Designing the Synthetic T^2 Quality Control Chart as a Multi-Objective Optimization Problem

FRANCISCO APARISI[1], MARCO DE LUNA[2]
[1]Departamento de Estadística e I. O. Aplicadas y Calidad
Universidad Politécnica de Valencia
Camino de Vera s/n
SPAIN

[2]Departamento de Ingeniería Industrial y Mecánica
ITESM
45140 Guadalajara
MEXICO

Abstract: - In some real applications of Statistical Process Control it is necessary to design a control chart to not detect small process shifts, but keeping a good performance to detect moderate and large shifts in the quality. In this work we develop a new quality control chart, the synthetic T^2 control chart, which can be designed to cope with this objective. A multi-objective optimization is carried out employing Genetic Algorithms, finding the Pareto-optimal front of non-dominated solutions for this optimization problem.

Key-Words: - Multi-objective optimization, Pareto-optimal front, Synthetic T^2 control chart, Genetic Algorithms.

1 Introduction

Nowadays it begins to be common to face problems or applications where the mathematical modeling produces an optimization problem with several objectives. The multi-objective optimization consists of optimizing simultaneously several objective functions. In many cases, some of the objective functions represent 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, xj, m restrictions and p objectives:

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

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 $F=\{ x \in R^n\text{: } g_i(x) \le 0,\ x_j \le 0, \forall i, j \}$ when g_i functions are the restrictions. In some cases, variables z_k are called objective functions or objectives.

In some real applications of Statistical Process Control it is necessary to design a control chart to not detect small shifts in the process mean, but keeping a good performance to detect moderate and large shifts. This design was first posed by Woodall (1985) and it is known as the design for In-control and Out-of-control regions. Although the idea is not new, it is quite difficult to design a quality control chart that can solve this problem. The typical control charts can not be adapted, and the optimization problem it is rather difficult.

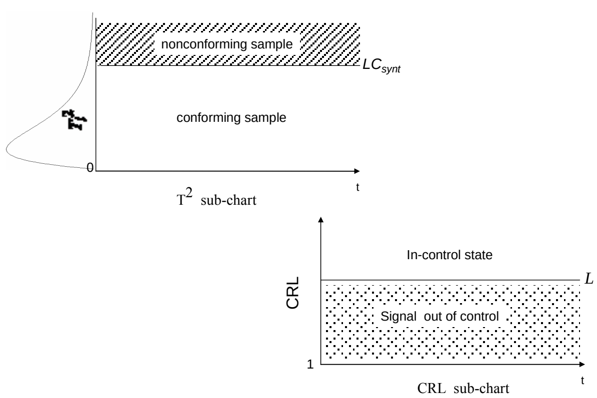

Fig. 1. The synthetic T^2 control chart.

In this work we develop a new chart, the synthetic T^2 control chart, an improvement of the standard Hotelling's T^2 chart. This new chart can be designed to solve this multi-objective problem. Therefore, the objective of this work is to apply Evolutionary Multi-Criterion Optimization to the design of the synthetic T^2 control chart to solve the problem of In-control and Out-of-control regions and to find the Pareto-optimal front.

2 The synthetic T^2 control chart

2.1 Defining the Synthetic T^2 Control Chart.

The main objective of quality control charts is to detect shifts in the production process that are due to assignable causes. Samples are taken from the process and the calculated statistic is plotted in a chart. If the point is plotted outside the control limit(s) we have to assume that there is an assignable cause of variation; we assume that there is a shift in the variable(s) we are monitoring, see [1]. A measure of performance of a control chart is the ARL (Average Run Length). ARL is the average number of points (or samples) that we have to plot until the chart signals. In the classic design of control charts, when there is really a shift in the process the ARL has to be as minimum as possible. However, when there is no a shift, the ARL must be as maximum as possible. In order to detect shifts in the process mean, the Shewhart control chart is the most widely used control chart. However, its performance to detect small shifts is not good (large ARL values).

The univariate synthetic chart (only one variable is monitored) was introduced in [2] as an alternative to improve the performance of the Shewhart control chart to detect small process shifts. It is the result of combining a Shewhart chart and a CRL chart (a chart originally designed to detect increments in the percentage of defective units). The synthetic-$\bar{X}$ chart shows better ARL values to detect process shifts, for any shift magnitude, than the $\bar{X}$ control chart. In some cases, especially for moderate and large shifts, the synthetic- chart has better performance than the EWMA control chart [2].

The synthetic chart has been also applied to monitor the variability of a process [3, 4] and the percentage of defective units of a process [5]. With the objective of improving the performance of the synthetic control charts, [4] apply variable sampling intervals (VSI). However, a multivariate synthetic control has not been developed and it is one the objectives of this work.

The synthetic T^2 control chart (synthetic-T^2) is definied as a chart for monitoring simultaneously **two or more** quality characteristics. It is the first time that this technique is applied to control several variables simultaneously. It consists of two sub-charts, a T^2 sub-chart and a CRL sub-chart. Fig. 1 shows the concept of the synthetic-T^2 chart. The T^2 sub-chart has a unique control limit, *LCsynt*. The CRL sub-chart has a low control limit, L, $L \geq 1$. The value of *LCsynt* is the criteria to classify a sample as *conforming* or *non-conforming*. The value of L is the criteria to decide if the process is in control or out of control.

The CRL chart was first proposed in [6]. In the synthetic-T^2 chart the value of CRL is defined as the number of inspected samples between two samples classified as non-conforming, including the last non-conforming sample. In Fig. 2 the white circles represent conforming samples and the black circles show non-conforming samples. In this Figure four values of CRL are shown, CRL1 = 5, CRL2 = 4, CRL3 = 7, CRL4 = 6, assuming that the sampling starts at t = 0.

The CRL concept assumes that in t = 0 there is a point above the *LCsynt* limit (a non-conforming sample in t = 0). This characteristic, called head start, is very important for the performance of the synthetic charts. When this assumption is ruled out, the performance of synthetic charts worsens [7]. The routine of the synthetic-T^2 chart follows:

1. A sample of size n is taken from the process at time i and the sample mean vector is computed, $\bar{\mathbf{X}}_i^T = (\bar{X}_1, \bar{X}_2, \bar{X}_3, ..., \bar{X}_p)$. The T_i^2 statistic is calculated.
2. The value of the T_i^2 statistic is plotted in the T^2 sub-chart. If $T_i^2 \leq LCsynt$ the sample is classified as *conforming* and we move back to point 1. Otherwise, if $T_i^2 > LCsynt$, the sample is classified as *non-conforming* and we continue to the next point.
3. It is counted the number of samples between this non-conforming sample and the last one. This number is called *CRL sample* and it is plotted in the CRL sub-chart.
4. If $CRL > L$ the conclusion is that the process is in an in-control state, and the control routine begins again in point 1. If $CRL \leq L$ the process is deemed as out of control.
5. The out of control signal is investigated. If no assignable cause is founded the signal is considered as

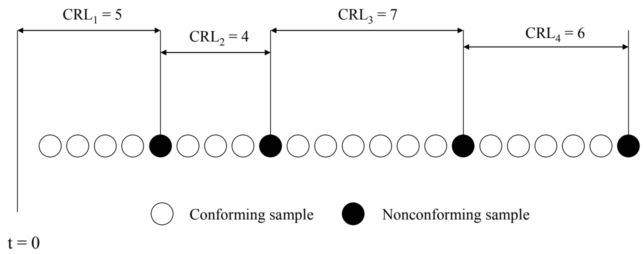

Fig 2. The CRL concept.

a false alarm and we continue to point 1. Otherwise, the assignable cause must be eliminated.

For a synthetic-T² chart, unlike the T² control chart, $T_i^2 > LCsynt$ does not mean that an out-of-control state has to be assumed, but the inspected sample must be classified as non-conforming.

The synthetic-T² chart shows an out-of-control signal, indicating the probably there is a shift in the process. However, this signal does not inform us about the variable or variables that have produced the shift. Several methods for the interpretation of the out-of-control signal of the T² control chart have been developed. These methods can also be applied to the synthetic-T² chart [8-12]. Applying these techniques the global performance of the synthetic-T² can be improved in comparison with other multivariate charts.

2.2 Obtaining the ARL values.

Two values of ARL are important for the design and performance of the synthetic-T² chart, the in-control ARL [$ARL_{S-T^2}(d=0)$] and the out-of-control ARL [$ARL_{S-T^2}(d \neq 0)$]. The value of the in-control ARL is selected taking into account the frequency of false alarms. The value of $ARL_{S-T^2}(d \neq 0)$ is important in order to rapidly detect a shift in the mean vector of magnitude *d*.

The value of ARL for a given shift of magnitude *d*, for whatever synthetic chart, is [4]:

$$ARL_s(\delta) = E\left[ARL_{CRL}\right] * E\left[CRL\right] = \frac{1}{1-(1-q)^L} * \frac{1}{q} \quad (1)$$

where *q* is the probability of a sample being non-conforming. The objective is to adapt the last formula for the synthetic-T² chart, obtaining:

$$ARL_{S-T^2}(d=0) =$$
$$= ARL_{CRL}(d \mid LCsynt, L) * ARL_{T^2}(d \mid LCsynt) * = \frac{1}{1-(1-q_0)^L} * \frac{1}{q_0} \quad (2)$$

where $q_0 = \mathrm{P}(T_i^2 > LCsynt)$.

Considering (1), when $d \neq 0$, the value of $ARL_{S-T}(d \neq 0)$ is obtained as:

$$ARL_{S-T^2}(d \neq 0) =$$
$$= ARL_{CRL}(d \mid LCsynt, L) * ARL_{T^2}(d \mid LCsynt) = \frac{1}{1-(1-q_d)^L} * \frac{1}{q_d} \quad (3)$$

where $q_d = 1 - \beta = 1 - \mathrm{P}(T_i^2 \leq LCsynt)$.

Therefore, the design of the synthetic T² control chart consist of fixing the following values: *n*, *L*, and $LCsynt$.

3 In-control and Out-of-Control regions

Woodall [13] 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 signal, this is regarded as a false alarm.

b) Out-of-control region [*B*, ∞[, 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.

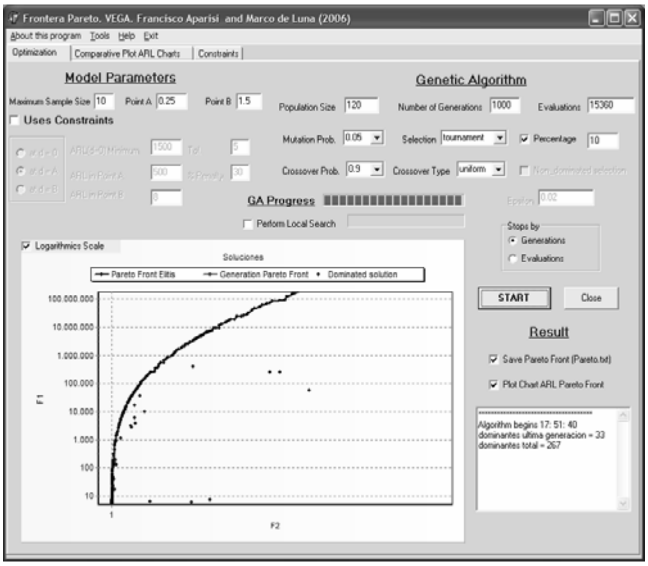

Fig. 3. Software solving the example of application.

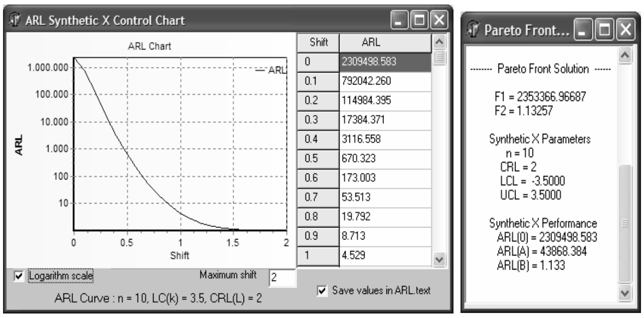

Fig. 4. Solution taken from the Pareto Front.

Therefore, this control chart design can be specified as a two objective problem. The first objective is to maximize the ARL in the in-control region. The second objective is to minimize the ARL is the out-of-control region. In order to formulate the optimization problem, the following objective functions have been selected:

$$Z_1 = ARL(d=0) + ARL(d=A) =$$
$$= ARL_{CRL}(d=0|LCsynt,L)\bullet ARL_{T^2}(d=0|LCsynt) +$$
$$+ARL_{CRL}(d=A|LCsynt,L)\bullet$$
$$\bullet ARL_{T^2}(d=A|LCsynt) = \frac{1}{1-(1-q_0)^L}\bullet\frac{1}{q_0} + \frac{1}{1-(1-q_{d=A})^L}\bullet\frac{1}{q_{d=A}} =$$
$$= \frac{1}{1-(\frac{1}{2^{p/2}\Gamma(p/2)}\int_0^{LCsynt} y^{\frac{p}{2}-1}e^{-y/2}dy)^L}\bullet\frac{1}{1-\frac{1}{2^{p/2}\Gamma(p/2)}\int_0^{x} y^{\frac{p}{2}-1}e^{-y/2}dy} +$$
$$\frac{1}{1-(1-q_{d=A})^L}\bullet\frac{1}{q_{d=A}}$$

$$Z_2 = ARL(d=B) = ARL_{CRL}(d=B|LCsynt,L)\bullet$$
$$\bullet ARL_{T^2}(d=B|LCsynt) =$$
$$= \frac{1}{(e^{\frac{-\lambda}{2}}\sum_{j=0}^{\infty}\frac{(\frac{1}{2}n\cdot d_B)^j}{j!}\cdot\frac{1}{2^{\frac{p}{2}+j}\Gamma(\frac{p}{2}+j)}\int_0^{LCsynt} y^{\frac{p}{2}+j-1}e^{\frac{-y}{2}}dy)^L}\bullet$$
$$\bullet\frac{1}{1-e^{\frac{-\lambda}{2}}\sum_{j=0}^{\infty}\frac{(\frac{1}{2}n\cdot d_B)^j}{j!}\cdot\frac{1}{2^{\frac{p}{2}+j}\Gamma(\frac{p}{2}+j)}\int_0^{LCsynt} y^{\frac{p}{2}+j-1}e^{\frac{-y}{2}}dy}$$

Hence, the goodness of each solution is checked using two values of the in-control region, d = 0 and d = A, and one value of the out-of-control region, d = B.

4 Pareto-optimal Front

The Pareto-optimal front is the set of non-dominated solutions for a given multi-objective optimization problem. Specifically, for the bi-dimensional problem stated here, a solution belongs to the Pareto-optimal front is there is no another solutions that dominates this solution. In our case, there is no solution that produces a higher value of Z_1 and, at the same time, a lower value of Z_2.

Therefore, the final user has to decide which is the synthetic T^2 control chart to employ, for a given industrial process to control, choosing one of the solutions of the Pareto-optimal front.

A real industrial problem is as follows. The user specifies the maximum sample size that can be employed to monitor the process. In addition, the size of the in-control and out-of-control regions must be specified, fixing the values A and B. The objective of the optimization consists of finding all the solutions (values of n, L, and $LCsynt$) that form the Pareto-optimal front.

5 Software Developed and Example of Application

We have developed software that easily finds the Pareto-optimal front, employing GAs [14], helping the user to choose one of the non-dominated solutions. As shown in Fig. 3, the user has to specify the maximum sample size, and the values of A and B that defines the in-control and out-of-control regions. The software returns the Pareto-optimal front, in less than 5 minutes.

If the user chooses one of the solutions of the front, clicking on its point in the front, two new windows appear, see Fig. 4. The first window shows the ARL curve for this solution. The second window shows the solution, and the associated values of ARL(d = 0), ARL(d = A), and ARL(d = B) and the values of the functions Z_1 and Z_2.

The algorithm used to find the Pareto-optimal front is the NSGA of Srinivas & Deb [14]. The front is well defined and the points in the front are well distributed. The Pareto front is obtained employing a population size of 120 and after 1000 generations, with mutation probability = 0.05, crossover probability = 0.9 and crossover type = uniform.

Let us solve now an example of application. A productive process needs to be monitored using a synthetic T^2 control chart. The maximum sample size is n = 10. The shifts less than A = 0.25 must not be detected, however, maximum performance is desired for shifts larger than B = 1.5. Fig. 3 shows the software solving this multi-objective optimization problem. After the Pareto-optimal front is found the user has to choose one of the solutions of this front. Fig. 4 shows one of these solutions, after clicking on it. In this case employing n = 10, CRL = 2 and $LCsynt$ = 3.5 the following performance is achieved: ARL(d = 0) = 2309498, ARL(d = A) = 43868 , and ARL(d = B) = 1.13.

Considering this results, it is possible to see that it is practically impossible to detect shifts lower that d = 0.25, one of the objectives of the optimization. On the other hand, the synthetic control chart will detect shifts $d \geq 1.5$ very quickly, the second objective of our optimization problem.

6 Conclusion

In this paper we have designed a new chart, the synthetic T^2 control chart. This chart can be designed to solve the optimization problem of in-control and out-of-control regions. The optimization can be solved successfully using GAs to find the Pareto-optimal front. Friendly software has been developed to help the user to select the most convenient solution from the Pareto front.

References:

[1] Montgomery, D. C. *Introduction to statistical quality control*, fifth edition, 2005, (New York: John Wiley & Sons.).

[2] Wu ,Z. & Speeding, T. A. “A synthetic control chart for detecting small shifts in the process mean”, *Journal of Quality Technology*, 32(1), 2000, pp. 32-38.

[3] Chen, F. L. & Huang, H. J. “A synthetic control chart for monitoring process dispersion with sample standard deviation”, Computers & Industrial Engineering, 49(2), 2005, pp. 221-240.

[4] Chen, F. L. & Huang, H. J. (2005b) A synthetic control chart for monitoring process dispersion with sample range, The International Journal of Advanced Manufacturing Technology, 26 (7-8), pp. 842-851.

[5] Wu, Z., Yeo, S. H. & Spedding, T. A. “A synthetic control chart for detecting fraction nonconforming increases”, Journal of Quality Technology, 33(1), 2001, pp. 104-111.

[6] Bourke, P. “Detecting a shift in fraction non-conforming using run length control chart with 100% inspection”, Journal of Quality Technology, 32, 1991, pp. 225-238.

[7] Davis, R. B. & Woodall, W. H. “Evaluating and improving the synthetic control chart”, Journal of Quality Technology, 34(2), 2002, pp. 200-208.

[8] Mason, R. L., Tracy, N. D. & Young, J. C. “A practical approach for interpreting multivariate T2 control chart signals”, Journal of Quality Technology, 29(4), 1997, pp. 396-406.

[9] Atienza, O. O., Tang, L. C. & Ang, B. W. “Simultaneous monitoring of univariate and multivariate SPC information using boxplots”, International Journal of Quality Science, 3(2), 1998, pp. 194-204.

[10] Mason, R. L. & Young, J. C. “Improving the sensitivity of the T2 statistic in multivariate process control”, Journal of Quality Technology, 31(2), 1999, pp. 155-165.

[11] Kourti, T. & Macgregor, J. F. “Multivariate statistical process control and industrial applications”, Quality Congress. ASQ World Conference on Quality and Improvement Proceedings, 58, 2004,pp. 449-452.

[12] Aparisi, F., Avendaño, G. & Sanz, J. “Interpreting T2 control signal: Effectiveness of MTY decomposition vs. a neural network”, IIE Transactions, 38(8), 2006, pp. 647-657.

[13] Woodall, W. H. “The statistical design of quality control charts”, The Statistician, 34, 1985, pp. 155-160.

[14] Deb, K. Multi-Objective Optimization using Evolutionary Algorithms. third edition, 2003, (New York: John Wiley & Sons.).