

A new alternative Hotelling's T^2 control chart using MEWMA to monitor minimal mean deviation

Kunyun Wang^{1,*}, Qianqian Li², and Guangdong Li³

¹Shanghai Aerospace Electronic Technology Institute, Shanghai, China.

²Zhongni Environment Technology and Development Co.,Ltd, Shanghai, China.

³Anhui University Hefei, China

Abstract. Hotelling T^2 control chart not only reflects the correlations between different quality characteristics but also has good efficiency on monitoring multivariate quality characteristics in production process. A new alternative control chart was constructed after the original products data are processed by using multivariate exponentially weighted moving average for cumulating failure effects because T^2 control chart is ineffective on detecting minimal mean deviations. Exemplified by bivariate quality characteristics, we compared the monitoring effects of Hotelling T^2 control chart and new control chart which is called as T^2_{MEWMA} control chart. Paper showed the improved T^2_{MEWMA} control chart has smaller average run length than Hotelling T^2 control chart on monitoring minimal mean deviation and that also studied the relationships between T^2_{MEWMA} control chart's forgetting factor, sample sizes N and type II error. It indicated the smaller forgetting factor is more sensitive to minimal mean value deviation and that average run length tended to become bigger gradually along with increase of sample sizes N when production process is out of control.

1 Introduction

Quality management is becoming increasingly important in manufacturing and other industries. Quality control is a key element of quality management. H. F. Dodge, H. G. Romig and W. A. Shewhart together created the statistical quality control which the statistics data mining techniques are used in industrial processes for sampling and quality control in the 1920s.

Quality control chart is a common statistical quality control theory and divided by the number of quality characteristics indicators into univariate and multivariate control charts. Although univariate control chart has matured, the actual production process usually accompanies two or more key quality characteristics, so the researches of creating multidimensional quality control charts are urgent. Currently multivariate control charts

* Corresponding author: yunviki@126.com

including MCUSUM, MEWMA, Hotelling T^2 and various optimized control charts based on all above researches [1-12]. Quality characteristic generally follow a given normal distribution in industrial production. Univariate control chart need to control the process mean and volatility simultaneously, multivariate control charts not only need to control both the mean and volatility of each variable but also need to control various correlations.

Hotelling T^2 control chart takes full advantage of the correlations and has a good effect on monitoring multiple quality characteristics, but Lin(2005) studied that hotelling T^2 control chart does not make use of products historical data effectively and detect process mean minimal deviation. Li and Qiu(2012,2014) found that Multivariate exponentially weighted moving average (MEWMA) detects minimal fault before the process mean shift, thereby leading to early prevention and trouble resolutions.

There are two types of errors of control charts. Type I error α , the probability of which control charts judge process unsteady when the production process is steady. Type II error β , the probability of which control charts judge process steady when the production process is unsteady. Generally, you want that Type I error α is as small as possible when process is under control, Type II error β is as small as possible when uncontrolled. In fact, the two types of errors can't become small simultaneously. We usually fix a type error to compare the other in order to determine the efficiency of control charts.

In this paper we proposed an alternative Hotelling's T^2 using MEWMA to detect minimal mean deviation. In Section 2, we introduced the MEWMA process in detail. Afterwards we constructed a new control chart to detect process minimal deviation quickly, namely T^2_{MEWMA} control charts. In Section 3, firstly, we compared two kinds of control chart's average run length by monte-carlo simulations when the process is out of control; secondly, we researched the relationships between λ_0 and detection efficiency of T^2_{MEWMA} control chart on mean minimal deviation; Finally, we studied the efficiency of T^2_{MEWMA} control chart by changing on the sample sizes N. In last Section, We came to some conclusions and considered a further study to be done.

2 A alternative Hotelling T^2 control chart

2.1 Multivariate exponentially weighted moving average

Products quality characteristics follow P multiv-ariate normal distribution, $X \sim N_p(\mu_p, \Sigma_p)$, with the process mean μ_0 and the covariance Σ_0 . The processed data is below according to multivariate exponentially weighted moving average (MEWMA), See formula (1).

$$Z_i = (I - \lambda)Z_{i-1} + \lambda X_i = \lambda \sum_{t=0}^i (I - \lambda)^t X_{i-t} \quad i = 1, \dots, n \tag{1}$$

where $Z_0 = (0, 0, \dots, 0)_{1 \times p}^T$, forgetting factor matrix λ is unknown. The larger λ , the smaller the influence whose historical data make on the current indicators.

$$\lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_p), \quad 1 \leq j \leq p \tag{2}$$

Let us consider the equation(1) can be simplified into $Z_i = (1 - \lambda_0)Z_{i-1} + \lambda_0 X_i$ under this condition if P quality characteristics priorities are equal value . At this moment, the average value and covariance of Z_i will depend on λ_0 .

$$E(Z_i) = E \left\{ \lambda_0 \sum_{j=1}^i (1 - \lambda_0)^{i-j} \right\} = \left[1 - (1 - \lambda_0)^{i-1} \right] \mu_0 \tag{3}$$

$$Var(Z_i) = Var \left\{ \lambda_0 \sum_{j=1}^i (1-\lambda_0)^{i-j} X_j \right\} = \left\{ \lambda_0 / (2-\lambda_0) \right\} \left[1 - (1-\lambda_0)^{2i} \right] \Sigma_0 \tag{4}$$

According to the property of independent normal random variable:

$$Z_i \sim N_p \left(\left[1 - (1-\lambda)^{i-1} \right] \mu_0, \left\{ \lambda / (2-\lambda) \right\} \left[1 - (1-\lambda)^{2i} \right] \Sigma_0 \right) \tag{5}$$

2.2 Group samples (n > 1)

Hotelling T² control chart gradually developed into a main multivariable quality monitoring tool by several years of development which is mentioned by Wang et al(2008) Quality characteristics data extracted by the quality inspector may be a group data or single-valued data. Hotelling(1947) found that the two different data patterns have different distribution characteristics.

Production quality characteristics X follow multivariate normal distribution, $X \sim N_p(\mu_0, \Sigma_0)$.

$$X^{(i)} = (X_1^{(i)}, X_2^{(i)}, \dots, X_n^{(i)}) \quad i = m \tag{6}$$

There are the statistics T².

$$T^2 = n \left(\overline{X^{(i)}} - \mu_0 \right)^T \Sigma_0^{-1} \left(\overline{X^{(i)}} - \mu_0 \right) \quad i = 1, \dots, m \tag{7}$$

where $\overline{X^{(i)}}$ is representative of cell mean, m is the number of independent sample groups, each element of X⁽ⁱ⁾ is also independent. When n > p,

$$T_i^2 / C_0(p, m, n) \sim F(p, mn - n - p + 1) \tag{8}$$

where, $C_0(p, m, n) = [p(m-1)(n-1)](mn - n - p + 1)^{-1}$.

2.3 Single-value sample (n = 1)

When the sample sizes is 1, that is n=1, statistics T² is saw equation (4).

$$T^2 = (X_i - \mu_0)^T \Sigma_0^{-1} (X_i - \mu_0) \quad i = 1, 2, \dots, m \tag{9}$$

We usually consider formula (4) as a special example of equation (2). In fact, statistics T² follow χ^2 distribution with P degrees of freedom, that is

$$T_i^2 = (x - \mu_0)^T \Sigma_0^{-1} (x - \mu_0) \sim \chi^2(p) \tag{10}$$

$$T_i^2 / C_1(m, p) \sim F(p, m - p) \tag{11}$$

where, $C_1(m, p) = p(m+1)(m-1)[m(m-p)]^{-1}$.

If production process is statistical process control in the initial stages, that is, multivariate process mean and covariance matrix are unknown, we should estimate overall mean μ_0 and covariance Σ_0 based on samples. Easy to understand, overall group average \overline{X} and group average covariance S_m respectively are the uniformly minimum variance unbiased estimate of sample population μ_0 and sample population Σ_0 .

$\overline{X^{(i)}}$ is the sample mean value of the group $\overline{X^{(i)}} = \frac{1}{n} \sum_{k=1}^n X_k^i$, Then, $S^{(i)}$ is the covariance array of the group i .

$$S^{(i)} = \{1/(n-1)\} \sum_{k=1}^n \left(X_k^{(i)} - \overline{X^{(i)}} \right)^T \left(X_k^{(i)} - \overline{X^{(i)}} \right) \quad (12)$$

$\overline{\overline{X}}$ is an average value of all of m groups mean value, and S_m is an average covariance of all of m groups average covariance.

$$\overline{\overline{X}} = \frac{1}{m} \sum_{i=1}^m \overline{X^{(i)}} \quad (13)$$

$$S_m = \frac{1}{m} \sum_{i=1}^m S^{(i)} \quad (14)$$

2.4 T_{MEWMA}^2 control chart

Because Bersimis, S. et al thought (2007) that hotelling T^2 control chart is insensitive to process average minimum deviation which., we achieve inspiration from MEWMA and construct a new control chart, namely T_{MEWMA}^2 control chart, which deals with samples and cumulates failure effects according to multivariate exponentially weighted moving average to detect process minimal deviation quickly.

$m \times n$ samples are extracted from production process. First of all, we should have a judgement whether the process is in control by using hotelling T^2 control chart. If you are uncertain on process mean's minimal deviation, the following procedure can solve the problem.

a) m single-value datas are constructed by multivariate exponentially weighted moving average, see equation (1).

b) T_{MEWMA}^2 control chart is constructed by using single-value data to decide whether process mean gets minimal deviation, see equation(9).(11).

c) μ_0 and Σ_0 are estimated based on single-value data, see equations(13).(14).

d) we should focus on the problem how to select the appropriate sample sizes N and forgetting factor matrix λ in the process of implementation of the T_{MEWMA}^2 control chart to make the control chart playing best efficiency.

3 Monte-Carlo simulation

Article compared two different control chart's average run length by monte-carlo simulation in order to confirm whether T^2 control chart is beneficial to detecting process mean slight shift when the process is out of control. In the end, we discussed the relationships between forgetting factor, sample sizes N and detection efficiency of mean minimal deviation about T_{MEWMA}^2 control chart, the article took bivariate quality characteristics for an example. .

Table 1. ARL_1 of T^2 and T^2_{MEWMA} control chart with $\lambda_0 = 0.99$, $n=10$, under different $K(L)$.

$k=l$	100		200		300		370		500	
	T^2	T^2_{MEWMA}	T^2	T^2_{MEWMA}	T^2	T^2_{MEWMA}	T^2	T^2_{MEWMA}	T^2	T^2_{MEWMA}
0	—	—	—	—	—	—	—	—	—	—
0.0001	43.47	6.17	111.11	7.46	200.00	8.62	250.00	9.09	333.33	9.90
0.0005	43.47	6.17	111.11	7.46	200.00	8.62	250.00	9.09	333.33	9.90
0.001	43.47	6.17	111.11	7.46	200.00	8.62	250.00	9.09	333.33	9.90
0.005	43.47	6.17	111.11	7.41	200.00	8.54	250.00	9.09	333.33	10.31
0.01	43.47	6.21	111.11	7.58	200.00	8.62	250.00	9.17	333.33	10.42
0.05	50.00	6.09	90.90	7.41	142.86	8.26	142.86	9.43	333.33	10.31
0.10	45.45	5.91	83.33	7.19	100.00	8.55	111.11	9.01	200.00	9.62
0.15	43.47	5.88	66.66	7.25	100.00	7.94	111.11	8.33	111.11	9.26
0.20	31.25	5.71	50.00	6.94	62.50	7.52	76.92	8.06	100.00	8.85
0.25	21.73	5.81	35.71	6.67	50.00	7.35	58.82	7.75	62.50	8.40
0.30	16.39	5.49	24.39	6.49	34.48	7.04	40.00	7.39	52.63	8.06

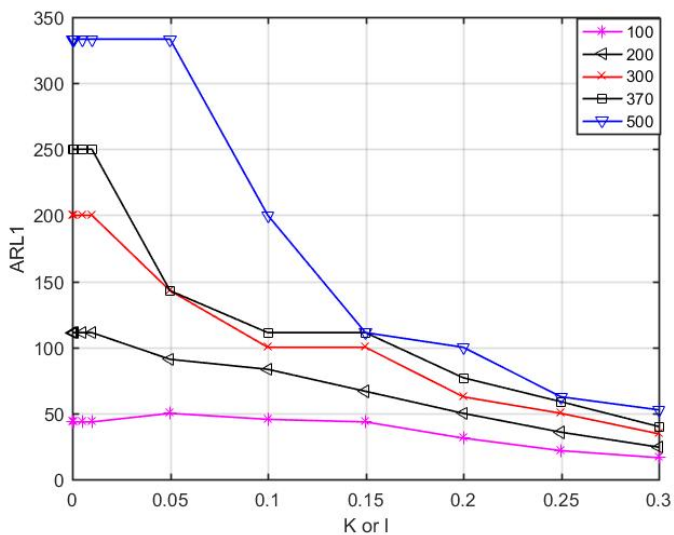


Fig.1. ARL_1 of Hotelling T^2 with ARL_0 and different $K(L)$.

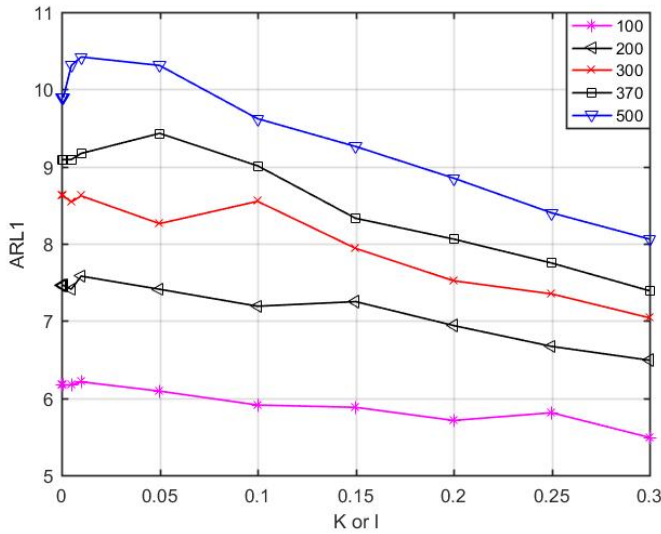


Fig. 2. ARL_1 of T^2_{MEWMA} with ARL_0 and different $K(L)$.

3.1 A comparative analysis of T^2 and T^2_{MEWMA} control chart

Here is the simulation process.

a) $m \times n$ (1000×10) random numbers following two dimensional normal distribution are created by Monte-Carlo simulation, where $\mu_0 = [0,0]$, $\Sigma_0 = [1,0.5; 0.5,1]$.

b) The average run length is 100, 200, 300, 370, 500 in turn when both control charts are in control. Minimal deviations of two quality indicators X_1 and X_2 appeared simultaneously, the offset $K = L$.

$$X'_1 = X_1 + K \cdot \sigma_1; X'_2 = X_2 + L \cdot \sigma_2$$

c) Mean minimal deviation sequentially are 0.0001, 0.0005, 0.0010, 0.0050, 0.0100, 0.0500, 0.1000, 0.1500, 0.2000, 0.2500, 0.3000. Those deviations are cumulated to the two quality indicators. Our simulations are conducted with C program written based on chapter 2. The program counts automatically once the statistic T^2 is less than upper control limit. In the last, the probability of Type 2 error β which the statistics is less upper control limit.

d) Supposing forgetting factor $\lambda_0 = 0.99$.

e) We get a ARL_1 after every simulation, where, $\alpha = 1/ARL_0$, $ARL_1 = 1/(1 - \beta)$, $\beta = P\{T_i^2 < UCL\}$. ARL_1 which is an average run length when the production process is unsteady will be expected smaller to detect process deviation as soon as possible.

f) Upper control limits vary with the hotelling T^2 control chart and T^2_{MEWMA} control chart.

g) According to the known distribution, we can get two kinds of upper control limits of hotelling T^2 control chart and T^2_{MEWMA} control chart.

$$UCL_{T^2} = C_0(p, m, n)F_{1-\alpha}(p, mn - n - p + 1)$$

$$UCL_{T^2_{MEWMA}} = C_1(m, p)F_{1-\alpha}(p, m - p)$$

Table 2. ARL_1 of T^2_{MEWMA} control chart with $\lambda_0 = 0.50, 0.80, 0.90$ under different $K(L)$.

λ_0	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
	50	80	90	50	80	90	50	80	90	50	80	90	50	80	90
$k=l$	100			200			300			370			500		
0	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
0.0001	5.	5.	6.	6.	7.	7.	7.	7.	8.	7.	8.	8.	8.	9.	9.
0.0005	26	78	06	37	46	41	35	87	20	69	26	62	33	01	80
0.001	5.	5.	6.	6.	7.	7.	7.	7.	8.	7.	8.	8.	8.	9.	9.
0.005	24	78	06	41	46	35	25	94	26	63	33	62	40	01	80
0.01	5.	5.	6.	6.	7.	7.	7.	7.	8.	7.	8.	8.	8.	9.	9.
0.05	21	75	06	37	46	30	19	87	47	63	40	62	47	01	80
0.10	5.	5.	5.	6.	6.	7.	6.	8.	8.	7.	8.	8.	7.	9.	9.
0.15	28	71	92	02	94	04	94	00	40	30	47	93	94	17	43
0.20	5.	5.	5.	5.	6.	7.	6.	7.	7.	6.	8.	8.	7.	9.	9.
0.25	08	59	71	88	67	19	54	63	81	90	47	70	52	43	71
0.30	4.	5.	5.	5.	6.	6.	6.	7.	8.	6.	7.	8.	6.	8.	9.
	83	32	65	71	54	75	25	52	00	41	94	47	80	70	01
	4.	5.	5.	5.	6.	6.	5.	7.	7.	6.	7.	8.	6.	8.	8.
	59	29	56	24	41	80	92	14	46	02	30	06	33	13	62
	4.	5.	5.	4.	6.	6.	5.	6.	7.	5.	6.	7.	5.	7.	7.
	33	21	43	98	06	54	35	71	19	52	99	30	99	46	75
	4.	5.	5.	4.	5.	6.	4.	6.	6.	5.	6.	6.	5.	6.	7.
	00	05	32	57	85	17	93	17	67	08	45	90	41	85	46
	5.	5.	6.	6.	7.	7.	7.	7.	8.	7.	8.	8.	8.	9.	9.
	26	75	06	37	46	41	35	94	20	69	26	62	33	01	80

Table 3. ARL_1 of T^2_{MEWMA} with $\lambda_0 = 0.50, K = L = 0.0001$ under different sample sizes N .

N	100	200	300	370	500
5	4.90	5.87	6.43	6.87	7.27
6	5.36	6.48	7.09	7.68	8.37
7	5.27	5.95	6.67	6.83	7.40
8	5.10	6.48	7.49	7.96	8.56
9	5.03	6.01	6.90	7.46	7.99
10	5.2	6.37	7.35	7.69	8.33
15	5.46	6.34	7.16	7.74	8.33
20	5.43	6.58	7.58	7.94	8.93
25	6.15	7.02	8.00	8.51	8.89
30	5.64	7.09	8.33	8.76	9.51

Legends in Figure 1 are match with Figure 2,3,4. Table 1 shows that T^2_{MEWMA} control chart has a smaller average run length than hotelling T^2 control chart on monitoring minimal mean shift when sample sizes $N=10$, forgetting factor $\lambda_0 = 0.99$, $ARL_0 = 100, 200, 300, 370, 500$. What's more, we find that T^2 control chart only has efficiency when the mean deviation gets big Figure 1 reflects that ARL_1 decreases rapidly by accident when mean deviation became

an uncertain bigger number no matter what ARL_0 is. Figure 2 reflects that ARL_1 decreases smoothly with increase of the mean deviation λ_0 regardless of ARL_0 . ger rather than smaller. In contrast, this problem does not appear in T_{MEWMA}^2 control chart.

3.2 Relationships between forgetting factor and mean minimal deviation detection efficiency about T_{MEWMA}^2 control chart.

T_{MEWMA}^2 control chart is implemented with new data which are produced by MEWMA based on the construction principle of T_{MEWMA}^2 . Formula.(1) shows that influences of historical data on the new data will reduce when forgetting factor λ_0 gradually close to 1; In contrast, influences will increase when gradually close to 0. What are the influences of forgetting factor on the new control chart? Answers will be revealed by the simulation. The research is similar to the simulation method in section 3.1, the forgetting factor is the only difference, $\lambda_0=0.50, 0.80, 0.90$ in turn.

Table 2 show, average run length (out of control) will reduce with the decrease of λ_0 when average run length (in control) and the mean deviation hold at the same value. Therefore, it indicates that the smaller forgetting factor λ_0 is more sensitive to mean minimal deviation for T_{MEWMA}^2 control chart. Figure 3 reveals the above phenomenon, data from the 5th row of Table 1 and 2 with mean minimal deviation 0.0001, $ARL_0=100, 200, 300, 370, 500$, λ_0 from left to right on the horizontal axis is 0.99, 0.90, 0.80, 0.50.

Actually, Figure 1-3 also demonstrates the relationship between two types of error in all of control charts. The smaller ARL_0 is, the larger type 1 error α is, because type 1 error α and type 2 error β cannot change toward the same direction at the same time, that is to say, the larger α is, the smaller β is; In addition, the production process and also change to the same direction simultaneously. Therefore, ARL_0 and ARL_1 change toward the same direction, the phenomenon in Figure1-3 that the whole curve strongly confirm the relationship between two types of error stay at the bottom when all average run length are 100.

3.3 Relationships between N and mean minimal deviation detection efficiency about control chart.

As previously described, T_{MEWMA}^2 control chart has a smaller average run length than T^2 control chart on detecting minimal mean value deviation, Then there is a problem whether sample sizes N affects the detection efficiency of T_{MEWMA}^2 . The research results are showed on Table.3 and Figure4, $K(L)=0.0001, \lambda_0=0.5$, Sample sizes N =5, 6, 7, 8, 9, 10, 15, 20, 25, 30 from left to right on the horizontal axis in Fig.4.

ARL_1 nearly has no change with the N variation from 5 to 30 regardless of ARL_0 on table 3. It reveals that significant change of ARL_1 cannot be caused with the variation in sample sizes N. However, ARL_1 tends to become bigger gradually along with the increase of sample sizes N, these conclusions are depicted in Fig.4.

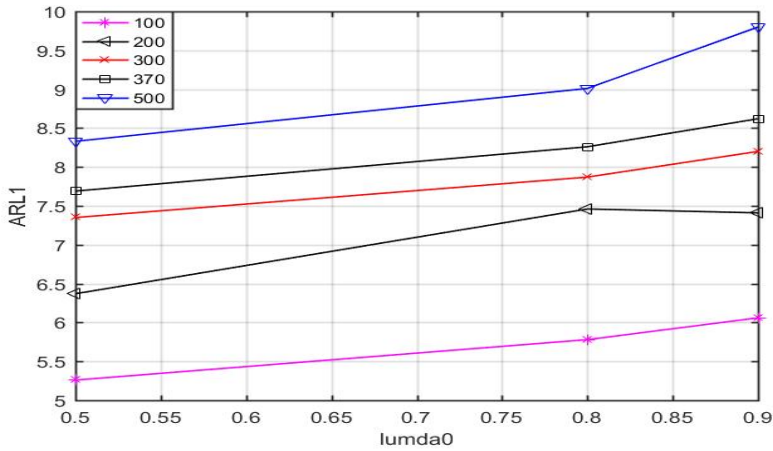


Fig. 3. ARL_1 depends on λ_0 with different ARL_0 .

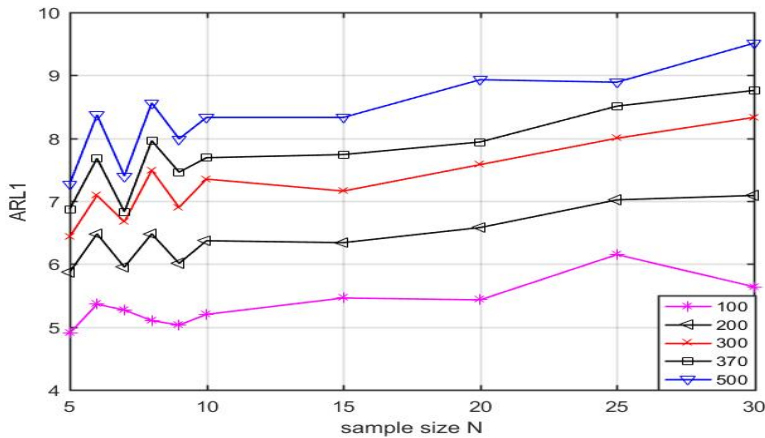


Fig. 4. ARL_1 depends on sample sizes N with different ARL_0 .

4 Conclusions

We proposed a new alternative multivariate control chart which is based on multivariate exponentially weighted moving average for monitoring process mean minimal deviation. The above study results showed:

(1) T^2_{MEWMA} control chart has a smaller average run length than Hotelling T^2 control chart on monitoring minimal mean deviation in statistical quality control. Hotelling T^2 control chart has lower efficiency when the mean deviation gets smaller rather than bigger.

(2) ARL_1 will reduce with the forgetting factor λ_0 decreasing when ARL_0 and mean value deviation hold at own same value. Therefore, it indicates that the smaller forgetting factor λ_0 is more sensitive to minimal mean value deviation about T^2_{MEWMA} .

(3) Significant changes in ARL_1 cannot be caused of the variation of sample sizes N, However, ARL_1 tends to become large gradually along with the increase in sample sizes N.

Meanwhile, it is worth noting that the study is based on two dimensional quality characteristics, a further study needs to be done on whether the research findings can be applied into high dimensional quality characteristics for multivariate statistical quality control. The selection of forgetting factor discussed with consumers and manufactures ought to be determined to achieve optimal economic benefits if the T_{MEWMA}^2 control chart will be used in statistical quality control.

References

1. Kang, C.W. & Kham, P.H. 2012 Basic Statistical Tools for Improving Quality, Wiley Online Library.
2. Kho, M.B. 2011. Optimal Designs of the Multivariate Synthetic Chart for monitoring the Process Mean Vector Based on Median Run Length. *Quality and Reliability Engineering International*, pp 981–997.
3. Farias, A & Sani gab, E.2011 Economic Statistical Design of a T^2 Control Chart with double Warning Lines. *Quality and Reliability Engineering International*, pp 125–139.
4. Wen, D. & Wu, Z.2012 A Comparison Study on Effectiveness and Robustness of Control charts for Monitoring Process Mean and Variance. *Quality and Reliability Engineering International*, pp 3–17.
5. Mestres, M.V.2014 Out-of-Control Signals in Three-Part Compositional T^2 Control Chart. *Quality and Reliability Engineering International*, pp 337–346.
6. Cheri, Z. T^2 Control Chart in a Multimodal Framework: Detection of Incipient Trends.
7. Haddad, F.S. Syed-Yahiya, S.S. & Alfaro, J.L. 2013 Alternative Hotelling T^2 Charts using Winsor zed Modified One-Step M-estimator. *Quality and Reliability Engineering International*, pp 583–593.
8. Alfaro, J.L & Ortega, J.F. 2008 A Robust Alternative to Hotelling's T^2 Control Chart Using Trimmed Estimators. *Quality and Reliability Engineering International*, pp 601–611.
9. Abu-Shawiesh, M.O.2014 A Robust Bivariate Control Chart Alternative to the Hotelling's T^2 Control Chart. *Quality and Reliability Engineering International*, pp 25–35.
10. Faraz, A. 2012 Optimal T^2 Control Chart with a Double Sampling Scheme: An Alternative to the MEWMA Chart. *Quality and Reliability Engineering International*, pp 751–760.
11. Charles W. Champ.2008 Double Sampling Hotelling's T^2 Charts, *Quality and Reliability Engineering International*, pp 153–166.
12. Celano, G. & Castagliola, P.2014. Design of a Phase II Control Chart for Monitoring the Ratio of two Normal Variables, *Quality and Reliability Engineering International*.
13. Lin, S.M. 2005 Small batch of standardized T^2 control chart. *Quality and Reliability*, pp 14–17.
14. Qiu, T, Bai, X.J, Zhu, X, Zheng, Q.Y. 2014. Incipient fault detection of multivariate exponentially weighted moving average principal component analysis. *Control Theory & Applications*, pp 19–26.
15. Li, J, 2012, Review of incipient fault diagnosis methods. *Control Theory & Application*, pp 1517–1529.

16. Wang, Z. J, Gong, Z. & Zou, C. L. 2008 The Calculation of Average Run Length: A Review, Tianjin, china.
17. Hotelling, Multivariate Quality Control Illustrated by the Air Testing of Sample Bomb Sights, Techniques of Statistical Analysis, Ch