AIIM

Association for Information and Image Management

1100 Wayne Avenue, Suite 1100
Silver Spring, Maryland 20910
301/587-8202

Centimeter

Inches

MANUFACTURED TO AIIM STANDARDS
BY APPLIED IMAGE, INC.

1 of 1

WSRC-MS-94-0370

A COMBINED SHEWHART-CUSUM PROCEDURE WITH APPLICATION TO GROUNDWATER MONITORING (U)

by
Harris, S.P.

Westinghouse Savannah River Company
Savannah River Site
Aiken, South Carolina 29808

Tuckfield, R.C.

A document prepared for TWENTY-FIRST ANNUAL NATIONAL ENERGY AND ENVIRONMENTAL QUALITY DIVISION CONFERENCE at Tucson, AZ, USA from 18 Sep-21 Sep 1994.

DOE Contract No. DE-AC09-89SR18035

DISCLAIMER

WSRC-MS-94-0370

Performance, Measurement and Control

A Combined Shewhart -CUSUM Procedure With Application To Groundwater Monitoring(U)

S.P. Harris, Ph.D.

Principal Scientist

R.C. Tuckfield, Ph.D.

Manager, Applied Statistics

Westinghouse Savannah River Site
Aiken, S.C. 29808

ABSTRACT

A combined Shewhart CUSUM statistical process control procedure is being implemented at Savannah River Site for monitoring groundwater effluent concentrations. The procedure is being developed for the Environmental Monitoring Section and will be available to yield information on remediation performance in real-time for improved decision making. The method is outlined and an application given to monitoring trichloroethylene concentrations at a Savannah River Site point of compliance well. The combined charting methods incorporate the strengths of both charting methods. This intra-well comparison method will provide a visual tool for detecting both abrupt changes and increasing or decreasing trends in constituent concentrations.

INTRODUCTION

Statistical process control(SPC) charts are widely used as a statistical tool for visually monitoring processes in various industrial applications. The SPC charts have found ready application in environmental monitoring areas for intrawell comparisons and comparisons to historically monitored background wells. The charts are most useful for those constituents with a high frequency of occurrence in monitoring wells, for example, some metals and geochemical monitoring parameters whose measurements are typically above the analytical detection limit.

The basic concept of an SPC chart is the graphical display of data according to the time sequence over which the data are collected. Sampling information, collected from one time period to another, plotted on an SPC chart is used to determine the state of the process. SPC charts are used to determine if the process is "stable" in terms of its' statistics. Specifically, SPC charts are used to determine if the process variability is constant or, more typically in environmental applications, if the mean has changed or if a trend is developing over time. When the sample statistic exceeds its' statistical limit(often called a control limit or an action limit), we have statistical evidence of a process shift.

The use of Shewhart control charts has been well documented in many industrial applications. Basically, the Shewhart charts are the simplest method of SPC that can be used for monitoring sampling results over time. The Shewhart charts are very sensitive for detecting large mean shifts. However, more sensitive methods are required for detecting moderate mean shifts on the order of one or two standard deviations, 1-2σ, from the "process aim"(1). Specifically, cumulative sum (CUSUM) control charts can be used, among other methods for this purpose. CUSUM is a more recent method in SPC used extensively in process industries to effect "on-aim" control, as distinct from Shewhart "within specifications" control. Currently, CUSUM is used in our Environmental Monitoring area to track well sampling data. The CUSUM results are reported quarterly to South Carolina Department of Health and Environmental Control(SCDEHEC).

A statistical procedure is being implemented for Environmental Monitoring Section using both methods of statistical control. Specifically, the combined Shewhart-Cumulative sum control chart , developed by Lucas (2), will be constructed for each constituent as recommended by the U.S.- E.P.A(3). The method will be used for each constituent on the Groundwater Protection Standard (GWPS) list that has yet to exceed the GWPS concentration limit for point of compliance wells.

The combined charting methods incorporates the strengths of both charting methods . This intra-well comparison method will be employed to provide a visual tool for detecting both abrupt changes and increasing or decreasing trends in constituent concentrations. A specific example of the combined Shewhart-CUSUM method will be illustrated.

CUSUM

The computational CUSUM procedure has been used for groundwater monitoring and corrective action over the current four year monitoring period for SRS's M-Area Hazardous Waste Management Facility(HWMF). The emphasis on our use of the CUSUM method for our corrective action program was to demonstrate statistically significant decreases in constituent concentrations as a result of remediation.

The use of CUSUM complements the small sample t-tests by providing information about trend. The CUSUM summarizes the past observations with the current one(utilizing information about trend contained in past observations). As the data accrue over time, the aggregated sums of deviations of the observations from the process mean, the background value, or the concentration limit are compared continuously against a test criterion (h) derived from two quantities:

1) the magnitude of the important difference to detect, δ (commonly one standard deviation)

2) a measure of the sensitivity, h, to that difference (4)

The size of the important difference (δ) and the sensitivity of the test to that difference (h) can be adjusted to bring the error rates into desired balance (K). Generally, h is between 3 and 5 and K is selected between 0.4 and 1.0, inclusive giving the CUSUM desirable average run lengths (ARL's). The ARL is the average number of samples needed to detect a significant difference from the process mean. The coarser the test mesh(i.e., the larger the δ and h), the more frequent a trend will slip through undetected (a Type II error); the finer the mesh, the more frequently an incorrect observation will be a false alarm (a Type I error).

Detailed Description of CUSUM

(i) Set the aim point μ_0(i.e., the process average, groundwater protection standard, background value, concentration limit or the detection limit).

(ii) Specify the departure from the aim point, δ, that the CUSUM should detect.

(iii) Determine h by setting the desired ARL to alarm when a trend is really present. Typically, K= $\delta/2$= 0.5 for balancing the Type I with Type II errors. ARL tables for selecting "h" are available in SAS-QC(5) or published tables(1) to select the resulting sensitivity.

(iv) Estimate the standard deviation, σ_p, of the individual plot points. The estimate of σ_p can be obtained from the mean sum squared for successive differences(MSSD),

$$sp = \mathrm{Sqrt}\left[\sum (\bar{Y}_{i+1} - \bar{Y}_i)^2 / (2^*(N-1)) \right]$$

where the $\bar{Y}_i$'s are the individual plot points. Although four replicates are taken each sampling period, they are not independent since they are taken over a very short time interval relative to the groundwater flow. Thus, the data are averaged and treated as "individual" measurements. If the σ_p is bigger than our estimated standard deviation, then many more false out-of-control signals will be received than would be expected. These false out-of-control signals can cause unnecessary work to be done. The estimate of sigma should contain inherent process variability plus measurement error.

(v) The computational form for implementing a combined Shewhart-CUSUM scheme uses the z-score statistic

$$\bar{Z}_i = (\bar{Y}_i - m) / \sigma_p$$

where m is the desired mean value(an estimate of μ_0), σ_p is the standard deviation on $\bar{Y}_i$ (assumed to be known) and $\bar{Y}_i$ is the i-th observation for the CUSUM scheme. The i-th observation represents a single reading in our groundwater application. The mean value, μ_0, is target value to which the process is to be controlled. The historical process average or the background well average are prime candidates for obtaining the estimate.

The CUSUM formulas are

$$S_{Hi} = \max[0, (Z_i - K) + S_{H(i-1)}]$$

for detecting positive mean shifts and

$$S_{Li} = \max[0, (-Z_i - K) + S_{L(i-1)}]$$

for detecting negative mean shifts where max[A,B] is the maximum of A and B. If either S_H or S_L becomes greater than the decision criteria h, this signals the process mean has shifted or is trending. Specifically, if $S_{Hi} > h$ the trend is increasing or if $S_{Li} > h$ the trend is decreasing ,otherwise no trend is detectable. The starting CUSUM value, S_0 is set to zero in a standard CUSUM scheme.

COMBINED SHEWHART-CUSUM

The CUSUM procedure can be improved and made competitive with an X-bar chart for detecting large mean shifts(2). This modification entails using X-bar chart limits in conjunction with a CUSUM scheme. This combined Shewhart-CUSUM quality control scheme possesses the key features of both the Shewnart and CUSUM control procedures. In this scheme, the CUSUM feature will quickly detect small shifts from the goal while the addition of Shewhart limits increases the speed of detecting large shifts. Implementing the Shewhart-CUSUM scheme require no additional storage and very little additional programming effort.

Using this combined Shewhart-CUSUM method, an out-of-control signal can be received not just from S_H or S_L but also from the z-score. The Shewhart modification to a CUSUM control scheme adds the additional criterion that if $|Z_i| >$ SCL (SCL denotes the Shewhart control limits) an out-of-control signal is also given. We could select SCL=3 as the threshold value. Then an out-of-control signal would be received if either the absolute value of z exceeds 3 or either S_H or S_L exceeds h.

A subgroup average will be outside the control limits for an Shewhart X-bar chart if z is either greater than 3 or less than -3. As a result of combining this criteria with the CUSUM criteria, the false alarm rate is naturally increased. In terms of the ARL, the in-control ARL for the combined Shewhart-CUSUM scheme must be less than the in-control ARL for an X-bar chart which is approximately 370.

It can be seen (2) that the in-control ARL for a CUSUM scheme with h=5, K=0.5, SCL=3 is 223.4. Consequently, a SCL value greater than 3 is preferable. Table 1 gives the ARL values for a basic CUSUM scheme and a combined Shewhart-CUSUM scheme with z=3.5 and 4.0. When the shift from the mean is very large and abrupt, the Shewhart chart will be sensitive (i.e., have very short ARL's) (4). In the first column of Table 1, the ARL's are tabulated in order of increasing displacement of the current process mean(1). The displacement ranges from zero to 5σ.

From Table 1, we see the effect of the Shewhart limits, especially in detecting shifts of at least 3-sigma. This improvement, however, is partially offset by a reduction in the in-control ARL. The main advantages of the Shewhart modifications are obtained when the SCL makes only a moderate change in the in-control ARL values. In this case, the ARL 's of the combined scheme look very much like the ARL 's of the standard CUSUM scheme except that large shifts are detected faster. Thus, the combined Shewhart-CUSUM scheme often provides real improvements in the ARL with little additional computational effort.

EXAMPLE

Table 2 contains data on trichloroethylene concentrations (µg/L) from Savannah River Site Well RWM1. The samples were taken quarterly from 2/83 to 4/93 and recorded in time sequence(Obs. Number 1 is 2Q83). Combined Shewhart-CUSUM charts are shown in Plot 1 with SCL = 3.5. The combined Shewhart-CUSUM procedure reveals the following. Both observation 3 and 21 are outside the control limits on the Shewhart Chart. Observation 3 is above the upper limit while observation 21 is below the lower limit. The CUSUM procedure in TABLE 2 signals a decreasing trend, starting at Observation 21, which is not detected using the SCL limits.

DESIGN OF SHEWHART-CUSUM SCHEMES

The following four parameter values are needed to implement the combined Shewhart-CUSUM control scheme(assuming a standardized scal^ σ=1).

h - decision interval value
K - reference value

S_{L0} ,S_{H0} - initial values
SCL - Shewhart control limits

The parameter K for the CUSUM scheme is obtained directly, K = $\delta/2$ where δ is a displacement that should be quickly detected(6). In practice, a K value 0.5 is often used. This allows a displacement of one standard deviation to be detected quickly. The h value is selected to give the desired in-control ARL. When K=0.5, h values of four or five often give an adequately long in-control ARL with quick detection of an out-of-control situation. The addition of Shewhart limits is then considered. For a K=0.5, h=4 or 5 CUSUM scheme, Shewhart limits of 3.5 or 4 cause only a small decrease to the in-control ARL and still give shorter ARL's for large shifts. A Shewhart limit of 3.0 decreases the in-control ARL too much.

BIBLIOGRAPHY

1. *Statistical Methods For Quality Improvement*. Ryan, T.P., John Wiley and Sons, Inc., New York, 1989.

2. "Combined Shewhart-CUSUM Quality Control Schemes." Lucas, J.M., *Journal of Quality Technology*, Vol. 14, No. 2, 1982, pp 51-59.

3. *Statistical Analysis of Groundwater Monitoring Data at RCRA Facilities. U.S. - EPA,* Interim Final Guidance document (EPA/530-SW-89-026), 1989.

4. *Product Quality Management*. Marquardt, D.W., editor, E.I. duPont de Nemours & Co., Inc., 1983, Ch 6.

5. Statistical Analysis System, CUSUMARL function, SAS Windows Version 6.08, SAS Institute Inc., Cary, N.C.

6. *Introduction to Statistical Quality Control* (2nd Ed.). Montogomery, D.C., John Wiley and Sons, Inc., New York. 1991.

TABLE 1

ARL Values for Various Two-Sided CUSUM Schemes
h=5, K=0.5

S_{L0} = 0, S_{H0} = 0

Mean Shift (Standard Units)	Basic CUSUM	Shewhart-CUSUM (SCL=3.5)	Shewhart-CUSUM (SCL=4.0)
0	465	391	459.0
0.25	139	130.9	139.0
0.50	38.0	37.2	38.0
0.75	17.0	16.8	17.0
1.00	10.4	10.2	10.4
1.50	5.75	5.58	5.74
2.00	4.01	3.77	3.98
2.50	3.11	2.77	3.05
3.00	2.57	2.10	2.43
4.00	2.01	1.34	1.59
5.00	1.69	1.07	1.16

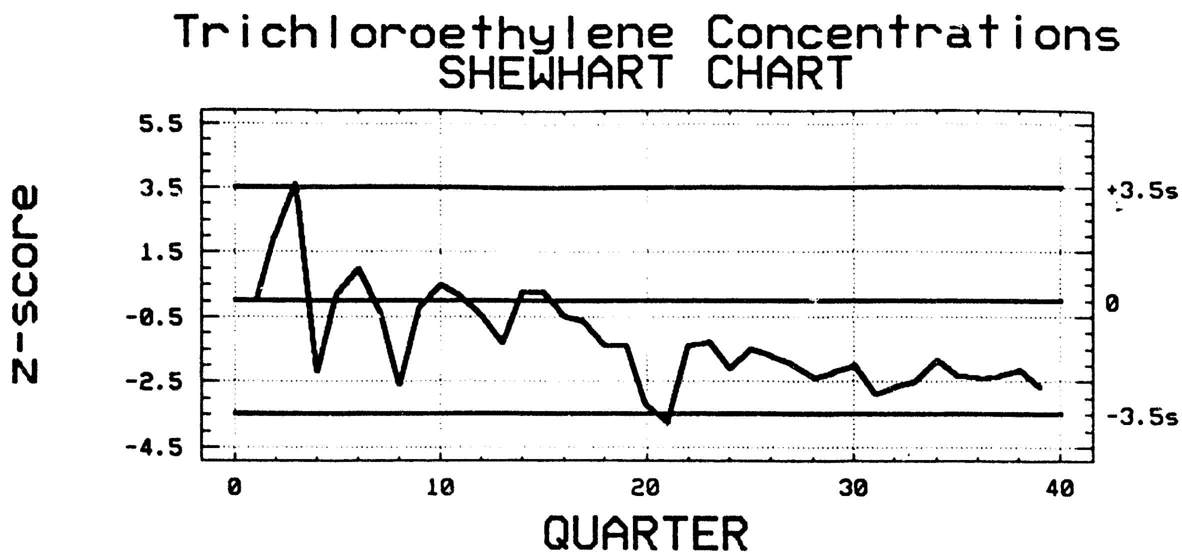
TABLE 2

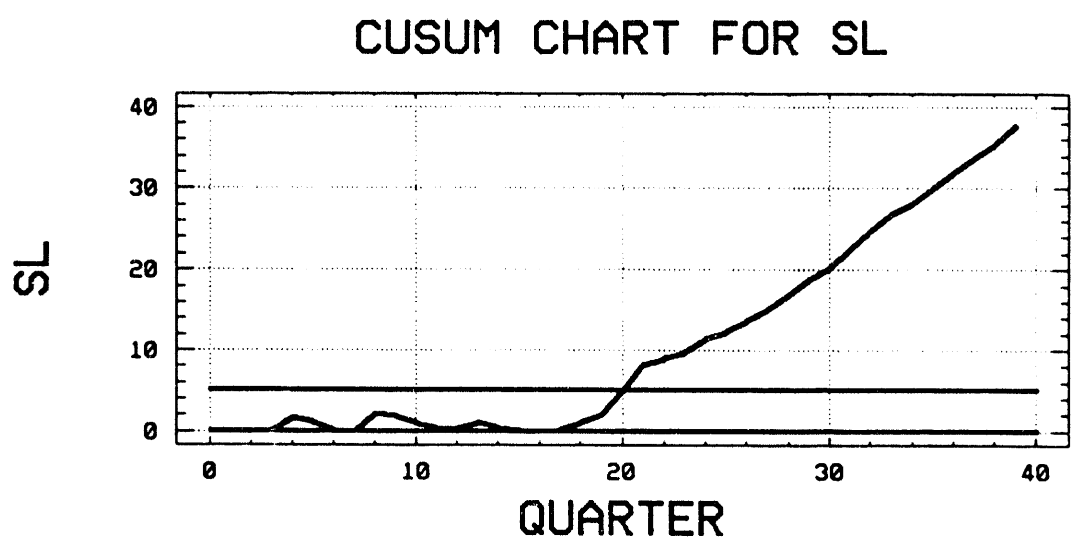
SRS WELL RWM 1
Trichloroethylene Concentrations (µg/L): 2/83 to 4/93, Quarterly
m = 68200, σ_p = 13378.9, h = 5, K = 0.5

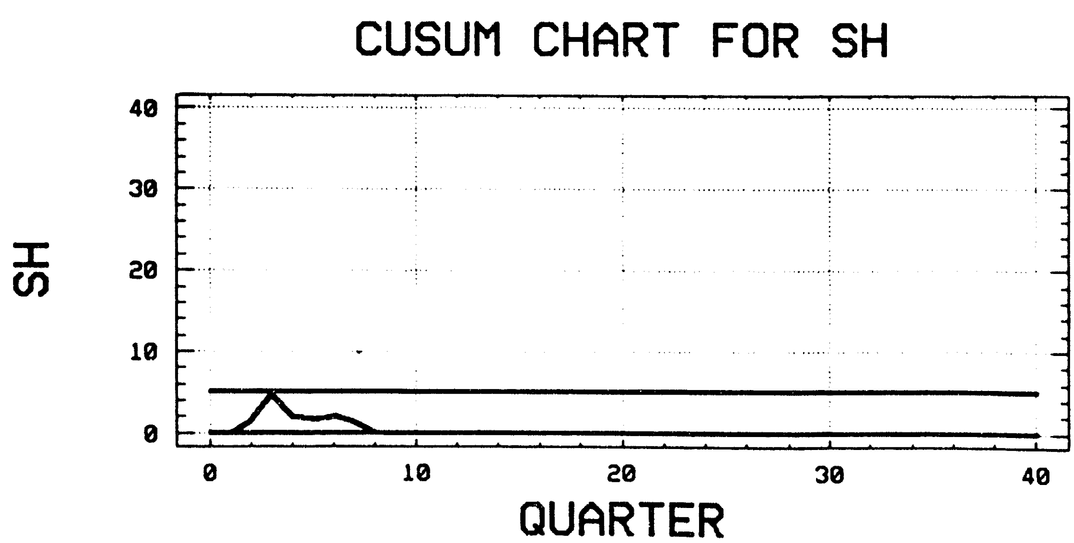
Obs Number	Observed Value	SHEWHART Z_i	CUSUM S_H	CUSUM S_L	ID
1	68200	0	0	0	
2	94980	2.00	1.50	0	
3	115967	3.57	4.57	0	SCL+
4	38950	-2.19	1.89	1.69	
5	70771	0.19	1.58	0.99	
6	81178	0.97	2.05	0	
7	63721	-0.33	1.21	0	
8	33300	-2.61	0	2.11	
9	65500	-0.20	0	1.81	
10	74543	0.47	0	0.84	
11	69267	0.08	0	0.26	
12	61383	-0.51	0	0.27	
13	50695	-1.31	0	1.07	
14	71700	0.26	0	0.31	
15	71100	0.22	0	0	
16	61467	-0.50	0	0	
17	59000	-0.69	0	0.19	
18	49700	-1.38	0	1.07	
19	48833	-1.45	0	2.02	
20	25155	-3.22	0	4.74	
21	17767	-3.77	0	8.01	BOTH-
22	49300	-1.41	0	8.92	CSUM-
23	51150	-1.27	0	9.70	CSUM-
24	40200	-2.09	0	11.29	CSUM-
25	48050	-1.51	0	12.29	CSUM-
26	44775	-1.75	0	13.55	CSUM-
27	41680	-1.98	0	15.03	CSUM-
28	35825	-2.42	0	16.95	CSUM-
29	38650	-2.21	0	18.66	CSUM-
30	41600	-1.99	0	20.14	CSUM-
31	29275	-2.91	0	22.55	CSUM-
32	32600	-2.66	0	24.71	CSUM-

33	34738	-2.50	0	26.72	CSUM-
34	43667	-1.83	0	28.05	CSUM-
35	36850	-2.34	0	29.89	CSUM-
36	35740	-2.43	0	31.82	CSUM-
37	36875	-2.34	0	33.66	CSUM-
38	39250	-2.16	0	35.32	CSUM-
39	32533	-2.67	0	37.49	CSUM-

PLOT 1


Trichloroethylene Concentrations
SHEWHART CHART
z-score
5.5
3.5
1.5
-0.5
-2.5
-4.5
+3.5s
0
-3.5s
0
10
20
30
40
QUARTER


CUSUM CHART FOR SL
SL
40
30
20
10
0
0
10
20
30
40
QUARTER


CUSUM CHART FOR SH
SH
40
30
20
10
0
0
10
20
30
40
QUARTER

DATE FILMED

9/14/94

END