



**HAL**  
open science

# ON THE RUN LENGTH OF A STATE-SPACE CONTROL CHART FOR MULTIVARIATE AUTOCORRELATED DATA

Manuel Vargas Vargas, Jose-Luis Alfaro Navarro, Jose Mondejar Jimenez

► **To cite this version:**

Manuel Vargas Vargas, Jose-Luis Alfaro Navarro, Jose Mondejar Jimenez. ON THE RUN LENGTH OF A STATE-SPACE CONTROL CHART FOR MULTIVARIATE AUTOCORRELATED DATA. *Communications in Statistics - Simulation and Computation*, 2009, 38 (09), pp.1823-1833. 10.1080/03610910903147771 . hal-00520669

**HAL Id: hal-00520669**

**<https://hal.science/hal-00520669>**

Submitted on 24 Sep 2010

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**ON THE RUN LENGTH OF A STATE-SPACE CONTROL CHART FOR MULTIVARIATE AUTOCORRELATED DATA**

Journal:	<i>Communications in Statistics - Simulation and Computation</i>
Manuscript ID:	LSSP-2008-0309.R2
Manuscript Type:	Review Article
Date Submitted by the Author:	13-May-2009
Complete List of Authors:	VARGAS VARGAS, MANUEL; UNIVERSIDAD DE CASTILLA-LA MANCHA, Facultad de Ciencias Económicas y Empresariales ALFARO NAVARRO, JOSE-LUIS; UNIVERSIDAD DE CASTILLA-LA MANCHA, Facultad de Ciencias Económicas y Empresariales MONDEJAR JIMENEZ, JOSE; UNIVERSIDAD DE CASTILLA-LA MANCHA, Facultad de Ciencias Sociales
Keywords:	Statistical Process Control, Multivariate Control Chart, State-Space Model, Average Run Length
Abstract:	The literature on statistical process control (SPC) describes the negative effects of autocorrelation in terms of false alarms. This has been treated by the individual modelling of each series or the application of VAR models. In the former case, the analysis of the cross correlation structure between the variables is altered. In the latter, the filtering process can modify the weakest relations. In order to improve these aspects, state-space models have been introduced in MSPC. This paper presents a proposal for building a innovations control chart, estimating its average run length to highlight its advantages over the VAR approach.



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

For Peer Review Only

1  
2  
3  
4  
5  
6  
7  
8  
9

## ON THE RUN LENGTH OF A STATE-SPACE CONTROL CHART FOR MULTIVARIATE AUTOCORRELATED DATA

10 Manuel Vargas<sup>1</sup>, José Luis Alfaro<sup>\*1</sup> and José Mondéjar<sup>2</sup>

11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22

<sup>1</sup> Facultad de Ciencias Económicas y Empresariales, Universidad de Castilla-La Mancha, Pza. de la Universidad nº 1, 02071, Albacete, España.

23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51

<sup>2</sup> Facultad de Ciencias Sociales, Universidad de Castilla-La Mancha, Avda. de los Alfares nº 44, 16002, Cuenca, España.

### ABSTRACT

26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51

The literature on statistical process control (SPC) describes the negative effects of autocorrelation in terms of the increase in false alarms. This has been treated by the individual modelling of each series or the application of VAR models. In the former case, the analysis of the cross correlation structure between the variables is altered. In the latter, if the cross correlation is not strong, the filtering process may modify the weakest relations. In order to improve these aspects, state-space models have been introduced in MSPC. This paper presents a proposal for building a control chart for innovations, estimating its average run length to highlight its advantages over the VAR approach mentioned above.

52  
53  
54  
55  
56  
57  
58  
59  
60

**KEY WORDS:** Statistical Process Control, Multivariate Control Chart, State-Space Model, Average Run Length.

---

\* Corresponding author.

## 1. INTRODUCTION

The main **aims** of statistical process control methods is to minimise the frequency of false alarms and quickly detect **special** sources of variation. **Shewhart control charts are designed for the first purpose**, but under the condition that the observations analysed are independent. However, in practice this is not a general condition, as there are a wide range of **autocorrelated** processes. This phenomenon **limits its use** and has led to the development of statistical models adapted to this phenomenon.

Traditionally, the treatment of **autocorrelation** focused on the use of empirical control limits or on the filtering of information using univariate time series models (Harris and Ross, 1991; Montgomery and Mastrangelo, 1991; Maragah and Woodall, 1992; Box and Luceño, 1997; Lu and Reynolds, 1999; Jiang, 2001; West and Jarrett, 2004). The earliest multivariate developments for treating **autocorrelation** consisted of the individual modelling of each series or the application of VAR modelling (Dyer *et al.*, 2003; Jiang, 2004; Kalgonda and Kulkarni, 2004; Noorossana and Vaghefi, 2006; Jarrett and Pan, 2007a, b; Pan and Jarrett, 2007. In the former case, the analysis of the structure of cross correlation between the variables, the great contribution of the multivariate approach, is altered. In the latter case, if the original cross correlation is not strong, the filtering process may modify the weakest relations. Moreover, the **autocorrelation** may reflect the existence of some “latent factor” or “inertia” in the process. In such a case, analysis of the residuals will reveal the “filtered” behaviour of the process, i.e., the behaviour of the non-foreseeable or innovational part, without going into an analysis of the “inertia factor”.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

On a simultaneous basis, the works of Negiz and Cinar (1997), Dorsey and Lee (2003), Lee and Dorsey (2004), Pan and Jarrett (2004), Alfaro (2005), Triantafyllopoulos (2006), Xie *et al.* (2006) and Zantek *et al.* (2007) introduced multiple time series state-space modelling in statistical process control. This modelling allows to eliminate the autocorrelation existing in the data, preserving a great amount of information in the residuals, that in this methodology are denoted as innovations, and also allowing to describe the existence of inertia in the process by building a control chart for the states variables, estimate with the state-space model.

This paper will propose state-space treatment for autocorrelated multivariate processes. The second section describes the new control chart for innovations that is more efficient than others existing in the literature. The third section includes an estimation of the average run length (ARL) for this chart, obtained by simulation for various scenarios established according to different levels of correlation and autocorrelation. Finally, the fourth section contains the general conclusions of this paper.

## 2. THE MULTIVARIATE STATE-SPACE CONTROL CHART

The interest in state-space representation began with Kalman's works (1960, 1963), although various approaches were subsequently developed for identifying and estimating these models. As well as using the maximum likelihood method, Akaike (1975, 1976) proposed an approach based on canonical correlation, which may lead to non-equivalent representations when the order of the endogenous variables was changed. Later on, Aoki (1987) developed an approach based on the decomposition of the estimated autocovariance matrix in singular values. Algorithms for identifying

1  
2  
3 subspaces have also been developed in several papers. For a complete reviewed of these  
4  
5 see Vargas (1999), García (2005) and the cited references.  
6  
7

8  
9 In statistical process control, the basic motivation for using state-space modelling  
10  
11 lies in the double reasons for sample autocorrelation. On the one hand, when the sample  
12  
13 procedure causes autocorrelation in the data, it is appropriated to apply a filter to  
14  
15 eliminate this, preserving the greatest possible amount of information and building a  
16  
17 control chart on the residuals. However, if the autocorrelation reflects the existence of a  
18  
19 “latent factor” or “inertia” inherent to the process, which provides basic information on  
20  
21 it, the residuals of a filtering operation are not statistically optimal, because of a  
22  
23 fundamental part of the process is omitted. This case requires an efficient estimation of  
24  
25 the “inertia factor”, as well as the building of a chart to control it. In this paper, we  
26  
27 focus on the control chart for innovations, comparing its ARL to that of charts obtained  
28  
29 on the basis of residuals obtained from VAR models.  
30  
31  
32  
33  
34

35  
36 Given the estimation of a state-space model, the resulting innovations ( $\varepsilon_t$ ) are  
37  
38 comparable to the residuals from VAR modelling. They describe the innovational or  
39  
40 non-foreseeable behaviour of the process, controlling the existence of special causes  
41  
42 whose effects are of great magnitude and short duration in time. The state-space  
43  
44 methodology allows to preserve the multivariate approach along with a more  
45  
46 parsimonious representation, an ideal situation in quality control.  
47  
48  
49

50  
51 In this framework, it is crucial to estimate “correctly” the state variables which are  
52  
53 sufficient statistics for the dynamics of the process and they determine the statistical  
54  
55 properties of the residuals. To estimate state-space models in statistical process control,  
56  
57 a number of approaches exist (Negiz and Cinar, 1997; Dorsey and Lee, 2003; Lee and  
58  
59  
60

Dorsey, 2004; Pan and Jarrett, 2004; Triantafyllopoulos, 2006; Xie *et al.*, 2006; Jarrett and Pan, 2007a; Zantek *et al.*, 2007), with different algorithms that provide different statistical behaviour.

Therefore, given a  $p$ -dimensional **autocorrelated** process,  $Y_t$ , its state-space representation:

$$\begin{aligned} X_{t+1} &= F_t X_t + G_t \varepsilon_t \\ Y_t &= H_t X_t + \varepsilon_t \end{aligned} \quad (1)$$

allows to estimate the innovations in the process, as explained in the appendix. Given the properties of the algorithm used, the innovations ( $\varepsilon_t$ ) are asymptotically normal and uncorrelated, with zero mean and variance  $\Psi$ . In these conditions, we can build a  $T^2$  Hotelling control chart using the estimation of the matrix  $\Psi$  provided by the state-space filtering process, by means of the expression:

$$T_{\varepsilon_i}^2 = \varepsilon_i' \hat{\Psi}^{-1} \varepsilon_i \quad (2)$$

When the innovations are uncorrelated, the  $T^2$  statistics follows the Beta distribution in the Phase I (and F distribution in the Phase II) and we can get the control limit as:

$$LSC(T_{\varepsilon_i}^2) = \frac{(m-1)^2}{m} \beta_{\alpha, p/2, (m-p-1)/2} \quad (3)$$

But, when we use the estimation  $\hat{\psi}$ , the innovations are uncorrelated only asymptotically. In this case, the corresponding control limit is:

$$LSC(T_{\varepsilon_i}^2) = \chi^2(p) \quad (4)$$

which is similar to the one obtained for the original series, as the number of variables is equal to the number of innovations and no observations are lost as a consequence of the filtering.

1  
2  
3 This control chart allows us to detect the special causes for variation that present  
4 an effect of relative magnitude and short duration in time. This chart is similar to the  
5 residuals control chart developed in Jarrett and Pan (2007b) and Pan and Jarrett (2007),  
6 although there are some aspects that are different due to the statistical methodology  
7 used in estimation. The algorithm for estimating the matrices in the model guarantees  
8 the minimality of the state vector, by balancing the grammians of the observability and  
9 controllability matrices, lending these characteristics to the estimated system  
10 (guaranteeing minimality). Also, the use of the Kalman filter ensures that the states  
11 constitute the optimal linear estimation (in the minimum quadratic sense of the  
12 expression) of the process dynamic. This all comes together in obtaining innovations  
13 that contain all of the complementary information, thereby allowing for thorough  
14 analysis.

15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33 As show Xie *et al.* (2006), the use of ARMA filters ensures the efficient  
34 elimination of autocorrelations, but the cross correlations still exist and the residuals are  
35 not independent. In contrast to this filter, the state-space model is enough to remove the  
36 auto and cross correlation. With the algorithm proposed, the statistical properties of the  
37 state variables ensures the “minimum” correlation in the residuals.

38  
39  
40  
41  
42  
43  
44  
45 To compare the statistical properties of the control charts, it is usual to examine  
46 their performance in terms of the ARL (Wardell *et al.*, 1994; Javaheri and Houshmand,  
47 2001; Jiang, 2001; Loredó *et al.*, 2002; Shu *et al.*, 2004; Snoussi *et al.*, 2005; Tsung *et*  
48 *al.*, 2006; Snoussi and Limam, 2007). Based on the ARL as criteria of control charts  
49 performance, we compare the proposed alternative control chart with the obtained by  
50 means of VAR modelling. We analyse the possible effect of cross and autocorrelation  
51 through the simulation under different levels of both.

### 3. SIMULATION STUDY

The simulation process has been developed using R-Program. More specifically, the algorithm used for simulated a multivariate normal process are developed in the function *rmvnorm* in the *mvtnorm* package written by Alan Genz in R software. For adjusted and estimated a VAR model by least square, we have used the function *estVARXls* in the *dse* package written in R by Paul Gilbert. All these packages are available from the R project website (<http://rprojects.org/>). The state-space model estimation, developed in the appendix, is also implemented in R with a function developed by the authors (available upon request)

To simplify, without losing generality, we opted to simulate three-dimension processes with the same level of correlation between the variables. We generated, under a normal multivariate distribution of parameters  $\mu_0$  and  $\Sigma_0$ , 150 in-control observations. Then, we generated five out-of-control observations ( $\mu = \mu_1 \neq \mu_0$ ), with changes in the mean vector of 0, 1, 2 and 3 standard deviations, changes usually considered in statistical process control. Finally, we generated another sequence of 5000 in-control observations to determine the run length.

To analyse the detection capability of the innovations control chart we have considered different scenarios by varying the levels of cross correlation and autocorrelation. In the first case, we focused on positive correlation, given it's the most usual in the production processes. Therefore, correlations of 0, 0.3, 0.5, 0.7 and 0.9 were considered, in order to detect if the level of correlation affects the detection capacity of control charts. These information are introduced in the Hotelling's  $T^2$  statistic via the covariance matrix.

1  
2  
3 Also, we analyse four possibilities of the level of autocorrelation, corresponding  
4 to coefficients of 0.3, 0.5, 0.7 and 0.9, given its greater impact on the properties of  
5 control charts in terms of false alarms. For each combination, the simulation was  
6 repeated 1000 times, which allows to obtain the ARL for the control charts exposed.  
7  
8  
9

10  
11  
12  
13  
14 Once the series had been simulated, we have adjusted and estimated two models:  
15 a VAR model by least square by means of the *estVARXls* function in the *dse* package in  
16 *R*, following the methodology used in Jarrett and Pan (2007b) and Pan and Jarrett  
17 (2007); and a state-space model using the proposed methodology, implemented in *R*.  
18 The control limits were estimated to attain a value of 280 for the in-control average run  
19 length ( $ARL_0$ ) and the run length was calculated for each series. The averages in each  
20 case are shown in tables 1 to 4, which depict the improvement, in terms of ARL, in the  
21 control chart proposed compared to the traditional one based on the VAR methodology.  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33

34 INSERT TABLE 1

35  
36  
37 The table 1 shows that if the process is in-control, both the VAR and the state-  
38 space approaches present similar behaviour, with non-significant differences in  $ARL_0$ .  
39 In this case, the greatest differences appear with high levels of correlation and  
40 autocorrelation and that the behaviour of the innovations control chart proposed is  
41 better.  
42  
43  
44  
45  
46  
47  
48  
49

50 INSERT TABLE 2

51  
52  
53 The table 2 shows that when the change in the means are smaller, the differences  
54 between the two methods are not statistically significant. The ARL takes values around  
55 140 in the both methods.  
56  
57  
58  
59  
60

1  
2  
3 The tables 3 and 4 show that when changes occur in the mean of the process, the  
4 state-space chart improves its behaviour, with an increase in relative efficiency as the  
5 magnitude of the change increases. Obviously, if we express these changes by means of  
6 the non-centrality parameter, affected by the existence of cross correlation between the  
7 variables, then we can not directly compare the values for the ARL. This fact manifests  
8 itself in tables 2, 3 and 4, where the run length of the models increases as the cross  
9 correlation increases, with the consequent reduction in the non-centrality parameter.

10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21 INSERT TABLES 3 and 4

22  
23  
24 Simulation results show that the state-space model formulation provides better  
25 ARL behaviour than residual charts based on VAR methodology (see figure 1). In  
26 general, the  $ARL_0$  is similar for the two control charts, as they originate from the  
27 innovations in two equivalent methodologies for modelling statistical processes. When a  
28 change occurs in the mean of the process, the behaviour of the control charts starts to  
29 diverge, whereas the ARL of the chart proposed is lower than the one that is built under  
30 the VAR approach. This difference is further accentuated as the non-centrality  
31 parameter increases, which is dependent on the distance between the real mean and the  
32 in-control mean.

33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46 Although the existence of autocorrelation alters the ARL of both charts, the  
47 results obtained with the methodology proposed are more homogeneous and stable,  
48 reflecting greater resistance to this problem. On the other hand, the existence of cross  
49 correlation presents no other effect than the reduction in the non-centrality parameter,  
50 with both charts equally affected.

51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971  
972  
973  
974  
975  
976  
977  
978  
979  
980  
981  
982  
983  
984  
985  
986  
987  
988  
989  
990  
991  
992  
993  
994  
995  
996  
997  
998  
999  
1000

INSERT FIGURE 1

#### 4. SUMMARY

The traditional proposals to eliminate the negative effect of autocorrelation in statistical quality control present problems related to the loss of the multivariate approach, the parsimony of adjusted models and the poor behaviour in the presence of positive autocorrelation. State-space modelling constitutes an alternative that allows to maintain the multivariate perspective, requiring the estimation of less parameters and with better behaviour in terms of ARL.

We developed a simulation to analyse the behaviour of the control chart proposed in the event of changes in the mean of the process, enabling comparison with the VAR control chart and the study of the effect of the non-independence between the observations.

The results of this indicate that the behaviour of the state-space innovations control chart is more adequate than the residuals control chart obtained by means of VAR modelling. To be specific, the innovations control chart presents better behaviour in terms of out-of-control ARL and greater resistance to the existence of positive autocorrelation. These advantages are accentuated the greater the importance of the change in the mean of the process, situations in which the methodology proposed in this paper is a significant improvement on the traditional VAR modelling.

This paper is a first approach to the behaviour of control chart based on innovations of a state-space model. As such, it has not approached the effect of changes on the correlation structure of the variables or on their variability. Similarly, future researches should analyse the sensitivity of the results of the simulation to changes in

1  
2  
3 the magnitude of the correlation between variables or combinations of different levels of  
4  
5 autocorrelation.  
6  
7

## 8 9 REFERENCES

10  
11 Akaike, H. (1975). Markovian Representation of Stochastic Processes by Canonical  
12  
13 Variables. *SIAM Journal on Control and Optimization* 13(1):162-173.  
14

15  
16 Akaike, H. (1976). *Canonical Correlation Analysis of Time Series and the Use of an*  
17  
18 *Information Criterion*. New York: Academic Press.  
19

20  
21 Alfaro, J. L. (2005). *Control estadístico de la calidad en procesos multivariantes*  
22  
23 *autocorrelacionados. Una aplicación en la industria cuchillera de Albacete.*  
24  
25 Universidad de Castilla-La Mancha.  
26  
27

28  
29 Aoki, M. (1987). *State-space modelling of time series*. New York: Springer-Verlag.  
30

31  
32 Bartlett, M. S. (1939). A note on tests of significance in multivariate analysis.  
33  
34 *Proceeding of the Cambridge Philosophical Society* 35: 180 - 185.  
35  
36

37  
38 Bauer, D. and Wagner, M. (2002). Estimating cointegrated systems using subspace  
39  
40 algorithms. *Journal of Econometrics* 111: 47-84.  
41  
42

43  
44 Box, G. E. P. and Luceño, A. (1997). *Statistical Control by Monitoring and Feedback*  
45  
46 *Adjustment*. New York: Wiley.  
47

48  
49 Dorsey, A. W. and Lee, J. H.(2003). Building Inferential Prediction Models of Batch  
50  
51 Processes Using Subspace Identification. *Journal of Process Control* 13(5): 397-  
52  
53 406.  
54  
55  
56  
57  
58  
59  
60

- 1  
2  
3 Dyer, J. N., Connerly, M. D. and Adams, B. M. (2003). A simulation study and  
4  
5 evaluation of multivariate forecast based control charts applied to ARMA  
6  
7 processes. *Journal of Statistical Computation and Simulation* 73: 709-724.  
8  
9
- 10  
11 García, A. (2005). *Identificación de Modelos para Series Temporales mediante Métodos*  
12  
13 *de Subespacios*. Tesis Doctoral. Universidad Complutense de Madrid. Available in:  
14  
15 <http://www.ucm.es/BUCM/tesis/cee/ucm-t28292.pdf>  
16  
17
- 18  
19 Harris, T. J. and Ross, W. H. (1991). Statistical Process Control Procedures for  
20  
21 Correlated Observations. *Canadian Journal of Chemical Engineering* 69: 48-57.  
22  
23
- 24  
25 Jarrett, J. and Pan, X. (2007a). Monitoring Variability and Analyzing Multivariate  
26  
27 Autocorrelated Processes. *Journal of Applied Statistics* 34(4): 459-469.  
28
- 29  
30 Jarrett, J. and Pan, X. (2007b). The quality control chart for monitoring multivariate  
31  
32 autocorrelated processes. *Computational Statistics & Data Analysis* 51(8): 3862-  
33  
34 3870.  
35
- 36  
37 Jiang, W. (2001). Average Run Length Computation of ARMA Charts for Stationary  
38  
39 Processes. *Communications in Statistics. Simulation and computation* 30(3): 699-  
40  
41 716.  
42
- 43  
44 Jiang, W. (2004). Multivariate control charts for monitoring autocorrelated processes.  
45  
46 *Journal of Quality Technology* 36: 367-379.  
47
- 48  
49 Kalgonda, A. A. and Kulkarni, S. R. (2004). Multivariate quality control chart for  
50  
51 autocorrelated processes. *Journal of Applied Statistics* 31: 317-327.  
52  
53
- 54  
55 Kalman, R. E. (1960). A New Approach to Linear Filtering and Prediction Problems.  
56  
57 *Journal of Basic Engineering. Transaction of ASME series D* 82: 35-45.  
58  
59  
60

- 1  
2  
3 Kalman, R. E. (1963). New Methods in Weiner Filtering Theory. In: Bogdanoff, J. L.  
4 and Kozin, F. (Eds), *Proceedings of the First Symposium of Engineering*  
5 *Applications of Random Function Theory and Probability*. New York: Wiley.  
6  
7  
8  
9  
10  
11 Lee, J. H. and Dorsey, A. W. (2004). Monitoring of batch processes through state-space  
12 models. *AIChE Journal* 50(6): 1198-1210.  
13  
14  
15  
16 Loredó, E. N., Jaerkpaporn, D. and Borrór, C. M. (2002). Model-based control chart for  
17 autoregressive and correlated data. *Quality and Reliability Engineering*  
18 *International* 18(6): 489-496.  
19  
20  
21  
22  
23  
24 Lu, C. W. and Reynolds, M. R. (1999). Control Charts for Monitoring the Mean and  
25 Variance of Autocorrelated Processes. *Journal of Quality Technology* 31: 259-274.  
26  
27  
28  
29 Maragah, H. O. and Woodall, W. H. (1992). The effect of Autocorrelation on the  
30 Retrospective X-Chart. *Journal of Statistical Computation and Simulation* 40(1):  
31 29-42.  
32  
33  
34  
35  
36  
37 Mondéjar, J. and Vargas, M. (2006). Análisis de tendencias comunes y cointegración en  
38 espacio de estados. *Contribuciones a la Economía* september. Available in:  
39 <http://www.eumed.net/ce/>  
40  
41  
42  
43  
44 Montgomery, D. C. and Mastrangelo, C. M. (1991). Some Statistical Process Control  
45 Methods for Autocorrelated Data. *Journal of Quality Technology* 23: 179-204.  
46  
47  
48  
49 Negiz, A. and Cinar, A. (1997). Statistical monitoring of multivariate dynamic  
50 processes with state-space models. *AIChE Journal* 43(8): 2002-2020.  
51  
52  
53  
54  
55 Noorossana, R. and Vaghefi, S. J. M. (2006). Effect of autocorrelation on performance  
56 of the MCUSUM control chart. *Quality and Reliability Engineering International* 2  
57 (2):191-197.  
58  
59  
60

- 1  
2  
3 Pan, X. and Jarrett, J. (2004). Applying State Space into SPC: Monitoring Multivariate  
4  
5 Time Series. *Journal of Applied Statistics* 31(4): 397-418.  
6  
7
- 8  
9 Pan, X. and Jarrett, J. (2007). Using vector autoregressive residuals to monitor  
10  
11 multivariate processes in the presence of serial correlation. *International Journal of*  
12  
13 *Production Economics* 106(1): 204-216.  
14  
15
- 16 Shu, L., Tsung, F., and Tsui, K. L. (2004). Run-Length Performance of Regression  
17  
18 Control Charts with Estimated Parameters. *Journal of Quality Technology* 36(3):  
19  
20 280-292.  
21  
22
- 23  
24 Snoussi, A., Ghourabi, M. E. and Limam, M. (2005). On SPC for short run  
25  
26 autocorrelated data. *Communications in Statistics: Simulation and Computation*  
27  
28 34(1): 219-234.  
29  
30
- 31 Snoussi, A. and Limam, M. (2007). The Change Point Model: SPC Method for Short  
32  
33 Run Autocorrelated Data. *Quality Technology & Quantitative Management* 4(3):  
34  
35 313-329.  
36  
37
- 38  
39 Triantafyllopoulos, K. (2006). Multivariate Control Charts Based on Bayesian State  
40  
41 Space Models. *Quality and Reliability Engineering International* 22: 693-707.  
42  
43
- 44 Tsung, F., Zhao, Y., Xiang, L., and Jiang, W. (2006). Improved Design of Proportional  
45  
46 Integral Derivative Charts. *Journal of Quality Technology* 38(1): 31-44.  
47  
48
- 49 Vargas, M. (1999). *Modelización de series temporales múltiples en espacio de estados.*  
50  
51 *Análisis de procesos no estacionarios y cointegración.* Universidad de Castilla-La  
52  
53 Mancha.  
54  
55
- 56  
57 Wardell, D. G., Moskowitz, H. and Plante, R. D. (1994). Run-length distribution of  
58  
59 special-cause control charts for correlated processes. *Technometrics* 36:1084-1105.  
60

1  
2  
3 West, D. and Jarrett, J. (2004). The Impact of First Order Positive Autoregression on  
4  
5 Process Control. *International Journal of Business and Economics* 3(1): 29-37.  
6  
7

8  
9 Xie, L., Zhang, J. and Wang, S. (2006). Investigation of Dynamic Multivariate  
10  
11 Chemical Process Monitoring. *Chinese Journal of Chemical Engineering* 14(5):  
12  
13 559-568.  
14  
15

16 Zantek, P., Li, S. and Chen, Y. (2007). Detecting multiple special causes from  
17  
18 multivariate data with applications to fault detection in manufacturing. *IIE*  
19  
20 *Transactions* 39(8): 771-782.  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## Tables

Table 1. In-control average run length ( $ARL_0$ )

			Auto-correlation level			
			0.3	0.5	0.7	0.9
Correlation level	0	VAR	267.774	274.172	277.545	286.145
		St-Sp	262.543	278.252	289.146	291.676
	0.3	VAR	271.775	267.401	274.946	279.617
		St-Sp	277.540	262.502	283.131	271.794
	0.5	VAR	278.312	275.072	274.208	265.448
		St-Sp	279.877	283.119	278.749	280.215
	0.7	VAR	264.815	280.609	267.846	270.358
		St-Sp	265.034	285.444	280.347	283.548
	0.9	VAR	280.033	279.645	285.984	267.749
		St-Sp	292.700	283.018	286.223	275.478

Table 2. ARL with changes in the mean vector of one standard deviations

			Auto-correlation level			
			0.3	0.5	0.7	0.9
Correlation level	0	VAR	116.962	124.928	120.474	125.021
		St-Sp	104.421	85.272	75.739	72.536
	0.3	VAR	138.554	128.560	136.514	134.893
		St-Sp	125.863	114.068	110.963	117.763
	0.5	VAR	147.012	135.150	136.581	131.092
		St-Sp	137.251	131.523	132.108	130.637
	0.7	VAR	138.834	137.614	145.510	147.812
		St-Sp	138.660	130.616	131.755	128.382
	0.9	VAR	137.131	143.609	147.686	141.313
		St-Sp	134.567	140.299	132.909	138.819

**Table 3. ARL with changes in the mean vector of two standard deviations**

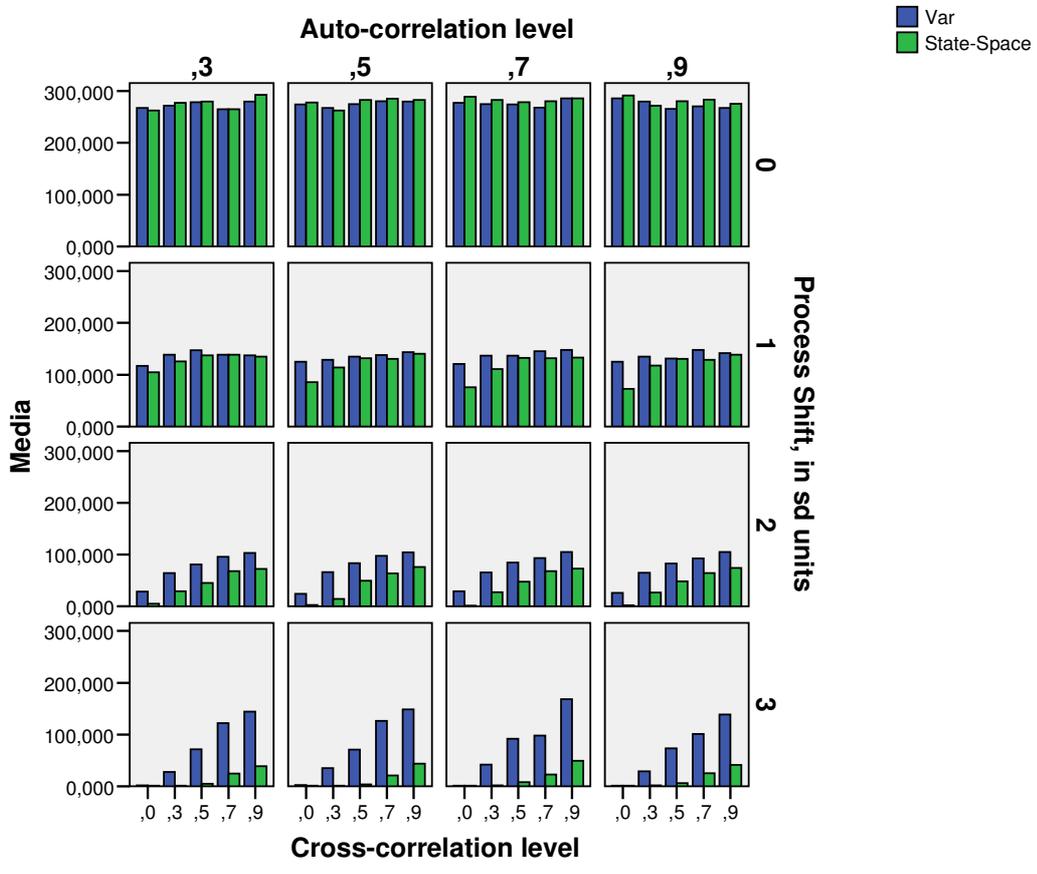
			Auto-correlation level			
			0.3	0.5	0.7	0.9
Correlation level	0	VAR	28.425	24.549	29.251	25.981
		St-Sp	5.509	2.739	1.782	2.116
	0.3	VAR	64.202	66.294	65.677	65.091
		St-Sp	29.183	14.305	27.284	26.932
	0.5	VAR	80.772	83.386	84.757	82.849
		St-Sp	45.516	49.375	47.983	48.514
	0.7	VAR	96.023	97.646	93.438	92.584
		St-Sp	68.132	64.029	68.272	64.416
	0.9	VAR	103.403	104.682	105.035	105.314
		St-Sp	72.352	76.127	72.733	74.181

**Table 4. ARL with changes in the mean vector of three standard deviations**

			Auto-correlation level			
			0.3	0.5	0.7	0.9
Correlation level	0	VAR	1.779	2.622	1.603	1.316
		St-Sp	1.049	1.070	1.116	1.200
	0.3	VAR	27.709	35.329	42.113	29.287
		St-Sp	1.469	1.423	2.115	2.002
	0.5	VAR	71.417	71.214	91.722	73.198
		St-Sp	5.327	3.880	8.015	5.983
	0.7	VAR	122.107	126.480	98.197	101.212
		St-Sp	25.137	21.126	23.130	25.506
	0.9	VAR	144.310	148.512	168.222	139.027
		St-Sp	39.168	43.686	49.456	41.334

FIGURE

Figure 1. ARL for VAR and State-Space residuals control charts



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## APPENDIX I: A Brief on Identification of State-Space Models

A discrete linear system in state-space may be defined by means of an innovational model:

$$\begin{aligned} X_{t+1} &= F_t X_t + G_t \varepsilon_t \\ Y_t &= H_t X_t + \varepsilon_t \end{aligned} \quad (\text{A-1})$$

where  $H_t$ ,  $F_t$ ,  $G_t$  are the real matrices of dimensions  $(p \times n)$ ,  $(n \times n)$  and  $(n \times p)$  respectively,  $p$  is the number of variables,  $n$  the number of observations and we assume, for simplicity, that  $E[Y_t] = 0$ . In this model,  $\varepsilon_t$  is the innovational process of  $Y_t$  that does not present autocorrelation and  $X_t$  is the state vector, a sufficient statistic for the dynamics of the system.

The specification for this model consists of three stages: determining the dimension of the system, estimating the matrices and finally, estimating the state vector. In statistical quality control, determining the dimension of the system is fundamental and therefore different algorithms based on canonical correlation coefficients or information criteria have been used. In this paper, we opted to use the block Hankel matrix, which in practice is usually composed of sample estimations of autocovariances  $\Gamma_\ell = E[Y_{t+\ell} Y_t']$  for  $\ell = 0, 1, 2, \dots$

$$\hat{H}_{N_f}^{N_p} = E[Y_{t,N_f}^+ Y_{t-1,N_p}^-] = \begin{pmatrix} \Gamma_1 & \Gamma_2 & \cdots & \Gamma_{N_p} \\ \Gamma_2 & \Gamma_3 & \cdots & \Gamma_{N_p+1} \\ \vdots & \vdots & \vdots & \cdots \\ \Gamma_{N_f} & \Gamma_{N_f+1} & \cdots & \Gamma_{N_p+N_f-1} \end{pmatrix} \quad (\text{A-2})$$

of dimension  $(pN_f \times pN_p)$ , and prefixed values  $N_p$  and  $N_f$ , and

$$\begin{aligned} Y_{t,N_f}^+ &= (Y_t', Y_{t+1}', \dots, Y_{t+N_f-1}')' \\ Y_{t-1,N_p}^- &= (Y_{t-1}', Y_{t-2}', \dots, Y_{t-N_p}')' \end{aligned} \quad (\text{A-3})$$

By estimating the above matrix by means of the sample auto-correlation matrices

$\hat{\Gamma} = n^{-1} \sum_{i=1}^n Y_{i+1} Y_i'$ , it is possible to decompose the matrix  $\hat{H}_{N_f}^{N_p} = \hat{U} \hat{\Sigma} \hat{V}'$  and use as the

estimation for its range the number of significantly positive singular values in the matrix  $\hat{\Sigma}$ , which will coincide with the dimension of the state vector. In this regard, it is possible to use an approximation of the statistical distribution of the canonical correlation coefficients proposed in Bartlett (1939) by means of the statistic:

$$B_{\hat{n}} = - \left\{ T - \frac{1}{2} [p(N_p + N_f) + 1] \right\} \ln \prod_{j>\hat{n}+1} (1 - \gamma_j^2) \rightarrow \chi_{(pN_f - \hat{n})(pN_p - \hat{n})}^2 \quad (\text{A-4})$$

where  $\gamma_j$  are the canonical correlation coefficients.

On the basis of this statistic, a sequential contrast is established for  $\hat{n}=1,2, \dots$  until the  $H_0$  is not refused, which establishes that there are only  $n$  strictly positive coefficients, in a value  $\hat{n}$  that is considered to be the dimension of the state-space model.

In order to obtain an estimation of the matrices in the system, Bauer and Wagner (2002) propose another approach based on the prior estimation of a regression equation. This algorithm presents the disadvantage that it tends to reduce the variability in series and may alter the determination of points out of control. This situation led us to use a modified version of the algorithm started in Aoki (1987), which may be consulted in its complete form in Vargas (1999) or Mondéjar and Vargas (2006).

As the states are not observable, they may be rotated without changing the algebraic validity of the model. To identify the coefficients by making the observability - controllability factorisation equal to the decomposition of the Hankel matrix in singular values means to implicitly select a system of coordinates for the state. This selection is called a balanced representation and it constitutes a major difference

1  
2  
3 between the developed algorithm and other procedures. This property guarantees that a  
4 size that is smaller than the correct size is specified for the state vector, the estimations  
5 for  $F$ ,  $H$  and  $\Omega$  would be an approximation of a lower order than the real matrices of the  
6 system. In this case, this lower dimension model includes the most important effects and  
7 estimates them on a consistent basis. On the other hand, if the dimension of the  
8 estimated model is higher than the correct dimension, some of the system matrices are  
9 redundant. The rest remain valid and thus it is not necessary to recalculate them.  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19

20 Once the order and the matrices of the system have been determined, the Kalman  
21 filter is used to generate the series of values for the state vector. This is a fundamental  
22 element in statistical quality control as it allows to obtain the value of innovations on a  
23 recursive and efficient basis.  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60