Working Paper Series, N. 19, Dicembre 2007

Bootstrap-based design of residual control charts

Giovanna Capizzi
Department of Statistical Sciences
University of Padua
Italy

Guido Masarotto
Department of Statistical Sciences
University of Padua
Italy

Abstract: One approach to monitoring autocorrelated data consists of applying a control chart to the residuals of a time series model estimated from process observations. Recent research shows that the impact of estimation error on the run length properties of the resulting charts is not negligible. In this paper a general strategy for implementing residual-based control schemes is investigated. The designing procedure uses the AR-sieve approximation assuming that the process allows an autoregressive representation of order infinity. The run length distribution is estimated using bootstrap resampling in order to account for uncertainty in the estimated parameters. Control limits that satisfy a given constraint on the false alarm rate are computed via stochastic approximation. The proposed procedure is investigated for three residual-based control charts: Generalized Likelihood Ratio (GLR), Cumulative Sum ($CUSUM$) and Exponentially Weighted Moving Average ($EWMA$). Results show that the bootstrap approach safeguards against an undesirably high rate of false alarms. In addition, the out-of-control bootstrap-charts sensitivity does not seem to be lower than that of charts designed under the assumption that the estimated model is equal to the true generating process.

Keywords: On-line quality control, Time series, Autocorrelated data, Uncertainty modeling, Control charts, Sieve bootstrap.

Contents

Department of Statistical Sciences
Via Cesare Battisti, 241
35121 Padova
Italy

tel: +39 049 8274168
fax: +39 049 8274170
http://www.stat.unipd.it

Corresponding author:
Giovanna Capizzi
tel: +39 049 827 4168
giovanna.capizzi@unipd.it

Bootstrap-based design of residual control charts

Giovanna Capizzi
Department of Statistical Sciences
University of Padua
Italy

Guido Masarotto
Department of Statistical Sciences
University of Padua
Italy

Abstract: One approach to monitoring autocorrelated data consists of applying a control chart to the residuals of a time series model estimated from process observations. Recent research shows that the impact of estimation error on the run length properties of the resulting charts is not negligible. In this paper a general strategy for implementing residual-based control schemes is investigated. The designing procedure uses the AR-sieve approximation assuming that the process allows an autoregressive representation of order infinity. The run length distribution is estimated using bootstrap resampling in order to account for uncertainty in the estimated parameters. Control limits that satisfy a given constraint on the false alarm rate are computed via stochastic approximation. The proposed procedure is investigated for three residual-based control charts: Generalized Likelihood Ratio (GLR), Cumulative Sum ($CUSUM$) and Exponentially Weighted Moving Average ($EWMA$). Results show that the bootstrap approach safeguards against an undesirably high rate of false alarms. In addition, the out-of-control bootstrap-charts sensitivity does not seem to be lower than that of charts designed under the assumption that the estimated model is equal to the true generating process.

Keywords: On-line quality control, Time series, Autocorrelated data, Uncertainty modeling, Control charts, Sieve bootstrap.

1 Introduction

Control charts are widely used for monitoring process and quality improvement (see Montgomery, 2004). Most statistical process control techniques assume that consecutive observations from a process are independent and identically distributed (i.i.d) over time. However, with the development of high sampling frequency in the data collection, observations are more likely to be autocorrelated. The run length, RL, properties of traditional control charts, like Shewhart, $CUSUM$ and $EWMA$, are strongly degraded by data autocorrelation. Thus, there has been a burst of research in recent years on designing procedures for handling autocorrelation. Assuming that the underlying time series model is known, two main approaches have emerged. In the first, the original data are directly monitored using a standard control chart whose control limits are adjusted to account for the autocorrelation (Vasilopoulos and Stamboulis, 1978; Yashchin, 1993; Schimd and Schöne, 1997; Runger, 2002). The second approach consists of monitoring the forecast errors (residuals) to identify unusual observations. When the time-series model is correctly specified, the residuals are i.i.d with mean zero for an in-control, IC, process. Consequently, it is possible to use traditional control schemes with well-understood run length properties (Alwan and Roberts, 1988; Harris and Ross, 1991; Montgomery and Mastrangelo, 1991; Wardell et al., 1994; Runger and Willemain, 1995; Runger et al., 1995; Lu and Reynolds, 1999a,b; Apley and Shi, 1999; Shu et al., 2002).

As mentioned above, both approaches rely on accurate process model knowledge. In practice, the structure of dependence and/or the time-series parameters have to be estimated on the basis of n observations from an IC process. The typical design procedure consists of controlling the false alarm rate. In the presence of modeling errors, the rate of incorrect signals is a random variable, being a function of the estimated model parameters. Thus, if the fitted model is inaccurate, the control limits of the modified and residual control schemes will fail to provide the desidered run length properties. Indeed, much of the recent research, that investigates the impact of estimation error, shows that even small errors in parameter estimates can significantly alter the RL characteristics (Adams and Tzeng, 1998; Boyles, 2000; Kramer and Schmid, 2000; Apley, 2002; Apley and Lee, 2003; Testik, 2005; Jensen et al., 2006). Further, the identification of an appropriate time-series model is sometimes difficult and requires skill obtained by experience.

Although the adverse impact of model uncertainty on the run length performance is well documented, only a few studies suggest practical guidelines to tackle this issue. A significant contribution to a robust design for dependent data is the pioneering work of Apley (2002) which provided a design method of the EWMA chart for ARMA processes. Since the proposed control limits are a function of the covariance matrix of the ARMA parameter estimates, the resulting chart is robust to parameter modeling errors. Apley and Lee (2003) also derived an approximate upper one-sided confidence interval for the standard deviation of the EWMA control statistic which can be used to widen the control limits by an aumont that depends on the level of model uncertainty and on how conservative the design practitioner is. Testik (2005), following an approach similar to Apley (2002), suggests another method to widen the residual EWMA control limits for a stationary first order autoregressive process. As a result of incorporating parameter uncertainty, all these control limit are wider than the standard control limits used when models are assumed perfect. Hence, such approaches clearly give some protection against an unacceptably rate of false alarms together with a certain amount of decrease in the EWMA out-of-control performance, as is the case with the more conservative procedure proposed by Apley and Lee (2003). Two drawbacks characterize these methods. First, a key step of these approaches consists on writing the residual EWMA control statistic as the output of a filter linear applied to an ARMA process. Then, approximated closed form expressions for the standard deviation of the EWMA chart statistic are used to derive the control charts limits. Hence, the designing procedure strictly depends on the EWMA chart characteristics and it is not obvious how to extend it to other control charts. Second, only estimation errors are considered assuming a complete knowledge of model structure. However, in practical situations, the order of the model is often unknown and the combined effect of model misspecification and parameter estimation should be addressed in designing and setting up control charts (Jensen et al., 2006).

This paper explores a general designing procedure for residual-based control charts in the presence of model uncertainty. This procedure is based on the very mild assumption that the true underlying process allows an autoregressive representation of order infinity with gaussian innovations. A design approach based on the $AR-$sieve bootstrap algorithm (Bühlmann, 1998a,b, 2002; Alonso et al., 2002, 2003) is used to take into account the effects of modeling errors. The control limits are computed via stochastic approximation (Ruppert, 1991; Kushner and Yin, 2003) so that a given constraint on the random false alarm rate is satisfied. The proposed designing procedure is illustrated for three control charts: the Generalized Likelihood Ratio, GLR (Willsky and Jones, 1976; Basseville and Nikiforov, 1993; Superville and Adams, 1994; Siegmund and Venkatraman, 1995; Apley and Shi, 1999; Lai, 2001) and the traditional $CUSUM$ and $EWMA$. We also compare the bootstrap control limits to the control limits suggested by Apley (2002), Apley and Lee (2003) and Testik (2005) for a residual EWMA control chart.

2 Framework

Assume that, when a system is under control, observations are generated by a Gaussian stationary process, x_t, that allows an autoregressive representation of order infinity, $AR(\infty)$,

$$x_t - \mu = \sum_{j=1}^{\infty} \phi_j (x_{t-j} - \mu) + \epsilon_t, \quad t \in Z,$$

where $\mu = E(x_t)$, ϕ_j are parameters such that $\sum_j \phi_j^2 < \infty$ and ϵ_t is an i.i.d. innovation sequence following a Gaussian distribution with $E(\epsilon_t) = 0$ and $E(\epsilon_t^2) = \sigma^2$. This class of models includes stationary and invertible autoregressive moving average models. We will denote with β the infinite dimensional parameter vector $(\mu, \sigma, \phi_1, \phi_2, \ldots)$. Observe that β completely determines the process probability distribution.

Suppose that a persistent shift in the mean occurs at some unknown time τ. Thus, the process data to be monitored are given by

$$y_t = \begin{cases} x_t & \text{if } t < \tau \\ x_t + \delta & \text{if } t \geq \tau \end{cases}$$

Other types of deviations, such as transient shifts or linear drifts, can also be considered.

Let

$$\hat{y}_t(\beta) = \begin{cases} \mu & \text{if } t = 1 \\ E(y_t | y_{t-1}, \ldots, y_1) & \text{if } t > 1 \end{cases}$$

be the best mean-squared predictor of y_t based upon $y_{t-1}, \ldots, y_1$ and

$$v_t^2 = E[(y_t - \hat{y}_t)^2]$$

the mean-squared prediction error computed under the hypothesis that $t < \tau$, i.e., assuming that the process is IC at time t. Since y_t is Gaussian, the best predictor is linear and $\hat{y}_t(\beta)$ and $v_t^2(\beta)$ can be computed by either the Durbin-Levinson or the innovation algorithms and, when $1 - \sum_j \phi_j z^j$ is rational in z, by using a Kalman filter approach (Brockwell and Davies, 1996).

Residuals control charts are based on the standardized one-step prediction errors

$$a_t(\beta) = \frac{y_t - \hat{y}_t(\beta)}{v_t(\beta)} = \tilde{x}_t + \delta \tilde{f}_\tau(t), \quad t = 1, 2, \ldots$$

where $\tilde{x}_t$ and $\tilde{f}_\tau(t)$ are the outputs of linear filter defining $a_t(\beta)$, when the input is x_t and $\mu + 1$, respectively. Since the sequence $\tilde{x}_t$ comprises the standardized one-step prediction errors of the x process, the $\tilde{x}_t$'s are i.i.d. random variables from a standard normal distribution. Hence, when the model is perfectly known, the residuals are normal and uncorrelated with a time-varying mean $\delta \tilde{f}_\tau(t)$. The value of $\tilde{f}_\tau(t)$, the so called *fault signature*, depends upon the autocorrelation structure of the data, but always $\tilde{f}_\tau(t) = 0$ when $t < \tau$.

A residual-based control chart can be summarized as follows: (i) at time $t = 1, 2, \ldots$, a control statistic $g_t(\beta) = g_t[a_t(\beta), \ldots, a_1(\beta)]$ is calculated from the process data; (ii) an out of control situation is signalled if $g_t(\beta) > h$, where h is the control limit.

In particular, windowed limited GLR, CUSUM, and EWMA control charts are based on the statistics

$$g_t(\beta) = \max_{j=0,\ldots,M-1} |T_j(t)| \text{ where } T_j(t) = \frac{\sum_{i=0}^{j} a_{t-i}(\beta) \tilde{f}_{t-j}(t-i)}{\sqrt{\sum_{i=0}^{j-1} \tilde{f}_{t-j}^2(t-i)}}, \qquad \text{(GLR)}$$

$$g_t(\beta) = \max(u_t^-(\beta), u_t^+(\beta)) \qquad \text{(CUSUM)}$$

where

$$u_t^-(\beta) = \max(0, u_{t-1}^-(\beta) - a_t(\beta) - k), \quad u_t^+(\beta) = \max(0, u_{t-1}^+(\beta) + a_t(\beta) - k)$$

with $u_0^-(\beta) = u_0^+(\beta) = 0$, and

$$g_t(\beta) = |u_t(\beta)| \sqrt{\frac{2-\lambda}{\lambda}} \text{ where } u_t(\beta) = u_{t-1}(\beta) + \lambda[a_t(\beta) - u_{t-1}(\beta)], \qquad \text{(EWMA)}$$

with $u_0(\beta) = 0$, respectively. Here, $M \in N$, $k > 0$ and $\lambda \in (0, 1]$ are suitable constants.

The run-length of a residuals chart can be formally expressed by the stopping rule

$$RL = \inf\{t : g_t(\beta) > h\}$$

Let $G(\cdot;\beta,\tau,\delta,h)$ be the distribution function of the RL, i.e., $G(rl;\beta,\tau,\delta,h) = P(RL \leq rl)$. In the following, we refer to $G_0(\cdot;\beta,h) = G(\cdot;\beta,\infty,0,h)$ as the in-control distribution function. When β is known without errors, the IC run-length distribution is completely known. This makes it easy to choose an appropriate value for the critical limit h fixing, in some way, the false alarm rate. The classical approach consists of determining h such that the IC average run length of the scheme, ARL_0, is equal to a prescribed high value. In passing, observe that, for traditional charts, like CUSUM and EWMA, the in-control run length distribution doesn't depend upon β and is equal to that of the i.i.d. context. Furthermore, the study of $G(\cdot;\beta,\tau,\delta,h)$, for some values of τ and δ, can be used to choose the other control charts constants, e.g. M, k and λ for the three charts described in the previous paragraph.

3 The effects of modelling errors

In practice, the time series model is rarely known. The standard approach consists of identifying the model from n in-control data, $y_1,\ldots,y_n$, and obtaining an estimate of β, denoted by $\hat{\beta}_n$. Then, the control charts use the estimated residuals $a_t(\hat{\beta}_n)$ instead of the true residuals $a_t(\beta)$. Let $H_n(\cdot;\beta,\tau,\delta,h)$ be the run-length distribution of the resulting charts. Observe that H_n depends implicitly on the method used to estimate the parameters. Due to the differences between β and $\hat{\beta}_n$, the $a_t(\hat{\beta}_n)$ are neither independent nor indentically distributed. Hence, $H_n(\cdot;\beta,\tau,\delta,h)$ is not equal to $G(\cdot;\beta,\tau,\delta,h)$, at least when n is finite. A *naive* and rather standard designing procedure neglects the fact that the model is estimated and designs control charts under the assumption of a perfect model, i.e. $\hat{\beta}_n = \beta$. Following this procedure the run-length distribution of the chart with estimated parameters is assumed to be equal to $G(\cdot,\hat{\beta}_n,\tau,\delta,h)$. Unfortunately, previous research shows that this *naive* approach can lead to a false alarm rate that is much higher than desired, even when substantial sample sizes are used.

In the following, a related example is used to illustrate why alternative measures to the ARL should be used when the known parameters are replaced with estimates. Suppose that a GLR, a $CUSUM$, and an $EWMA$ control chart, designed to give an in-control ARL of 1000 in the i.i.d. case, are applied to the forecast residuals from the $AR(1)$ model $x_t = 0.75x_{t-1} + \epsilon_t$, where $\epsilon_t \sim N(0,1)$. Tables 1, 2, and 3 give moments and quantiles of the corresponding in-control run-length distributions, estimated using 100000 Monte Carlo replicates. Dealing with time series modeling three cases will be distinguished:

A. the underlying time series model is known *a priori*;
B. the model order is specified, but process parameters are unknown. The unknown parameters must be estimated from an in-control reference sample of size n. In particular, we make use of Burg's estimation method (Brockwell and Davies, 1996).
C. both the model order and time series parameters are unknown. Here, we apply the $AICc$ criterion (Burnham and Anderson, 2002) to select the order of the model while the other parameters of interest are estimated as in case B. In the

Table 1: Performance of a GLR control chart ($M = 20$ and $h = 3.6751$) when the underlying process is $x_t = 0.75x_{t-1} + \epsilon_t$. ARL_0, σ_{RL_0} and Q_p denote the average, the standard deviation and the $p-$th quantile of the in-control run-length distribution.

	A: completely	B: partially known model			C: completely unknown model		
	known model	$n = 100$	$n = 200$	$n = 300$	$n = 100$	$n = 200$	$n = 300$
$Q_{0.01}$	16	6	9	11	5	9	10
$Q_{0.05}$	57	18	29	35	15	25	32
$Q_{0.10}$	110	33	55	68	26	47	59
$Q_{0.25}$	292	94	152	184	71	127	162
$Q_{0.50}$	695	286	422	490	219	357	433
$Q_{0.75}$	1383	865	1070	1159	677	924	1036
$Q_{0.90}$	2293	2315	2358	2358	1862	2069	2131
$Q_{0.95}$	2978	4126	3728	3530	3390	3294	3206
$Q_{0.99}$	4590	12139	8441	7217	10231	7597	6632
ARL_0	999.17	1053.12	987.65	981.67	864.44	865.03	886.08
σ_{RL_0}	993.78	3504.77	1877.72	1531.18	3166.39	1700.59	1421.71

following, the value of the maximum order is fixed at $10\log_{10}(n)$.
Under the assumption that there are no modeling errors, case A, the residuals are i.i.d. and the standard control limits provide the desired ARL_0. With model uncertainty, cases B and C, the residuals are autocorrelated and the same value of h may fail to provide the specified ARL_0. For $n = 100$, for both cases B and C, the 10-th and 50-th quantiles of the run-length distribution are roughly one-fourth and one-third as those of the distribution with known parameters. However, the 99-th quantiles are from two to five times larger than in the parameters known case. Although the GLR test seems to be less affected by the modeling uncertainty, in all cases the estimated in-control run length distribution, appears to be shifted to the left (lower values) as result of a large percent of earlier false alarms. This increased rate of shorter runs between alarm signals is not captured by the expected value of the run length that is affected by the presence of a few extremely long runs.

Since the ARL is not able to reflect the whole run-length performance, it may be interesting to investigate alternative measures of the control chart sensitivity in the presence of model uncertainty. A reasonable criterion might consist of determining the control limit h so that the probability of a false alarm within some specified value, N_0, is equal to p_0

$$H_{0,n}(N_0; \beta, h) = P(RL \leq N_0) = p_0 \tag{1}$$

where $H_{0,n}(\cdot; \beta, h) = H_n(\cdot; \beta, \infty, 0, h)$ is the in-control distribution. For instance, with $N_0 = 200$ and $p_0 = 0.20$, the design would consist of finding h such that only one in five false alarms should be given before the 200-th observation.

Observe that as more data become available the designing procedure may be optimized through a regular updating of the h estimate. In this case N_0 may be set equal to the time until the next update and p_0 equal to an acceptable rate of false alarms for this time interval.

Table 2: Performance of a $CUSUM$ control chart ($k = 0.5$ and $h = 5.7573$) when the underlying process is $x_t = 0.75x_{t-1}+\epsilon_t$. ARL_0, σ_{RL_0} and Q_p denote the average, the standard deviation and the $p-$th quantile of the in-control run-length distribution.

	A: completely	B: partially known model			C: completely unknown model		
	known model	$n = 100$	$n = 200$	$n = 300$	$n = 100$	$n = 200$	$n = 300$
$Q_{0.01}$	18	8	11	12	7	10	12
$Q_{0.05}$	59	18	29	35	16	26	32
$Q_{0.10}$	112	32	53	65	28	47	60
$Q_{0.25}$	293	87	143	175	73	127	161
$Q_{0.50}$	696	272	408	477	227	365	440
$Q_{0.75}$	1384	909	1128	1210	753	1010	1117
$Q_{0.90}$	2290	2893	2811	2710	2435	2525	2499
$Q_{0.95}$	2974	5995	4856	4361	5061	4374	4033
$Q_{0.99}$	4553	23747	13648	10604	20826	12432	9852
ARL_0	999.69	1686.41	1265.17	1166.58	1461.17	1152.59	1079.95
σ_{RL_0}	990.03	9192.80	3628.33	2506.23	8544.90	3587.21	2311.72

Table 3: Performance of a $EWMA$ control chart ($\lambda = 0.1$ and $h = 3.0586$) when the underlying process is $x_t = 0.75x_{t-1}+\epsilon_t$. ARL_0, σ_{RL_0} and Q_p denote the average, the standard deviation and the $p-$th quantile of the in-control run-length distribution.

	A: completely	B: partially known model			C: completely unknown model		
	known model	$n = 100$	$n = 200$	$n = 300$	$n = 100$	$n = 200$	$n = 300$
$Q_{0.01}$	19	8	11	13	8	11	13
$Q_{0.05}$	60	19	29	35	17	27	33
$Q_{0.10}$	114	33	52	65	28	48	60
$Q_{0.25}$	294	86	140	172	74	127	160
$Q_{0.50}$	694	260	392	464	224	359	433
$Q_{0.75}$	1383	843	1073	1165	726	985	1093
$Q_{0.90}$	2288	2727	2694	2620	2328	2474	2463
$Q_{0.95}$	2971	5818	4756	4245	4972	4367	4014
$Q_{0.99}$	4567	27301	14350	10672	22710	13251	10110
ARL_0	999.71	1873.26	1288.50	1151.94	1602.94	1183.93	1085.99
σ_{RL_0}	989.63	12321.28	4620.76	2735.01	11059.75	4260.14	2558.53

4 An $AR-$sieve bootstrap design

Since $H_n(\cdot;\beta,\tau,\delta,h)$ is a function of the unknown parameter β, its exact computation is not possible. However, the run length distribution function can be estimated by using a bootstrap method for time series that is known as the $AR-$sieve bootstrap. This approach is based on a sieve finite approximation that, in the present case, consists of the following steps:

a) given a Phase I sample, $y_1, ...y_n$, identify a finite order autoregressive approximation for the data-generating process. Let p_n be the selected order of the approximating model M_n and $\hat{\beta}_n = (\hat{\mu}, \hat{\sigma}, \hat{\phi}_1, \ldots, \hat{\phi}_{p_n}, 0, \ldots)$, the Burg estimate of the corresponding parameter vector β.

Automatic criteria, such as AIC, $AICc$ or BIC (see Shibata, 1980; Hurvich and Tsai, 1989; Burnham and Anderson, 2002), can be used to select the model order in a fully automatic fashion. In particular, we use the $AICc$ criterion (Burnham and Anderson, 2002). As previous works outline (Hurvich and Tsai, 1989; Alonso et al., 2002), this criterion is preferred because the true model can be complex and not of finite dimension, and also because it is less affected than other methods by changes in the maximum order considered (here fixed at $10\log_{10}(n)$);

b) use the fitted model M_n to generate pseudo-data y_t^* and x_t^* according to

$$y_t^* = \begin{cases} x_t^* & \text{if } t < \tau + n \\ x_t^* + \delta & \text{if } t \geq \tau + n \end{cases}, \quad x_t^* - \hat{\mu} = \sum_{j=1}^{p_n} \hat{\phi}_j (x_{t-j}^* - \hat{\mu}) + \epsilon_t^*, \quad t = 1, 2, \ldots \tag{2}$$

with $\epsilon_t^* \sim N(0, \hat{\sigma}^2)$.

In order to obtain a pseudo-stationary in-control sequence x_t^*, we generate the time series starting from $t = -100$ and setting $x_t^* = 0$ when $t < -100$. Then we discard the first 100 observations;

c) use the first n observations from (2), $(y_1^*, \ldots, y_n^*)$, to select the order $\hat{p}_n^*$ of an approximating model M_n^*. Then, compute the estimates of the autoregressive parameters $\hat{\beta}_n^* = (\hat{\mu}^*, \hat{\sigma}^*, \hat{\phi}_1^*, \ldots, \hat{\phi}_{\hat{p}_n^*}^*, 0, \ldots)$, as done in step a).

Observe that for each bootstrap replicate a model M_n^* is re-fitted. Since for each replication both the autoregressive order and parameters are re-estimated, the proposed procedure takes into account for uncertainty in the parameter estimates as well as in the choice of an approximating order;

d) apply the control chart to the sequence of residuals, $a_{n+1}^*(\hat{\beta}_n^*), a_{n+2}^*(\hat{\beta}_n^*), \ldots$, computed from the model M_n^*;

e) record the run length $RL^* = T - n$, where T is the first time at which the control chart gives an out-of-control signal;

f) repeat steps b)-e) a large number of times and use the empirical distribution of the run-lengths RL^* to estimate the unknown distribution function $H_n(\cdot; \beta, \tau, \delta, h)$.

Note that, as the number of boostrap replications goes to infinity, the empirical distribution of the run-lengths RL^* tends to $H_n(\cdot; \hat{\beta}_n, \tau, \delta, h)$ that is hence used to estimate the unknown distribution function $H_n(\cdot; \beta, \tau, \delta, h)$.

Then, according to (1), a suitable value of the control limit can be obtained by solving for h the following equation

$$H_{0,n}(N_0; \hat{\beta}_n, h) = p_0 \tag{3}$$

where $H_{0,n}(N_0; \hat{\beta}_n, h)$ is the bootstrap estimated probability of a false alarm before some pre-assigned value N_0.

As Yashchin (1993) pointed out, stochastic approximation may be an appropriate procedure for an iterative search of the control limit able to satisfy the given constraint on the frequency of false alarms. The main drawback which the author himself ascribes to the considered approach is that it can be computationally demanding. However, our experimental results show that stochastic approximation may be successfully used to compute the solution of equation (3) in particular when an efficient scheme like the Polyak-Ruppert algorithm is used (see Ruppert, 1988; Polyak, 1990; Ruppert, 1991; Polyak and Juditsky, 1992). Following this approach, which is even simple to implement, the estimate at the $s-$th step of the approximation algorithm is given by

$$\hat{h}_s = \frac{1}{s - s_0} \sum_{i=s_0+1}^{s} h_i$$

where the h_i values are generated by the recursion

$$\hat{h}_{i+1} = \max[0, h_i + Ai^{-\alpha}(I(RL^* \leq N_0) - p_0)], \quad i = 1, 2, \ldots$$

Here, $\{RL_i^*\}$ are independent random variables with distribution $H_{0,n}(\cdot; \hat{\beta}_n, h_i)$, $I(p)$ is the indicator function which yields 1 if the proposition p is true and 0 otherwise, h_1 is an initial guess of h, while $A > 0$, $s_0 < s$ and $0.5 < \alpha < 1$ are suitable constants. Practical and theoretical considerations, for instance that

$$E[H_{0,n}(N_0; \hat{\beta}_n, h_s) - p_0]^2 \approx \frac{p_0(1 - p_0)}{s - s_0}$$

for s sufficiently large, suggest that setting $A = 3$, $s_0 = 100$, $s = 10000$ and $\alpha = 0.9$ yields results of acceptable accuracy, at least for $p_0 \in (0.05, 0.30)$.

Once an estimate of h has been obtained, the study of $H_n(\cdot, \hat{\beta}_n, \tau, \delta, h)$, for some values of τ and δ, could be used to choose the other control chart constants.

5 Simulation Results

A Monte Carlo simulation experiment is conducted on time series generated from the following in-control models:
$AR1 : x_t = 0.75x_{t-1} + \epsilon_t;$
$MA1 : x_t = \epsilon_t + 0.75\epsilon_{t-1};$
$MA4 : x_t = \epsilon_t + 0.6\epsilon_{t-4};$
$ARMA22 : x_t = 1.4x_{t-1} - 0.5x_{t-2} + \epsilon_t - 0.7\epsilon_{t-2},$
with $\epsilon_t \sim N(0, 1)$.

Observe that only $AR1$ is a finite order autoregressive model while the other considered models only allow an infinite autoregressive representation.

The proposed designing procedure is here investigated for three residual-based control charts:
$C1$: a GLR chart, with $M = 20$;
$C2$: a $CUSUM$, with reference value $k = 0.5$;
$C3$: an $EWMA$, with $\lambda = 0.1$.
For each time series model, 2000 sequences of size n are generated under the hypothesis of no shift in the process mean. From $t = n + 1$, 2000 pseudo-random

continuations for each of the original sequences are simulated for both in control and out of control conditions. In particular, for the out-of-control scenario, a shift of two times the standard deviation of the in-control process is added to the process mean starting at $t = n + 201$.

The first n data are used to fit a suitable stochastic model and compute the *naive* and the *bootstrap* decision interval h. In particular, if $N_0 = 200$ and $p_0 = 0.20$, these two methods consist of obtaining the control limit h as solution of

$$G_0(200; \hat{\beta}_n, h) = 0.2 \tag{4}$$

and

$$H_{0,n}(200; \hat{\beta}_n, h) = 0.2, \tag{5}$$

respectively. Note that equation (4) assumes that the true parameters coincide with their estimates, i.e. $\beta = \hat{\beta}_n$, whereas, via the AR-sieve bootstrap approach, equation (5) is able to take into account the sampling variability in both the model order and the parameter estimates.

Recorded the resulting run lengths, control-charts performance may be discussed in terms of frequency of erroneous and correct signals. In particular, we here estimate the following performance measures

i) $FS(\hat{\beta}_n)$: the probability of a False Signal before $t = 200$;

ii) $TS(\hat{\beta}_n)$: the probability of a True Signal between $t = 201$ and $t = 220$,

of residual control charts based on $a_t^*(\hat{\beta}_n)$. Observe that, since $FS(\hat{\beta}_n)$ and $TS(\hat{\beta}_n)$ depend on $\hat{\beta}_n$, i.e. on the sample $(y_1, \ldots, y_n)$, the probabilities of false and true alarms are stochastic. Thus, for evaluating the extent to which the bootstrap and the *naive* procedures are able to achieve desired performances, the distribution of these probabilities may be investigated with respect to summary values such as the mean, standard deviation or some upper and lower percentiles.

Tables 4, 5, and 6 give the mean and standard deviation of the control limit, h, and of $FS(\hat{\beta}_n)$ and $TS(\hat{\beta}_n)$, for Phase I samples of size $n = 50, 100, 200$ and 300. Since each entry in the tables has been estimated using 2000 time-series and 2000 replicates of the run length for each time series, the number of simulated run lengths for each entry is equal to 4000000. The control limit when the underlying model is perfectly known, i.e. h_∞ such that

$$G_0(200; \beta, h_\infty) = 0.2,$$

is also included. All the critical values have been computed by stochastic approximation setting $h_1 = 3.5$ for GLR and EWMA and $h_1 = 5$ for CUSUM.

Observe that the h values determined from equations (4) and (5) attempt to satisfy the following constraint

$$E[FS(\hat{\beta}_n)] = 0.2 \tag{6}$$

on the rate of false alarms. However, while the bootstrap-based control limits approximately guarantee conditions (6), at least when $n \geq 100$, the *naive* control limits fail to yield the nominal rate of false alarms, i.e. $p_0 = 0.20$. In particular, using the *naive* approach, the probability of a false alarm within 200 observations is greater

Table 4: Mean and standard deviation of the control limits ($\hat{h}$) and of the probabilities of false and true alarms, $FS(\hat{\beta}_n)$ and $TS(\hat{\beta}_n)$, for a residual-based GLR control chart ($M = 20$)

			bootstrap			naive		
model	n		$\hat{h}$	$FS(\hat{\beta}_n)$	$TS(\hat{\beta}_n)$	$\hat{h}$	$FS(\hat{\beta}_n)$	$TS(\hat{\beta}_n)$
AR1	50	mean	5.169	0.221	0.364	3.647	0.639	0.292
		sd	0.486	0.325	0.253	0.034	0.335	0.260
	100	mean	4.352	0.201	0.523	3.644	0.497	0.420
		sd	0.136	0.187	0.262	0.029	0.303	0.233
	200	mean	3.986	0.191	0.626	3.641	0.378	0.534
		sd	0.064	0.183	0.131	0.023	0.227	0.171
	300	mean	3.864	0.193	0.654	3.640	0.323	0.586
		sd	0.038	0.144	0.095	0.019	0.177	0.127
	∞		3.631	0.203	0.696			
MA1	50	mean	4.997	0.234	0.671	3.668	0.673	0.323
		sd	0.401	0.321	0.319	0.011	0.316	0.310
	100	mean	4.380	0.217	0.759	3.670	0.549	0.449
		sd	0.168	0.249	0.240	0.010	0.286	0.283
	200	mean	4.046	0.201	0.795	3.672	0.416	0.583
		sd	0.067	0.172	0.169	0.008	0.220	0.219
	300	mean	3.929	0.201	0.797	3.671	0.354	0.645
		sd	0.043	0.144	0.143	0.008	0.178	0.177
	∞		3.672	0.201	0.800			
MA4	50	mean	5.302	0.295	0.608	3.640	0.757	0.242
		sd	0.747	0.344	0.333	0.018	0.377	0.276
	100	mean	4.553	0.222	0.766	3.647	0.610	0.389
		sd	0.205	0.268	0.261	0.016	0.287	0.287
	200	mean	4.129	0.208	0.790	3.651	0.476	0.523
		sd	0.091	0.191	0.190	0.012	0.227	0.227
	300	mean	3.986	0.210	0.789	3.652	0.410	0.589
		sd	0.056	0.159	0.158	0.011	0.193	0.192
	∞		3.658	0.198	0.804			
ARMA22	50	mean	5.230	0.213	0.711	3.673	0.698	0.299
		sd	0.407	0.311	0.304	0.028	0.314	0.309
	100	mean	4.463	0.185	0.809	3.671	0.561	0.439
		sd	0.136	0.233	0.229	0.021	0.283	0.282
	200	mean	4.079	0.193	0.807	3.668	0.434	0.566
		sd	0.061	0.176	0.175	0.017	0.221	0.221
	300	mean	3.948	0.193	0.807	3.667	0.366	0.634
		sd	0.042	0.141	0.141	0.015	0.184	0.184
	∞		3.667	0.199	0.805			

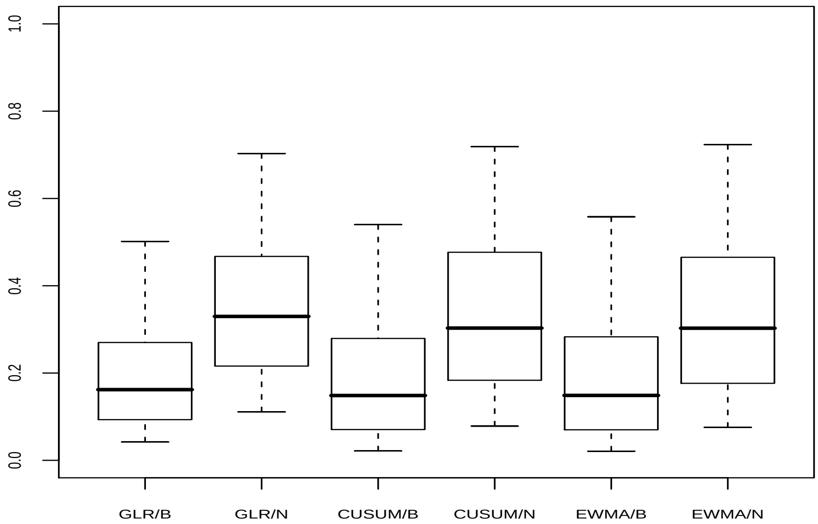
Table 5: Mean and standard deviation of the control limits ($\hat{h}$) and of the probabilities of false and true alarms, $FS(\hat{\beta}_n)$ and $TS(\hat{\beta}_n)$, for a residual-based $CUSUM$ control chart ($k = 0.5$)

			bootstrap			naive		
model	n		$\hat{h}$	$FS(\hat{\beta}_n)$	$TS(\hat{\beta}_n)$	$\hat{h}$	$FS(\hat{\beta}_n)$	$TS(\hat{\beta}_n)$
AR1	50	mean	12.383	0.254	0.290	5.633	0.625	0.271
		sd	1.440	0.341	0.251	0	0.351	0.244
	100	mean	8.396	0.225	0.436	5.633	0.486	0.386
		sd	0.621	0.284	0.223	0	0.318	0.221
	200	mean	6.931	0.198	0.556	5.633	0.373	0.496
		sd	0.261	0.207	0.153	0	0.245	0.166
	300	mean	6.457	0.198	0.592	5.633	0.323	0.544
		sd	0.134	0.168	0.112	0	0.199	0.128
	∞		5.633	0.200	0.661			
MA1	50	mean	11.132	0.259	0.637	5.633	0.636	0.359
		sd	1.085	0.339	0.340	0	0.334	0.339
	100	mean	8.270	0.235	0.741	5.633	0.517	0.480
		sd	0.727	0.279	0.271	0	0.264	0.310
	200	mean	6.995	0.201	0.795	5.633	0.395	0.604
		sd	0.285	0.200	0.197	0	0.172	0.241
	300	mean	6.548	0.198	0.800	5.633	0.340	0.659
		sd	0.171	0.169	0.167	0	0.134	0.199
	∞		5.633	0.199	0.799			
MA4	50	mean	12.988	0.325	0.547	5.633	0.728	0.271
		sd	2.730	0.366	0.305	0	0.351	0.306
	100	mean	9.516	0.233	0.747	5.633	0.589	0.410
		sd	0.967	0.298	0.289	0	0.279	0.304
	200	mean	7.478	0.210	0.788	5.633	0.458	0.541
		sd	0.421	0.210	0.216	0	0.192	0.248
	300	mean	6.864	0.210	0.789	5.663	0.400	0.599
		sd	0.229	0.184	0.183	0	0.152	0.214
	∞		5.633	0.199	0.797			
ARMA22	50	mean	12.493	0.242	0.516	5.633	0.676	0.304
		sd	1.408	0.338	0.342	0	0.332	0.307
	100	mean	8.795	0.209	0.704	5.633	0.548	0.440
		sd	0.699	0.272	0.254	0	0.309	0.294
	200	mean	7.245	0.197	0.773	5.633	0.425	0.567
		sd	0.280	0.209	0.194	0	0.245	0.236
	300	mean	6.708	0.197	0.785	5.633	0.366	0.627
		sd	0.169	0.171	0.160	0	0.206	0.200
	∞		5.633	0.199	0.799			

Table 6: Mean and standard deviation of the control limits ($\hat{h}$) and of the probabilities of false and true alarms, $FS(\hat{\beta}_n)$ and $TS(\hat{\beta}_n)$, for a residual-based $EWMA$ control chart ($\lambda = 0.1$)

			bootstrap			naive		
model	n		$\hat{h}$	$FS(\hat{\beta}_n)$	$TS(\hat{\beta}_n)$	$\hat{h}$	$FS(\hat{\beta}_n)$	$TS(\hat{\beta}_n)$
AR1	50	mean	4.649	0.238	0.268	3.008	0.625	0.259
		sd	0.529	0.331	0.259	0	0.348	0.234
	100	mean	3.817	0.219	0.386	3.008	0.495	0.367
		sd	0.159	0.277	0.217	0	0.316	0.213
	200	mean	3.419	0.200	0.508	3.008	0.381	0.475
		sd	0.069	0.202	0.155	0	0.242	0.163
	300	mean	3.280	0.201	0.551	3.008	0.330	0.522
		sd	0.038	0.169	0.117	0	0.202	0.126
	∞		3.008	0.201	0.638			
MA1	50	mean	4.386	0.252	0.618	3.008	0.623	0.323
		sd	0.437	0.330	0.334	0	0.337	0.328
	100	mean	3.774	0.221	0.728	3.008	0.510	0.486
		sd	0.190	0.271	0.261	0	0.306	0.302
	200	mean	3.425	0.203	0.789	3.008	0.389	0.610
		sd	0.079	0.198	0.192	0	0.238	0.237
	300	mean	3.299	0.200	0.796	3.008	0.338	0.661
		sd	0.051	0.169	0.167	0	0.200	0.199
	∞		3.008	0.201	0.795			
MA4	50	mean	4.766	0.309	0.528	3.008	0.703	0.296
		sd	0.796	0.359	0.348	0	0.319	0.316
	100	mean	4.004	0.225	0.728	3.008	0.565	0.434
		sd	0.238	0.292	0.280	0	0.	0.313
	200	mean	3.541	0.218	0.777	3.008	0.449	0.551
		sd	0.111	0.220	0.216	0	0.192	0.251
	300	mean	3.389	0.211	0.786	3.008	0.388	0.611
		sd	0.070	0.187	0.186	0	0.152	0.216
	∞		3.008	0.198	0.802			
ARMA22	50	mean	4.672	0.232	0.497	3.008	0.636	0.339
		sd	0.512	0.329	0.336	0	0.345	0.314
	100	mean	3.899	0.201	0.677	3.008	0.518	0.467
		sd	0.179	0.263	0.249	0	0.309	0.292
	200	mean	3.484	0.195	0.760	3.008	0.402	0.587
		sd	0.078	0.208	0.189	0	0.248	0.238
	300	mean	3.337	0.196	0.779	3.008	0.352	0.639
		sd	0.049	0.170	0.156	0	0.205	0.198
	∞		3.008	0.200	0.794			

(a) False signal



(b) True signal

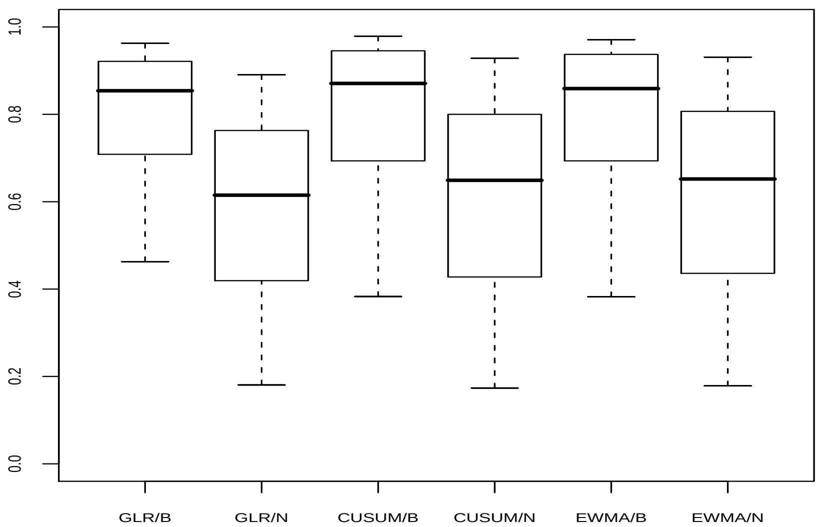


Figure 1: Distribution of $FS(\hat{\beta}_n) = P(1 \leq RL \leq 200|\hat{\beta}_n)$ and $TS(\hat{\beta}_n) = P(201 \leq RL \leq 220|\hat{\beta}_n)$, for the $MA1$ model and $n = 200$.

than 62% when $n = 50$, and greater than 32% when $n = 300$, for all the considered control charts and time series models.

Obviously, as we pointed out, the probability of false and true signals depends on the estimated parameters. This unavoidable variability is captured by the standard errors of $FS(\hat{\beta}_n)$ and $TS(\hat{\beta}_n)$ listed in Tables 4-6. Observe that, the *naive* approach mostly leads to a higher variability of the probabilities, whereas the bootstrap-based procedure seems to provide a better robustness with respect to modeling errors.

In contrast, as the bootstrap approach uses wider control limits to account for uncertainty in the estimated parameters, the probability of detecting a mean shift before the 20th time step after its occurrence, conditioned on no previous false alarm, i.e.

$$P(RL \leq 220|RL > 200) = \frac{P(RL > 200 \cap RL \leq 220)}{P(RL > 200)} = \frac{E[TS(\hat{\beta}_n)]}{1 - E[FS(\hat{\beta}_n)]} \qquad (7)$$

will be inevitably reduced. For example, for the model $AR1$ and $n = 200$, in Table 4 the probability in (7) is roughly equal to 77% with the bootstrap and 86% with the *naive* approach. This in-control versus out-of-control trade-off has been widely discussed in the previous related literature (see Schimd and Schöne, 1997; Apley, 2002; Apley and Lee, 2003; Jones et al., 2001; Jones, 2002). However, when the costs of frequent false alarms are also considered and the interest is in the probability of a correct signal after the beginning of the monitoring, i.e. in the values of $TS(\hat{\beta}_n)$, on average the bootstrap approach leads to higher values of the probability to signal real out-of-control conditions.

In order to better emphasize the impact of *naive* and the bootstrap procedures on the estimated control chart performance, we graphically show the distribution of the probabilities $FS(\hat{\beta}_n)$ and $TS(\hat{\beta}_n)$, when residual-control charts are designed for the $MA1$ model. In particular, for a reference sample of $n = 200$, Figures 1(a) and 1(b) show boxplots of the probabilities $FS(\hat{\beta}_n)$ and $TS(\hat{\beta}_n)$, with whiskers drawn to the 5-th and 95-th percentiles, respectively. Note that, compared to the *naive* approach, the use of the resampling techniques leads, for all the residual-control charts, to boxplots showing i) a median strictly close the nominal value $p_0 = 0.20$, under the in-control situation; ii) larger values of all the position values in the out-of-control case; iii) a smaller variation under the in-control and out-of-control scenarios.

We conclude that, for each of models $AR1$ through $ARMA22$, the bootstrap designing procedure is able to guarantee the prescribed rate of false alarms when n is as low as 100, whereas even when $n = 300$ the *naive* approach leads to a substantial increase in the expected number of false alarms. In addition, for some specific autocorrelation structure, an effective bootstrap-based scheme seems to be designable even using $n = 50$. Anyway, also for this smaller value of the reference sample, the bootstrap charts widely outperform the naive charts in getting the desired rate of false alarms. Further, in all cases shown, the out-of control performance of the bootstrap control scheme is at least comparable to that of the naive chart. Note also that, since similar results have been obtained for other models and for different choices of the constants N_0 and p_0 the proposed design procedure seems to have a wide applicability.

Table 7: Average values and root mean square errors of the residual EWMA control limits when the underlying process is $x_t = 0.75x_{t-1} + \epsilon_t$; $\lambda = 0.1$ For the Apley and Lee (2003) control limits an approximate 90% confidence interval for the standard deviation of the EWMA control statistic is used.

n		h_{true}	Bootstrap	Apley(2002)	Testik(2005)	Apley and Lee (2003)
100	mean	3.682	3.606	3.079	3.470	3.670
	rsme		0.110	0.603	0.216	0.041
200	mean	3.341	3.329	3.045	3.335	3.497
	rsme		0.037	0.296	0.020	0.157
300	mean	3.226	3.227	3.033	3.273	3.413
	rsme		0.024	0.193	0.048	0.187

Although the performance comparisons of these charts is beyond the scope of this paper, it is interesting to observe that our exercise confirms the results of Apley and Shi (1999)) and points to a greater efficiency of the GLR chart even when the adjusted limits are used.

Finally, we compare the bootstrap control limits with the decision interval proposed by Apley (2002), Apley and Lee (2003) and Testik (2005). In order to make possible a direct comparison, the results refer to a residual EWMA control chart designed for the model AR1, under the assumption that the model order is known but that the parameters must be estimated from a reference sample of size n. Chosen the smoothing constant λ and a suitable constraint on the false alarm rate, the extent to which the EWMA control limits are widened depend on the different expressions of the estimated standard deviation of the EWMA control statistic that account for the random variations of parameter estimates. Given $\lambda = 0.1$, let $h = 3.008$ be the decision interval that satisfies the constraint $P(rl \leq 200) = 0.20$ on the false alarm rate assuming that the estimated model is perfect, i.e. as $n \to \infty$. Table 7 shows the averages and the root mean square errors of 2000 values control limits obtained using the bootstrap approach and EWMA variance equations suggested by Apley (2002), Testik (2005) and Apley and Lee (2003). Table 7 also contains the decision interval, h_{true}, that satisfies the given constraint on the false alarm rate, when a sample of size n is used. Note that a larger value of the control limit corresponds to a lower rate of false alarms but also in a slower reaction to real changes in the mean process. On the other hand, a smaller value of the decision interval leads to an increase of the expected number of signals under the in-control and out-of-control scenarios. Results show that on the whole bootstrap critical values are less biased and performs better than the modified EWMAs suggested by Apley (2002) and Testik (2005). As expected, the bootstrap control limits are slightly smaller than the EWMA control limits, proposed by Apley and Lee (2003), which seem to provide a viable but conservative approach. Thus, the proposed method seems to be robust to parameter uncertainty, at least at a comparable level than the worst-case EWMA design, without depending on the analytical properties of the charted statistic.

6 Conclusions

The adverse impact of model uncertainty on the performance of residual-based control charts is well known and documented in the literature. However, a general strategy for implementing a residual control chart, in the presence of modeling errors, is still lacking.

In this paper we have investigated an $AR-$sieve bootstrap method for designing residual-based control charts when the underlying time series model is unknown.

Results suggest that the proposed method is able to guarantee a prescribed rate of false alarms for all the investigated time series model when a Phase I sample size is as low as 100. Further, for some specific autocorrelation structure, an effective bootstrap-based chart seems to be designable even using a reference sample of size 50. If compared with the naive approach the bootstrap designing procedure exhibits a much better in-control performance and at least a similar out-of-control performance.

Further, since it is based on automatic identification, it can be also used when only a limited time series modeling experience is available.

Future research will include extension of the presented approach to other residual control charts (e.g. charts for a joint monitoring of a process mean and a variance and/or charts for multivariate processes) and the investigation of sieve approximations based on classes of dynamic models different from autoregressive. Finally, it seems worth exploring a similar procedure for designing control charts for the original correlated observations, and not, as here, for the one step ahead forecast errors.

References

Adams, B. and Tzeng, I. (1998). Robustness of forecast-based monitoring schemes. *Journal of Quality Technology*, 30:328–339.

Alonso, A., Pẽna, D., and Romo, J. (2002). Forecasting time series with sieve bootstrap. *Journal of Statistical Planning and Inference*, 100:1–11.

Alonso, A., Pẽna, D., and Romo, J. (2003). On sieve bootstrap prediction intervals. *Statistics & Probability Letters*, 65:13–20.

Alwan, L. and Roberts, H. (1988). Time-series modelling for statistical process control. *Journal of Business and Economic Statistics*, 6:87–95.

Apley, D. (2002). Time series control charts in the presence of model uncertainty. *Journal of Manifacturing Science and Engineering*, 124:891–898.

Apley, D. and Lee, H. (2003). Design of exponentially weighted moving average control charts for autocorrelated processes with model uncertainty. *Technometrics*, 45:187–198.

Apley, D. and Shi, J. (1999). The GLRT for statistical process control of autocorrelated processes. *IIE Transactions*, 31:1123–1134.

Basseville, M. and Nikiforov, I. (1993). *Detection of Abrupt Changes, Theory and Applications*. Prentice Hall, Englewood Cliffs, N.J.

Boyles, R. (2000). Phase I analysis for autocorrelated processes. *Journal of Quality Technology*, 32:395–409.

Brockwell, P. J. and Davies, R. A. (1996). *Time Series: Theory and Methods*. Springer, New York, 2nd edition.

Bühlmann, P. (1998a). Sieve bootstrap for smoothing in non-stationary time series. *Annals of Statistics*, 26:48–83.

Bühlmann, P. (1998b). Sieve bootstrap for time series. *Bernoulli*, 3:123–148.

Bühlmann, P. (2002). Bootstraps for time series. *Statistical Science*, 17:52–72.

Burnham, K. and Anderson, D. (2002). *Model Selection and Multimodel Inference: A Practical-Theoretic Approach*. Springer, New York.

Harris, T. and Ross, W. (1991). Statistical process control procedures for correlated observations. *Canadian Journal of Chemical Engineering*, 69:48–57.

Hurvich, C. and Tsai, C. (1989). Regression and time series model selection in small samples. *Biometrika*, 76:297–307.

Jensen, W., Jones-Farmer, L., Champ, C., and Woodall, W. (2006). Effects of parameter estimation on control charts properties: a literature review. *Journal of Quality Technology*, 38:349–364.

Jones, L. (2002). The statistical design of EWMA control charts with estimated parameters. *Journal of Quality Technology*, 34:277–288.

Jones, L., Champ, C., and Rigdon, S. (2001). The performance of exponentially weighted moving average charts with estimated parameters. *Technometrics*, 42:156–167.

Kramer, H. and Schmid, W. (2000). The influence of parameter estimation on the ARL of Shewhart type charts for time series. *Statistical Papers*, 41:173–196.

Kushner, H. and Yin, G. (2003). *Stochastic Approximation and Recursive Algorithms and Applications*. Springer, New York.

Lai, T. (2001). Sequential analysis: Some classical problems and new challenges. *Statistica Sinica*, 11:303–408.

Lu, C. and Reynolds, M. (1999a). Control charts for monitoring the mean and variance of autocorrelated processes. *Journal of Quality Technology*, 31:259–274.

Lu, C. and Reynolds, M. (1999b). EWMA control charts for monitoring the mean of autocorrelated processes. *Journal of Quality Technology*, 31:166–188.

Montgomery, D. (2004). *Introduction to Statistical Quality Control*. Wiley, New York, 5th edition.

Montgomery, D. and Mastrangelo, C. (1991). Some statistical process control methods for autocorrelated data (with discussion). *Journal of Quality Technology*, 23:179–274.

Polyak, B. (1990). New method of stochastic approximation type. *Automatic Remote Control*, 51:937–946.

Polyak, B. and Juditsky, A. (1992). Accelaration of stochastic approximation by averaging. *SIAM Journal of Control Optmization*, 30:838–855.

Runger, G. (2002). Assignable causes and autocorrelation: Control charts for observations or residuals? *Journal of Quality Technology*, 34:165–170.

Runger, G. and Willemain, T. (1995). Model-based and model-free control of autocorrelated processes. *Journal of Quality Technology*, 27:283–292.

Runger, G., Willemain, T., and Prabhu, S. (1995). Average run lengths for CUSUM control charts applied to residuals. *Communications in Statistics-Theory and Methods*, 24:273–282.

Ruppert, D. (1988). Efficient estimators from a slowly convergent Robbins-Monro process. Technical Report 781, School of Operations Research and Industrial Engineering, Cornell University, Itacha, New York.

Ruppert, D. (1991). Stochastic approximation. In Ghosh, B. and Sen, P., editors, *Handbook of Sequential Analysis*, pages 503–529. Marcel Dekker, New York.

Schimd, W. and Schöne, A. (1997). Some properties of the EWMA control chart in the presence of autocorrelation. *The Annals of Statistics*, 25:1277–1283.

Shibata, R. (1980). Asymptotically efficient selection of the order of the model for estimating parameters of a linear model. *The Annals of Statistics*, 8:147–164.

Shu, L., Apley, D., and Tsung, F. (2002). Autocorrelated process monitoring using triggered cuscore charts. *Quality and Reliability Engineering International*, 18:411–421.

Siegmund, D. and Venkatraman, E. (1995). Using the generalized likelihood ratio statistic for sequential detection of a change-point. *Annals of Statistics*, 23:255–271.

Superville, C. and Adams, B. (1994). An evaluation of forecast-based quality control schemes. *Communication in Statistics: Simulation and Computation*, 23:645–661.

Testik, M. (2005). Model inadequacy and residual control charts for autocorrelated processes. *Quality and Reliability Engineering International*, 21:115–130.

Vasilopoulos, A. and Stamboulis, A. (1978). Modification of control chart limits in the presence of data correlation. *Journal of Quality Technology*, 10:20–30.

Wardell, D., Moskowitz, H., and Plante, R. (1994). Run length distributions of special cause control charts for correlated processes. *Technometrics*, 36:3–27.

Willsky, A. and Jones, H. (1976). A generalized likelihood ratio approach to detection and estimation of jumps in linear systems. *IEEE Transactions of Automatic Control*, 21:108–112.

Yashchin, E. (1993). Performance of CUSUM control schemes for serially correlated observations. *Technometrics*, 35:35–52.

Acknowledgements

This research was partially funded by Italian MIUR-Cofin 2004 and 2006 grants.