Daniela Cocchi, Michele Scagliarini

Control Charts and the Effect of the Two-Component Measurement Error Model

Quaderni di Dipartimento

Serie Ricerche 2007, n. 2

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Dipartimento di Scienze Statistiche "Paolo Fortunati"

Abstract

Monitoring algorithms, such as the Shewhart and Cusum control charts, are often used for monitoring purposes in the chemical industry or within an environmental context. The statistical properties of these algorithms are known to be highly responsive to measurement errors. Recent studies have underlined the important role played by the two-component measurement error model in chemical and environmental monitoring. In the present work, we study the effects of the two-component error model on the performance of the $\overline{X}$ and S Shewhart control charts. Results reveal that gauge imprecision may seriously alter the statistical properties of the control charts. We propose how to reduce the effects of measurement errors, and illustrate how to take errors into account in the design of monitoring algorithms.

Keywords: Average run length, calibration curve, constant measurement error, Monte Carlo study, proportional measurement error, repeated measurements, Shewhart control charts.

1. Introduction

Statistical monitoring algorithms, such as Shewhart control charts and cumulative sums control charts (CUSUM), are widely used in many industries, including the chemical and pharmaceutical industries Montgomery (2005). They are also employed in the environmental context to monitor, for example, pollutants Gibbons (1999), pollutant measurement devices Bordignon and Scagliarini (2000) and biological populations Anderson and Thompson (2004).

In the industrial field, it is widely acknowledged that measurement errors may significantly alter the performance of statistical process control methodologies, as has been confirmed by the works of several authors including Kanzuka (1986), Mittag and Stemann (1998), Linna and Woodall (2001a), and Linna *et al*. (2001b). Since the pioneering research of Bennnet (1954), who studied the effects of measurement errors in chemical process control problems, further studies by Rocke *et al*. (2003) and Gibbons and Coleman (2001), have established that there are two types of measurement errors in chemical and environmental monitoring. The measurement error is constant over a range of concentrations close to zero; at higher concentrations, the measurement error is observed to be proportional to the concentration of the analyte, pollutant or toxic substance. Moreover, many measurement technologies require a linear calibration curve in order to estimate the actual concentration of a substance (pollutant, analyte, etc.) in a sample. This leads to the following model (Rocke and Lorenzato (1995))

$$Y_{\mu}^{e} = \alpha + \mu\beta e^{\eta} + \varepsilon \tag{1}$$

where: Y_{μ}^{e} is the response at concentration μ; α and β are the parameters of the calibration curve; $\eta \sim N(0, \sigma_{\eta}^{2})$ represents the proportional error that is always present but is only noticeable at concentrations significantly above zero; and $\varepsilon \sim N(0, \sigma_{\varepsilon}^{2})$ represents the additive error that is always present but is only really noticeable at near zero concentrations.

In (1) μ is non-random. This assumption is acceptable for a laboratory sample, but in a real world monitoring situation, one needs to assume that the concentration of a given substance, or a suitable transformation, constitutes a random effect that we assume to be normally distributed $X \sim N(\mu, \sigma^2)$. Thus, we have

$$Y^e = \alpha + \beta X e^{\eta} + \varepsilon \tag{2}$$

Therefore the observable response Y^e is a random variable consisting of the sum of a normal variable, ε, and the product of a normal variable, *X*, multiplied by a log-normal variable η. The mean and standard deviation of Y^e are, respectively:

$$E(Y^e) = \alpha + \beta\mu\sqrt{e^{\sigma_\eta^2}} \tag{3}$$

and

$$\sigma_{Y^e} = \left(\beta^2\left[\sigma^2 e^{\sigma_\eta^2} + \mu^2\left(e^{\sigma_\eta^2}(e^{\sigma_\eta^2}-1)\right) + \sigma^2\left(e^{\sigma_\eta^2}(e^{\sigma_\eta^2}-1)\right)\right] + \sigma_\varepsilon^2\right)^{1/2} \tag{4}$$

The present paper aims to study the effect of the two-component measurement error model on the Shewhart Control Charts used to monitor the level ($\overline{Y}^e$-chart) and the variability (S_{Y^e}-chart) of the observable response Y^e. First of all we outline the performance of the monitoring algorithms in the presence of this kind of error. Then we study ways of reducing the effects of measurement errors, and of taking errors into account when designing the control chart.

The paper is organised as follows. In the next section we study the behaviour of the usual Shewhart control chart used to monitor the process level when the two-component error model holds. In the third section the impact of the error model on the S_{Y^e}-chart is analyzed. In section 4 we suggest a way of taking account of measurement errors when evaluating the performance of control charts. Finally, section 5 contains the concluding remarks.

2. The $\overline{Y}^e$ Control Chart

In an error-free situation, the response $Y = \alpha + \beta X$ is normally distributed with mean and standard deviation given, respectively, by

$$E(Y) = \alpha + \beta\mu \tag{5}$$

and

$$\sigma_Y = \beta\sigma \tag{6}$$

In order to monitor the mean level of response Y at time *i*, a sample of *n* observations from Y is collected and the mean of these measurements, $\overline{Y}_i$, is plotted on the control chart (3σ-limits):

$$CL = \alpha + \beta\mu \tag{7}$$

$$UCL = CL + 3\frac{\sigma_Y}{\sqrt{n}} \tag{8}$$

$$LCL = CL - 3\frac{\sigma_Y}{\sqrt{n}} \tag{9}$$

Linna and Woodall (2001a) studied the effects of an additive error ε on the performance of the above control chart. They noted that one of the effects of measurement errors is a loss of power in detecting parameter shifts.

In the two-component error case, in order to monitor the mean level of response Y^e, the control chart is given by:

$$CL^e = \alpha + \beta\mu\sqrt{e^{\sigma_\eta^2}} \tag{10}$$

$$UCL^e = CL^e + 3\frac{\sigma_{Y^e}}{\sqrt{n}} \tag{11}$$

$$LCL^e = CL^e - 3\frac{\sigma_{Y^e}}{\sqrt{n}} \tag{12}$$

where σ_{Y^e} comes directly from (4).

In order to study the effect of the two-component error model, we examined the chart's behaviour when the process is under statistical

control (H_0), and in the out-of-statistical-control state (H_1). Since the distribution of Y^e cannot be treated analytically, we adopted a simulation approach.

In designing the Monte Carlo study, we considered the case of a pseudo-population representing a real situation by developing the toluene example reported in Rocke and Lorenzato (1995), where $\alpha = 11.51$, $\beta = 1.524$, $\sigma_\varepsilon = 5.698$, $\sigma_\eta = 0.1032$, and concentration ranges from 5 picograms to 15 nanograms in 100mL. The under control (H_0) and out of control (H_1) conditions were simulated for several values of the coefficient of variation (c.v.(X)) and several values of the mean ($E(X) = \mu$) of the unobservable X. In this way, for a given precision structure (σ_ε and σ_η fixed) we reproduce a set of different working conditions for the measurement device. We fixed the samples size n=5 and generated 10000 samples for each condition. We built the mean control chart with the 3-sigma limits. For the out of control situations we considered shifts of standardized magnitude equal to ±0.5, for which the theoretical value of the Average Run Length (ARL) function is 33.4. The standard deviation chart has been designed with a false alarm probability of 0.01. The out of control situations were studied for a relative variation in standard deviation 1.1, for which the theoretical value of the ARL function is 37.2.

2.1. State of Statistical Control (H_0)

We first examined the effects of measurement errors on the false-alarm rate. The results are summarized in Figures (1) and (2), where the bold line, denoted as "no errors", corresponds to $0.00135 = Pr\left(\overline{Y} \le LCL \mid H_0\right) = Pr\left(\overline{Y} \ge UCL \mid H_0\right)$, *i.e.* the probability, under H_0 and in the error-free case, of a signal below the LCL (or above the UCL).

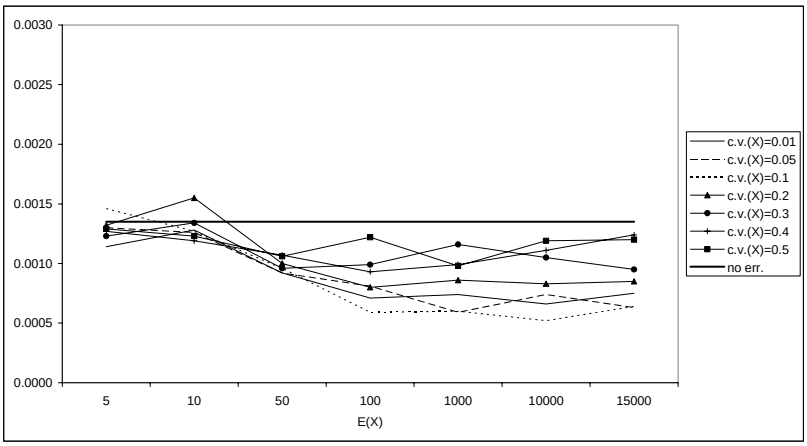

Figure 1. False alarm rates below the LCL (σ_ε =5.698, σ_η =0.1032)

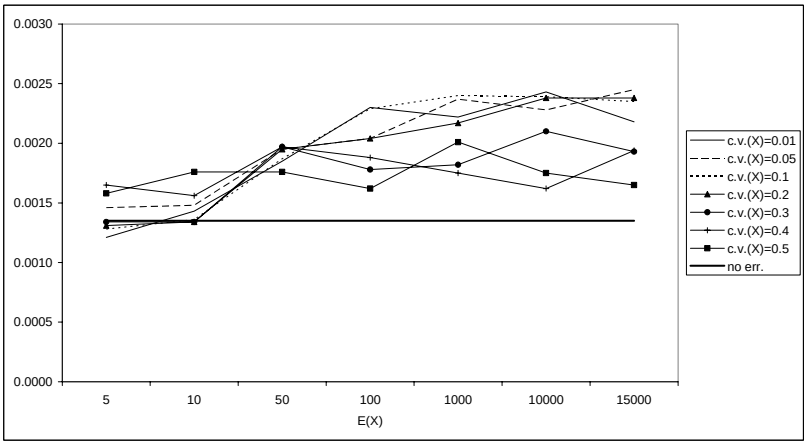

Figure 2. False alarm rates above the UCL (σ_ε =5.698, σ_η =0.1032)

The control chart reveals asymmetric behaviour: the false alarm rates below the LCL tend to be lower than the false alarm rate in the error-free case (0.00135), while the false alarm rates above the UCL show a clear pattern of values slightly greater than 0.00135.

2.2. Out of Control Situation (H_1)

Let us consider the case of a shift in the mean of the non-observable X from μ to μ_1, which corresponds to a standardized shift of magnitude

$$\delta = \frac{E(X \mid H_1) - E(X \mid H_0)}{\sigma} = \frac{\mu_1 - \mu}{\sigma} \tag{13}$$

Under this hypothesis the expected value of the response Y^e is

$$E(Y^e \mid H_1) = \alpha + \beta\mu_1\sqrt{e^{\sigma_\eta^2}} \tag{14}$$

and the corresponding standardized shift in the monitored Y^e is given by

$$\delta_{Y^e} = \frac{E\left(Y^e \mid H_1\right) - E\left(Y^e \mid H_0\right)}{\sigma_{Y^e}} = \\ = \frac{\delta}{\sqrt{1 + \frac{\mu^2}{\sigma^2}\left(e^{\sigma_\eta^2} - 1\right) + \left(e^{\sigma_\eta^2} - 1\right) + \frac{\sigma_\varepsilon^2}{\beta^2\sigma^2 e^{\sigma_\eta^2}}}} \tag{15}$$

We note that, for non-zero σ_η and σ_ε, the denominator in (15) is greater than 1, therefore $\delta_{Y^e} < \delta$. The consequence is that measurement errors lead to a smaller shift in the observed response Y^e, which means that the change is more difficult to detect. Table (1) shows the values of δ_Y^e corresponding to a shift $|\delta|$=0.5 (in the mean of X) for the *E(X)* and c.v.*(X)* values considered in the simulations.

μ	c.v.(X)						
	0.01	**0.05**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**
5	0.007	0.033	0.066	0.129	0.185	0.235	0.277
10	0.013	0.064	0.125	0.230	0.306	0.359	0.394
50	0.039	0.182	0.308	0.420	0.458	0.474	0.482
100	0.045	0.207	0.336	0.436	0.467	0.480	0.486
1000	0.048	0.217	0.346	0.442	0.470	0.482	0.487
10000	0.048	0.217	0.347	0.442	0.470	0.482	0.487
15000	0.048	0.217	0.347	0.442	0.470	0.482	0.487

Table 1. Values of δ_Y^e corresponding to a standardized shift of magnitude $|\delta|=0.5$

Table 1 reveals that the reduction in the shift is particularly evident for small values of *E(X)* and c.v.(*X*).

For the purpose of evaluating the statistical properties of the control chart in the out of control situation, one of the most commonly-used measures is the ARL. For the $\overline{Y}^e$-chart the theoretical ARL for shifts of magnitude δ_Y^e can be calculated as

$$ARL(\delta_{Y^e}) = \left(\Phi\left(-3+\delta_{Y^e}\sqrt{n}\right)+\Phi\left(-3-\delta_{Y^e}\sqrt{n}\right)\right) \qquad (16)$$

Table 2 shows the ARL values corresponding to the δ_Y^e in Table 1.

μ	c.v.(X)						
	0.01	**0.05**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**
5	369.99	360.58	334.17	260.05	193.02	145.49	114.00
10	368.87	335.93	264.18	149.80	96.04	71.14	58.34
50	356.90	196.11	95.02	50.74	41.44	38.13	36.60
100	352.51	170.99	81.13	46.50	39.51	37.04	35.90
1000	350.35	161.07	76.27	45.08	38.87	36.68	35.67
10000	350.33	160.96	76.23	45.07	38.87	36.68	35.66
15000	350.33	160.96	76.22	45.07	38.87	36.68	35.66

Table 2. Theoretical values of ARL for the δ_Y^e in Table 1

Expression (16) is based on the normality assumption of the sample statistic used in the chart. However, we noticed that measurement errors

lead to departures from the Normal distribution. Therefore, in order to study the effects of the two-component measurement error, we chose to perform a simulation study of the out-of-control situations: for known change instants t_0 we imposed shifts of magnitude $\delta=\pm 0.5$ (n=5) in the variable X, and we observed the alarm times in the $\overline{Y}^e$-control-chart. The simulations were designed to evaluate the mean detection delay time, which is an approximation of the ARL function. Once again, each condition was replicated 10000 times and the results thereof summarized in Table 3, in case of a positive shift, and in Table 4 in cases of a negative shift.

μ	c.v.*(X)*						
	0.01	**0.05**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**
5	359.83	350.34	330.29	254.49	182.68	138.77	107.29
10	353.57	321.53	254.37	138.03	86.45	63.53	53.09
50	320.56	144.59	69.10	41.63	35.08	32.84	32.95
100	287.46	116.04	58.18	37.36	33.65	32.07	32.63
1000	264.98	104.01	53.91	35.84	32.27	32.35	32.15
10000	272.10	105.93	53.84	35.95	32.74	32.06	32.08
15000	272.05	106.07	54.21	35.88	32.36	32.17	31.49

Table 3. Mean detection delay for δ=0.5

μ	c.v.*(X)*						
	0.01	**0.05**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**
5	369.04	355.33	335.56	266.23	197.05	148.33	118.01
10	366.76	349.75	275.15	161.83	101.10	76.01	61.96
50	377.04	265.54	128.30	61.74	47.98	41.88	39.61
100	383.25	263.54	119.70	59.65	46.13	41.04	38.95
1000	384.25	271.74	114.67	57.21	44.97	40.32	38.81
10000	375.74	262.79	119.19	56.99	44.59	40.80	38.01
15000	379.37	265.84	118.42	56.37	45.21	41.44	38.10

Table 4. Mean detection delay for δ=-0.5

Tables 3 and 4 confirm the asymmetric effect of the measurement errors. The mean detection delays for positive shifts are smaller than the corresponding mean detection delays for negative shifts. As one would

have expected, the values in Tables 3 and 4 often differ remarkably from the theoretical values of ARL reported in Table 2.

3. The S_Y^e-chart

The test statistic is the sample standard deviation

$$S_Y = \sqrt{\frac{\sum_{i=1}^{n}\left(Y_i - \overline{Y}\right)^2}{n-1}} \quad (17)$$

in a sample of size n. It follows that the control limit of the unilateral control chart used to monitor the variability of Y in the error-free case is

$$UCL_Y = \sigma_Y\sqrt{\frac{\chi^2_{n-1,1-\alpha}}{n-1}} = \beta\sigma\sqrt{\frac{\chi^2_{n-1,1-\alpha}}{n-1}} \quad (18)$$

where $\chi^2_{df,1-\alpha}$ is the α^{th} percentile of the chi-square distribution with df degrees of freedom, and α=Pr($S_Y \geq UCL_Y|H_0$). Detailed studies of the effect of the additive measurement error on the S-chart have been made by Mittag and Stemann (1998) and Linna and Woodall (2001a).

In the two-component measurement error case, the S_Y^e-chart has the same structure as (18)

$$UCL_{Y^e} = \sigma_{Y^e}\sqrt{\frac{\chi^2_{n-1,1-\alpha}}{n-1}} = \beta\sigma_{Y^e}\sqrt{\frac{\chi^2_{n-1,1-\alpha}}{n-1}} \quad (19)$$

but with standard deviation σ_Y^e given by (4).

3.1. Effects of the two-component error model under H_0

We first checked the effects of measurement errors on the false alarm rate. The results are summarized in Figure 3, where the bold line, denoted

as "no errors", corresponds to α=0.01, *i.e.* the probability, under H_0 and in the error-free case, of a signal above the UCL.

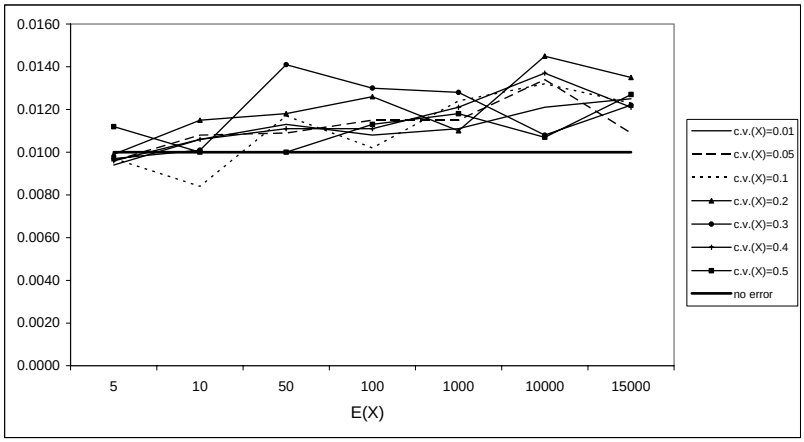

Figure 3. S_Y^e -Chart False alarm rates above the UCL (σ_ε =5.698, σ_η =0.1032)

Increasing values of *E(X)* cause an increasing measurement error effect on the false-alarm rate, which tends to assume values greater than false-alarm probability in the error-free case.

3.2. The effects of the error model under H_1

If the variability measure σ of X shifts to σ_1 ($\sigma_1 > \sigma$), we can define the relative variation as

$$\gamma = \sqrt{\frac{Var(X \mid H_1)}{Var(X \mid H_0)}} = \frac{\sigma_1}{\sigma} \tag{20}$$

In the presence of a relative shift γ in the standard deviation of X, the standard deviation of Y^e becomes

$$\sigma_{1Y^e} = \sqrt{Var\left(Y^e \mid H_1\right)}$$
$$= \gamma \left(\beta^2 \left[\sigma^2 e^{\sigma_\eta^2} + \frac{\mu^2 \left(e^{\sigma_\eta^2} (e^{\sigma_\eta^2} - 1) \right)}{\gamma^2} + \sigma^2 \left(e^{\sigma_\eta^2} (e^{\sigma_\eta^2} - 1) \right) \right] + \frac{\sigma_\varepsilon^2}{\gamma^2} \right)^{1/2} \quad (21)$$

and the correspondent relative variation is given by

$$\gamma_{Y^e} = \sqrt{\frac{Var\left(Y^e \mid H_1\right)}{Var\left(Y^e \mid H_0\right)}} = \frac{\sigma_{1Y^e}}{\sigma_{Y^e}} = \frac{A\gamma}{\sigma_{Y^e}} \quad (22)$$

where

$$A = \left(\beta^2 \left[\sigma^2 e^{\sigma_\eta^2} + \frac{\mu^2 \left(e^{\sigma_\eta^2} (e^{\sigma_\eta^2} - 1) \right)}{\gamma^2} + \sigma^2 \left(e^{\sigma_\eta^2} (e^{\sigma_\eta^2} - 1) \right) \right] + \frac{\sigma_\varepsilon^2}{\gamma^2} \right)^{1/2} \quad (23)$$

Since $A \leq \sigma_Y^e$, it follows that $\gamma_Y^e \leq \gamma$. Once again, the presence of measurement errors lead to a reduction in the observed shifts in the variability, making such changes more difficult to detect. For a relative shift of magnitude γ=1.1, Table 5 shows the corresponding values of γ_Y^e for the *E(X)* and (c.v.(*X*)) values used in the Monte Carlo studies.

μ	c.v.*(X)*						
	0.01	**0.05**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**
5	1.000	1.000	1.002	1.007	1.014	1.023	1.032
10	1.000	1.002	1.007	1.022	1.039	1.053	1.064
50	1.001	1.014	1.039	1.072	1.085	1.091	1.094
100	1.001	1.018	1.047	1.078	1089	1.093	1.096
1000	1.001	1.020	1.050	1.080	1.090	1.094	1.096
10000	1.001	1.020	1.050	1.080	1.090	1.094	1.096
15000	1.001	1.020	1.050	1.080	1.090	1.094	1.096

Table 5. Values of γ_Y^e corresponding to γ=1.1

The ARL function of the S_Y^e-chart, assuming the normality of Y^e, is

$$ARL = \left(1 - Ch\left(\frac{\chi^2_{n-1,1-\alpha}}{\gamma^2_{Y^e}} \middle| n-1 \right) \right)^{-1} \tag{24}$$

where $Ch(x|p)$ denotes the distribution function of a χ^2 with p degrees of freedom evaluated in x.

The ARL computations corresponding to the γ_Y^e values in Table 5 are reported in the following Table 6, and can be used as a benchmark for the chart's performance studied by means of simulation.

μ	c.v.*(X)*						
	0.01	**0.05**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**
5	99.978	99.462	97.898	92.326	84.942	77.325	70.430
10	99.918	98.008	92.693	78.134	65.493	56.864	51.291
50	99.254	85.355	65.186	47.484	42.187	40.098	39.085
100	99.002	81.740	60.624	45.163	40.983	39.382	38.615
1000	98.877	80.123	58.853	44.355	40.576	39.143	38.459
10000	98.876	80.105	58.834	44.347	40.572	39.141	38.457
15000	98.876	80.105	58.834	44.347	40.572	39.141	38.457

Table 6. Theoretical values of ARL for the γ_Y^e in Table 5

Also for the S_Y^e-chart we simulate the out of control situations for a relative variation γ in the variability of X; the mean detection delays obtained are summarized in Table 7.

μ	c.v.*(X)*						
	0.01	**0.05**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**
5	100.04	99.10	97.09	90.69	83.19	75.47	67.45
10	99.36	96.72	91.07	77.88	63.13	55.33	48,39
50	88.16	74.70	56.59	41.31	37.91	36.15	34.84
100	81.45	67.13	50.28	39.86	36.43	35.24	34.89
1000	79.65	63.91	48.39	38.26	35.04	35.02	34.12
10000	79.17	63.78	48.17	38.18	35.27	34.23	34.22
15000	78.22	63.11	48.70	38.65	35.54	34.74	34.96

Table 7. Mean detection delay for γ=1.1

The results given in Table (7) indicate that the effects of measurement errors are particularly evident for small values of c.v.(X) and large values of E(X). In such situations, the mean detection delays are much smaller than the corresponding ARL values. This result is in keeping with the increased false alarm rates observed under H_0.

4. Dealing with Measurement Errors

The results seen in the previous sections show that the presence of measurement errors causes dramatic modifications in the detection ability of the control charts, leading to difficulties in the practical uses of such monitoring algorithms.

The present section is going to show how to reduce the effects of errors, and how to take errors into account when designing control charts.

As suggested by Linna and Woodall (2001a), repeated measurements enable us to reduce the effects of measurement errors. Let us suppose that for each sampled unit (i=1,2,…,n) we take k measures; as a consequence, the generic observation is

$$Y_{ij}^{e} = \alpha + \beta X_i e^{\eta_{ij}} + \varepsilon_{ij} \tag{25}$$

with j=1,2,…,k.

Therefore the monitored variable is the mean of k-repetitions of the i-measure:

$$\overline{Y}_{i.}^{e} = \frac{1}{k}\sum_{j=1}^{k} Y_{ij}^{e} \tag{26}$$

with expected value and standard deviation given by

$$E\left(\overline{Y}_{i.}^{e}\right) = \alpha + \beta\mu\sqrt{e^{\sigma_\eta^2}} \tag{27}$$

and by

$$\sigma_{\overline{Y}^e_{i.}} = \left(\frac{\beta^2}{(k)^2} Var\left(X_i \left(\frac{1}{k} \sum_{j=1}^{k} e^{\eta_{ij}} \right) \right) + \frac{1}{(k)^2} Var\left(\sum_{j=1}^{k} \varepsilon_{ij} \right) \right)^{1/2} =$$

$$= \left(\beta^2 \sigma^2 e^{\sigma_\eta^2} + \frac{\beta^2 \mu^2 e^{\sigma_\eta^2} \left(e^{\sigma_\eta^2} - 1 \right)}{k} + \frac{\beta^2 \mu^2 e^{\sigma_\eta^2} \left(e^{\sigma_\eta^2} - 1 \right)}{k} + \frac{\sigma_\varepsilon^2}{k} \right)^{1/2} \tag{28}$$

respectively.

4.1. Mean Control Chart

We examined the effects of errors under H_1 first. When monitoring the mean level of $\overline{Y}^e_{i.}$, the sample statistic is

$$\overline{\overline{Y}}^e_{..} = \frac{1}{n} \sum_{i=1}^{n} \overline{Y}^e_{i.} \tag{29}$$

and the "new" control chart is

$$CL^e = \alpha + \beta\mu\sqrt{e^{\sigma_\eta}} \tag{30}$$

$$UCL^e = CL^e + 3\frac{\sigma_{\overline{Y}^e_{i.}}}{\sqrt{n}} \tag{31}$$

$$LCL^e = CL^e - 3\frac{\sigma_{\overline{Y}^e_{i.}}}{\sqrt{n}} \tag{32}$$

Therefore a shift of standardized magnitude δ in X corresponds, in the monitored variable $\overline{Y}^e_{i.}$, to a shift of magnitude

$$\delta_{\overline{Y}^e_{i.}} = \frac{\delta}{\left(1 + \frac{\mu^2 \left(e^{\sigma_\eta^2} - 1 \right)}{k\sigma^2} + \frac{\left(e^{\sigma_\eta^2} - 1 \right)}{k} + \frac{\sigma_\varepsilon^2}{k\beta^2\sigma^2 e^{\sigma_\eta^2}} \right)^{1/2}} \tag{33}$$

Equation (33) shows that repeated measurements, $k>1$, increase the shift $\delta_{\overline{Y}^e_{i.}} > \delta_{Y^e}$ which means a quicker change detection. As an example,

the values of $\delta_{\bar{Y}_{i.}^{e}}$ for k=4 corresponding to a shift of magnitude $|\delta|$=0.5 in the variable X are reported in Table (8). In comparing Table (8) with Table (1), we notice that the effect of k is particularly evident for small values of $E(X)$ and c.v.(X).

μ	c.v.(X)						
	0.01	**0.05**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**
5	0.013	0.066	0.129	0.235	0.312	0.364	0.399
10	0.026	0.125	0.230	0.360	0.420	0.450	0.466
50	0.078	0.309	0.421	0.476	0.488	0.493	0.495
100	0.089	0.336	0.438	0.482	0.491	0.495	0.496
1000	0.095	0.347	0.444	0.483	0.492	0.495	0.497
10000	0.095	0.347	0.444	0.483	0.492	0.495	0.497
15000	0.095	0.347	0.444	0.483	0.492	0.495	0.497

Table 8. Values of $\delta_{\bar{Y}_{i.}^{e}}$ corresponding to a standardized shift of magnitude $|\delta|$=0.5

In order to evaluate the improving effect of repeated measurements in the shift-detection ability of the control chart, we repeated the simulation for both positive and negative shifts (δ=±0.5) for k=4.

μ	c.v.(X)						
	0.01	**0.05**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**
5	360.37	337.91	270.40	151.98	97.33	72.75	58.72
10	357.57	285.12	163.42	74.09	52.36	45.98	41.25
50	351.78	109.86	54.41	39.30	36.51	36.10	35.52
100	346.21	95.16	51.41	39.21	36.78	35.19	34.90
1000	345.04	88.83	49.18	38.14	35.58	35.79	35.10
10000	343.46	88.05	49.56	38.18	36.40	35.32	35.23
15000	343.44	85.74	49.76	39.10	36.12	35.63	35.19

Table 9. Mean detection delay for $\delta = 0.5$ and $k = 4$

μ	c.v.*(X)*						
	0.01	**0.05**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**
5	355.98	319.96	248.40	141.07	92.76	68.83	56.43
10	347.99	247.70	141.06	68.31	49.45	42.94	40.07
50	281.46	84.88	48.24	36.97	34.85	34.82	33.95
100	268.67	72.01	44.48	35.78	35.07	34.54	33.98
1000	265.16	68.80	42.36	35.19	34.53	34.31	33.10
10000	264.71	69.50	42.28	35.52	34.74	34.29	33.63
15000	267.60	68.42	43.26	35.93	34.77	34.09	34.30

Table 10. Mean detection delay for $\delta = -0.5$ *and* $k = 4$

The results summarized in Tables (9) and (10) indicate that in the case of repeated measurements, the differences between performance in cases of a positive shift and that in cases of a negative shift, are less evident than the differences between Tables (3) and (4). This would suggest the feasibility of normal approximation when repeated measurements are adopted.

Table (11) shows ARL values, calculated using

$$ARL(\delta_{\bar{Y}^e_{i.}}) = \left(\Phi\left(-3 + \delta_{\bar{Y}^e_{i.}}\sqrt{n}\right) + \Phi\left(-3 - \delta_{\bar{Y}^e_{i.}}\sqrt{n}\right) \right)^{-1} \tag{34}$$

that correspond to the shifts in Table (8).

μ	c.v.*(X)*						
	0.01	**0.05**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**
5	368.79	334.17	260.01	145.31	92.91	68.98	56.76
10	364.38	264.14	149.63	70.85	50.74	43.29	39.79
50	322.03	94.77	50.41	37.78	35.42	34.59	34.20
100	308.39	80.85	46.17	36.70	34.93	34.31	34.02
1000	302.03	75.99	44.75	36.34	34.77	34.22	33.97
10000	301.96	75.94	44.74	36.34	34.77	34.22	33.97
15000	301.96	75.94	44.74	36.34	34.77	34.22	33.97

Table 11. Theoretical values of ARL for the $\delta_{\bar{Y}^e_{i.}}$ *in Table 8*

If we compare these values with the estimates in Tables (9) and (10), we will now see that the differences are small. More precisely, for positive shift, the mean of the relative absolute difference between the estimated mean detection delay, Table (9), and the theoretical ARL values, Table (11), is 6.93%. In the case of a negative shift, the mean of the relative absolute difference falls to 3.92%.

Therefore, repeated measurements can be said to constitute a robust approach: after calculating the standardized shift $\delta_{\overline{Y}_{i.}^{e}}$ the adoption of the normal approximation enables approximate assessments of the chart's performance (useful for practical purposes) to be made.

Multiple measurements reduce the effects of measurement error also under H_0 . Figures (4) and (5) summarise the false alarm rates, for k=4, of a signal below the LCL (or above the UCL). The plotted points now follow a random pattern around the theoretical value 0.00135.

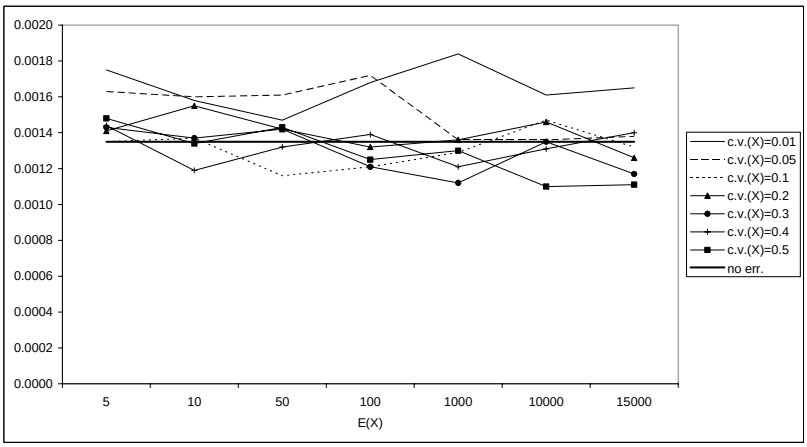

Figure 4. False alarm rates below the LCL (σ_ε =5.698, σ_η =0.1032)

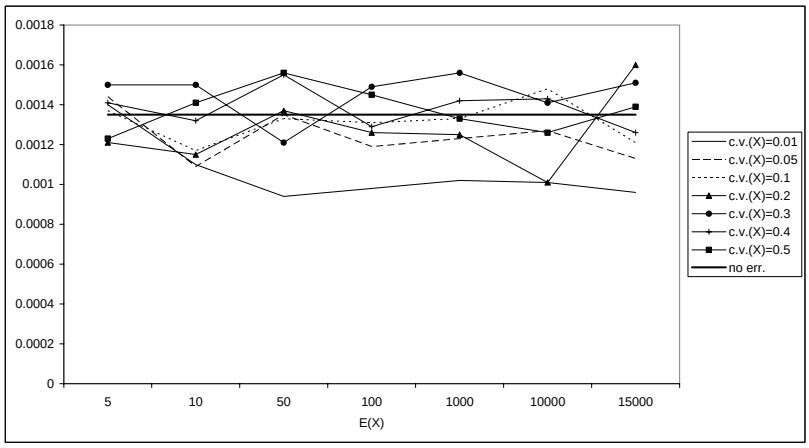

Figure 5. False alarm rates above the UCL (σ_ε=5.698, σ_η=0.1032)

4.2. The $S_{\overline{Y}^e_{i.}}$-Control Chart

In the case of an increase of σ to $\sigma_1=\gamma\sigma$, the relative variation in the monitored $\overline{Y}^e_{i.}$ is defined by

$$\gamma_{\overline{Y}^e_{i.}} = \frac{{}_1\sigma_{\overline{Y}^e_{i.}}}{\sigma_{\overline{Y}^e_{i.}}} = \left(\frac{\beta^2\gamma^2\sigma^2 e^{\sigma_\eta^2} + \frac{\beta^2\mu^2 e^{\sigma_\eta^2}\left(e^{\sigma_\eta^2}-1\right)}{k} + \frac{\beta^2\gamma^2\sigma^2 e^{\sigma_\eta^2}\left(e^{\sigma_\eta^2}-1\right)}{k} + \frac{\sigma_\varepsilon^2}{k}}{\beta^2\sigma^2 e^{\sigma_\eta^2} + \frac{\beta^2\mu^2 e^{\sigma_\eta^2}\left(e^{\sigma_\eta^2}-1\right)}{k} + \frac{\beta^2\sigma^2 e^{\sigma_\eta^2}\left(e^{\sigma_\eta^2}-1\right)}{k} + \frac{\sigma_\varepsilon^2}{k}} \right)^{1/2} \tag{35}$$

It is easy to see that $\gamma_{\overline{Y}^e_{i.}} > \gamma_{Y^e}$ for k>1. This result once again confirms that multiple measurements reduce the effects of measurement errors. Table (12) gives the values of $\gamma_{\overline{Y}^e_{i.}}$ corresponding to γ=1.1, while Table (13) summarises the results of the Monte Carlo study for the $S_{\overline{Y}^e_{i.}}$-Chart with k=4, the focus being on out of control situation.

μ	c.v.*(X)*						
	0.01	**0.05**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**
5	1.000	1.002	1.007	1.023	1.040	1.054	1.065
10	1.000	1.007	1.022	1.053	1.072	1.082	1.088
50	1.003	1.039	1.072	1.091	1.096	1.098	1.098
100	1.003	1.047	1.078	1.093	1.097	1.098	1.099
1000	1.004	1.050	1.080	1.094	1.097	1.098	1.099
10000	1.004	1.050	1.080	1.094	1.097	1.098	1.099
15000	1.004	1.050	1.080	1.094	1.097	1.098	1.099

Table 12. Values of $\gamma_{\bar{Y}_{i.}^{e}}$ corresponding to γ=1.1

μ	c.v.*(X)*						
	0.01	**0.05**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**
5	98.46	96.00	92.02	77.41	65.50	56.18	50.62
10	98.97	92.92	78.69	57.94	47.43	43.87	41,92
50	97.32	65.89	47.13	40.63	37.87	38.02	38.04
100	96.00	60.22	44.45	38.27	38.24	38.25	37.71
1000	94.77	57.19	44.84	39.44	38.51	37.98	37.67
10000	94.61	57.40	44.03	38.81	37.66	37.49	37.68
15000	93.74	58.05	43.59	39.17	38.47	38.06	37.67

Table 13. Mean detection delay for γ=1.1 and k=4

Once again, in order to provide a suitable tool with which to evaluate the chart's performance, we have adopted the normal approximation approach. The theoretical ARL values

$$ARL = \left(1 - Ch\left(\frac{\chi^2_{n-1,1-\alpha}}{\gamma^2_{\bar{Y}_{i.}^{e}}}\middle| n-1\right)\right)^{-1} \tag{36}$$

are shown in Table (14).

μ	c.v.*(X)*						
	0.01	**0.05**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**
5	99.91	97.91	92.38	77.43	64.74	56.21	50.77
10	99.68	92.74	78.25	56.97	47.72	43.55	41,40
50	97.14	65.31	47.55	40.12	38.53	37.96	37.69
100	96.21	60.74	45.22	39.40	38.20	37.77	37.57
1000	95.75	58.97	44.41	39.16	38.09	37.71	35.53
10000	95.75	58.95	44.40	39.16	38.09	37.71	35.53
15000	95.75	58.95	44.40	39.16	38.09	37.71	35.53

Table 14. Theoretical values of ARL for the $\delta_{\overline{Y}_{i.}^{\varepsilon}}$ *in Table 13*

The differences between Table (13) and Table (14) are negligible: the mean of the relative absolute difference is 0.949%.

Once again, multiple measurements reduce the measurement error effect also under H_0 . Figure (6) summarises the false alarm rates, for k=4, of a signal above the UCL.

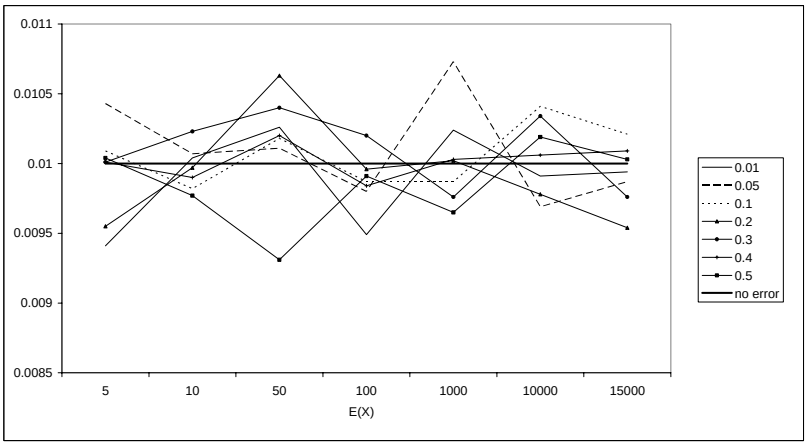

Figure 6. False alarm rates above the UCL for k=4 (σ_ε *=5.698,* σ_η *=0.1032)*

5. Discussion

The two-component error model is particularly well-suited to chemical and environmental monitoring activities, since it provides an accurate approximation of error across the entire usable range of a measurement technology. Hence the importance of studying the role of this error structure in monitoring algorithms.

In the present work, we have studied the effects of the two-component error model on the Shewhart control charts. We noted that the two-component measurement error model leads to a reduction in the detection ability of the control charts. While the additive error basically inflates the variance of the observable response Y^e, the proportional error component η leads to a remarkable asymmetry in the performances of the mean control chart. Under H_0, the false alarm rate above the UCL is greater than the theoretical false alarm probability, while the false alarm rate below the LCL is smaller. This is an important issue, since false alarms could lead to a series of expensive, unnecessary actions or regulations. Moreover, asymmetry can complicate monitoring management. The effects of the two-component error model are also evident under H_1: the mean detection delay for negative shifts is always greater than the corresponding mean detection delay for positive shifts, and in some cases is twice as large. In monitoring chemical plant or environmental phenomena, the consequences of such delays can be very serious.

As far as variability is concerned we have considered a unilateral *S*-chart, since it is usually important to avoid any increase in the spread of the monitored variable. Once again, in this case we discovered increases in false-alarm rates and altered performances for the detection of increments in the variability of the variable. Another important point, which can be put down to the nature of the two-component error model, is that the performance of the control charts is strongly responsive to the values of E(*X*).

We have also suggested how to reduce the effects of measurement error to provide robustness, and how to take errors into account when designing control charts, showing that multiple measurements reduce the asymmetric effects of measurement errors. Therefore, once obtained an evaluation of the shift magnitude, $\delta_{\overline{Y}^e_{i.}}$ or $\gamma_{\overline{Y}^e_{i.}}$, it is possible to get

approximate assessments, through (34) and (37), of control chart performance which may be of use for practical purposes.

Acknowledgments

This work was partly funded by a 2006 MIUR grant (Sector 13: Economics and Statistics, Project Title: Modelli statistici per il supporto alla produzione di dati ambientali).

References

Anderson, M.J., Thompson, A.A. (2004). Multivariate control charts for ecological and environmental monitoring. *Ecological Applications* 14, pp.1921–1935.

Bennet, C.A. (1954). Effect of measurement error on chemical process control. *Industrial Quality Control* 10, pp. 17–20.

Bordignon, S., Scagliarini, M. (2000). Monitoring algorithms for detecting change in the ozone concentrations. *Environmetrics* 11, pp 125–137.

Gibbons, R.D. (1999). Use of combined shewhart-cusum control charts for ground water monitoring applications. *Ground Water* 37, pp 682–691.

Gibbons, R.D., Coleman, D.E. (2001). *Statistical Methods for Detection and Quantification of Environmental Contamination*. John Wiley and Sons, New York.

Kanazuka, T. (1986). The effects of measurement error on the power of X-r charts. *Journal of Quality Technology* 18, pp 91–95.

Linna, K.W., Woodall, W.H. (2001a). Effect of measurement error on shewhart control charts. *Journal of Quality Technology* 33, pp 213–222.

Linna, K.W., Woodall, W.H., Busby, K.L. (2001). The performances of multivariate effect control charts in presence of measurement error. *Journal of Quality Technology* 33, pp 349–355.

Montgomery, D.C. (2005). *Introduction to statistical quality control*, Fifth Edition. John Wiley and Sons, New York.

Mittag, H.J., Stemann, D. (1998). Gauge imprecision effect on the performance of the X−S control chart. *Journal of Applied Statistics* 25(3) pp 307–317.

Rocke, D.M., Lorenzato, S. (1995). A two-component model for measurement error in analytical chemistry. *Technometrics* 37, pp 176–184.

Rocke, D.M., Durbin, B., Wilson, M., Kahn, H.D. (2003). Modeling uncertainty in the measurement of low-level analytes in environmental analysis. *Ecotoxicology and Environmental Safety* 56, pp. 78–92.