Setup error adjustment: sensitivity analysis and a new MCMC control rule

Z. Lian[1], B.M. Colosimo[2] and E. del Castillo[1*]

[1]Department of Industrial & Manufacturing Engineering
The Pennsylvania State University, University Park, PA 16802, USA

[2] Dipartimento di Meccanica - Politecnico di Milano
Piazza Leonardo da Vinci 32, 20133 Milano, Italy

Summary

In this paper, we focus on performance of adjustment rules for a machine that produces items in batches and that can experience errors at each setup operation performed before machining a batch. The adjustment rule is applied to compensate for the setup offset in order to bring back the process to target. In particular, we deal with the case in which no prior information about the distribution of the offset or about the within-batch variability is available. Under such conditions, adjustment rules that can be applied are Grubbs' rules [1], the EWMA controller [2] and the MCMC adjustment rule, based on a Bayesian sequential estimation of unknown parameters that uses Markov Chain Monte Carlo (MCMC) simulation [3]. The performance metric of the different adjustment rules is the sum of the quadratic off-target costs over the set of batches machined. Given the number of batches and the batch size, different production scenarios (characterized by different values of the lot-to-lot and the within-lot variability and of the mean off-set over the set of batches) are considered. The MCMC adjustment rule is shown to have better performance in almost all the cases examined. Furthermore, a closer study of the cases in which the MCMC policy is not the best adjustment rule motivates a modified version of this rule which outperforms alternative adjustment policies in all the scenarios considered.

Key Words: Statistical Quality Control, Markov Chain Monte Carlo, Setup adjustment, Bayesian hierarchical models, Random effects model.

*corresponding author

1 Introduction

The quality of discrete-part manufactured products can significantly worsen when parts are processed off-target, i.e. the mean of the quality characteristic shifts to values different from its target. Such shifts frequently arise when parts are machined in batches because setup operations are often manually executed given the intrinsic complexity of these procedures (e.g. the fixturing of workpieces can be automated just for very simple part geometry). Even in cases when automatic setup operations are available, batches of the same part type are often affected by different initial offsets because conditions of the machine, materials, and operators vary from time to time. In this case, procedures designed to automatically compensate for the initial offset or setup error can significantly improve the outgoing quality of the parts processed. Assuming a quadratic off-target cost function, Grubbs [1] presented two adjustment rules aimed at adjusting for a potential initial offset present at setup of a single batch (what we will refer to as Grubbs' "harmonic" rule) and over a set of batches (what we will call Grubbs' "extended" rule), respectively [4]. The second rule, in particular, is optimal for a quadratic off-target cost function when parameters characterizing both within-batch and between-batches variability are known in advance. Del Castillo, Pan and Colosimo [5] show how Grubbs' extended rule has a Bayesian interpretation based on a Kalman filter when *prior* knowledge of parameters characterizing both batch-to-batch and within-batch distributions is required. Hence, Grubbs' extended rule can be applied just after a set of batches have been already machined in order to have accurate estimates of the required parameters. To start adjusting when no estimates are available, e.g., when a new product has to be processed or a new process is installed, simpler adjusting rules such as the harmonic rule and a discrete integral controller or EWMA controller [2], can perform better than the extended rule as shown by Del Castillo *et al.* [6].

Recently, a different adjustment rule was proposed [3] for situations in which no prior knowledge on parameters characterizing offset and process distributions is available. This rule is based on computing sequential Bayesian estimates of the unknown parameters using Markov Chain Monte Carlo (MCMC) methods.

In this paper, an in-depth comparison of the performance of setup adjustments rules that can be applied when no parameter estimates are available, namely, Grubbs' harmonic rule, the EWMA controller, and the MCMC policy, is provided. Performance will be evaluated considering different production scenarios characterized by the following set of parameters: a) the number of lots, b) the number of parts in each lot, c) the variability within and between-batches and d) the presence of a systematic error affecting the mean offset over a set of lots. Not all parameters are equally uncertain in practice. In fact, the first two parameters, i.e., the number of lots and the lot size, can be assumed known even when a new product and/or a new process is considered. Given these two parameters, the better rule will be considered as the one which determines enhanced performance over all the alternative scenarios associated with a set of possible values of the remaining unknown parameters (i.e., the within-batch and between-batches variability and the mean offset over a set of lots).

It will be shown that the MCMC-based approach outperforms the other rules in almost all the cases. A closer study of situations in which performance of the MCMC policy is actually worse has motivated a modified version of this rule, which, as we will show, has an enhanced performance in all the cases examined.

Rules for adjusting an initial offset over a set of batches when parameters are unknown

When discrete parts are processed in batches and no adjustment rule is adopted, a common model used in the statistical literature to describe the quality characteristic observed after each part is machined is the random effects model [7], given by:

$$\begin{gathered} Y_{ij} = \theta_i + v_{ij} \\ \theta_i \sim N(\mu, \sigma_\theta^2) \\ v_{ij} \sim N(0, \sigma_v^2) \quad ; \end{gathered} \tag{1}$$

where:

- $i = 1, ..., I$ is the batch index;
- $j = 1, ..., J$ is the part index;
- Y_{ij} is the deviation from the nominal value for the quality characteristic observed at the j^{th} part of the i^{th} batch;
- θ_i is the (unknown) mean in each batch, here representing the initial offset due to setup errors. This is assumed to be normally distributed with mean μ and variance σ_θ^2;
- v_{ij} represents the random error due to the combined effect of the intrinsic variability in the machining process and the variability in the measurement system. It is assumed to be normally distributed with mean equal to zero and variance σ_v^2.

The mean offset over a set of batches μ can be interpreted as the systematic error in the set-up operations, while the two variance components σ_v^2 and σ_θ^2 represent the within and between-batches variability, respectively.

To improve the quality of parts processed, adjustment rules should be adopted to eliminate the initial offset of parts processed in each lot by means of a compensatory variable U_{ij} $(i = 1, ..., I; j = 1, ..., J)$. When the adjustment rule is applied, the mean of the quality characteristic of the j^{th} part processed in the i^{th} lot is given by:

$$\theta_{ij} = \theta_{ij-1} + \nabla U_{ij-1} \tag{2}$$

where $\theta_{i0} = \theta_i$, $\nabla U_{ij-1} = U_{ij-1} - U_{ij-2}$ is the magnitude of the adjustment, and the quality characteristic can be thus expressed as:

$$Y_{ij} = \theta_{ij} + v_{ij} \quad . \tag{3}$$

In view of equation (2), the "ideal" adjustment would be given by $\nabla U_{ij-1} = -\theta_{ij-1}$, but the actual adjustment should be based on an estimate of θ_{ij-1} since it is unknown. Solving recursively equation (2) and considering that $\nabla U_{ij} = U_{ij} - U_{ij-1}$, the quality characteristic at the j^{th} part in the i^{th} lot can be rewritten as:

$$Y_{ij} = \theta_i + U_{ij-1} + v_{ij} \tag{4}$$

and the adjustment problem can be reformulated as selecting U_{ij-1} $(j = 2, .., J)$ to be closer as possible to $-\theta_i$.

If parameters μ, σ_θ, σ_v characterizing the distribution of the quality characteristic Y_{ij} (equation 1) are unknown, traditional adjustment rules that can be applied in this case are:

- Grubbs' harmonic rule [1] where $\nabla U_{ij} = -\frac{1}{j} Y_{ij}$ $(U_{i0} = 0 \ \ \forall i)$. We note that Grubbs extended rule, characterized by $\nabla U_{ij} = Y_{ij}/(j + \sigma_v^2/\sigma_\theta^2)$, is not applicable in practice as it requires knowledge of $\sigma_v^2/\sigma_\theta^2$;
- the integral or EWMA controller [2], where $\nabla U_{ij} = -\lambda Y_{ij}$ $(U_{i0} = 0 \ \ \forall i)$ and λ is some weight parameter (the integral constant) that needs to be chosen a priori;
- the sequential Bayesian adjustment rule [3], in which required parameters are sequentially predicted or estimated using Markov Chain Monte Carlo (MCMC) as detailed in the Appendix. In particular:
  1. before processing a new lot i (where $i \geq 3$ because at least two lots must be processed to estimate the variance between lots σ_θ^2), the initial set-point U_{i0} that has to be set on the machine before processing lot i is based on the predictive distribution of the offset;
  2. after observing at least one part in a lot, adjustments ∇U_{ij} $(j = 1, .., J)$ are based on the posterior distribution of the offset in the current lot, as seen from equation (4).

The Appendix describes the hierarchical normal means model used to estimate/predict required parameters and details on computation performed using the Gibbs Sampler method. Although the Gibbs sampler method is not the only approach available to compute the Bayesian analysis required, it has been selected because of its simplicity and flexibility. As Gelfand *et al.* [8] pointed out "the efficiency of other approach is at the expense of detailed sophisticated applied analysis [7] or tailored "one-off" numerical tricks or sophisticated adaptive quadrature methodology, in conjuction with subtle sensitivity of parameterization issues". Furthermore, Gibbs sampler approach can be easily extended to deal with a wide range of problems (i.e., unbalanced data, non homogeneous variances, missing data) that, although not considered in the paper, represent further directions of future research in this area.

Evaluation of different adjustment rules with reference to different performance indexes

Colosimo, Pan and del Castillo [3] compare the performance of the three adjustment rules mentioned above under the usual quadratic cost function, given by:

$$C = \sum_{i=1}^{I} \sum_{j=1}^{J} Y_{ij}^2. \tag{5}$$

Two classical variance components examples presented in [7] were used for evaluation purposes. Percentage advantages in quadratic cost determined by the MCMC Bayesian procedure over Grubbs' rule and EWMA controllers varied from 27 % to 63 % for the first example and from 18 % to 50 % for the second one [3].

To deepen comparison between adjustment rules applicable when no knowledge of parameters is *a priori* available, we refer to a manufacturing process where parts are processed in lots. To represent a wide range of production situations, we considered as initial reference the second example presented in [7], further studied in [8] (characterized by $I = 6$, $J = 5$, $\sigma_v = 4$, $\mu = 4$, $\sigma_\theta = 2$) and we perturbed all the parameters to generate different scenarios, as reported in Table I. The number of lots (I) was considered fixed and equal to 20, since performance for a number of lots $I < 20$ can be computed from partial results of the complete 20-lot simulations. For the other four parameters affecting the performance of the rules, the set of cases studied in this paper was derived considering all the possible combinations of the parameters at two values, as reported in Table II. Three replications were conducted for each scenario, thus the total number of simulations in this analysis was equal to $2^4 * 3 = 48$ runs.

insert Tables I and II about here

For each scenario, percentage savings in quadratic costs (equation 5) obtained using the MCMC-based approach instead of Grubbs' harmonic rule and two EWMA controllers (with $\lambda = 0.4$ and $\lambda = 0.1$) were computed. Nonparametric statistical analysis was used to analyze results because of the lack of normality of percentage savings obtained in simulated scenarios. Figure 1 reports the interquartile (IQ) and the confidence interval (CI) boxplots of the percentage savings induced by the MCMC-based adjustment rules over competitor rules. In particular, each CI boxplot is plotted inside the IQ one, and refers to the 95% confidence interval on the median. IQ and CI boxplots are reported in Figure 1 as a function of: the number of lots (I); the number of parts in lots (J); the ratio of the mean offset to the within-lot standard deviation (μ/σ_v); the ratio of the between to the within-lot standard deviation (σ_θ/σ_v).

insert Figure 1 about here

As it can be observed, the MCMC approach outperforms the competitor adjustment rules in almost all the scenarios, inducing sometimes significant advantages. In particular, actual applications of the adjustment rules should be based on the behavior of the percentage savings with respect to the number of lots (I) and the number of parts in each lot (J). In fact, when a new product is considered or a new process is introduced I and J are the only parameters that can be assumed known. Table III reports the 95% confidence interval on the median of the percentage savings induced by the MCMC policy over the competitor rules, as a function of the number of lots I and the number of parts in each lot J. These confidence intervals represent possible advantages induced by the MCMC policy if all the scenarios simulated for the other set of parameters (μ, σ_θ and σ_v) can be considered *a priori* equally likely. We note that the lower confidence intervals on the percentage savings induced by the MCMC policy is almost always greater than zero, except for cases in which the number of lots processed (I) and the lot size (J) are small. Thus, for example, when $I = 5$ there is no significant advantage in adopting the MCMC rule instead of Grubbs' rule, because the 95% confidence interval on the median of percentage savings is given by (-1.9% , 5.8%) and contains 0%. A similar comment applies to the case when the lot size is small, i.e. $J = 5$, where there is no significant advantage compared to Grubbs' rule and the EWMA controller with $\lambda = 0.4$.

insert Table III about here

To better investigate the performance of the different methods, the quadratic cost function reported in equation (5) was expanded into its different components. The total sum of squares criterion contains the part-to-part errors v_{ij} that are not controllable. Therefore, a more informative evaluation of the performance of any adjustment rule is to discount the variability induced by the v_{ij}. This alternative criterion is evidently more preferable the large σ_v^2 is. To do this, we partitioned the total quadratic cost function as follows:

$$\begin{aligned} C &= \sum_{i=1}^{I}\sum_{j=1}^{J} Y_{ij}^2 = \sum_{i=1}^{I}\sum_{j=1}^{J}(\theta_i + U_{ij} + v_{ij})^2 = \\ &= \sum_{i=1}^{I}\sum_{j=1}^{J}(\theta_i + U_{ij})^2 + \sum_{i=1}^{I}\sum_{j=1}^{J} 2(\theta_i + U_{ij})v_{ij} + \sum_{i=1}^{I}\sum_{j=1}^{J} v_{ij}^2 \end{aligned} \tag{6}$$

Since $v_{ij} \sim \text{NID}(0, \sigma_v^2)$ and is independent of $(\theta_i + U_{ij})$, the second term in equation (6) $(\sum_i \sum_j 2(\theta_i + U_{ij})v_{ij})$ will be close to zero for relative large I and J. The third term, which is independent of U_{ij}, is not controllable. The first term $\sum_i \sum_j (\theta_i + U_{ij})^2$, the only part left in the total cost function that could be improved by an adjustment rule, will be refereed to in what follows as the *quadratic bias* cost function.

The quadratic bias directly measures how well U_{ij} converges to $-\theta_i$, where the gap between U_{ij} and $-\theta_i$ is the bias. Notice that, for an ideal adjustment rule without any bias, i.e. when U_{ij} always equals $-\theta_i$ (a situation impossible to achieve in practice), the corresponding saving rates would be 100%.

Considering this new performance index, IQ and CI boxplots of percentage savings induced by the MCMC policy are shown in Figure 2. Focusing just on parameters known before processing, Table IV shows the 95% confidence interval on the median of the percentage savings induced by the MCMC policy as a function of the number of lots I and the size of the lots J. Although basic considerations drawn for the quadratic cost function hold in this case too, the adoption of the quadratic bias cost function better outlines the advantages and disadvantages implied by the use of the MCMC policy. In particular, considering just the parameters known before processing lots (I and J), the new performance index confirms that when few lots are processed (i.e., $I = 5$) there is no statistical evidence to assess that the MCMC policy performs better than Grubbs' rule. Similarly, when lot size is small (i.e., $J = 5$) the performance of the MCMC policy is equivalent to that of an EWMA controller with $\lambda = 0.4$ (note that this conclusion was reached also using the quadratic cost function, but using such function the performance of Grubbs' rule was also found to be equivalent). However, the relative performance of the EWMA controller is very sensible to a proper selection of the weight λ. In fact, when $\lambda = 0.1$ in the EWMA controller, savings induced by the MCMC policy are significantly greater than zero even for a small lot size ($J = 5$) since in this case the median of the percentage savings is equal to 45.7% and the 95% confidence interval is given by (37.5% , 55.8%). It is important to emphasize that in practice, there is no way to determine λ appropriately in an EWMA controller if the process parameters are unknown.

insert Figure 2 about here

insert Table IV about here

Motivation for a modified rule

As shown in the previous section, the MCMC method can result in significant savings over the other three adjustment methods in most, but not all, cases. To better explore situations in which the MCMC adjustment policy induces worst performance, a particular study was conducted to better investigate motivation for this behavior. The two cases where the MCMC policy results in worse performance compared to competitor rules were studied more in detail. Table V reports the mean of the percentage savings implied by using the MCMC policy for the three simulations in each scenario. From it, cases 16 ($J = 5$; $\mu = 0$, $\sigma_v = 2$ $\sigma_\theta = 4$) and 9 ($J = 5$; $\mu = 4$, $\sigma_v = 4$ $\sigma_\theta = 2$) were considered as representative of the worst and the best performance of the MCMC policy. Behavior of adjustment rules in one simulation from each of these cases was specifically analyzed.

insert Table V about here

Figure 3 shows the sum of quadratic costs induced over the set of lots I by all the competitor rules as a function of the number of parts in the lot, i.e., $\sum_{i=1}^{20} Y_{ij}^2$ for $j = 1, 2, ..., 5$. In particular, graphs (a) and (b) in Figure 3 refer to simulation of case 16 (worst performance case) and 9 (best performance case), respectively. As it can be observed in Figure 3 (a), the worst performance of the MCMC policy is mainly due to the first part ($j = 1$) processed in the lot. For parts other than the first one (i.e., $j = 2, ..., 5$), costs of the MCMC approach are almost identical to the one obtained with Grubbs' rule, which is the best approach in this case. On the contrary, Figure 3(b) shows that the MCMC policy outperforms the other rule starting from the first part ($j = 1$) processed in each lot.

insert Figure 3 about here

This behavior could be better explained considering the two components of the MCMC policy:

1. the initial setpoint U_{i0} which is computed starting at the third lot ($i \geq 3$) and is based on the posterior predictive distribution of a future offset θ_i' given all the data collected from previously produced lots (equation 12 in the Appendix);

2. the adjustments ∇U_{ij} for $j = 1, .., J$ based on the posterior distribution of the offset θ_i and which affect the quality characteristic of parts other than the first (equations 11 and 14 in the Appendix).

Figure 3 clearly shows how the first component of the MCMC policy, the initial setpoint (and initial adjustment since $\nabla U_{i0} = U_{i0}$) does not perform well for some cases. From equation (12) in the Appendix, the initial setpoint U_{i0} set on the machine before processing the first part in lot i, is given by:

$$U_{i0} = -E(\theta_i' | \boldsymbol{x^{i-1\ J}}, M_1) \quad \text{for} \quad i = 3, ..., I \quad , \tag{7}$$

where $\boldsymbol{x^{i-1\ J}}$ represents information (data and adjustments) available at this point in time and M_1 refers to the hierarchical normal means model used in the MCMC simulation (see the Appendix for further details). Figure 4 shows a plot of the estimated $\bar{\theta'}_i$ and of the accuracy of this estimate, measured by the interval $\bar{\theta'}_i \pm \hat{\sigma}_{\theta_i'}$, as a function of lots processed. Both the sample mean $\bar{\theta'}_i$ and the sample standard deviation $\hat{\sigma}_{\theta_i'}$ were obtained from the MCMC simulation.

insert Figure 4 about here

As it can be seen, the sample mean converges to the true value of μ (equal to 0 and 4 for cases represented in Figure 4 (a) and (b), respectively) as the number of lots processed increases. Thus, the initial set-point U_{i0} tends to be set to the proper value $-\mu$. That is, once the steady-state behavior is achieved, the MCMC policy allows to correct a possible systematic error before observing parts in a lot. However, there is some slowness in converging to the right value, as it is clear from Figure 4 (a). When $\mu = 0$, Grubbs' rule and the EWMA controllers perform better because they do not include the predictive feature in the control rule, i.e., $U_{i0} = 0$ is set by default, and this happens to be the optimal selection for the initial set-point U_{i0}. Figure 4 (a) and (b) show that also the accuracy of the estimate depends on the specific scenario examined. Hence, a new modified version of the MCMC adjustment rule should take into account the disadvantage of using the initial adjustment when there is no systematic error (i.e., when $\mu = 0$) and when the uncertainty in the predicted lot mean θ_i' is large. This is discussed next.

Conditional First Adjustment Rule based on MCMC Approach

It is clear from the preceding discussion that applying the first adjustment in each lot $\nabla U_{i0} = U_{i0}$ can lead to an advantage or a disadvantage when using the MCMC policy, depending on whether there is a systematic error ($\mu \neq 0$) or not ($\mu = 0$). However, the actual value of μ can not be known in advance, due to our assumption on unknown parameters. Therefore a different MCMC approach should be based on selecting "online" the best strategy for the U_{i0}'s. This can be done taking into account the accuracy of the estimate $\hat{\sigma}_{\theta_i'}$. Figure 4 shows that the width of the interval $\bar{\theta'}_i \pm \hat{\sigma}_{\theta_i'}$ reduces as the number of lots processed increases. Hence, problems related to the slow convergence of $\bar{\theta'}_i$ to μ can be overcome by including information on estimate accuracy in the adjustment rule.

A new "Conditional First Adjustment" MCMC method is thus proposed as follows:

1. At the beginning of each lot i (with $i \geq 3$), substitute equation (12) of the original approach with:

$$\begin{cases} \text{adjust} \quad (U_{i0} = -\bar{\theta}'_i) & \text{if } |\bar{\theta}'_i| > k\hat{\sigma}_{\theta'_i} \ ; \\ \text{do not adjust} \quad (U_{i0} = 0) & \text{otherwise.} \end{cases} \tag{8}$$

2. Use the original MCMC method (equations 11 and 14) for all the other adjustments.

Here, k is a tuning constant of the method that we shall discuss later and $\bar{\theta}'_i$ and $\hat{\sigma}_{\theta'_i}$ are respectively the mean and standard deviation of the posterior predictive distribution of the offset, conditional on the data observed before lot i. When the ratio $|\bar{\theta}'_i|/\hat{\sigma}_{\theta'_i}$ is relatively high, i.e., the estimate has a high precision compared with its magnitude, we have a strong belief of the estimate. This is similar to a test of significance for a normal population mean. While strictly speaking we should use a Bayesian factor for testing significance since the posterior distribution is not necessarily normal, the simpler conditional rule works very well, as will be shown below. Therefore, we decided to keep its simplicity instead of the rigor of Bayesian factors.

As $k \to 0$, the conditional first adjustment rule becomes the same as the original MCMC method. As k increases, the percentage savings in the cases where the MCMC performs well start to drop, while those in the cases with poor performances become higher. When $k \to \infty$, the conditional rule becomes the same as an MCMC rule that always omits the first adjustment in each lot.

Figure 5 shows how the maximum and minimum of the percentage savings among the 16 cases vary as k is changed in the Conditional First Adjustment rule. In particular, this Figure indicates that k should be chosen to have a proper trade-off between reduction of the maximum savings (i.e., the best performance of the MCMC policy in Table V) and improvement of the minimum savings (i.e., the worst performance of the MCMC policy in Table V). The figure indicates that the best performance deteriorates very little while the worst performance improves considerably for the first few values of k that were tried. From this it was concluded that the value of $k = 0.64$ showed the best performance and was used in what follows.

insert Figure 5 about here

Figures 6 and 7, report IQ and CI boxplots of the percentage savings obtained with the Conditional First Adjustment MCMC rule ($k = 0.64$) using the quadratic and the quadratic bias cost functions, respectively. Compared with previous boxplots obtained adopting the original MCMC approach (Figures 1 and 2, respectively), these graphs show that the new policy induces advantages over the competitor rules. In particular, when the quadratic cost function is used as performance index, the 95% CI boxplots obtained with the new rule shifts all above the line at 0% (Figure 6).

insert Figures 6 and 7 about here

Focusing only on the performance related to parameters that are always known (I and J), Tables VI and VII report the 95% confidence interval on the median of savings induced by the Conditional First Adjustment MCMC method with $k = 0.64$ (using the quadratic and quadratic bias costs, respectively). Comparing these two tables with Tables III and IV (obtained with the original MCMC method), we find that the new policy is preferable to other competitor rules for all the values of I and J, since the left confidence interval limits are now always greater than 0%. In terms of the medians of the percentage savings induced by this new rule, (Tables VI and VII), it should be noticed that the median of advantages ranges from 2.8% to 29.0% when using the quadratic cost function and from 9.6% to 57.6% when using the quadratic bias cost function. Therefore, we can conclude that the Conditional First Adjustment MCMC rule should be preferred to competitor rules when new products or new processes need to be introduced.

insert Tables VI and VII about here

Conclusions

In this paper, we dealt with the problem of adjusting initial offsets for quality characteristics of discrete-parts manufactured in batches. At the beginning of each batch, a set-up error can cause the mean of the quality characteristic to be off-target and the adjustment procedure is designed to compensate for this initial off-set as the number of parts processed in the lot increases. In particular, we focused on the case in which off-target costs are quadratic, adjustment costs can be neglected and no previous knowledge on parameters characterizing the off-set distribution and the process intrinsic variability is available. This situation can properly model adjustment rules designed for new products or newly installed processes, in which there is no previous experience on set-up operations. We compared different adjustment rules that can be applied in this case: Grubbs' rule, the integral or EWMA controller, and the adjustment rule based on a Bayesian sequential estimation of unknown parameters using MCMC simulation (MCMC adjustment rule). Considering different production scenarios, the last rule was shown to outperform other existing rules unless very few lots has to be machined or lots size is particularly small. A further study of performance of the MCMC adjustment rules in these last cases motivated a revised version of the rule, that has consistent advantages over Grubbs' rule and the EWMA controller in all the cases examined.

Acknowledgements

This research was partially funded by NSF grant DMI 0200056.

References

[1] Grubbs FE. An Optimum Procedure for Setting Machines or Adjusting Processes. *Industrial Quality Control.* July, 1954; reprinted in *Journal of Quality Technology* 1983 **15**, 4: 186–189.

[2] Box GEP, Luceño A. *Statistical Control by Monitoring and Feedback Adjustment*; John Wiley & Sons: New York, NY, 1997.

[3] Colosimo BM, Del Castillo E, Pan R. A Sequential Markov Chain Monte Carlo Approach to Setup Adjustment of a Process Over a Set of Lots. 2003. To appear in *Journal of Applied Statistics.* Technical Paper, Engineering Statistics Laboratory, Department of Industrial & Manufacturing Engineering, Penn State University. http://lb350e.ie.psu.edu/publications.htm.

[4] Trietsch D. The Harmonic Rule for Process Setup Adjustment with Quadratic Loss. *Journal of Quality Technology* 1998 **30**, 1: 75–84.

[5] Del Castillo E, Pan R, Colosimo BM. An Unifying View of Some Process Adjustment Methods. *Journal of Quality Technology* 2003 **35**, 3: 286–293.

[6] Del Castillo E, Pan R, Colosimo BM. Small sample comparison of some process adjustment methods. *Communications in Statistics, Simulation and Computation* 2003, **32**, 3: 923–942.

[7] Box GEP, Tiao GD. *Bayesian Inference in Statistical Analysis*, John Wiley and Sons, Inc, 1973.

[8] Gelfand AE, Hills SE, Racine-Poon A., Smith AFM. Illustration of Bayesian Inference in Normal Data Models Using Gibbs Sampling. *Journal of the American Statistical Association* 1990 **85**, 412: 972–985.

[9] Gelfand AE, Smith AFM. Sampling-based Approaches to Calculating Marginal Densities. *Journal of the American Statistical Association* 1990 **85** 410: 398–409.

[10] Spiegelhalter DJ, Thomas A, Best NG and Gilks WR. *BUGS: Bayesian inference Using Gibbs Sampling, version 0.30*, Cambridge: Medical Research Council Biostatistics Unit, 1994.

[11] Gelman A, Carlin JB, Stern HS, Rubin DB. *Bayesian Data Analysis*, Chapman & Hall, 1995.

[12] Brooks SP, Roberts GO. Assessing Convergence of Markov Chain Monte Carlo Algorithms. *Statistics and Computing* 1998 **8**: 319–335.

[13] Cowles MK, Carlin BP. Markov Chain Monte Carlo Convergence Diagnostics: A Comparative Review. *Journal of the American Statistical Association* 1995 **91**: 883–904.

[14] Gelman A. Inference and monitoring convergence, in W.R. Gilks, S. Richardson and D.J. Spiegelhalter (eds.), *Markov Chain Monte Carlo in Practice*, London, UK: Chapman & Hall, 1996; pp 131–140.

[15] Raftery AE, Lewis SM. Implementing MCMC, in W.R. Gilks, S. Richardson, and D.J. Spiegelhalter (eds.), *Markov Chain Monte Carlo in Practice*, London, UK: Chapman & Hall, 1996; pp 115–130.

[16] Gelman A, Rubin D. Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science* 1992 **7**: 457–511.

Appendix: A Bayesian adjustment rule based on MCMC simulation

When no prior knowledge on parameters is available, an adjustment rule can be based only on the history of data acquired which includes the response values and the previous adjustments. Under this assumption, a transformed variable X_{ij} computed at the time Y_{ij} is observed (since at that time U_{ij-1} is known) is completely defined as:

$$X_{ij} = Y_{ij} - U_{ij-1} = \theta_i + v \tag{9}$$

and allows to derive a one-way random effects model for the adjusted process. Adopting a Bayesian perspective, the one-way random effects model is a special case of a *hierarchical model*, where the first stage models the distribution of observable data conditionally on unknown parameters, the second stage in the hierarchy specifies the prior distribution of these unknown parameters which can depend on some hyperparameters, which in turn have a prior at a third stage. The three-stage hierarchical model (that we will call M_1) is hence given by:

$$\begin{array}{rcl} \textit{first stage: } X_{kp}|\theta_k,\sigma_v^2 & \sim & N(\theta_k,\sigma_v^2) \;\; (k=1,...,i \;\text{ and }\; p=1,...,J); \\ \textit{second stage: } \theta_k|\mu,\sigma_\theta^2 & \sim & N(\mu,\sigma_\theta^2) \;\; (k=1,...,i) \;\; , \\ \sigma_v^2|a_1,b_1 & \sim & IG(a_1,b_1) \;\; ; \\ \textit{third stage: } \mu|\mu_0,\sigma_0^2 & \sim & N(\mu_0,\sigma_0^2) \;\; , \\ \sigma_\theta^2|a_2,b_2 & \sim & IG(a_2,b_2) \;\; ; \end{array} \tag{10}$$

where conjugacy has been used at each step of the hierarchical model (a common choice for the random effects model [8], [9]) and where IG represents an Inverse-Gamma distribution. Parameters $\mu_0, \sigma_0^2, a_1, b_1, a_2, b_2$ are assumed known. In particular, they were selected according to values suggested in the literature [10] to model "vague" prior distributions, i.e., $\mu_0 = 0$, $\sigma_0^2 = 1.0E+10$, $a_1 = b_1 = a_2 = b_2 = 0.001$. Denoting with $\boldsymbol{x^{ij}} = \{x_{11}, x_{12}, ..., x_{1J}, ..., x_{i1}, ..., x_{ij}\}$ all (transformed) data observed after the j^{th} part in the i^{th} lot has been just machined, the adjustment $\nabla U_{ij} = U_{ij} - U_{ij-1}$ $(j > 0)$ can be computed using the posterior distribution $(\theta_i|\boldsymbol{x^{ij}}, M_1)$ when at least one part in the lot has been processed. The predictive distribution $\theta'_{i+1}|\boldsymbol{x^{iJ}}, M_1$ [11] can be used to select the initial set-point $U_{i+1\ 0}$ that has to be set on the machine *before* processing the first part in the next lot (lot $i+1$), i.e.:

$$U_{ij} = -E(\theta_i|\boldsymbol{x^{ij}}, M_1) \;\; i=2,...,I \;\text{ and }\; j=1,...,J, \tag{11}$$

$$U_{i+1\ 0} = -E(\theta'_{i+1}|\boldsymbol{x^{iJ}}, M_1) \;\;\text{ for }\; i=2,...,I \tag{12}$$

where the index i is set to consider that adjusting can be performed only if σ_θ^2 can be estimated (i.e., when at least two parts from different lots have been already processed) and $\theta'_{i+1}|\boldsymbol{x^{iJ}}, M_1$ is the posterior predictive distribution [11] of a future offset given all the data collected from previously produced lots.

Adjustments for parts in the first lot can be similarly derived considering a "reduced" two-stage hierarchical model M_2 given by:

$$\begin{array}{rcl} \textit{first stage: } X_{1j}|\theta_1,\sigma_v^2 & \sim & N(\theta_1,\sigma_v^2) \;\; , \\ \textit{second stage: } \theta_1|\mu,\sigma_\theta^2 & \sim & N(\mu,\sigma_\theta^2) \\ \sigma_v^2 & \sim & IG(a_1,b_1) \;\; . \end{array} \tag{13}$$

Since at least two parts have to be processed to estimate the within-lot variance σ_v^2, the adjustments ∇U_{1j} for $j = 2, .., J$ can be computed using the hierarchical model M_2 while a trivial estimator (as in Grubbs' harmonic rule) can be used to start adjusting just after the first part in the first lot has been machined , i.e.:

$$\begin{array}{rcl} U_{11} & = & -x_{11} \;\; , \\ U_{1j} & = & -E(\theta_1|\boldsymbol{x^{1j}}, M_2) \;\; j=2,3,...,J \;\; . \end{array} \tag{14}$$

As each observation becomes available, a Gibbs Sampler is run to estimate the parameters of the hierarchical normal means model given the observations up to that point in time. The current lot mean estimate is then used for adjustment. The MCMC simulation coded in the *Bugs* (***B****ayesian inference* ***U****sing* ***G****ibbs* ***S****ampling*, [10] language was used to perform Gibbs Sampling. Following the literature on convergence diagnostic ([12], [13], [14]) both the algorithms of Raftery and Lewis [15] and Gelman and Rubin [16] were used within the MCMC simulation for assessing convergence of the chains (for details on the integration of software packages used, the interested readers can refer to [3].

Tables and Illustrations

parameters	Value in example 2 by Box and Tiao ([7] p. 247)	Perturbed value
I	6	20
J	5	20
μ	4	0
σ_θ	2	4
σ_v	4	2

Table I: Parameters characterizing different scenarios examined.

Case	J	σ_v	μ	σ_θ	Case	J	σ_v	μ	σ_θ	Case	J	σ_v	μ	σ_θ	Case	J	σ_v	μ	σ_θ
1	20	4	4	2	5	20	2	4	2	9	5	4	4	2	13	5	2	4	2
2	20	4	4	4	6	20	2	4	4	10	5	4	4	4	14	5	2	4	4
3	20	4	0	2	7	20	2	0	2	11	5	4	0	2	15	5	2	0	2
4	20	4	0	4	8	20	2	0	4	12	5	4	0	4	16	5	2	0	4

Table II: Set of parameters characterizing cases studied in this paper.

		Grubbs			EWMA (0.1)			EWMA (0.4)		
		median	lower CI 95%	upper CI 95%	median	lower CI 95%	upper CI 95%	median	lower CI 95%	upper CI 95%
I	5	2.7%	-1.9%	5.8%	18.1%	5.4%	29.3%	9.0%	3.9%	11.4%
	10	3.7%	0.4%	6.7%	20.6%	13.9%	32.8%	9.7%	6.7%	13.5%
	20	6.8%	3.6%	9.5%	24.4%	18.3%	33.8%	10.7%	7.4%	14.0%
J	5	5.1%	-1.6%	11.1%	28.5%	20.1%	34.0%	-0.4%	-3.8%	8.8%
	20	3.7%	2.6%	6.5%	16.7%	9.1%	25.5%	11.7%	9.7%	13.7%

Table III: Confidence intervals (at level 95%) on the median of the percentage savings in total quadratic costs induced by the MCMC method compared to each alternative adjustment rule.

		Grubbs			EWMA (0.1)			EWMA (0.4)		
		median	lower CI 95%	upper CI 95%	median	lower CI 95%	upper CI 95%	median	lower CI 95%	upper CI 95%
I	5	7.3%	-2.3%	21.6%	40.4%	13.1%	55.6%	24.2%	9.5%	35.3%
	10	15.2%	8.2%	28.3%	47.8%	32.7%	65.3%	38.3%	22.1%	46.6%
	20	25.4%	16.0%	36.4	58.3%	38.8%	68.2%	39.3%	28.3%	48.4%
J	5	14.0%	0.3%	27.0%	45.7%	37.5%	55.8%	2.0%	-4.7%	17.9%
	20	20.2%	11.1%	27.9%	55.5%	31.6%	65.9%	41.6%	38.8%	49.3%

Table IV: Confidence intervals (at level 95%) on the median of the percentage savings in quadratic bias costs induced by the MCMC method.

					5 lots			10 lots			20 lots		
Case	J	σ_v	μ	σ_θ	Grubbs	EWMA (0.1)	EWMA (0.4)	Grubbs	EWMA (0.1)	EWMA (0.4)	Grubbs	EWMA (0.1)	EWMA (0.4)
1	20	4	4	2	5%	12%	12%	9%	20%	17%	11%	20%	15%
2	20	4	4	4	-2%	22%	4%	3%	20%	9%	4%	20%	11%
3	20	4	0	2	7%	-1%	16%	4%	-2%	13%	7%	-1%	13%
4	20	4	0	4	-1%	1%	5%	1%	5%	6%	2%	5%	7%
5	20	2	4	2	6%	44%	21%	9%	46%	23%	11%	47%	23%
6	20	2	4	4	5%	36%	14%	9%	47%	19%	13%	51%	23%
7	20	2	0	2	3%	-4%	9%	0%	8%	8%	2%	10%	9%
8	20	2	0	4	-4%	13%	2%	-2%	29%	8%	0%	30%	9%
9	5	4	4	2	24%	34%	22%	26%	36%	25%	27%	38%	25%
10	5	4	4	4	-16%	-2%	-24%	-1%	18%	-3%	6%	24%	4%
11	5	4	0	2	18%	-20%	-6%	18%	-11%	-3%	15%	-7%	0%
12	5	4	0	4	-11%	19%	-5%	-5%	23%	-1%	0%	23%	-1%
13	5	2	4	2	27%	44%	21%	26%	50%	26%	33%	59%	36%
14	5	2	4	4	0%	43%	8%	-2%	37%	5%	8%	55%	25%
15	5	2	0	2	-12%	8%	-14%	-9%	16%	-11%	0%	14%	-5%
16	5	2	0	4	-24%	36%	-5%	-20%	24%	-14%	-11%	29%	-6%

Table V: Mean percentage savings in total quadratic costs induced by the MCMC method in simulated scenarios.

		Grubbs			EWMA (0.1)			EWMA (0.4)		
		median	lower CI 95%	upper CI 95%	median	lower CI 95%	upper CI 95%	median	lower CI 95%	upper CI 95%
I	5	2.8%	1.5%	5.9%	22.8%	13.3%	28.8%	9.4%	6.6%	11.2%
	10	3.8%	1.7%	6.6%	24.4%	16.9%	33.8%	10.6%	7.3%	13.8%
	20	5.8%	2.7%	8.5%	26.5%	19.7%	32.0%	10.5%	7.5%	13.8%
J	5	4.2%	2.0%	8.4%	29.0%	24.2%	32.5%	5.7%	2.1%	8.5%
	20	4.1%	2.6%	6.0%	18.3%	10.7%	26.1%	12.1%	10.7%	14.4%

Table VI: Confidence intervals (at level 95%) on the median of the percentage savings in total quadratic costs induced by the Conditional First Adjustment MCMC method.

		Grubbs			EWMA (0.1)			EWMA (0.4)		
		median	lower CI 95%	upper CI 95%	median	lower CI 95%	upper CI 95%	median	lower CI 95%	upper CI 95%
I	5	9.6%	2.6%	21.2%	44.2%	27.4%	57.0%	27.1%	15.4%	32.6%
	10	19.1%	8.3%	26.5%	55.5%	37.7%	64.7%	35.3%	27.9%	40.9%
	20	21.6%	17.6%	33.2%	57.3%	38.6%	64.9%	37.5%	24.5%	46.5%
J	5	18.7%	6.9%	24.0%	50.4%	38.9%	57.6%	10.9%	0.1%	18.2%
	20	19.4%	10.3%	25.9%	57.6%	37.3%	65.5%	41.3%	37.2%	48.7%

Table VII: Confidence intervals (at level 95%) on the median of the percentage savings in quadratic bias costs induced by the Conditional First Adjustment MCMC method.

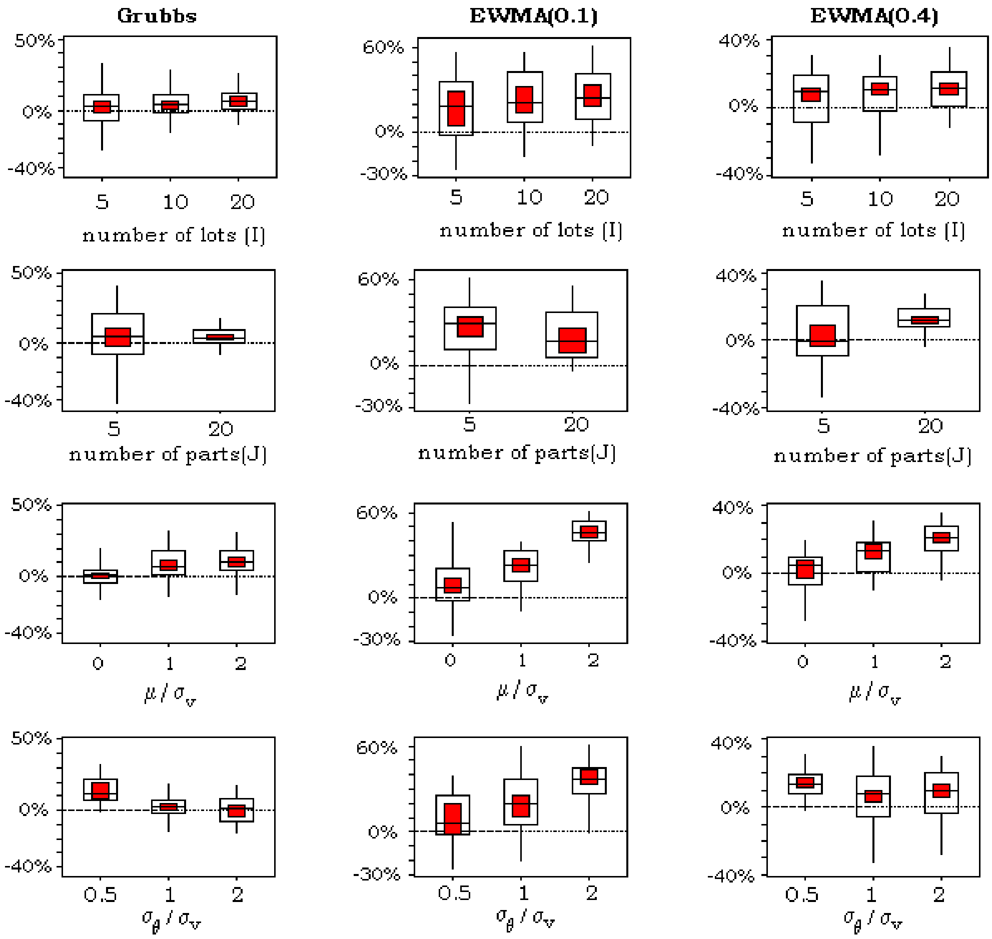

Figure 1: Boxplots of the percentage savings (using the quadratic cost function) induced by the MCMC adjustment rule over Grubbs' harmonic rule and the two EWMA controllers ($\lambda = 0.1$ and $\lambda = 0.4$) as a function of different parameters: the number of lots (I); the number of parts in lots (J); the ratio of the mean offset to the within-lot standard deviation (μ/σ_v); the ratio of the between to the within lot standard deviation (σ_θ/σ_v).

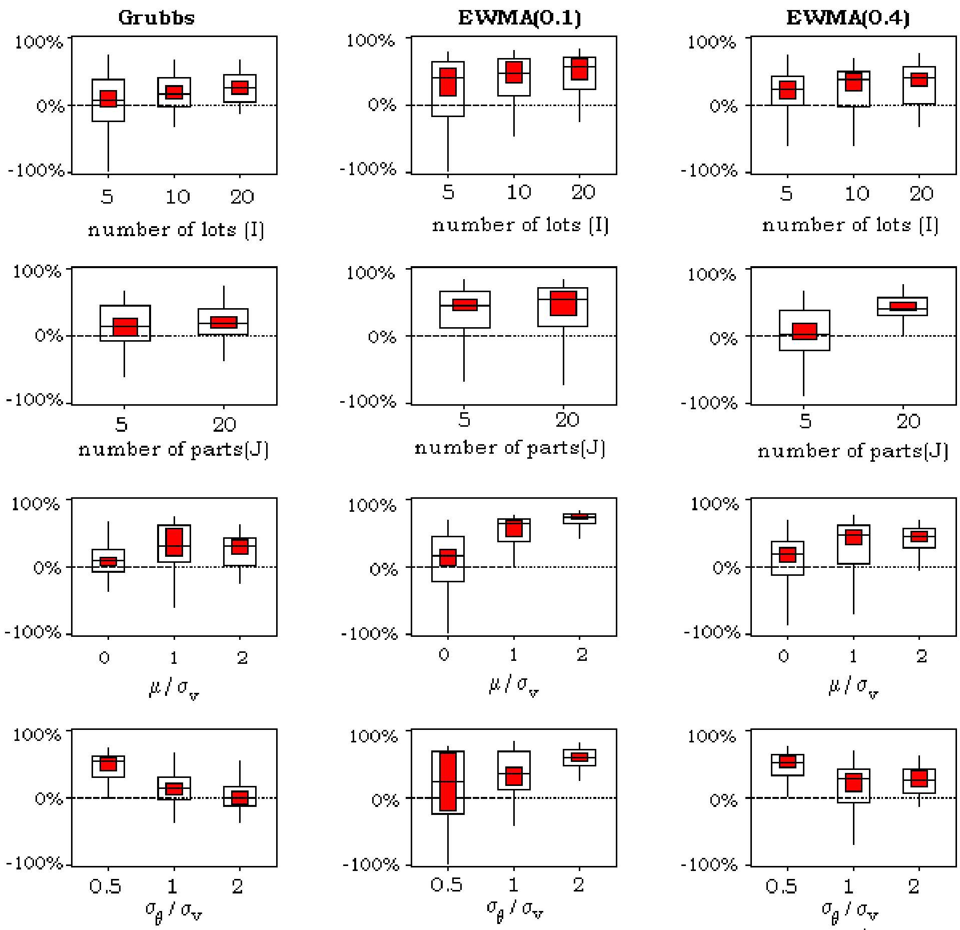

Figure 2: Boxplots on the percentage savings (using the quadratic bias cost function) induced by the MCMC adjustment rule over Grubbs' harmonic rule and the two EWMA controllers ($\lambda = 0.1$ and $\lambda = 0.4$) as a function of different parameters: the number of lots (I), the number of parts in lots (J), the ratio of the mean offset to the within-lot standard deviation (μ/σ_v), and the ratio of the between to the within-lot standard deviation (σ_θ/σ_v).

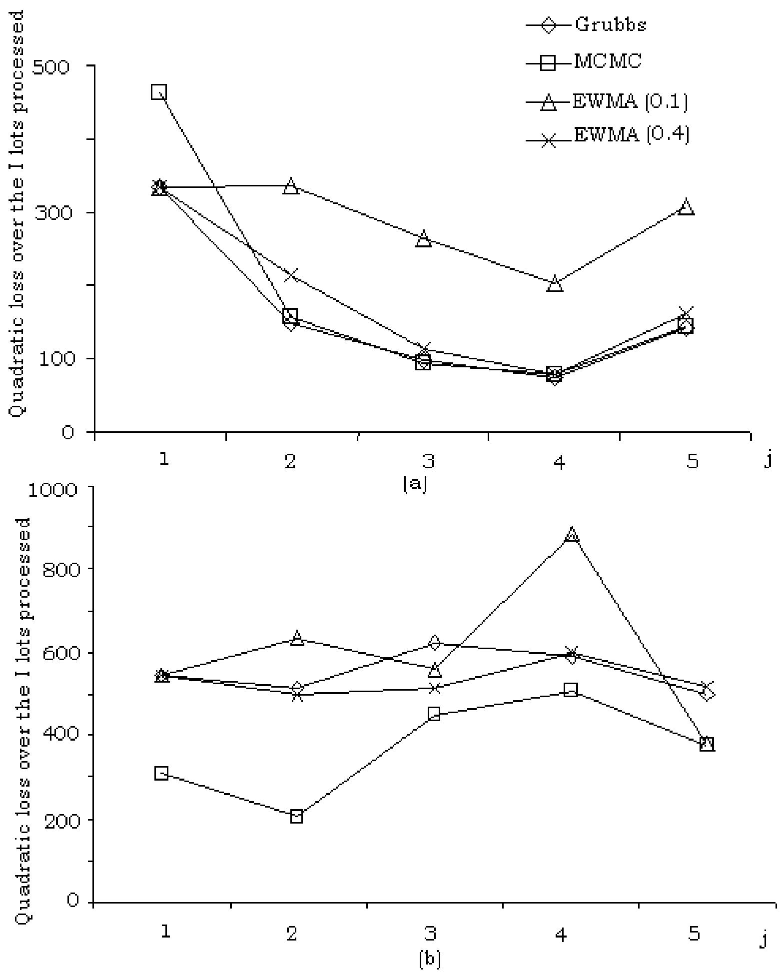

Figure 3: Quadratic costs in two simulation runs (a: case 16 and b: case 9) as a function of the number of parts processed in the lot.

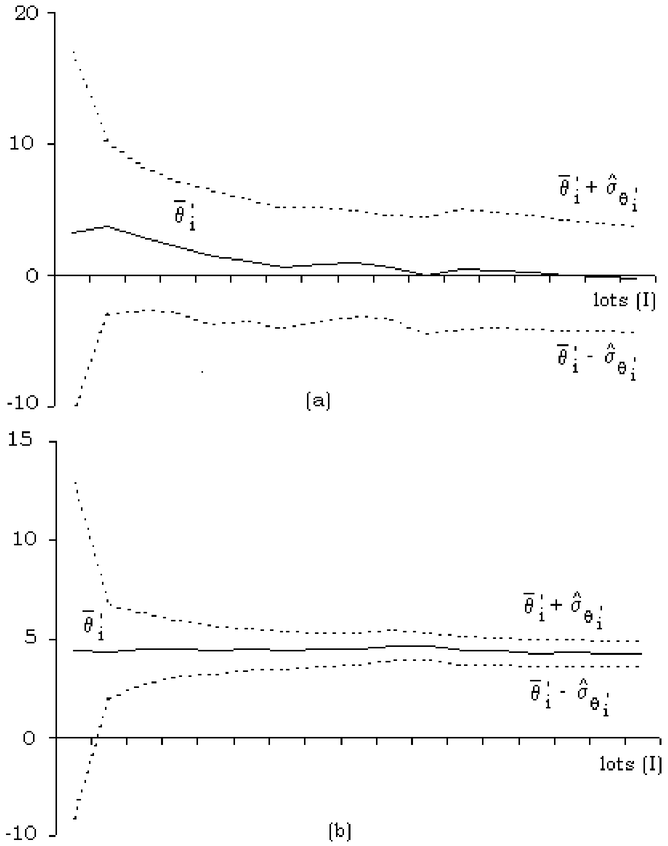

Figure 4: Plot of the interval $\bar{\theta}'_i \pm \hat{\sigma}_{\theta'_i}$, as a function of lots processed.

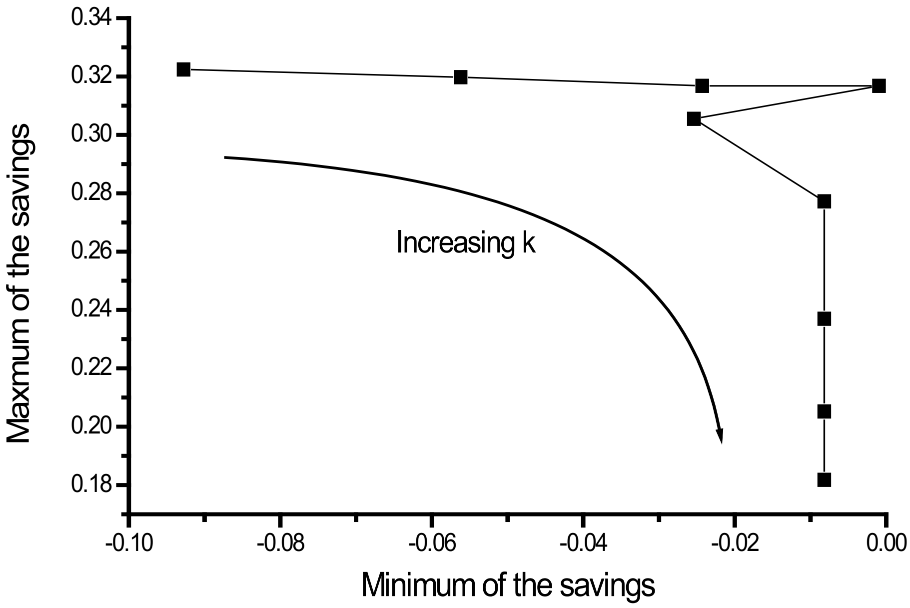

Figure 5: The maximum savings vs. the minimum savings among the 16 cases on the Tables as a function of the adjustment limit k when applying the Conditional First Adjustment MCMC rule.

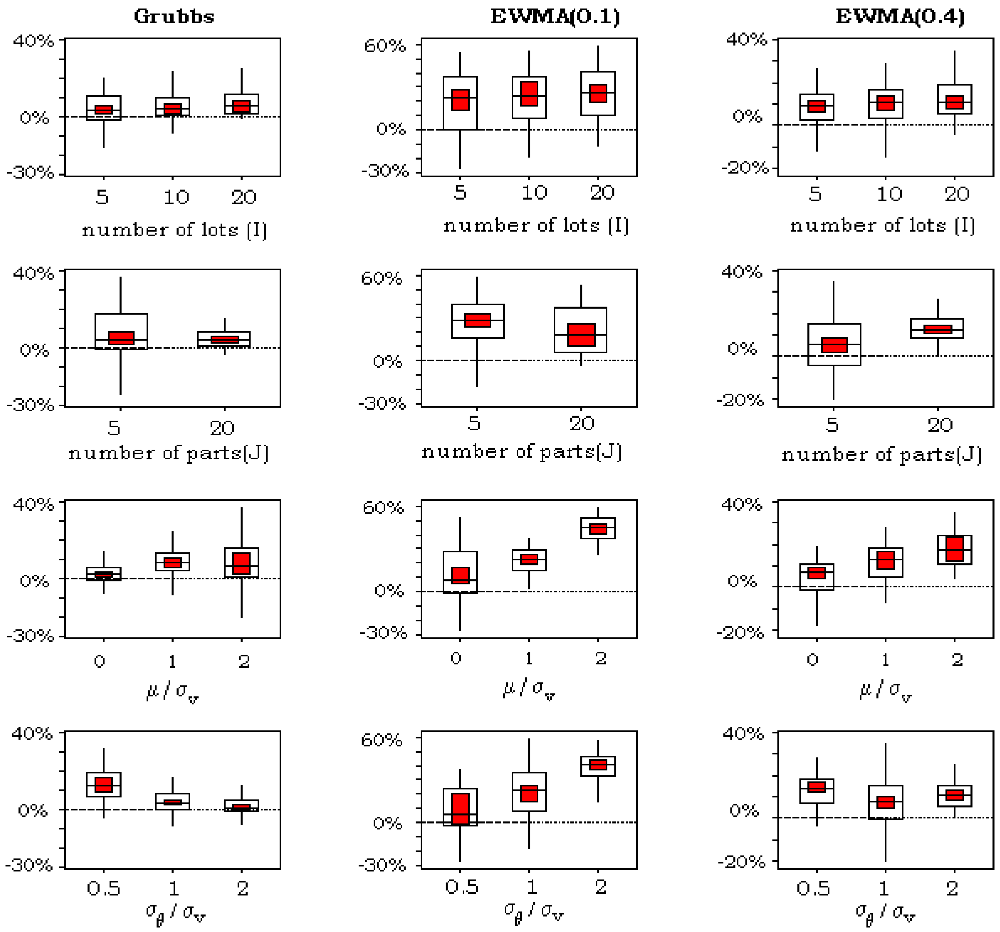

Figure 6: Boxplots on the percentage savings (using the quadratic cost function) induced by the Conditional First Adjustment MCMC policy over Grubbs' harmonic rule and the two EWMA controllers ($\lambda = 0.1$ and $\lambda = 0.4$) as a function of different parameters: the number of lots (I), the number of parts in lots (J), the ratio of the mean offset to the within-lot standard deviation (μ/σ_v), and the ratio of the between to the within-lot standard deviations (σ_θ/σ_v).

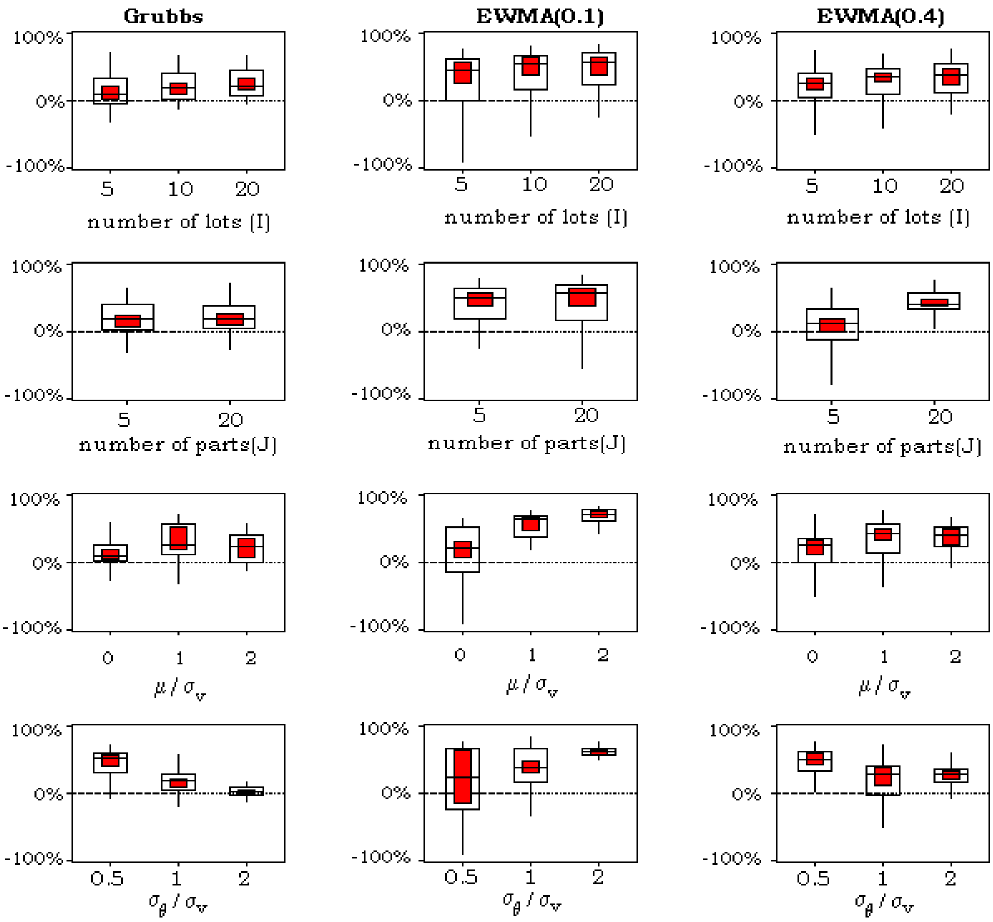

Figure 7: Boxplots on the percentage savings (using the quadratic bias cost function) induced by the Conditional First Adjustment MCMC policy over Grubbs' harmonic rule and the two EWMA controllers ($\lambda = 0.1$ and $\lambda = 0.4$) as a function of different parameters: the number of lots (I), the number of parts in lots (J), the ratio of the mean offset to the within-lot standard deviation (μ/σ_v), and the ratio of the between to the within-lot standard deviations (σ_θ/σ_v).