INTEGRATION OF CONTROL CHARTS AND DATA MINING FOR PROCESS CONTROL AND QUALITY IMPROVEMENT

E.V.Ramana[1] and P. Ravinder Reddy[2]
[1]Deptt. of Mech. Engg., Shadan College of Engineering & Technology, Hyderabad, India
[2]Deptt. of Mech. Engg., Chaitanya Bharati Institute of Technology, Hyderabad, India

ABSTRACT

Controlling the quality of materials, components, batches, products and assemblies during the course of manufacture is probably the most popularly recognized Quality Control activity. A charting and pattern recognition tool has been developed using Visual Basic and MS-Excel to automate the drawing of control charts. This can also enable better visualization of process data and provide automated recognition of basic unnatural patterns responsible for process to go out of control. Data mining has emerged as an important tool in extracting the knowledge from process databases. An out of control condition can trigger an appropriate data mining technique if any unnatural pattern is found in any control chart for a particular quality characteristic of any stage of the process. Process data can be analyzed by independent data mining technique to discover hidden patterns in the parameters that control the manufacturing process to improve the quality of products. Expert system database of the organization may be updated by the extracted knowledge from various stages of the process. An algorithm shall be developed that can control input parameters to various stages of the process so as to achieve its over all performance by making use of comprehensive expert system database of the organization along with the integration of control charts and data mining.

KEYWORDS: *Data mining, control chart, Pattern recognition, data visualization*

I. INTRODUCTION

Data mining is a collection of tools that explore data in order to discover previously unknown patterns. [16] Data mining applications include manufacturing process control, quality control, condition monitoring and fault diagnosis, optimization of manufacturing yield etc. It is a blend of techniques and concepts from statistics, AI, machine learning, neural networks, data visualization, pattern recognition etc.

Control chart may be defined as "a graphical method for evaluating whether a process is or is not in a state of statistical control". It is a chronological (hour-by-hour, day-by-day) graphical comparison of actual product, part or other unit quality characteristics with limits reflecting the ability to produce as shown by experience on the unit-quality characteristics.[5] This paper focuses on application of control charts in monitoring the out of control condition that may occur at any stage of manufacturing process. If any out of control condition is found the reasons for the variation outside the stable patterns shall be discovered by appropriate data mining techniques like association rule, decision tree, regression, neural network etc and corrections are applied to input process parameters to control the product quality. This paper consists of six sections. In Section 2, definition of quality and causes for variation in quality are presented. In Section 3, types of control charts are described. Interpretation of control charts and basic tests that can be considered for pattern analysis on $\bar{X}$ chart are presented in Section 4. In Section 5, the methodology of integrating control charts and data mining for process control is presented. Finally the conclusions of this paper are presented in Section 6.

II. QUALITY

Quality is defined simply as meeting the requirements of the customer and this has been expressed in many ways

- Fitness for purpose or use
- The totality of features and characteristics of a product or product that bear on its ability to satisfy stated or implied needs

2.1 Variation in Quality

Every process and its output are subjected to variation. No two things can be made exactly alike. There are two types of causes responsible for change of dimensions (quality characteristics) on any component manufactured.

- Inherent or Random (chance) sources of variation: This is caused by large number of chance factors and depends on rigidity of machine tool and its process capability etc.
- Assignable causes of variation: This is caused by human factors, tool wear, machine wear, tool getting loose, difference in raw materials, environments which can be eliminated.

Quality control uses inspection as a valuable tool. In quality control programme inspection data are effectively used to take prompt corrective actions to arrest the incidence of defects in the job on hand and to plan the prevention of similar defects in future conceptually. The broad areas of application of quality control are incoming material control, process control and product control. [1, 13, 18]

III. TYPES OF CONTROL CHARTS

There are two fundamental types of control charts [5]

- Measurement or variables charts(of which the most popular are $\bar{X}$,R, S charts) for use when actual readings are taken
- Charts for use with go and no go or attributes are called p, np, c and u charts

Once the quality characteristic to be studied has been determined; data is collected and samples must be properly selected by the principle of rational sampling. Rational samples are groups of measurements, the variation among which is attributable to one system of causes. Examples of non rational sampling include sampling from parallel machines, sampling over extended period of time and sampling from products of several sources.
If the quality related problems are not easily solved in numerical form, then Attribute Control Charts (ACC) are useful. ACC tools are effective for detecting quality improvement or degradation. ACC is used to evaluate the variation in a process where measurement is an attribute i.e. is in discrete or count (e.g. pass/fail, number of defects). In ACC defining sample size is a problem. A fuzzy based model for ACC in multistage processes was developed and the model is solved by GAs (Genetic Algorithms) and the sample size is suggested. [19]

IV. INTERPRETATION OF CHARTS

- When interpreting charts, it is important to start with the R chart and get it under statistical control first because the limits of $\bar{X}$ chart depend on the magnitude of the common cause variation of the process measured by $\bar{R}$. If some points on the R chart are initially out of control, the limits on $\bar{X}$ chart will be inflated.
- The occurrence of trend or recurring cycles in the data pattern can indicate that system experiencing a drift or cyclical change with respect to its mean or range
- Run of points above or below the center line may be indicating small shifts in the mean or level of variability

When ever an out of control condition is indicated, it is important to determine the basic process fault that is producing it. Some useful generic conditions to look for include the following:

- Trends/cycles :- Systematic change in the process environment, worker fatigue, maintenance schedules, wear conditions, accumulation of waste materials and contamination

- Sudden shifts in level :- New machine, die or tooling, new batch of material, change in measurement system and change in production method

The unusual patterns were fairly obvious even through the most careful examination of the charts. Often, special causes of variation produce patterns that are less obvious. Therefore more rigorous pattern examination should generally be conducted. Several useful tests may be performed by dividing the distance between upper and lower control limits into six zones as shown in Fig-1. Each zone is one standard deviation wide. The zones are termed as Zone A, Zone B, Zone C from the upper control limits toward the central line. The lower side is reflected as mirror images. The probabilistic basis for the tests to be discussed using the zones is derived from the normal distribution. Basic tests that can be considered for pattern analysis on $\bar{X}$ chart are presented in Fig.2. How ever some tests are applicable for attribute charts that don't follow the normal distribution. Test 1 works for all types of attribute charts. More over Test 2, Test 3 and Test 4 are applicable for p chart. [8]

LEGEND

UCL: Upper Control Limit

UWL: Upper Warning Limit

CL: Center Line

LCL: Lower Control Limit

LWL: Lower Warning limit

+SL: Upper Sigma Limit

-SL: Lower Sigma Limit

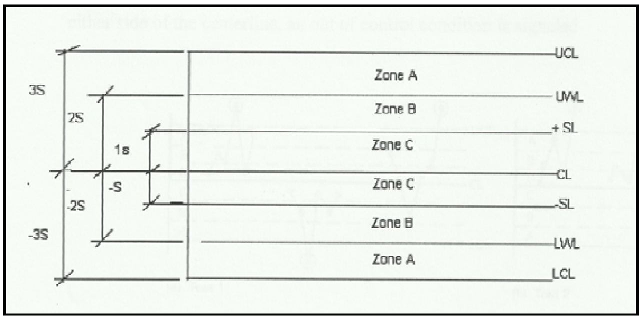


Fig-1 Zone Rules for control chart analysis

Test 1: The existence of a single point beyond Zone A signals the presence of an out of control condition. (Extreme points)
Test 2: Two out of three points in Zone A or beyond signals the presence an out of control condition
Test 3: Four out of five points in zone B or beyond signals the presence of an out of control condition
Test 4: Seven or more successive points either strictly above or strictly below the center line (this rule is applicable to both $\bar{X}$ and R charts)
Test 5: When six successive points on either $\bar{X}$ or R charts show a continuing increasing or decreasing trend, a systematic trend in the process is signaled
Test 6: When 14 successive points, oscillate up and down on either $\bar{X}$ or R charts, a systematic trend in the process is signaled
Test 7: When eight successive points, occurring on either side of the center line with none in Zone C, an out of control condition is signaled
Test 8: When fifteen successive points on the $\bar{X}$ chart fall in Zone C only, to either side of the center line, an out of control condition is signaled

Although these tests can be considered as basic, they are not totally comprehensive. Analyst should be alert to any patterns of points that might indicate the influence of other special causes in the particular process. Tests 1, 2, 5 and 6 are separately applied to the upper and lower halves of the chart. Tests 3, 4, 7 and 8 are applied to the entire chart ($\bar{X}$ and R charts). Points can contribute to more than one test. [1, 8, 13, 18]

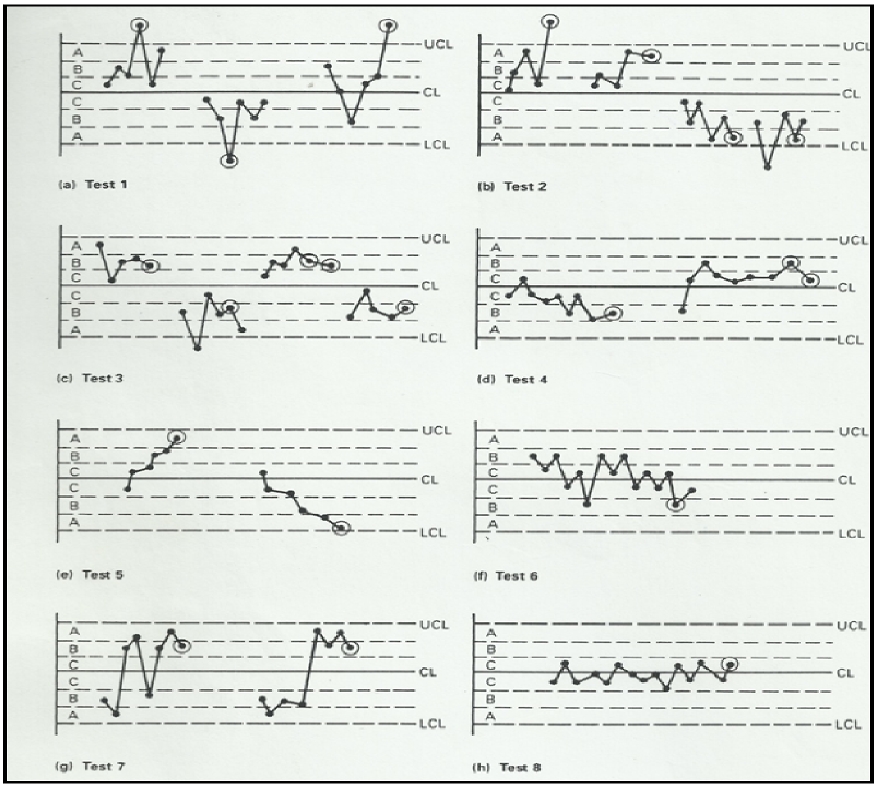
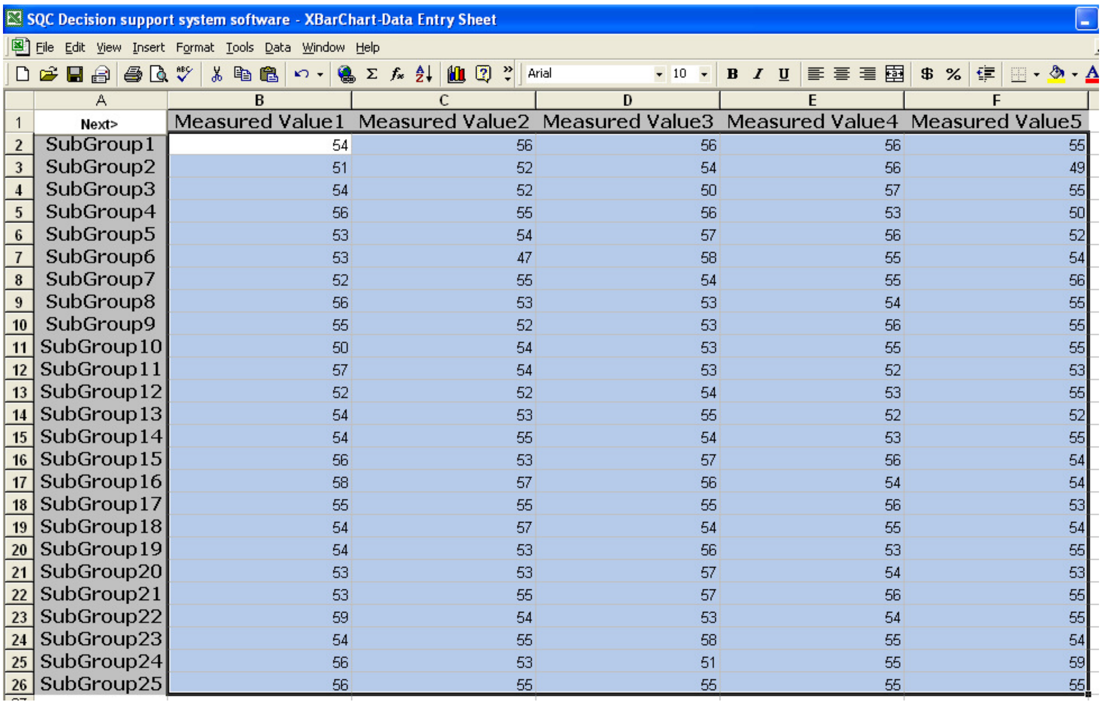


Fig-2 Pattern analysis of $\bar{X}$ charts

V. INTEGRATION OF CONTROL CHARTS AND DATA MINING FOR PROCESS CONTROL

To plot control charts and drawing conclusions from them requires collection and organizing the process data, determining control limits, recognition and analysis of unusual patterns exhibited by the process. More over plotting the charts manually for different quality characteristics becomes tedious work and error prone. This consumes more time and requires greater attention. A charting and pattern recognition tool has been developed using Visual Basic and MS-Excel [7, 9] to automate the drawing of control charts. This tool provide better visualization of process data and enable automated recognition of basic unnatural patterns responsible for process to go out of control.

Typical data entry sheet for manual collection of sample data to draw $\bar{X}$ and R charts is shown in Fig-3. $\bar{X}$ and R charts generated by the charting tool is presented in Fig-4 for typical data. (XDoubleBar: Arithmetic mean of the sample averages, RBar: Arithmetic average of sample ranges) Pattern recognition tool performs tests on collected data to discover basic unusual patterns responsible for unstable process. Passing of a test indicates absence of unusual pattern linked with that test and vice versa. A typical output showing the outcome of tests performed on $\bar{X}$ and R charts is given in Fig-5.


SQC Decision support system software - XBarChart-Data Entry Sheet

Next>	Measured Value1	Measured Value2	Measured Value3	Measured Value4	Measured Value5
SubGroup1	54	56	56	56	55
SubGroup2	51	52	54	56	49
SubGroup3	54	52	50	57	55
SubGroup4	56	55	56	53	50
SubGroup5	53	54	57	56	52
SubGroup6	53	47	58	55	54
SubGroup7	52	55	54	55	56
SubGroup8	56	53	53	54	55
SubGroup9	55	52	53	56	55
SubGroup10	50	54	53	55	55
SubGroup11	57	54	53	52	53
SubGroup12	52	52	54	53	55
SubGroup13	54	53	55	52	52
SubGroup14	54	55	54	53	55
SubGroup15	56	53	57	56	54
SubGroup16	58	57	56	54	54
SubGroup17	55	55	55	56	53
SubGroup18	54	57	54	55	54
SubGroup19	54	53	56	53	55
SubGroup20	53	53	57	54	53
SubGroup21	53	55	57	56	55
SubGroup22	59	54	53	54	55
SubGroup23	54	55	58	55	54
SubGroup24	56	53	51	55	59
SubGroup25	56	55	55	55	55

Fig-3 Data entry sheet for $\bar{X}$ and R charts

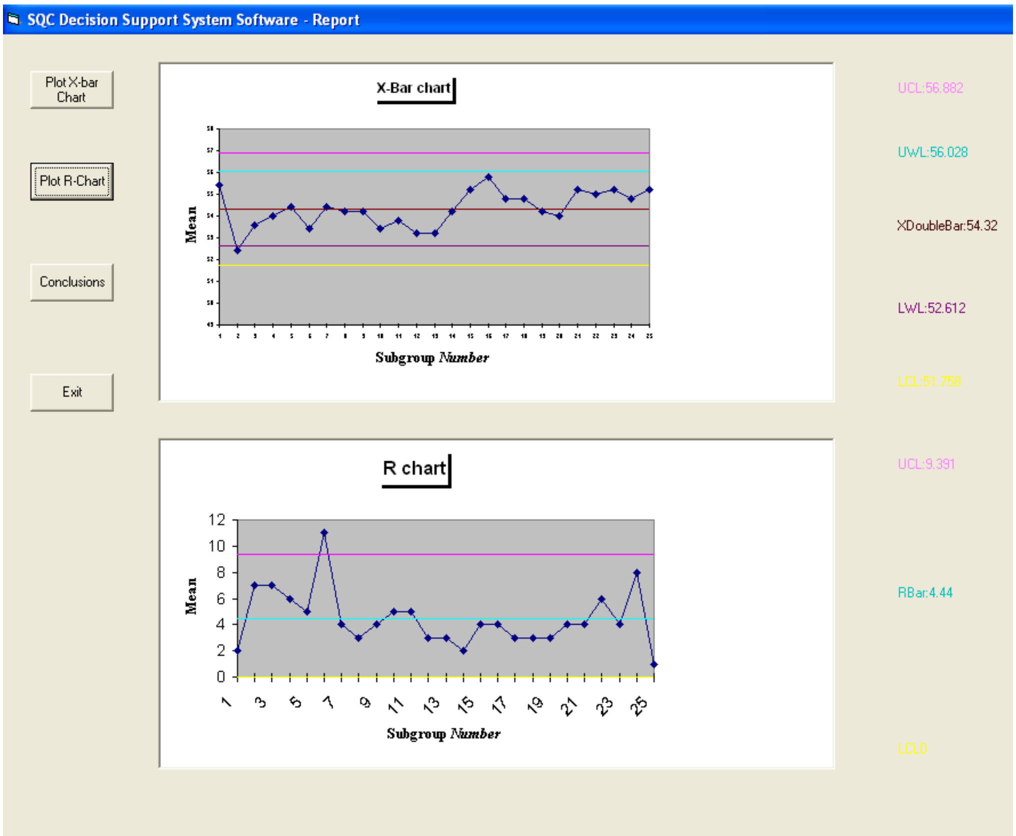


Fig-4 $\bar{X}$ and R charts generated by the charting tool

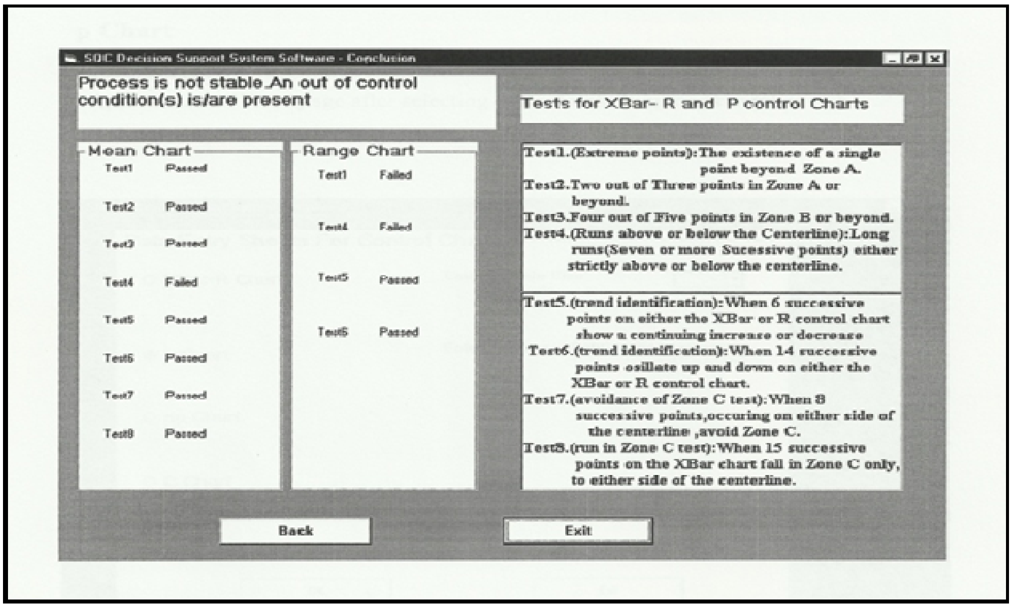


Fig-5 Outcome of tests performed on $\bar{X}$ and R charts

Quality troubles often gradually "drift" into a process. Tool wear deflection, tool temperature or an improperly ground tool may result in a trend which finally leads to production of non-conforming parts. A chart which compares actual production variations with control limits may flag the entrance into the process of this sort of quality trouble before scrap or rework is actually caused. Unusual variation caused by assignable causes may represent temporary difficulties that can be eliminated with out much expenditure. [5]

Classical methods like control charts aim to monitor the process not to infer the relationship between target attribute and input attributes. [16].The goal of mining quality related data is to find the relation between the quality measure (target attribute) and input attributes (manufacturing process data). The idea is to predict the quality measure of a certain product/ batch based on its process parameters.

The control signature is a set of feature values or ranges that lead toward an expected output. Rough set theory was used by Kusiak to determine the association between control parameters and product quality in the form of decision rules and generated control signature from those rules.[15] The manufacturing parameters include the characteristics of the production line such as machine that has been used in each step, how machine has been set up, raw material in the process, environment (moisture, temperature), operators on the production line (experience level of worker, which have been assigned on each machine in the line), the shift number and other significant factors. [16]

Data is collected electronically and stored in databases in modern plants. The volume of collected data makes it impractical to explore and detect intricate relations between parameters of different processing steps using standard statistical procedures. Also comprehensive manual analysis of voluminous data becomes prohibitive, especially on-line monitoring and control of manufacturing processes are considered.

Data can be analyzed by data mining techniques to identify hidden patterns in the parameters that control manufacturing processes or to determine and improve quality of products. [2] "Interestingness" (of pattern) is usually taken as an overall measure of pattern value, combining validity and novelty, usefulness, simplicity and understandability. The measures of interestingness can be classified into objective measure and subjective measure. An objective measure depends on structure of pattern and underlying data used. Subjective measure depends on the class of users who examine the patterns. These are based on two concepts. [17] Unexpectedness (a pattern is interesting if it is unexpected) and actionability (a pattern is interesting if the user can use it to his or her advantage) of the pattern.

Study of entire process data/factory data altogether to discover the problem areas instantly affecting subsequent processes requires consideration of hundreds of attributes simultaneously. During the product design and development process data mining can be used in order to determine "internal" factors at each stage and "external" factors at consecutive and previous stages. Data mining can be seen as a supporting vehicle for determining causal relationships among "internal" factors associated with manufacturing processes and external factors related to the competitiveness of the manufacturing company. Since manufacturing have an inherently temporal context, time factors may be taken into process in order to correctly interpret the collected data. (e.g., from a certain production date the number of defects is much higher than normal)

The deviation of process data from the normal and/or discernible pattern detected over time may have an adverse impact on current and subsequent stages. Extracted knowledge may include classification to predetermined types of deviation from the norm and causal relationships among temporally oriented events. [6]

Different data mining techniques may be applied with independent goals at each stage of the process. The failed batch / yield identified by a date can help in identifying the cause relating to faulty equipment or tool or material etc. Condition based monitoring information could come from either or both on-line condition monitoring data and historical records. Process analysis is concerned with optimizing inter related processes with in an enterprise. Identifying inter relationships among these processes has proven a valuable field of data mining technology. The knowledge extracted by mining process databases at various stages can be integrated with existing expert systems of the organization. Various decision making tools can be networked in support of globally optimal decisions. [3] Data mining tools have shown a profound impact on informed, transparent and autonomous decision making. An algorithm can be developed to integrate overall system performance with the component systems based on extracted knowledge.

Control chart helps in monitoring a particular stage of a process and necessary for the development of process know how. It can trigger a specific data mining technique applicable for that stage if any out of control pattern is detected and expert system database of the organization can be updated with the extracted knowledge. Fig-6 shows the integration of control charts and data mining for process control.

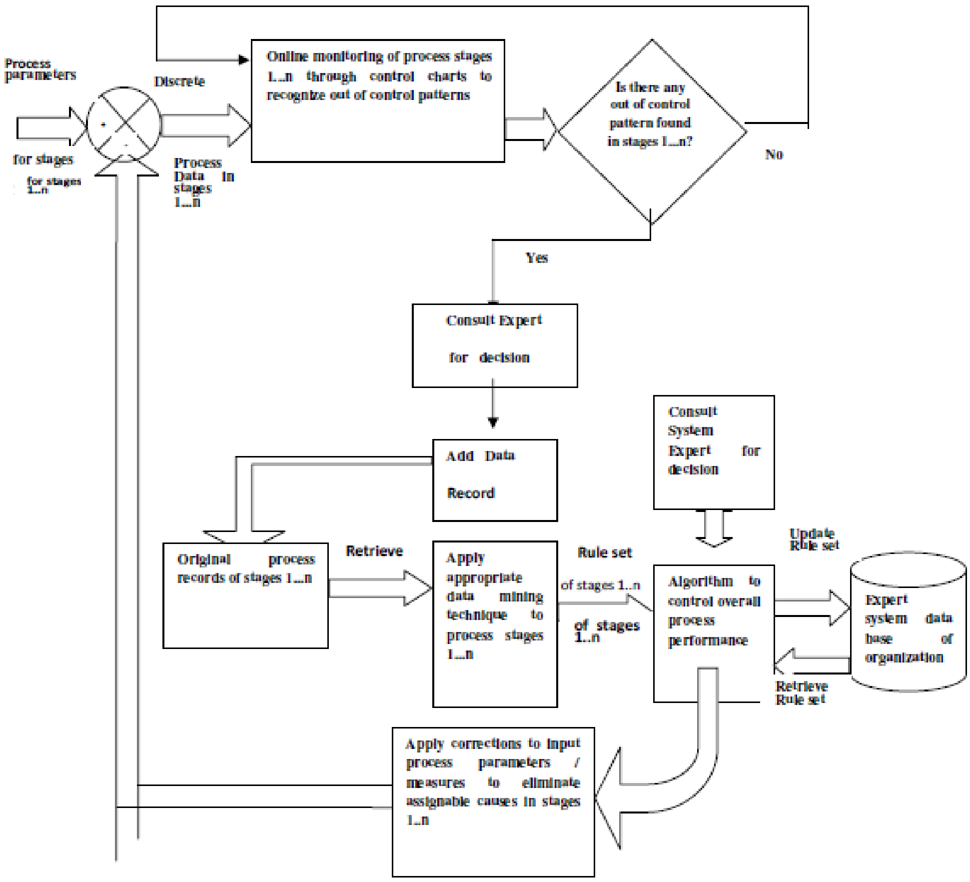


Fig-6 Integration of control charts and data mining for process control

VI. CONCLUSIONS

This paper presented the methodology for the integration of control charts with data mining techniques to achieve process control. Knowledge may be extracted by applying data mining techniques on temporal data collected at various stages of the process; condition based monitoring data of equipment, tools etc and analysis of defects in products/components produced.

Expert system database of the organization may be updated by the extracted knowledge from various sources as mentioned above. An algorithm can be developed to integrate over all process performance (e.g. performance parameters, production indices, yield, company goals etc) with its component stages through comprehensive expert system database. The algorithm is expected to provide optimal process parameters (input) to each stage of the process so as to improve over all process performance.

References

[1] "ASM Handbook: Non destructive evaluation and quality control", ASM International, pp728-732, 1989

[2]A.K.Choudhary J.A.Harding, M.K.Tiwari, Data mining in manufacturing: a review based on kind of knowledge, Journal of Intelligent Manufacturing, Vol. No: 20, pp501-521, 2008

[3]A.Kusiak, "Data mining: manufacturing and service applications", International Journal of Production Research, Vol. No: 44, pp4175-4191, 2006

[4]Alex G.Buchner, Sarabjot S.Anand and John G.Hugghes, "Data Mining in Manufacturing Environments: Goals, Techniques and Applications", University of Ulster, Northern Ireland, UK

[5]Armand V.Feignbaum, "Total Quality Control", McGraw-Hill International Editions, pp 395-409, 1991

[6]Dan Braha (Ed), Data mining for Design and Manufacturing: methods and applications, Kluwer Academic Publishers, 2002, ISBN: 1-4020-0034-0

[7] Evangelos Petroutsos , "Mastering Visual Basic 6.0", BPB Publications, pp10-65,2001

[8]Gavriel Salvendy, "Hand book of Industrial Engineering", John Wiley & Sons, pp2259, 1992

[9]Gene Weisskopf, "Excel 2000 No Experience Required", BPB Publications, pp54-81, 2003

[10]Hovhannes Sadoyan, Armen Zakarian, Pravansu Mohanty, "Data mining algorithm for manufacturing process control", The International Journal of Advanced Manufacturing Technology, Vol. No: 28, pp342-350, 2006

[11]J.A. Harding, M.Shahbaz, Srinivas, A.Kusiak, "Data mining in Manufacturing: A Review," Journal of Manufacturing Science and Engineering, Vol. 128, pp968-976, 2006

[12] Jing Rong Li, Li Pheng Khoo, Shu Bengtor, "RMINE: A rough set based data mining prototype for the reasoning of incomplete data in condition based fault diagnosis", Journal of Intelligent Manufacturing, Vol. No: 17, pp 163-176, 2006

[13]John S.Oakland, "Statistical Process Control" ,Butterworth-Heinemann Ltd, 1999

[14]Kescheng Wang, "Applying data mining to manufacturing: the nature and applications", Journal of Intelligent Manufacturing, Vol. No: 18, pp487-495, 2007

[15] Kusiak, A., "Data mining approach for generation of control signatures" ASME Journal of Manufacturing science and Engineering, Vol. No: 124, pp 923-926, 2002

[16]Lior Rokach, Oded Maimon, "Data mining for improving the quality of manufacturing: a feature set decomposition approach", Journal of Intelligent Manufacturing, Vol. No: 17, pp285-299, 2006

[17]Silberschatz, A., Tuzlhilin, A., "What makes patterns interesting in knowledge discovery systems", Knowledge and Data Engineering, vol. No: 8, pp970-974, 1996

[18]Tom Drozda, Charles Wick, John T. Benedict, Raymond F.Villeux, "Tool and Manufacturing Engineers Handbook("Quality control and assembly)", Society of Manufacturing Engineers, pp2-17, 1986

[19] Orhan Engin , Ahmet C¸ elik , I˙hsan Kaya, "A fuzzy approach to define sample size for attributes control chart in multistage processes: An application in engine valve manufacturing process", Applied Soft Computing, Vol:8, 2008, pp1654-1663

Authors

E. V. Ramana is working as a Professor& Head in the Department of Mechanical Engineering, Shadan College of Engineering & Technology, Hyderabad. He received his first M.Tech degree in Energy Systems and second M.Tech degree in CAD/CAM in the years 1992 and 1997 from JNT University, Hyderabad. He possesses 14 years of experience in teaching and 6 years in industry and research. He is currently pursuing the PhD degree from JNTUH, Hyderabad, India. He has considerable experience in developing mechanical engineering related application soft wares. His active area of research is Data Mining in Manufacturing.

P. RAVINDER REDDY is born in 1965 and he is working as a Professor and Head of Mechanical Engineering, Chaitanya Bharathi Institute of Technology, Hyderabad. He is having 22 Years of Teaching, Industrial and Research experience. Taught Postgraduate and under graduate Engineering subjects. Published Research Papers over 132 in International and national Journals, and Conferences. Guided 5 Ph.Ds and 6 Ph.D scholars submitted their thesis. Guided over 250 M.E/M.Tech Projects and carried out research and consultancy to a tune of Rs. 1.9 Cr sponsored by BHEL, AICTE, UGC, NSTL and other industries. Organized 23 Refresher/STTPs/ workshops, one international conference and delivered 63 invited/ keynote/ special lecturers. Received "UGC Fellowship" award by UGC (1999). Raja Rambapu Patil National award for promising Engineering Teacher by ISTE for the year 2000 in recognition of his outstanding contribution in the area of Engineering and Technology. Excellence "A" Grade awarded by AICTE monitoring committee for the MODROB project sponsored by AICTE in 2002. "Engineer of the year Award-2004" for his outstanding contribution in Academics and research by the Govt. of Andhra Pradesh and Institution of Engineers (India), AP State Centre on 15th September 2004 on the occasion of 37th Engineer's Day. Best Technical Paper Award in the year Dec. 2008 by National Governing Council of Indian Society for Non Destructive Testing.