

PROACTIVE vs. REACTIVE SPC Analysis at LTV Steel

Paul M. Centric, LTV Steel

Robert J. Scharl, LTV Steel

Abstract

Current Statistical Process Control (SPC) uses historic data to try and correct the process. Control would be impossible when the data is made available for analysis after the product has been produced. A better method would be to use real time analysis. Move from a **reactive** mode into a **proactive** mode when real time data is available during production. The use of multiple regression trend analysis with the clustering of variables to predict a potential production condition(s) that require(s) immediate attention is favorable to a single alarm after alert after the process has already gone out of control.

Introduction

LTV Steel began using the principles of SPC in 1982 with formal training of management progressing to all hourly workers. The use of both automated and manual control procedures are being utilized today. Automated controls are being done by process control computers. The control computers monitor the operation and make adjustments to produce a product within the control model. The control model still makes mistakes yielding rejected material. Manual control chart procedures are used for off-line, non automated procedures. It is the automated process control computers that is the focus of real time SPC control moving from a **reactive** mode a day after the product is produced to a **proactive** mode to correct potential problems before becoming a disaster producing defective material.

A extensive computer process control upgrade program began in 1989. Prior data consisted of only the most basic for the physical characteristics of the product, (i.e. thickness, width and temperatures, etc.). The upgrade made available all the data as used by the on line programmable logic controllers. The total number a variables increased to over 1000 for one department. The process control computer summarized the incremental time data elements gathered into one (1) reading per produced unit (coil). The summarized data is transferred immediately to an in plant mini computer and then transferred to the corporate mainframe. The data is 'stored' in a temporary daily file. Once a day, usually between 11:00 PM and 7:00 AM, the temporary data is transferred to a sequential file. Then the SASTM user community transforms yesterdays data into an indexed SASTM data set for statistical analysis.

SAS Institute has developed a procedure under SAS/QCTM(TSHEWHART)TM to analyze real time data. If data is available as the product is being manufactured, then individual process variables or clusters of process variables, can be controlled during manufacture. Various types of alarm systems can signal when the process variable (cluster) is approaching an out-of-control condition. With the potential of making minute process changes, a more uniform and consistent product is the end result.

Methodology

The first requirement is to have process control data that is not captive to the process control computer. This data must be readily available after the process control computer has processed the observation (record). It is possible to work with either RAW or SUMMARIZED data. RAW data will be defined as each and every data point in time sequence as generated by the process control computer. SUMMARIZED data will be defined as the average for that production unit. This analysis is being done with summarized data with an observation (record) being a coil. Our production rate for this study averages 600 records/day, or a record every 2.4 minutes. The data is representative of both physical characteristics and machine control variables. Physical characteristics can be thickness, width and length. Machine control variables can be electrical power draw, bending forces, bearing forces, water volume, water pressure, etc. The machine control variables have a direct influence on the production of the product.

The process control computers will only control the process. No analysis can be done at this level. Their only responsibility is to control all the Programmable Logic Controllers. Data can be 'summarized' from the instantaneous messages into one message per logical process unit (record).

Communications and Hardware terminology

Computers

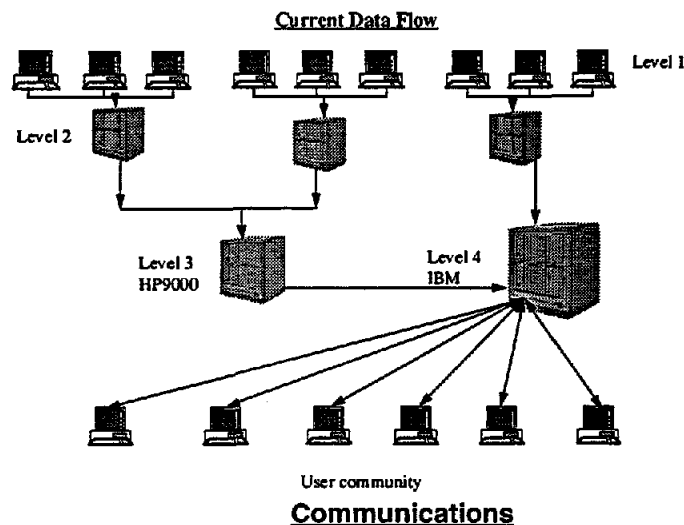
A Level 1 is a Programmable Logic Controller. There are many Level 1 computers/controllers that determine the functions of the mill. They monitor and adjust mechanical or electrical input to the mill and make adjustments as required. No Statistical analysis is done on this level machine.

A Level 2 controls all Level 1 units. A DEC/VAX is the manufacture. It will send instructions to Level 1's

and monitor their feedback for minute changes during production. The Level 2 has the process model for the operation. It can store a limited amount of data and can do a limited amount of data crunching. No statistical analysis is done on this level machine. Some data is stored at this level for a short period of time mainly for temporary backup procedures in case of a temporary communications failure. Long term storage is not provided.

A Level 3 will send and receive messages from Level 2 or Level 4 units. It is typically an HP3000 or HP9000. This machine acts as a communications buffer to convert Level 2 messages into a format readable by Level 4. Statistical analysis can be done on this machine. The Level 3 used is an HP9000, model E55 with 64 MB RAM and 4 GB hard drive.

A Level 4 computer is referred to as the mainframe and is a IBM 3090 series. It has control of the legacy systems such as ORDER ENTRY, CLAIMS, SHIPMENTS, etc. All historic data is maintained in systems as CICS, VSAM, IMS and sequential files. Reporting and analysis functions are executed using SAS or other mainframe languages.



The wide area network (WAN) is a 100 MB FDDI fiber optics utilizing SynOptics concentrators and Wellfleet routers. Primary communication protocol is DECNET with DECmessageQ for message handling.

Local area networks (LAN) is Ethernet. Transmission speed is 10 MB twisted pair copper. Communications protocol is TCP/IP.

The weak link in our system is from the Level 3 to the Level 4. It is a T1 carrier using SNA at 256 KB.

Communications from the Level 4 to individual users, either PC's or mainframe terminals varies between 19.2 KB and 256 KB.

Current Data

The Level 1 programmable logic controllers respond to input signals received from the Level 2 computer to adjust electromechanical devices. They will also feed back output signals for the instantaneous reading for a particular device. There is a constant 'talking' during production between the Level 1 and Level 2 computers. The Level 2 computer manipulates the instantaneous reading taken at a set time interval into an average for the produced unit (coil or record). It may also calculate other basic statistics for the coil such as minimum, maximum, range and standard deviation. The compiled record is given short term storage on the Level 2 in case of a communication problem between Level 2 and Level 3. Level 2 can also do long term storage on tape of the raw data. This is also done strictly for backup or recovery. Level 2 will send the data in one (1) logical record occupying as many bytes as required.

The Level 3 receives the message and may convert it into messages no longer than 1014 bytes. A 'header' record is added so the many split records per coil can be reassembled into a logical record at the mainframe (Level 4). The records are aging temporarily stored until a draw off occurs placing the records into a sequential file, either disk or tape. When the records are finally in a sequential file, the SAS™ community is able to access for production reporting placing the data into an indexed SAS data file(s).

All analysis is done after the fact when defective material has already been generated and the engineers must explain why defective material was produced. Purely a 'REACTIVE' mode for operation. Our goal is 'PROACTIVE' process control.

Real Time Data

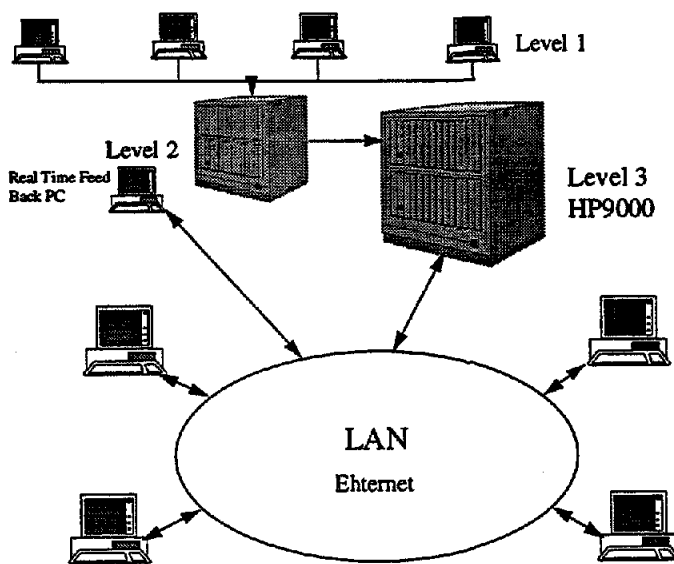
The current Level 1 and Level 2 computer will continue to operate in the same manner.

The duties of the Level 3 computer will be expanded greatly.

- SAS™ will be installed. The packages will be BASE™, CORE™, SAS/CONNECT™, SAS/STAT™ and SAS/QC™.
- A message handling procedure will be installed. The procedure will be written by our local Information Systems. It will have a dual function.

- A message will continue to be sent to Level 4 for next day business analysis, 'Reactive Mode Mentality'.
- A message will be sent to a TCP/IP socket for 'REAL TIME' analysis.
- When the message is received at the TCP/IP socket, a alert is received by the SAS™ processor that new data is being received and an additional observation is to be processed.
- A conventional SAS™ data step is processed for calculation of the new process variables to plot via SAS/QC™, TSHEWHART™.
- TSHEWHART™ is executed calling upon the newly created data set. The options used and putting the graphics to a output catalog are the most critical. The critical options are the 'TESTS=', 'TESTOVERLAP' and 'OUTTABLE='.
- The 'OUTTABLE' data set is used for further process analyze.
- There are two (2) variables of particular interest, **_TESTS_** and **_SUBI_** (for IRCHART™) or **_SUBX_** (for BOXCHART™, XCHART™, XRCHART™ and XSCHART™). The last five (5) observations are kept. If the **_TESTS_** has a non-zero entry, a **RED** alert is immediately sent to a PC in process control room for the Level 2. A traffic light analogy is used for the alert. If there is a continuous positive or negative change in the **_SUBI_** for the last 3 to the 5 observations, the alert is changed from **GREEN** to **YELLOW**, indicating a process shift is occurring and the process may trigger a statistical unbalance and out-of-control product at the next observation.

Test Data Flow



Real Time Display

There can be as many Real Time Display monitors as desired. The screen is divided into a X-Y matrix determined by how many output variables are to be monitored. Each individual X-Y coordinate of the matrix has a three level display similar to traffic light with the red, yellow green lights. Sitting below each cluster of lights is a SAS/AF™ program. When any light is pressed, a PROC GREPLAY™ program is activated to replay the control chart that is associated with that button. Along with the previously mentioned TSHEWHART™ control chart, a Pareto chart is displayed for the individual variables for the cluster variable. The value of the bar is represented by its contribution to the output value. The Pareto bars set for 'HOT SPOT' keys for further analysis of the control chart for that variable.

Proactive vs. Reactive

It is argued that when a computer is controlling the process, then no human intervention is required. The computer has the process completely under control. If the process is in complete control, and the setup model used to control the operations is infallible, then how can defective material be produced? The setup model is not perfect and a signal alarm **Reactive** mentality prevails. Upper and lower process control parameters are pre programmed to alert when they have been exceeded. The control limits can be either set by absolute specifications or by the process standard deviation. If the absolute process limits are chosen, they may be set by the process average plus/minus three (3) multiples of the process standard deviation. To use the process standard deviation, the process must exhibit a normal distribution and be in statistical control as established by a control chart method. Absolute specification controls are probably wider than those established by the process average plus/minus three (3) multiples of the process standard deviation.

To move into the **Proactive** mode, a control chart process must be used. Control charting will calculate a estimated process standard deviation, σ which is based on the equation $\bar{R}/d_2$. In most cases σ will be less than the standard deviation of the individuals, σ_0 which is based on the equation

$$\sqrt{\sum (x - \bar{x}_i)^2 / (n - 1)}$$
. Control charting of a single variable with σ is the a good beginning solution. A measurable output characteristic for the process can most likely be related to the variables that are being captured by a Level 2 computer. A **PROC CORR™** can be used to find process variables that are dependent on the measurable output characteristic. Once the dependent variables are found, another **PROC CORR™** can be run to establish if the is a correlation between

variables. If a strong correlation is found between variables, a **PROC PLOT™** or **PROC GPLOT™** is run to determine the polynomial relationship. A corresponding regression equation is calculated and used as another control variable for further analysis. All these control variables from correlations or plot routines are passed through **PROC REG™** and a multiple regression 'equation' is calculated. The determined 'equation' is the input for **REAL TIME DISPLAY**.

Another possible **Proactive** mode is to use 'set up' aims and the measured output for that aim. It could be thought as another way of calculating a measurable reference point. The equation to use is identical to that for the standard deviation of the individuals,

$\sigma_0 = \sqrt{\sum (x_i - \bar{x})^2 / (n-1)}$. x_i is the measured output characteristic and $\bar{x}$ is the set aim value for variable x_i .

Practical Example

For a simple yet very graphic example of cluster analysis versus single process alerts is possible using a data set supplied by **SAS™** Institute, **FITNESS**. **SAS** has already used this data set for examples in **REGRESSION** and **CORRELATION**. The data values have not been changed, the order of occurrence of the later observations has been changed to illustrate the principle.

As with any analysis, it should be determined if the process is in a steady state of statistical control. If the process is in statistical control as evidenced by a control chart, the capabilities or frequency distributions can be analyzed. For the statistical purest, these processes are not in complete statistical control. One variable shows stratification and others have points either above or below the control limits. **PROC SHEWHART™** **XCHART™** with the **TESTS™** option is used for illustration.

To establish single variable alarms, the process model uses the population average and the standard deviation of the individuals. The alarms are set at average plus or minus three (3) times the standard deviation. **PROC CAPABILITY™** can be used for the calculation. The limits established are quite liberal. If these control limits are used for feed back alarms, none of the observations will send an alert for an out-of-control situation using the standard deviation of the individuals

How does clustering affect the feed back signals. The formula, $\sigma_0 = \sqrt{\sum (x_i - \bar{x})^2 / (n-1)}$, is

used for the cluster model. Until observation 24, the process does not violate any of the eight (8) test for statistical control. Point 25 is at a low and an upward trend is beginning. At point 30, **TEST3** is noted for six (6) steadily increasing. With **REAL TIME**, point 28 would be flagged and send an alert back changing the traffic light from **GREEN** to **YELLOW**. The **YELLOW** would remain until point 30 when it is changed to **RED** because the trend of six (6) points increasing has been violated. Points 31 and 32 are above the upper control limit, therefore the **RED** light will remain lit. The operator can touch the display screen on either the **YELLOW** or **RED** light and the corresponding control and pareto charts are displayed.

Looking at the individual control charts, nothing seems amiss until observation 32 for variable **RUNPLUSE**. Going to the correlations, **RUNPLUSE** and **MAXPLUSE** have a strong correlation. the control charts for both variables have a large difference between the average and the actual value. Adding the larger deviations of **RUNPLUSE**, **MAXPLUSE** and **WEIGHT** yields the last points out of control for the calculated value. This is analogous to the individual control variables being in control at the time of manufacture with enough variability in the process to cause the product to be rejected later in the manufacturing process.

REFERENCES

SAS/QC Software, Version 6, First Edition, Cary, N.C., SAS Institute, Inc. 1990

SAS/GRAPH Software, Version 6, First Edition, Cary, N.C., SAS Institute, Inc. 1990

ACKNOWLEDGMENTS

SAS is a registered trademark of **SAS Institute**. in the USA and other countries. **™** indicates the USA registration.

Figure 1.

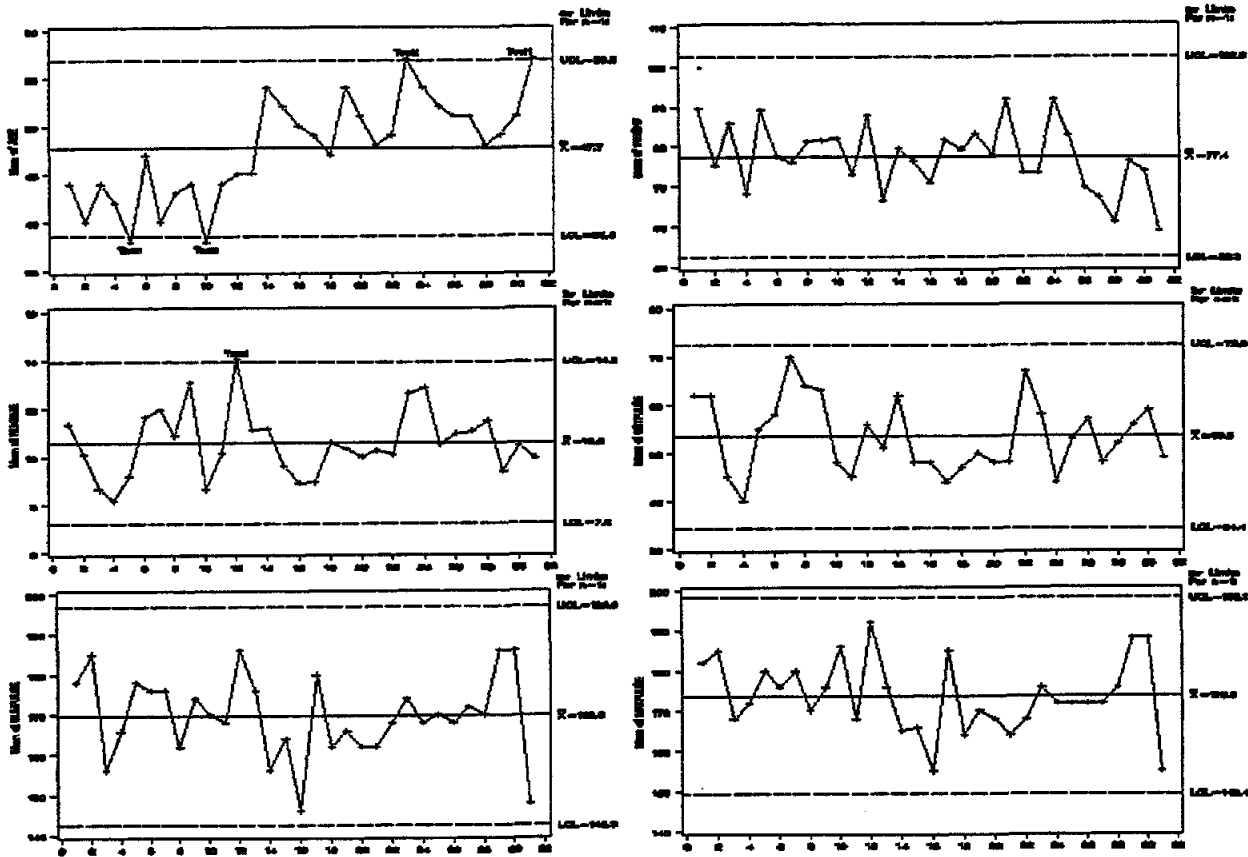


Figure 2.

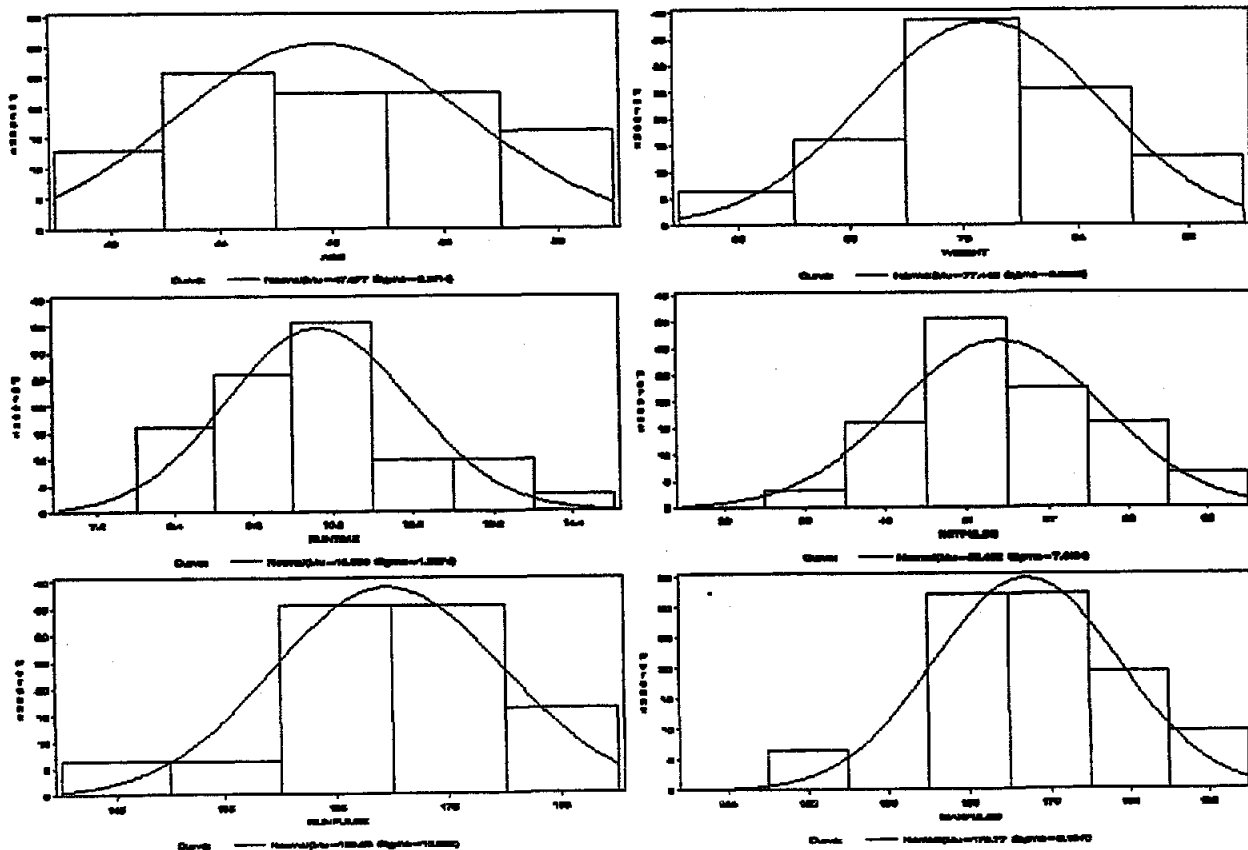


Figure 3.

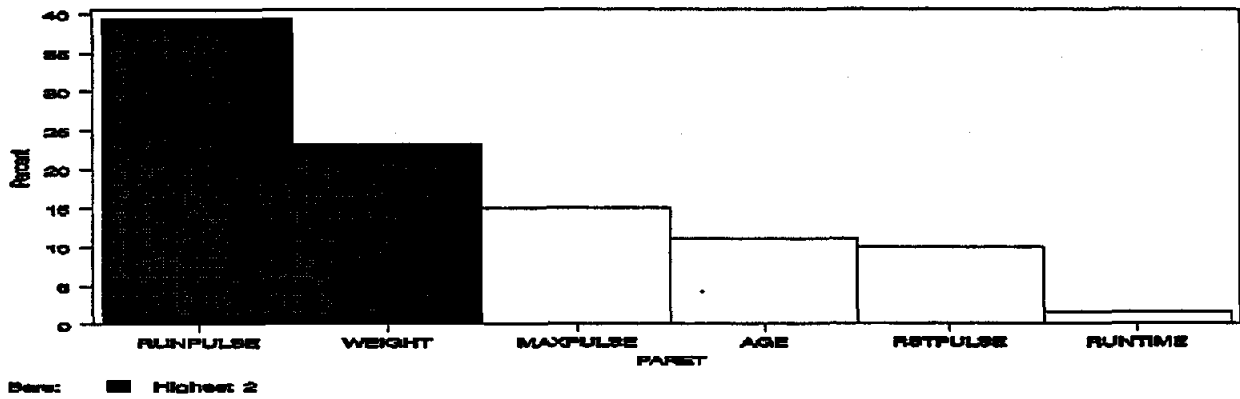
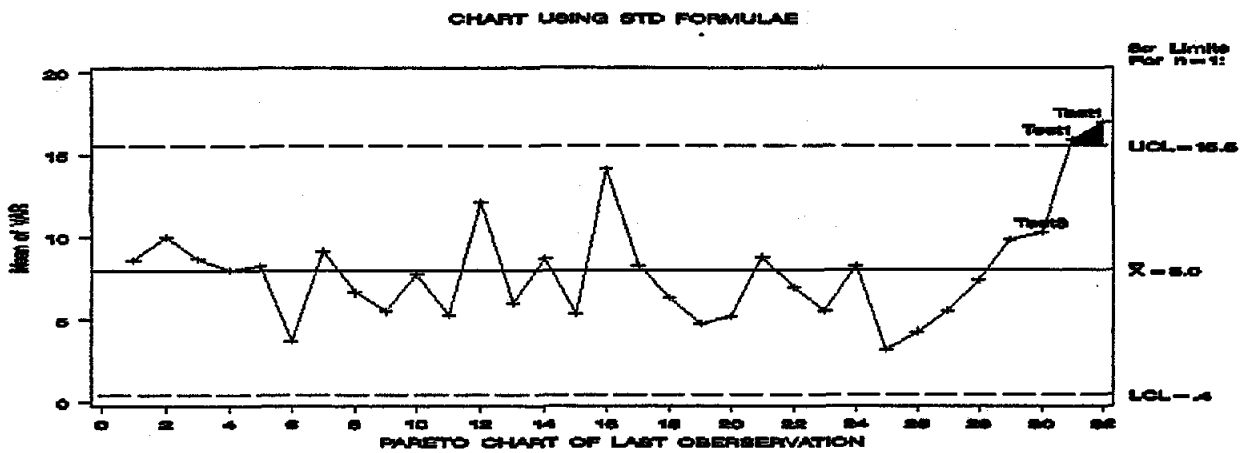


Figure 4.

