on the same side of the center line on the x-bar chart or the range median of the R chart. Further analysis caused by the run detection is based on two parameters: slope and rms. A line is fitted through the subgroup points and the slope of this line is obtained. Then, the rms value of the detected run is calculated. A decision is made on whether or not to notify the engineer and to its prioritization (process shutdown, immediate notification or daily/weekly report).

Hardware: CAM stations.

3.9. *Kuo (1989)*

This prototype expert system was developed for diagnosing continuous processes if the assumption of normality doesn't exist. In the case of the existence of non-normality in the data, time series is very useful to describe the data dependency. Another advantage is that time series analysis gives precise information on the length of periodicity and the slope of the trend. This feature provides quantitative information which leads to more accurate interpretation and diagnosis.

A two-step algorithm was developed to interpret the control charts. Supplemental runs rules (Nelson, 1984) are first employed to locate out-of-control conditions. If any out-of-control occurs, time-series analysis is then used to obtain quantitative information about the out-of-control process. Parameters for time-series analysis models are estimated by the FORTRAN program, developed by Pandit and Wu (1983). Estimated parameters are then entered manually into the expert system for analysis. Assignable causes are acquired by utilizing information from time-series analysis and generic information provided by Nelson (1985).

The advantage of this two-step algorithm is that the substantial computational efforts required by time-series analysis are avoided, while quantitative information about the process is secured.

Knowledge representation: production rule system.

Shell used: M1.

Hardware: IBM PC/XT/AT or compatible.

3.10. *Rowan (1989)*

Du Pont has determined that process control systems are good for on-line real-time expert systems application in several ways: sensor validation, data reconciliation, process and equipment diagnosis, and model-based expert systems. Process control systems receive data primarily from on-line process sensors and respond in real time to process problems. Individual measurements from sensors are critical because they supply feedback and closed-loop control. The expert system is segmented into a monitoring component and a rule-base component. The monitoring segment continuously scans on-line data for suspicious behavior, including the signal stability of the measurement and reference signal, and the characteristics of internal signals, such as the automatic gain control. With a short-term history of process data, suspicious sensor behavior is detected by specially developed algorithms which compute the rate of change, moving average and standard deviation, and other variables. If a fault has occurred, the sensor validation expert system alerts the operator. The diagnostic expert system reads a 'snapshot' of information from the on-line scanner, processes the data and collects the additional information necessary to diagnose the problem.

The objective of using expert systems for sensor validation is to improve the measurement system's overall reliability and to improve the overall availability of the process measurement. By providing trouble-shooting assistance to the maintenance technician, the mean time to repair is reduced.

Drifting of individual sensors may occur over a long period. Although it may not be detectable by sensor-validation techniques, it can often be detected by data-reconciliation techniques.

Model-based expert systems involve deep knowledge about the process. Models are derived from first-principle relationships of physics and chemical engineering and from empirical models based on statistical regression of process data. Incorporating this type of deep knowledge into an expert system allows for a very compact, precise system design.

Hardware: DEC VAX

Software: G2 (Gensym Corp.)

3.11. *Frerichs (1990)*

This is an application of SPC methods to continuous processes which integrates a real-time process diagnostic expert system with an on-line SPC package. The functionalities of SPC include determination of sample size and frequency, Pareto analysis, and pattern alarms (Nelson, 1984). Rules governing process operation are formulated and form the nucleus of the diagnostic expert system that can both advise the operator and take corrective action during a process change.

Knowledge representation: production rule system.

Shell: EXPERT 90.

Hardware: distributed control system microprocessor.

3.12. *Pandelidis and Kao (1990)*

A knowledge-based system (DETECTOR) for injection modeling diagnostics has been developed and implemented in the field. For representing inexact and incom-

Table 1. Features of existing QCESs

	Selection	*Control limits*	*Design*		*Interp.*				*Diagnosis*
			Stat.	*Econo.*	*A*	*B*	*C*	*D*	
Alexander and Jagannathan	y								
Gipe and Jasinski									
Scott and Elgomayel		y	y		y	y			y
Dagli and Stacey	y		y	y					
Evans and Lindsay					y				y
Hosni and Elshennawy	y				y	y			y
Love and Simaan		y			*				y
Brillhart and Wible		y			y	y	y		
Kuo		y			y			y	y
Rowan					†				y
Frerichs					y				y
Pandelidis and Kao					‡				y

A: Runs rules.
B: C_p.
C: Regression.
D: Time-series analysis.
* Nonlinear filtering techniques (peaks, steps and ramps).
† Closed-loop control systems (rate of change, moving average, σ)
‡ Heuristic field knowledge.

plete information, the authors show a general knowledge representation scheme which is based on fuzzy set theory. All the defects are represented as d_j, $j = 1, 2, \ldots n$, where d_j is one of the possible defects. All possible causes are represented as c_i, $i = 1, 2 \ldots n$, where c_i is one of the possible causes. The term $<c_i, d_j>$ then represents 'c_i can cause d_j'. Note that $<c_i, d_j>$ does not imply that c_i always occurs when d_j is present, but only that c_i may occur. Uncertain information is handled by fuzzy set theory which relates causes and defects to the grade of membership. The fuzzy relationship grade may range from 0 to 1.

The first step in the diagnostic process is to accept the set of observed defects from the user, and associate appropriate weights for their respective severity. The second step is to determine the associated set of possible causes to the observed defects and obtain their associated weight in a matrix form. The minimum cover criterion is used to determine the smallest possible set of causes explaining the defects. The most obvious advantage of this is to narrow the search space in the specified domain and to produce the sequential problem-solving paradigm to seek further information for the human diagnostician.

An example of the syntax of DETECTOR knowledge is given below:

cause_of ([short, shot], modeling, [mold, temperature], low, 0.5)

This expression means
Defect: short shot
Category: molding
Causes: mold temperature
Description: low
Initial weight: 0.5.

The DETECTOR knowledge base consists of five modules: a trouble-shooting guide, inference rules, the construction of plastic variables and machine variables, a material database, and a machine database.
Knowledge representation: PROLOG.
Hardware: IBM PS/2, IBM PC/XT, AT (compatible).

3.13. *Summary*

The features of existing QCESs are summarized in Table 1. The abbreviated column headings are explained as follows: (1) selection, i.e. control charts selection; (2) control limits, i.e. capability to establish control limits from raw data for processes to be controlled; (3) design (i) Stat., i.e. capability to statistically design control chart parameters, such as control limits, sample size and sample frequency, (ii) Econo., i.e. economically design control chart parameters, and (4) Interp., i.e. interpretation. The authors of each system are listed in the first column of each row. Entries in Table 1 indicate that the particular system has the corresponding feature. For instance, Alexander's system has the capability to select only the control chart.

4. Illustration

A typical QCES interactive consultation session is presented (system response is emboldened):

Control Charts Interpretation System

How many subgroups do you have for initial state?
QC pattern >25
Please input data of sample group number 1.
QC pattern: >[228.0, 224.0, 220.0, 231.0]
Please input data of sample group number 25.
QC pattern>[226.0, 225.0, 221.0, 233.0]
Ranges are out of control at initial state
Please input coefficients of ARMA (2,1).
QC pattern>[1.05, −0.05, 1.0]
A stochastic seasonality with period of 9 detected.
Possible causes:

1. **Tool wear**
2. **Depletion of chemical bath**
3. **Deteriorating maintenance**
4. **Improvement of skill**

5. Conclusions

Expert systems can be an effective tool for quality control. However, for an expert system to be useful, it should at least have the following characteristics:

(1) It should select suitable control chart(s) for the process of interest. For example, if the data collected from product manufacturing industries are statistically independently distributed, traditional Shewhart control charts might be a good choice. For industrial processes, however, data are probably statistically dependently distributed and, therefore, multivariate control charts should be considered. The same consideration should be given to interpretation and diagnosis;

(2) The parameters of the selected control chart should be economically designed. It should also be possible to use the QCES to verify existing process parameters;

(3) The QCES should act as an experienced quality control engineer to monitor the process and perform the diagnostics;

(4) Training in quality control techniques should be possible if requested by the user;

(5) The effort needed to fine tune an existing QCES should be minimized by providing an effective utility tool for the user.

References

Alexander, S. M. and Jagannathan, V. (1986) Advisory system for control chart selection. *Computers and Industrial Engineering*, **10**(3), 171–177.

ANSI/ASQC (1985) American National Standards Z1.1, Z1.2, Z1.3.

Banks, J. (1989) *Principles of Quality Control*, John Wiley & Sons, New York.

Barr, A. and Feigenbaum, E. (1981) *The Handbook of Artificial Intelligence*, Vol. 1, William Kaufmann, CA.

Braun, R. J. (1990) Turning computers into experts. *Quality Progress*, Feb., 71–75.

Brillhart, D. C. and Wible, S. F. (1989) An expert system for real-time process characterization and control, in *IEEE/CHMT IEMT Symposium*, pp. 76–83.

Champ, C. W. and Woodall, W. H. (1987) Exact results for Shewhart control charts with supplementary runs rules. *Technometrics*, **29**(4), 393–399.

Dagli, C. and Stacey, R. (1988) A prototype expert system for selecting control charts. *International Journal of Production Research*, **26**(5), 987–996.

Duncan, A. J. (1956) The economic design of $\bar{x}$ charts used to maintain current control of a process. *Journal of American Statistical Association*, **51**, 228–242.

Evans, J. R. and Lindsay, W. M. (1988) A framework for expert system development in statistical quality control. *Computers and Industrial Engineering*, **14**(3), 335–343.

Frerichs, D. K. (1990) Integrating a diagnostic expert system with statistical process control in a modern distributed control system. *ISA*, 1981–1987.

Gipe, J. P. and Jasinski, N. D. (1986) Expert system applications in quality assurance. *ASQC Quality Congress Transactions*, **40**, 280–284.

Hosni, Y. A. and Elshennawy, A. K. (1988) Quality control and inspection: knowledge-based quality control system. *Computers and Industrial Engineering*, **15**(1–4), 331–337.

Kuo, T. Y. (1989) Control charts interpretation system – a prototype expert system for patterns recognition on control charts, Master's Thesis, Ohio University, Athens, OH.

Love, P. L. and Simaan, M. (1988) Automatic recognition of primitive changes in manufacturing process signals. *Pattern recognition*, **21**(4), 333–342.

Love, P. L. and Simaan, M. (1989) A knowledge-based system for the detection and diagnosis of out-of-control events in manufacturing processes, in *Proceedings of 1989 American Control Conference* Vol. 3, Published by IEEE, pp. 2394–2399.

MacKertich, N. A. (1990) Precontrol vs. control charting: a critical comparison. *Quality Engineering*, **2**(3), 253–260.

Montgomery, D. C. (1985) *Introduction to Statistical Quality Control*, John Wiley & Sons, New York.

Nelson, L. S. (1984) The Shewhart control chart – tests for special causes. *Journal of Quality Technology*, **16**(4), 237–239.

Nelson, L. S. (1985) Interpreting Shewhart X-BAR control charts. *Journal of Quality Technology*, **17**(2), 114–116.

Ott, E. R. (1975) *Process Quality Control: Troubleshooting and Interpretation of Data*, McGraw-Hill, New York.

Pandelidis, I. O. and Kao, J. F. (1990) DETECTOR: a knowledge-based system for injection modeling diagnostics. *Journal of Intelligent Manufacturing*, **1**(1), 49–58.

Pandit, S. M. and Wu, S. M. (1983) *Time Series and System Analysis, with Applications*, John Wiley & Sons, New York.

Rowan, D. A. (1989) On-line expert systems in process industries. *AI Expert*, Aug., 30–38.

Scott, L. L. and Elgomayel, J. I. (1987) Development of a rule based system for statistical process control chart interpretation. *American Society of Mechanical Engineers, Production Engineering Division (PED)*, **27**, 73–91.

Wadsworth, H. M., Stephens, K. S. and Godfrey, A. B. (1986) *Modern Methods for Quality Control and Improvement*, John Wiley & Sons, New York.

Western Electric (1956) *Statistical Quality Control Handbook*, AT&T, Princeton, New Jersey.