Inverse Problems, Design and Optimization Symposium
Rio de Janeiro, Brazil, 2004

A MULTIOBJECTIVE OPTIMIZATION FOR THE EWMA AND MEWMA QUALITY CONTROL CHARTS

Francisco Aparisi
Departamento de Estadística e Investigación Operativa Aplicadas y Calidad.
Universidad Politécnica de Valencia.
46022 Valencia (Spain).
faparisi@eio.upv.es

J. Carlos García-Díaz
Departamento de Estadística e Investigación Operativa Aplicadas y Calidad.
Universidad Politécnica de Valencia.
46022 Valencia (Spain).
juagardi@eio.upv.es

ABSTRACT

The Multivariate EWMA control chart, MEWMA, Lowry, Woodall, Champ and Ridgon [1] and its univariate version EWMA, may be designed to efficiently detect small shifts in the mean vector of a set of *p* quality characteristics of a production process. However, this work presents a method for the optimal design of MEWMA and EWMA charts parameters to control processes where it is not convenient to detect small magnitude shifts and, at the same time, powerful enough to detect shifts considered important. This problem can be considered as a multiobjective optimization.

Woodall [2] studied the statistical design of control charts and recommended choosing the magnitude of the shift that it is important to detect as a design criterion for control charts. For this purpose, he suggested defining three regions: in-control, indifferent, and out-of-control. These regions will be delimited by two values (*A* and *B*).

The main objective of this paper is to find the best MEWMA and EWMA quality control charts given the previous regions, where the requirements for each region has to be balanced to decide which solution is better. For this purpose, friendly Windows software has been developed to optimize this problem, using Genetic Algorithms.

A comparison is made among the EWMA chart designed employing this software, the typical design of a EWMA chart and the Shewhart control chart. Results show that the design using our approach outperforms the other designs.

NOMENCLATURE

ARL. Average run length.
ARL_0. ARL when the process is in-control.
ARL_A. ARL for point $d = A$.
ARL_B. ARL for point $d = B$.
ARLmin. ARL minimum desired for $d = 0$.
ARL_{PA}. ARL desired for point $d = A$.
d. Mahalanobis' distance.
L. EWMA or MEWMA control limit.
n. Sample size.
r. Smooting constant for EWMA or MEWMA control charts.

INTRODUCTION

Nowadays it begins to be common to face problems or applications where the mathematical modelling produces a optimization problem with several objectives. The multiobjective optimization consists of optimizing simultaneously several objective functions. In many cases, some of the objective functions represent more or less conflicting criteria. Obviously, in these cases no unique solution can be found because the entire objective functions cannot be optimized (maximized or minimized) without considering the effect of the experimental changes in the other response functions.

In general terms, the optimization problem can be formulated as follows, being *n* the number of decision variables, x_j, *m* restrictions and *p* objectives:

Find $\boldsymbol{x}$ $(x_1, x_2, \ldots, x_n)$ that
Maximize / minimize $Z = (z_1(\boldsymbol{x}), z_2(\boldsymbol{x}), \ldots, z_n(\boldsymbol{x}))$
Subject to $\boldsymbol{x} \in F$

With $F \subset R^n$, F feasible region of solutions space R^n and $Z = z(F) \subset R^p$, Z feasible region of objectives space R^p. Many times the set F can be written as $F=\{ \boldsymbol{x} \in R^n: g_i(\boldsymbol{x}) \leq 0,\ x_j \leq 0, \forall i, j \}$ when g_i functions are the restrictions. In some cases, variables z_k are called objective functions or objectives.

One feasible solution $\boldsymbol{x}$ is efficient, no dominated or Pareto optimum, if there is no another feasible solution $\boldsymbol{x^*}$ that improves the values of one objective without worsening one of the other objectives. The set of all the efficient solutions is called efficient set or Pareto front.

Conventionally multiobjective optimization problems have been tackled trying to find a single optimum solution, using the so-called preference-based methods, which assume, explicitly or implicitly, a hierarchy in the objectives importance. In the best case, these approaches conduct to a single optimum solution located in the Pareto front, although no information about the Pareto front is produced.

One of the most used methods is minimizing weighted sums of functions.

Mathematically, this method is expressed as:

$$\text{Maximize } z(\boldsymbol{x}) = \sum_{k=1}^{p} w_k z_k(x)$$

Subject to $\boldsymbol{x} \in F$

where $w_k \geq 0$ is the weight corresponding to objective $z_k(\boldsymbol{x})$ and can be interpreted as the importance of objective k in comparison with the rest of objectives. Now the problem is reduced to find $P(w)$ where $w = (w_1, w_2, \ldots, w_p)$. Hence, the multiobjective problem is now reduced to a unique optimization problem.

The objective of this paper is to apply multiobjective optimization to the design of MEWMA and EWMA quality control charts.

EWMA AND MEWMA CONTROL CHARTS

The statistical design of a quality control chart like EWMA o MEWMA consists of selecting three parameters. The power of the chart (measured through ARL) depends on these parameters, sample size, n, position of control limits, L, and a smoothing constant r.

EWMA (Exponentially Weighted Moving-Average) control charts were introduced by Roberts [3] as an alternative to Shewhart control charts for the detection of small shifts in the process. However, Shewhart control chart only takes into account the present information of the process and does not detect quickly changes smaller than $2s$. EWMA control charts take into account present and past information and therefore they are more efficient (fast) in detecting small shifts (Montgomery [4]). A widely used measurement of the efficiency of a process statistical control method is the ARL (Average Run Length). The ARL is the average number of samples to take (points in the chart) until an out-of-control-signal appears.

In the case of EWMA, the statistical data to chart Z_i, to be compared with control limits at instant i, is obtained as a weighted average value according to parameter r between the observed value $\overline{X}_i$ and the smoothed value Z_{i-1} , following expression:

$$Z_i = r\overline{X}_i + (1-r)Z_{i-1} \qquad (1)$$

As it can be observed, weighting is done with parameter r so that the smaller the parameter is, the greater the influence of past observations as weight decreases geometrically in function of r. This is the reason why EWMA control charts are said to have "human memory" since they provide weights to data exponentially: assigning more weight to present data which decreases as data are far back in the past. When $r = 1$, then the average value is represented by the Shewhart control chart, and when $r = 0$, Z_i is a constant equal to m_0.

EWMA control chart sensitivity to detect changes in the process depends on the value of r. When r tends to 1 EWMA values will depend on the most recent observations and the behaviour of the control chart is similar to that of the Shewhart control chart. However, as r tends to 0, the historical behaviour of the process gets more weight, and then it approaches the behaviour of normal CUSUM charts. A recommended value for r is 0.2 (Hunter [5]). For Z_0 the value adopted is the nominal average value m_0 or the sampling average value in in-control processes.

Some authors (Hunter [5], Crowder [6] and Lucas and Saccucci [7]) have studied the properties of this chart for the statistical control of industrial processes.

Let's analyze the design of the chart. If the quality variable to control is distributed according to $N(m_0, s_0)$ in in-control processes and the observations are independent, therefore, the control limits of the EWMA control chart are calculated with the approximate expression:

$$\text{UCL} = m_0 + \text{L} \cdot s_0 / \sqrt{n} \cdot \sqrt{\frac{r}{2-r}} \qquad (2)$$

$$\text{LCL} = m_0 - \text{L}\cdot s_0/\sqrt{n}\cdot\sqrt{\frac{r}{2-r}}$$

where L and r are selected to get a given in-control ARL and n is the size of the subgroup. A typical value of L is 3, following the criterion $3s$, of the Shewhart control chart. If we want to obtain an in-control ARL, ARL_0, of 370.4 (α= 0.0027), then we should fix the value of r = 0.25 and L = 2.898.

The first reference on multivariate EWMA (MEWMA) control charts corresponds to Lowry, Woodall, Champ and Rigdon [1] who define MEWMA as an extension of the univariate EWMA. Hotelling's T^2 multivariate control chart only takes into account current process data, whereas MEWMA chart also includes past data, thereby it being more powerful to detect small changes in the process.

Univariate systems only controlled one quality variable or characteristic. In multivariate systems a set of p interrelated variables will be controlled. In this latter case, $\vec{\bar{X}}_1, \vec{\bar{X}}_2$..., are run length vectors p which represent the sampling average values of the process. Let random vectors $\vec{\bar{X}}_i$ be independent and equally distributed following a p-variate normal variable of vector $\vec{m}$ and covariance matrix Σ, $\vec{\bar{X}}_i$ iid$\approx N_p(\vec{m}, \Sigma)$. The process will be under control if $\vec{m} = \vec{m}_0$ and out of control in the opposite case.

Vector $\vec{Z}_i$ is defined as

$$\vec{Z}_i = r\vec{\bar{X}}_i + (1-r)\vec{Z}_{i-1}\,, \quad i \geq 1 \tag{3}$$

the starting vector being $\vec{Z}_0 = \vec{m}_0$ since the process is under control $E(\vec{Z}_i) = \vec{m}_0$ and covariance matrix of $\vec{Z}_i$ is Σ_{Z_i} whose expression is given here below. $\vec{\bar{X}}_i$ is the vector of the sampling data and r is a scalar value between 0 and 1. If r =1 we will obtain Hotelling's T^2 control chart. The statistical data charted T_i^2 is defined as

$$T_i^2 = \vec{Z}_i{}'\Sigma_{Z_i}^{-1}\vec{Z}_i \tag{4}$$

where $\Sigma_{Z_i}^{-1}$ is the inverse of the variance-covariance matrix of Z_i. The covariance matrix of Z_i is expressed by:

$$\Sigma_{Z_i} = \frac{r\left[1-(1-r)^{2i}\right]}{2-r}\Sigma \tag{5}$$

The measurement of vector shift (or distance between two vectors) used in multivariate analysis is Mahalanobis' distance. In our case, the distance between the original mean vector and the new mean vector is $d = \sqrt{(m_i - m_0)'\,\Sigma^{-1}(m_i - m_0)}$. The ARL performance of the MEWMA chart depends only on the noncentrality parameter $l = nd^2$, where n is sample size (Lowry, Woodall, Champ and Rigdon [1] and Lowry and Montgomery [8]).

For the design of the MEWMA control chart, the asymptotic covariance matrix can be used, given by:

$$\Sigma_Z = \lim_{i\to\infty} \Sigma_{Z_i} = \left(\frac{r}{2-r}\right)\Sigma \tag{6}$$

similarly to what happened in univariate systems for individual observations. For sample size other than 1, equation (6) corrected by n will be obtained (Rigdon [9])

$$\Sigma_Z = \left(\frac{r}{(2-r)\,n}\right)\Sigma \tag{7}$$

The chart displays an out-of-control signal when T_i^2 > h, where h is the control limit selected to obtain a given value of ARL for in-control processes (ARL_0).

The comparison of the ARLs (obtained through simulation) presented in Rigdon's work shows that MEWMA, using the exact covariance matrix given by equation (5), is somewhat higher than MCUSUM, specially when the average value vector presents great changes.

Woodall [2] studied the statistical design of control charts and recommended choosing the magnitude of the shift that it is important to detect as a design criterion for control charts. For this purpose, he suggested defining three regions: in-control, indifferent, and out-of-control. These regions will be limited by two values (A and B), as follows:

a) In-control region [0, A]. This region corresponds to a state equivalent to one in-control and is made up of a shift change that ranges from

d = 0 to d = A. No shift detection is required in this region. A maximum ARL is needed. If the chart shows an out-of-control sign, this is regarded as a false alarm.

b) Out-of-control region [B, 8 [, corresponding to the shift value $d > B$. Maximum detection power is required from this area. A minimum ARL is needed.

c) Indifferent region,]B, A[, covering $d > A$ and $d < B$. This region is indifferent if the process shift is detected or not.

Therefore, giving A and B values and the number of variables to control simultaneously we desire to find the parameters of EWMA and MEWMA control charts (r, L and n) that satisfy the Woodall regions. In adittion, a minimum in-control ARL (ARL_0) is spcefied and the ARL for $d = A$ has to be equal to a given one, ARL_A.

This value ARL_A is arestriction that will help to make comparisons against $\overline{X}$ chart Hotelling's T^2 control chart, as shown in figure 1.

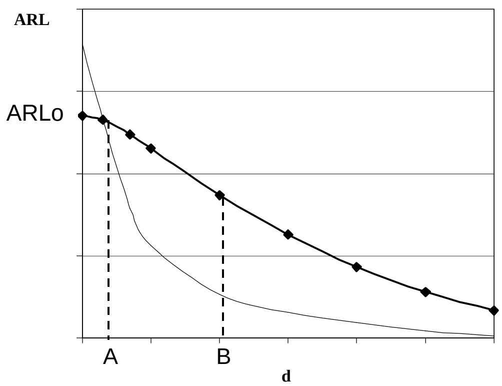

Figure 1. ARL curve.

In this work, the additive utility function method has been employed. This procedure converts the multiobjective problem into a optimization problem with only one objective. This methos is based on defining a function that combines the different objectives, using weights that show the relative importance of each objective for the user. Once this function is obtained, the uni-objective problem is solved.

In this case we have two objectives *(p = 2)* to optimize:

$$z_1(x) = ARLo - ARL_{\min} \quad \text{and} \quad z_2(x) = -ARL_B$$

where $z_1(x)$ is a objective to maximize, as it is desired to have control charts that satisfy $ARLo \geq ARL_{\min}$ and $z_2(x)$ has a negative values because ARL_B has to be minimum. Finally, our optimization problem is:

$$\text{Maximize} \quad z(\boldsymbol{x}) = \sum_{k=1}^{p} w_k z_k(x) = w_1 z_1(x) + w_2 z_2(x) = w_1 \cdot (ARLo - ARL_{\min}) - w_2 \cdot ARL_B$$

Subject to

$$|ARL_{PA} - ARL_A| \leq tol \quad \text{and} \quad ARLo \geq ARL_{\min}$$

where ARL_0 is the in-control ARL for $d = 0$, ARL_B is the ARL for point B, ARL_A is the real ARL for point A, ARL_{PA} is the ARL desired (user input) in point A, ARL_{min} is the desired minimum ARL for $d = 0$ and w_1, w_2 are the weights. In our case we have employed the weights $w_1 = 1$ and $w_2 = 50$.

OPTIMUM SEARCH USING GENETICS ALGORITHMS.

Genetic Algorithms (GA) are optimization algorithms based on the natural evolution of the species (Holland [10], Goldberg [11]). The search for the global optimum value in an optimization problem is carried out when an initial population (generation) of individuals passes to a new population (next generation) through the application of genetic operators. In the original population, each individual represents a possible solution to the optimization problem, that is, a population of individuals consists of a set of possible solutions to the problem to optimize. The principles, implementation and applications of GA can be followed in Bäck [12], Chambers [13] and Michalewicz [14] .

The assessment function, referred to as fitness function, assigns to each individual of the population (set of possible solutions to the problem to optimize) the fitness value, which indicates the fitness of that individual with respect to the other individuals of the population. The "fitness" value is a quality value of the individual and the only data processed by GA to search for the best solution to the problem. Its correct definition allows for a better operation of the algorithm since to find the global optimum value the search is exclusively guided by the "fitness" value of the possible solutions.

Prior to the application of the genetic algorithm we have to code the solutions, that is, it is necessary to define how to better represent each possible solution to the problem, an aspect that is essential for the design and efficiency of the GA. The genetic algorithm operates on a coded representation of the solutions, equivalent to the genetic material of an individual, and not directly on the solutions. These parameters known as genes, form chains referred to as chromosomes. In this paper the following crossover mechanisms has been employed to these chromosomes: 1 point, 2 points and uniform (Bäck [12], Beasley et al [15, 16]), obtaining the best results with the 2 points operator.

During the last years, many researchers have paid attention to the problems involved with multiobjective optimization (Schaffer [17], Tabucanon [18], Fonseca and Fleming [19], Zitzler et al. [20], Coello et al. [21]). The fisrt multiobjective GA was the named Vector Evaluated Genetic Algorithm (VEGA), Schaffer [17]. Recently, more prefectioned GA has been proposed,. The most importants are: the Multiobjetive Genetic Algorithm (MOGA), Fonseca and Fleming, [22], the Niche Pareto Genetic Algorithm (NPGA), Horn et al. [23], the Nondominated Sorting Genetic Algorithm (NSGA), Srinivas and Deb [24], the Strength Pareto Evolutionary Algorithm (SPEA), Zitzler and Thiele [25] and the Pareto-Archivied Evolutionary Strategy (PAES), Knowles and Corne [26].

To solve this multiobjective optimization friendly Windows software has been developed.

RESULTS.EXAMPLE OF APPLICATION.

We will now move on to the optimum design of the EWMA control chart, with a sample size not fixed previously, by using the software developed. We call the EWMA chart found by the software developed in this work "EWMA-Regions". We wish to compare the chart obtained with an $\overline{X}$ chart for the ARL in-control (d = 0) of 500, presenting an ARL in $d = A = 0.25$ of 373.88.

The programme entry data are: minimum desirable ARL_0 of 1500, and ARL_{PA} of 373.88. Once the programme has been run we obtain the "EWMA-Regions" optimum control chart (r = 0.91, L = 3.4, n = 5).

In Figure 2, the ARL values for the chart "EWMA-Regions" is compared to the ARL for the $\overline{X}$ chart. Also we include the EWMA control chart optimum to detect a shift of size $d = B = 1.5$. This EWMA chart is called "EWMA-Point", because it is optimum for only this point. Running the software developed by Aparisi and García-Díaz [27] the "EWMA-point" control chart is characterised by the parameters L = 3.09 and r = 0. 85 for n = 5 and an ARLo = 500. A copy of this software can be downloaded at http://ttt.upv.es/~faparisi.

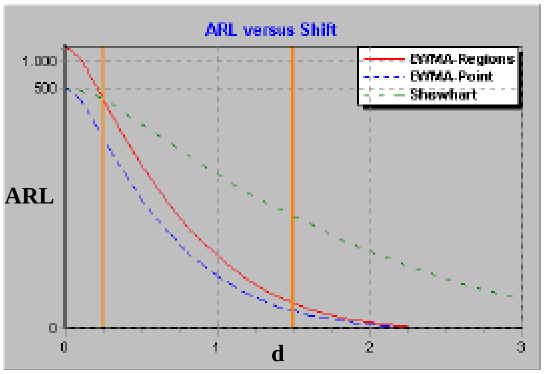

Figure 2. ARL comparison.

It can be seen how the "EWMA-Region" and $\overline{X}$ control charts have the same ARL in point A. This defines the region of shifts we are not interested in detecting (d < 0.25). As we commented in Section 3, the best control scheme would be one where the region d < 0.25 has the largest ARL value (lowest power), and presents the smallest ARL value (maximum power) for shift magnitudes d > 1.5.

Compared to the "EWMA-Point", the "EWMA-Regions" chart offers the advantage of producing lower probability of false alarms. This is because its ARL is much higher in the region of shifts not to be detected, d < 0.25. On the other hand, it can be seen that for shifts that are genuinely important to detect, d > 1.5, both EWMA charts show very similar powers, although the EWMA-point is slightly more powerful. In the indifferent region, 0.25 < d < 1.5, the optimum EWMA is more powerful in d = 1.5, although, as it was discussed before, the ARLs in this region are not important.

The final conclusion to be drawn is that the EWMA chart obtained using the software developed in this paper practically have the same power for detecting genuinely important shifts than the EWMA chart that is more efficient at detecting shifts in d = 1.5. However, the "EWMA-Regions" charts has a very low probability of false alarm in the in-control region.

CONCLUSIONS.

In view of the results discussed in this work, we may draw the following conclusions.

The genetic algorithms technique has proved to be a suitable method for the optimisation of the EWMA and MEWMA control charts using the regions defined by Woodall [2]. An easy-to-use software programme has been developed in a Windows environment, enabling the optimum parameters of these charts to be obtained. These are used for controlling processes where minimum power is required for detecting extremely small shifts, and maximum power for detecting genuinely important ones.

It is possible to design EWMA and MEWMA charts that may reveal a very low false alarm probability and that are, at the same time, genuinely powerful in detecting shifts considered important. The use of these charts would represent an extremely significant control tool in both practical situations as well as capable processes, processes hard to adjust, or whose cost of adjustment is high.

Acknowledgements

The authors acknowledge the financial support of the Ministry of Science and Technology of Spain, Research Project Reference DPI2002-03537 and European FEDER funding.

REFERENCES

1. Lowry CA, Woodall WH, Champ CW, Rigdon SE. A Multivariate Exponentially Weighted Moving Average Control Chart. Technometrics 1992; 34: 46-53
2. Woodall, W.H. (1985). The statistical design of quality control charts. *The Statistician* , 34, 155-160.
3. Roberts, S. W. (1959). Control chart tests based on Geometrics Moving Averages. *Technometrics*, 1, 239-250.
4. Montgomery, D.C. Introduction to statistical quality control. New York: 4end edn, John Wiley, 2001.
5. Hunter JS. The exponentially weighted moving average. Journal of Quality Technology 1986; 18: 155-162
6. Crowder SV. Design of Exponentially Weighted Moving Average Schemes. Journal of Quality Technology 1989; 21: 155-162
7. Lucas JM, Saccucci MS. Exponentially Weighted Moving Average Control Schemes: Properties and Enhancements. Technometrics 1990; 32: 1-29.
8. Lowry CA, Montgomery DC. A review of multivariate control charts. IIE Transactions 1995; 27: 800-810.
9. Rigdon SE. An integral equation for the in-control average run length of a multivariate exponentially weighted moving average control chart. Journal of Statistical Computation and Simulation 1995a; 52: 351-365.
10. Holland J. Adaptation in natural and artificial systems. Ann Arbor: University of Michigan Press, 1975.
11. Goldberg DE . Genetic Algorithms in Search, Optimization, and Machine Learning. Addison Wesley, 1989.
12. Bäck T. Evolutionary algorithms in theory and practice. Evolutionary strategies, evolutionary programming, genetics algorithms. New York: Oxford University Press, 1996.
13. Chambers LD. Practical Handboock of Genetic Algorithms: Applications, Vol.1. CRC Press, 1995.
14. Michalewicz Z. Genetic algorithms + data structures = evolution programs. Berlin: Springer, 1996.
15. Beasley D, Bull D, Martin R. An overview of genetic algorithms: Part 1. Fundamentals. University Computing, 1993a; 15: 58-69.
16. Beasley D, Bull D, Martin R. An overview of genetic algorithms: Part 2. Fundamentals. University Computing, 1993b; 15: 170-181.
17. Schaffer JD. Multiple objective optimization with vector evaluated genetic algorithms. Genetic Algorithms and Their Applications: Proc. Ist. Int. Conf. Genetic Algorithms. Hillsdale, NJ: Lawrence Erlbaum, 1985: 93-100.
18. Tabucanon MT. Multiple Criteria Decision Making in Industry. Elsevier Science Publishers, 1989.
19. Fonseca CM, Fleming PJ. An Overview of Evolutionary Algorithms in Multiobjective Optimization. Evolutionary Computation 1995; 3:1-16.

20. Zitzler E, Deb K, Thiele L, Coello Coello CA, Corne D. Eds, Evolutionary multi-criterion optimization . Proc. Ist. Int. Conf. (EMO'01): Springer, Berlin, 2001, vol. 1993. Lecture Notes in Computer Science.

21. Coello Coello CA, Van Veldhuizen DA, Lamont GB. Evolutionary Algorithms for Solving Multi-Objective Problemns.New York: Kluwer, 2002.

22. Fonseca CM, Fleming PJ. Genetic algorithms for multiobjective optimization: Formulation, discussion and generalization. Proc. 5th Int. Conf. Genetic Algorithms. Ed. San Mateo, CA: Morgan Kaufmann, 1993: 416-423.

23. Horn J, Nafpliotis N, Goldberg DE. A niched pareto genetic algorithm for multiobjective optimization. Proc. 1st IEEE Conf. Evolutionary Computation, IEEE World Congress on Computational Intelligence. Piscataway, NJ: IEEE Service Center, 1994; 1: 82-87.

24. Srinivas N, Deb K. Multiobjective Optimization Using Nondominated Sorting in Genetic Algorithms. Evolutionary Computation 1995; 2: 221-248.

25. Zitzler E, Thiele L. Multiobjective evolutionary algorithms: A compative case study and the strength pareto approach. IEEE Trans. Evol. Comput,1999; 3: 257-271.

26. Knowles J, Corne DC. Aproximating the Nondominated Front Using Pareto Archivied Evolutionary Strategy. Evolutionary Computation, 2000; 8: 149-172.

27. Aparisi F, García-Díaz, JC. Optimization of univariate and multivariate exponentially weighted moving-average control charts using genetic algorithms. Computers and Operations Research, 2004 (in press).