Available online at www.sciencedirect.com
ScienceDirect
Statistics & Probability Letters 77 (2007) 843–852
STATISTICS & PROBABILITY LETTERS
www.elsevier.com/locate/stapro

Adaptive deadband control of a drifting process with unknown parameters

Zilong Lian, Enrique del Castillo*

Department of Industrial and Manufacturing Engineering, The Pennsylvania State University, University Park, PA 16802, USA

Received 8 April 2005; received in revised form 21 September 2006; accepted 19 December 2006
Available online 20 January 2007

Abstract

Adjusting a drifting process to minimize the expected sum of quadratic off-target and fixed adjustment costs is considered under unknown process parameters. A Bayesian approach based on sequential Monte Carlo methods is presented. The benefits of the resulting "deadband" adjustment policy are studied.

Keywords: Fixed adjustment cost; Sequential Monte Carlo methods; Random walk; Bounded adjustment

1. Introduction

Consider a production process that drifts randomly if left uncontrolled. The process means can be observed only under measurement error at a sequence of stages. There exists a chance to adjust the process mean back to target at each stage by manipulating a controllable variable. It is assumed that a fixed adjustment cost, independent of the magnitude of the adjustment, exists. In addition, there is a symmetric quadratic off-target cost.

This problem has been considered by Box and Jenkins (1963) and Crowder (1992) based on the assumption that the variance parameters, i.e. the variance of the random drift from stage to stage and the variance of the observation errors, are known. In particular, Box and Jenkins (1963) solved the infinite-horizon version of this problem by minimizing the long-run expected cost. The solution to the finite-horizon version of the problem was given by Crowder (1992) through a dynamic programming formulation. Both solutions are of the form of a *deadband*, where the process is adjusted only if the response is predicted to be far enough from target, with the deadband width denoting the "adjustment limits" that depend on the costs involved.

In this paper, we reconsider the adjustment problem studied by Box and Jenkins (1963) and Crowder (1992) for the case that the parameters are *unknown*. The proposed adjustment method is based on the Bayesian estimates of the process parameters and also has a deadband form. Since no conjugate prior distribution is available, dynamic programming cannot be formulated in this case. Therefore, an adaptive method is

*Corresponding author. Fax: +1 814 863 4745.
E-mail address: exd13@psu.edu (E. del Castillo).

0167-7152/$ - see front matter
doi:10.1016/j.spl.2006.12.006

proposed, in which the action limits are obtained by substituting the Bayesian estimates into Box and Jenkins' (1963) solution.

Due to the complexity of this problem, the Bayesian computation has to be carried out by numerical methods. Such computation is required to be done between stages to allow for an on-line adjustment operation. Some computational methods such as Markov chain Monte Carlo (MCMC) methods are too time-consuming for this purpose (Lian et al., 2006). In this paper we investigate the use of sequential Monte Carlo (SMC) methods. A detailed description about the development and the advantages of the SMC method in Bayesian inference can be found in Doucet et al. (2001), see also Lian and del Castillo (2006).

The remainder of this paper is organized as follows. Section 2 introduces the process model and the Bayesian estimation method based on the SMC technique. In Section 3, an adaptive deadband adjustment method is proposed to control the process under a fixed adjustment cost, and an example is given for illustration. More simulation results are given in Section 4. Finally, conclusions and a summary are given in Section 5.

2. Process model and parameter estimation

The assumed process is

$$\theta_i = \theta_{i-1} + U_{i-1} + v_i, \tag{1}$$

$$Y_i = \theta_i + \varepsilon_i, \quad i = 1, 2, \ldots, N, \tag{2}$$

where $v_i \overset{\text{iid}}{\sim} \mathrm{N}(0, \sigma_v^2)$, $\varepsilon_i \overset{\text{iid}}{\sim} \mathrm{N}(0, \sigma_\varepsilon^2)$ are two random sequences that are independent of each other. θ_i is the process mean at stage i, which follows a random walk if the process is not adjusted (i.e., if $U_i = 0$ for $i = 1, 2, 3, \ldots, N$). The process mean can be linearly adjusted by U_i units at each stage to keep it close to its target, where the target can be assumed to be 0 without loss of generality. The process is assumed to start on target, i.e. $\theta_0 = 0$. If the process is not adjusted, the observations are a random walk with observational noise, a stochastic process with autocorrelation structure equivalent to the ARIMA$(0,1,1) =$ IMA$(1,1)$ process, popular in the industrial time series literature (Box et al., 1994).

We first consider a symmetric quadratic loss function

$$L^q = \sum_{i=1}^{N} \theta_i^2. \tag{3}$$

We will discuss the addition of fixed adjustment costs in the next section. To minimize the total loss in (3), an adjustment needs to be made at each stage by the amount

$$U_i = -\hat{\theta}_i, \tag{4}$$

where $\hat{\theta}_i$ is an estimate of θ_i obtained at stage i after Y_i is observed. When the adjustments are carried out at every stage, the loss function in (3) not only evaluates the performance of an adjustment rule but also reflects the accuracy of the process mean estimate. Traditionally, this estimate is calculated through a Kalman filter (KF). Given the observations from the beginning until stage i, $y^i = \{Y_1, Y_2, \ldots, Y_i\}$, the estimate of θ_i at stage i is given by

$$\begin{aligned} \hat{\theta}_i &= \hat{\theta}_{i-1} + U_{i-1} + \omega[Y_i - (\hat{\theta}_{i-1} + U_{i-1})] \\ &= (1-\omega)(\hat{\theta}_{i-1} + U_{i-1}) + \omega Y_i, \end{aligned} \tag{5}$$

where $\hat{\theta}_0 = 0$ and

$$\omega = \frac{1 + \sqrt{1 + 4\sigma_\varepsilon^2/\sigma_v^2}}{1 + \sqrt{1 + 4\sigma_\varepsilon^2/\sigma_v^2} + 2\sigma_\varepsilon^2/\sigma_v^2}. \tag{6}$$

Notice that the KF estimate is a weighted average of the old estimate $\theta_{i-1} + U_{i-1}$ and the new observation Y_i (Duncan and Horn, 1972). The weight ω is uniquely determined by the ratio of two variances $\sigma_\varepsilon^2/\sigma_v^2$. In the case that $\sigma_\varepsilon^2/\sigma_v^2$ is large, which implies relatively large observation errors, a small weight ω will be assigned to the new observation. When observations are more precise, i.e. when $\sigma_\varepsilon^2/\sigma_v^2$ is small, the weight for the new observation will be larger. Duncan and Horn (1972) show how the Kalman filter estimate provides the minimum mean square error estimator of θ_i. This assumes the error distributions are normal with the ratio $\sigma_\varepsilon^2/\sigma_v^2$ known.

In the case when σ_ε^2 and σ_v^2 are unknown, KF estimation cannot be applied. If biased estimates of σ_ε^2 and σ_v^2 are used, KF estimation may lead to a poor performance of the adjustment methods based on it. In the case that these two process parameters are unknown, a new adjustment method is proposed in this paper, which also has the same form as in (4). However the estimates of θ_i are obtained based on a different Bayesian model in which the process parameters σ_ε^2 and σ_v^2 are unknown. In this model, specifications of the prior distributions of σ_ε^2 and σ_v^2 are needed before the process starts. At each stage when a new observation is obtained, the posterior distribution of the process mean and of the parameters σ_ε^2 and σ_v^2 can be computed, according to which the adjustment can be made. The steps to update the posterior distributions and adjust the process can be summarized as follows:

- Start with prior distributions

$$\begin{aligned} &\sigma_\varepsilon^2 \sim \mathrm{LN}(\mu_1, \sigma_1^2) \\ &\sigma_v^2 \sim \mathrm{LN}(\mu_2, \sigma_2^2), \end{aligned} \tag{7}$$

  where LN represents a log-normal distribution (a justification of these priors is given below). The predictive distribution of θ_1 is

$$\theta_1 | \sigma_v^2, \theta_0 \sim \mathrm{N}(\theta_0, \sigma_v^2), \tag{8}$$

  where $\theta_0 = 0$ is assumed. The joint distribution of σ_ε^2, σ_v^2 and θ_1 before the first observation is obtained is denoted as $\pi(\sigma_\varepsilon^2, \sigma_v^2, \theta_1 | y^0)$, where y^0 is an empty data set.
- At stage i, $i = 1, 2, 3, \ldots, N$, when the new observation Y_i is obtained,
  - Update the posterior distribution:

$$\begin{aligned} p(\sigma_\varepsilon^2, \sigma_v^2, \theta_i | y^i) &\propto \pi(\sigma_\varepsilon^2, \sigma_v^2, \theta_i | y^{(i-1)}) \times \text{likelihood } (\sigma_\varepsilon^2, \sigma_v^2, \theta_i | Y_i) \\ &\propto \pi(\sigma_\varepsilon^2, \sigma_v^2, \theta_i | y^{(i-1)}) \times \frac{1}{\sigma_\varepsilon} \exp\left[-\frac{(Y_i - \theta_i)^2}{2\sigma_\varepsilon^2} \right]. \end{aligned} \tag{9}$$

  - Adjust the process by

$$U_i = -\hat{\theta}_i = -E(\theta_i | y^i). \tag{10}$$

  - Compute the new predictive distribution

$$\begin{aligned} \pi(\sigma_\varepsilon^2, \sigma_v^2, \theta_{(i+1)} | y^i) &= \int p(\sigma_\varepsilon^2, \sigma_v^2, \theta_i | y^i) \times p(\theta_{(i+1)} | \sigma_v^2, \theta_i) \, \mathrm{d}\theta_i \\ &= \int p(\sigma_\varepsilon^2, \sigma_v^2, \theta_i | y^i) \times \frac{1}{\sqrt{2\pi\sigma_v^2}} \exp\left[-\frac{(\theta_{(i+1)} - \theta_i)^2}{2\sigma_v^2} \right], \end{aligned} \tag{11}$$

    where $p(\bullet | y^i)$ is the posterior distribution computed in (9).
  - Iterate on i.

The posterior distribution computation is carried out by a SMC method, specifically, the one-pass particle filter (1PFS) algorithm (Balakrishnan and Madigan, 2004). The "1PFS" algorithm was modified in this research for computing the posterior distributions of those parameters that have a bounded domain (i.e., variance parameters). Details about the 1PFS algorithm are shown in Appendix A.

SMC methods do not require conjugate priors, so various forms of priors can be applied. The log-normal distribution was chosen in this paper because its convenience in incorporating the prior information.

Table 1
Top: Scenarios with different prior information; Bottom: processes with different parameters

Scenario	b_ε	b_v
1	0.5	0.5
2	0.5	1
3	0.5	1
4	1	0.5
5	1	1
6	1	2
7	2	0.5
8	2	1
9	2	2
Case	σ_ε	σ_v
1	2	4
2	2	2
3	4	2

In contrast, the inverse gamma distribution, which is conjugate in simple normal data problems, is actually difficult to setup as a ‘‘non-informative’’ prior, a fact that has not been recognized until recently (Spiegelhalter et al., 2004, Section 5.7.3.).

We now compare the performance of the adaptive SMC method with that of a fixed-estimates KF for a quadratic off-target cost. This allows to study the advantages in estimation resulting from the ability to adapt bad priors on the parameters. Evidently, if the KF is provided with the true parameters, it is optimal for a quadratic cost function like (3), see Duncan and Horn (1972).

To compare numerically the proposed adjustment method (SMC method) and the adjustment method based on KF method, the nine scenarios with different prior information shown on the top of Table 1 were investigated. In each scenario, suppose the prior estimates of the parameters σ_ε and σ_v are available, and they are $b_\varepsilon\sigma_\varepsilon$ and $b_v\sigma_v$, respectively. These estimates may be accurate ($b_\varepsilon, b_v = 1$) or biased ($b_\varepsilon, b_v = 0.5, 2$). The estimates on the ratio $\sigma_\varepsilon^2/\sigma_v^2$ are given by $b_\varepsilon^2/b_v^2 \times \sigma_\varepsilon^2/\sigma_v^2$, which may also be biased. If the KF method is applied, these estimates are plugged into Eqs. (5)–(6) to estimate the process mean. In the SMC method, on the other hand, these estimates are only used as the means of the corresponding prior distributions. To represent little confidence on them, we set the variances to be equal to 4 times of the square of their distribution means ($4b_\varepsilon^4\sigma_\varepsilon^4$ and $4b_v^4\sigma_v^4$, respectively).

Three cases of process parameters were considered as shown on the botton of Table 1. Each case was studied for 500 stages or time units. For each case on the right of Table 1, common random numbers were used to simulate the adjusted process under each of the two adjustment methods (the KF method and the SMC method). All combinations of scenarios and process parameters on Table 1 were tried. The total quadratic losses were calculated for both methods, and the performance was evaluated by the percentage savings induced by the SMC method over the KF method:

$$S^q = \frac{L_k^q - L_s^q}{L_k^q}, \tag{12}$$

where L_k^q is the loss under the KF adjustments and L_s^q is the loss under the SMC adjustments.

For each combination of case and scenario (Table 1), 20 replications were made, based on which the average savings and the frequentist 95% confidence interval were computed. The results are shown in Fig. 1.

As it can be seen, in scenarios 1, 5 and 9, where the ratios used in the KF method happen to be equal to the true values (i.e., $b_\varepsilon/b_v = 1$, the KF method provides the best estimates, as expected. However, a more surprising result is that the negative savings incurred by the SMC method are only slightly below 0. Thus, when the SMC estimation starts from unbiased priors, the performance compared to the optimal KF approach is only slightly worse, even when started with relatively flat priors.

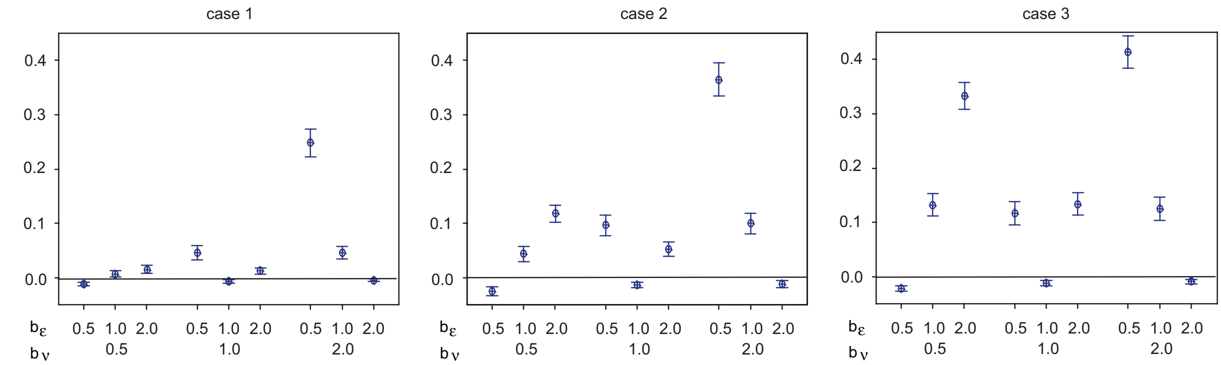

Fig. 1. Average savings induced by SMC method vs. the KF method for each combination of a prior scenario and a process case, and their 95% confidence intervals (no fixed adjustment cost).

In other scenarios, where an incorrect ratio was used in the KF method, the advantage of using the SMC method turns out to be very significant. In particular, the ratios are underestimated in scenarios 2, 3 and 6, such that the KF method gives too much weight to new observations that have considerable observation error. In scenarios 4, 7 and 8, overestimated ratios are used in the KF method so that the errors in the new observations are exaggerated and too little weight is assigned to the new observations. In these cases, although the SMC method also starts with biased prior information, the effect of the prior distributions decays very soon as more data is available, given that the priors are not very informative. Thus, in a situation where the prior information is likely to be biased, the SMC method can achieve a more robust estimation and hence better adjustment performance if started with a relatively flat prior.

We now turn to the case when fixed adjustment costs need to be considered.

3. Fixed adjustment cost and adaptive deadband adjustment method

Consider now the following total loss function:

$$L^t = \sum_{i=1}^{N} \theta_i^2 + c \sum_{i=1}^{N} \delta(U_i), \tag{13}$$

where the function $\delta(x) = 1$ if $x \neq 0$ and $\delta(x) = 0$ otherwise. This loss function includes the total quadratic loss in (3) and the sum of fixed adjustment costs. Constant c is the relative cost of adjusting to the cost of a unit deviation from target. The presence of a fixed adjustment cost generally implies an adjustment policy with a "deadband" form (Box and Jenkins, 1963; Crowder, 1992; Jensen and Vardeman, 1993), i.e. instead of adjusting at each step as in (4), adjustments are carried out only when the magnitude of the predicted process mean is large enough to justify economically the adjustment:

$$U_i = \begin{cases} -\hat{\theta}_i & \text{if } |\hat{\theta}_i| > \alpha_i \\ 0 & \text{if } |\hat{\theta}_i| \leqslant \alpha_i, \end{cases} \tag{14}$$

where α_i is the adjustment limit or the half width of the "deadband" at stage i. Adjustment methods with such form have been proposed for the known parameters case, to control the process described in (1)–(2) under the loss function in (13) based on a KF estimator (Crowder, 1992). Box and Jenkins (1963) considered the case when the length of the process is infinite ($N = \infty$). In their solution, α_i is constant for all stages and is uniquely determined by σ_v for a fixed adjustment cost c, according to

$$\alpha_i = \alpha = [(6c/\sigma_v^2)^{1/4} - 0.63]\sigma_v. \tag{15}$$

We will refer to this method as the constant deadband adjustment method. Crowder (1992) gave an optimal solution to the known-parameters problem for the finite horizon case using dynamic programming, in which the optimal α_i is nondecreasing with i. Specifically, when $(N - i) \to \infty$, α_i will converge to the constant adjustment limit given by (15). When the process is approaching the end, i.e. when $i \to N-1$, the action limit α_i will become larger until it reaches $\sqrt{c}$ at the last stage $N-1$. As shown by Crowder (1992), this "funneling effect" is significant only for very short run processes or for an extremely large adjustment cost. For a moderately long process, the optimal adjustment limits at most stages can be well approximated by Eq. (15). Hence, the constant deadband adjustment method works well in most finite-horizon, known-parameters problems.

Both the finite and infinite horizon solution to this problem require the prior knowledge of σ_ε^2 and σ_v^2, which is used in the KF estimates and in determining the adjustment limits. When such knowledge is unavailable, the SMC method described in Section 2 can be utilized to estimate the process mean. The question that remains is how to determine the adjustment limits. When σ_ε^2 and σ_v^2 are unknown, the optimal adjustment limits should depend on the posterior distributions of θ_i, σ_ε^2 and σ_v^2, in both the infinite-horizon and finite-horizon problem. It is very difficult to find such optimal limits for this problem, in which the posterior distributions can only be numerically computed based on a Monte-Carlo method.

As a practical solution to this problem, we propose an adaptive deadband adjustment method, in which the adjustment limits are approximated based on the sequential Bayesian estimate of the process parameter σ_v^2. In particular, at each stage the SMC method gives the estimates of θ_i and σ_v^2, which are the means of their posterior distributions. The estimate of σ_v^2 can be substituted into (15) to calculate the adjustment limit for the current stage. This limit, together with the estimate of the process mean, can be used in (14) to determine if an adjustment is needed. When more observations are available, the limits calculated in this way converge to the optimal constant limit for the infinite-horizon problem. An example is shown next to illustrate the calculation of the adaptive deadband.

Example. Consider a process that will operate for 500 stages or time units and process parameters equal to $\sigma_\varepsilon = 4$ and $\sigma_v = 2$. The fixed adjustment cost is assumed to be $c = 100$. Suppose the process can be adjusted under the following three methods: (1) constant deadband adjustment method with the true values of the process parameters known (this will give optimal control); (2) constant deadband adjustment method with biased process parameters $\sigma'_\varepsilon = 1$ and $\sigma'_v = 4$; and (3) adaptive deadband adjustment method with prior information also biased such that the means of the prior distributions of σ_ε^2 and σ_v^2 equal to σ'^2_ε and σ'^2_v, respectively. To reflect little a priori information, the variances of the prior distributions of both parameters were set at 4 times the square of the means.

The gains obtained because of a better estimation of the process parameters using the SMC method, which is used in the adaptive deadband method, were shown in the previous section. Now we investigate what additional benefits are obtained due to better (adaptive) adjustment limits. The calculated adjustment limits/half widths are shown in Fig. 2.

Since the adaptive deadband method was started with biased prior distributions, the calculated adjustment limits at the beginning are close to those in the constant deadband method where the same biased information was used. As the process continues, better estimates of σ_v^2 are obtained through the SMC method and they are used in the adaptive deadband method. The adjustment limits converge to the constant adjustment limit using the true process parameters. The total costs are calculated for each method by simulating the process using common random numbers. The total cost for method (1), which should be the best method in this case, is 13 347; the total cost for method (2) is 15 127; and the total cost for method (3) is only 13 605.

It is remarkable that the adaptive deadband adjustment method with biased prior information induces a cost only slightly larger than the cost caused by the best method. This is evidently due to the relative non-informative priors that were used, but the proximity to optimality is notable. When it is compared to the constant deadband method with the same biased prior information, the savings are significant, as expected. We know present a more exhaustive simulation study of the performance of the proposed adaptive SMC method.

4. Performance analysis

In order to investigate the performance of the proposed adjustment method, more cases were investigated in this section. For the different processes in Table 1, each prior information scenario was solved using both the

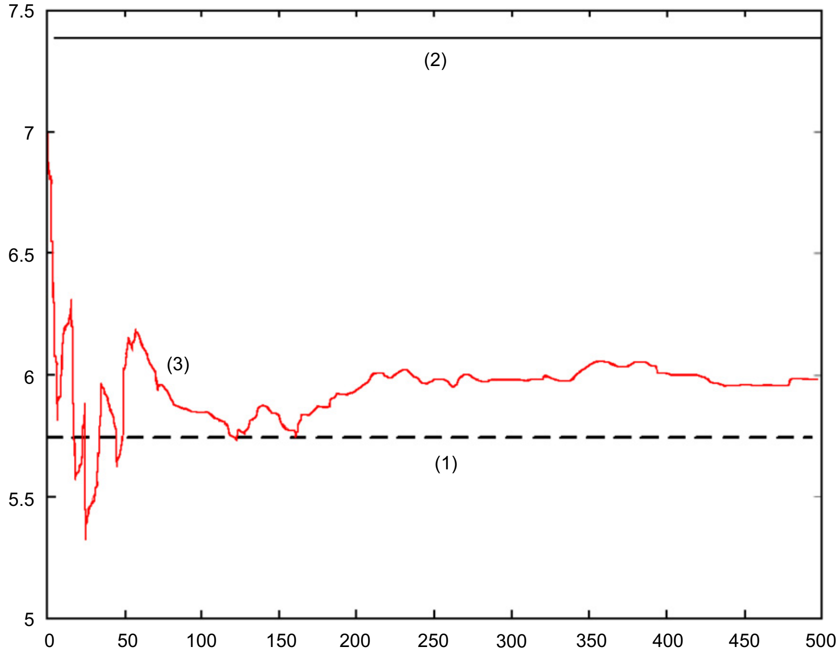

Fig. 2. The adjustment limits/half widths for the (1) constant deadband method with the true process parameters, (2) constant deadband method with the wrong process parameters, (3) adaptive deadband method with biased prior information.

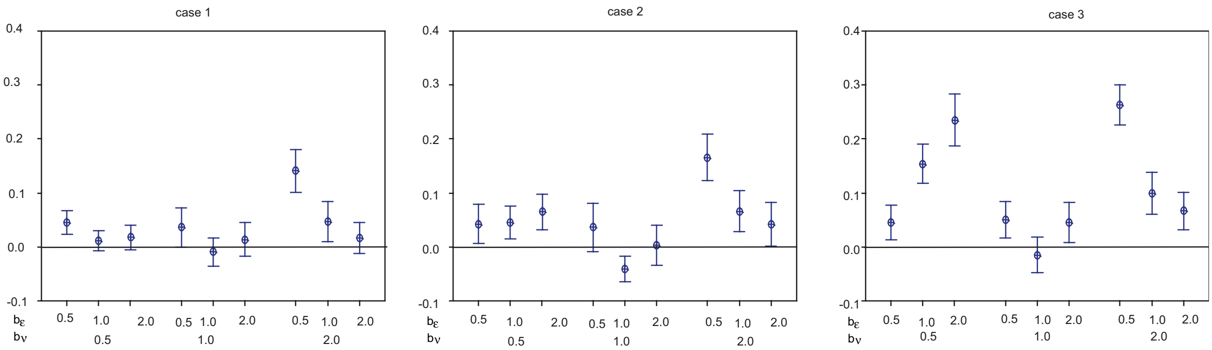

Fig. 3. Average savings induced by the adaptive deadband method vs. the constant deadband method for each combination of a prior scenario and a process case, and their 95% confidence intervals, when $c = 100$.

constant deadband method and the adaptive deadband method. The fixed adjustment cost was assumed to be $c = 100$ in all cases. The variances of the prior distributions of the process parameters were always equal to 4 times of the square of their means. The relative savings induced by the adaptive deadband method were calculated with respect to the constant deadband method with the same prior information scenario using common random numbers. Twenty replications were made to calculate the average savings and their 95% confidence intervals for each combination of a process case and a prior scenario. The results are shown in Fig. 3.

The only prior information scenario that has negative average savings is scenario 5, where the true values of the parameters were used in the constant deadband method and hence it is the best method for both estimation and determination of the adjustment limits. However, the difference between the best adjustment method and the adaptive deadband method is not significant. For scenarios 1 and 9, where the Kalman filter estimates are accurate since the ratios $\sigma_\varepsilon^2/\sigma_\nu^2$ that were used happened to be equal to the true values,

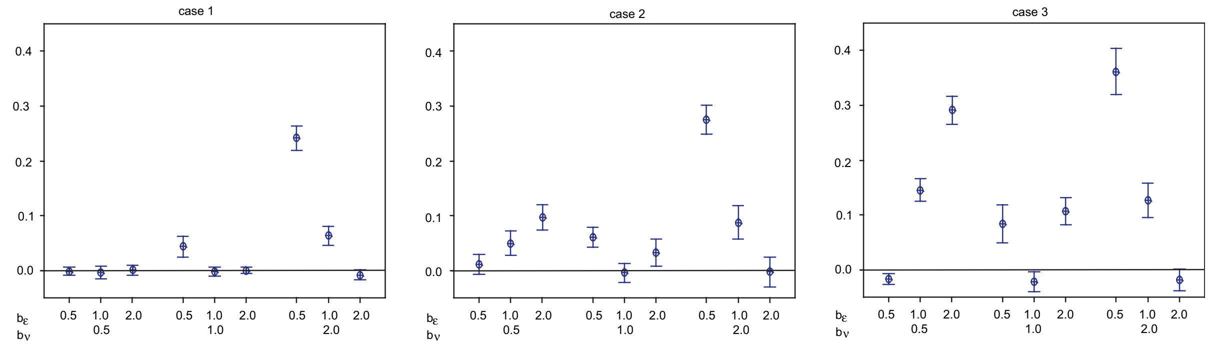

Fig. 4. Average savings induced by the adaptive deadband method vs. the constant deadband method for each combination of a prior scenario and a process case, and their 95% confidence intervals, when $c = 16$.

the positive savings are mainly due to the improved adjustment limits obtained in the adaptive deadband method. For scenarios 2 and 8, where the constant adjustment limits used in the constant deadband method are optimal since the true value of σ_v^2 was used to determine the limits, the savings made by the adaptive deadband method are mainly due to the better process mean estimates obtained in the SMC method. For other scenarios, the savings can be considered to be due to a combination of better adjustment limits and better process mean estimation.

In general, when the adjustment cost c is large, the adjustment limits will be wide and there will be fewer adjustments. Hence having good process mean estimates becomes less important, as adjustments are based on such estimates. Contrarily, when c is small, having good process mean estimates becomes more critical compared to the effect of having very precise adjustment limits (which depend in turn on the σ_v^2 estimate only, see Eq. (15)).

Fig. 4 shows the average savings and their 95% confidence intervals induced by the adaptive deadband method versus the constant deadband method for the same scenarios and cases when a smaller adjustment cost $c = 16$ was considered. The savings for $c = 16$ are closer to those in Fig. 1, where the savings are purely due to the process mean estimation. Specifically, for the scenarios when the Kalman filter estimates are more accurate (e.g., scenarios 1 and 9), the savings for $c = 16$ drop compared to those for $c = 100$. For the scenarios when the Kalman filter estimates are not accurate (e.g. scenarios 3 and 7), the savings for $c = 16$ are more significant than those for $c = 100$.

5. Conclusion

As demonstrated in Section 2, the Bayesian method using a SMC technique provides better estimates of the process means than the Kalman filter (KM) estimates do, in the cases when the prior information of the variance ratio $\sigma_\varepsilon^2/\sigma_v^2$ is not available or is biased. Such advantage in estimation exists in the process adjustment problems regardless of whether or not the adjustment cost is considered.

An adaptive deadband adjustment method was proposed in Section 3 based on the Bayesian estimation method. In contrast to the constant deadband method, the adaptive deadband method calculates the deadband limits at each step based on the on-line Bayesian estimate of the variance of the state equation σ_v^2. The main value of the proposed method is its ability to "recover" from an initially biased prior on σ_v^2 and obtain more adequate adjustment limits that result in significant savings. Compared to the optimal KF solution obtained when the parameters are known, the proposed approach, which assumes unknown parameters, has a surprisingly close performance.

A MATLAB computer program that implements the adaptive deadband adjustment method developed in this research can be downloaded from: http://www2.ie.psu.edu/Castillo/research/EngineeringStatistics/.

With this program, at each stage i, the observed values of the quality characteristic, Y_i, need to be entered. The program sequentially updates the parameters estimates with the algorithm described in Appendix A, and computes the adjustment using expression (14). Before starting a control session, the total number of stages

(N) and the relative fixed adjustment cost (c) need to be specified at the beginning of the `AdaptiveAdjust.m` program.

Appendix A. Algorithm for estimating the process parameters θ_i, σ_v^2 and σ_ε^2

At the beginning of the process:

Draw M random numbers $\theta_0^{(m)}$ from the prior distribution of θ_0 (in this paper, $\theta_0^{(m)}=0$), $m=1,2,\ldots,M$;
Draw M random numbers $\sigma_\varepsilon^{2(m)}$ from the prior distribution of σ_ε^2, $m=1,2\ldots,M$;
Draw M random numbers $\sigma_v^{2(m)}$ from the prior distribution of σ_v^2, $m=1,2\ldots,M$;
Create an initial weight vector $(w_1,w_2,w_3,\ldots,w_M)$, where $w_m=1/M$ for all m.

For $n=1,2,3,\ldots,N$ (i.e., iterating over all stages):

Observe Y_n
Generate 1 random number $\theta_n^{(m)}$ from the distribution $\mathrm{N}(\theta_{n-1}^{(m)},\sigma_v^{2(m)})$ for each $m=1,\ldots,M$.
Calculate the likelihood of the mth "particle" (i.e., a combination $(\theta_i^{(m)}$, $\sigma_\varepsilon^{2(m)}$,$\sigma_v^{2(m)})$):
Let $L_m=(1/\sqrt{\sigma_\varepsilon^{2(m)}})\exp\{-(Y_i-\theta_i^{(m)})^2/2\sigma_\varepsilon^{2(m)}\}$
Update the new weight vector and normalize it

$$w_m \leftarrow w_m \times L_m \text{ then } w_m \leftarrow w_m \Big/ \sum\nolimits_{m=1}^{M} w_m.$$

Obtain the new parameter estimators:

$$\hat{\theta}_n|D_n=\sum\nolimits_{m=1}^{M} w_m\theta_n^{(m)},$$
$$\hat{\sigma_\varepsilon^2}|D_{ij}=\sum\nolimits_{m=1}^{M} w_m\sigma_\varepsilon^{2(m)},$$
$$\hat{\sigma_v^2}|D_{ij}=\sum\nolimits_{m=1}^{M} w_m\sigma_\theta^{2(m)}.$$

Calculate the effective sample size (ESS) factor:

$$\mathrm{ESS}=\frac{M}{1+M^2 var(w_m)}.$$

If ESS$<pM$, where p is a specified level between 0 and 1 ($p=0.5$ was used in our computations), rejuvenate M particles using the "1PFS" algorithm (Balakrishnan and Madigan, 2004).

References

Balakrishnan, S., Madigan, D., 2004. A one-pass sequential Monte Carlo method for Bayesian analysis of massive datasets. Technical Paper, ⟨http://www.stat.rutgers.edu/~madigan/papers/⟩.

Box, G.E.P., Jenkins, G.M., 1963. Further contributions to adaptive quality control simultaneous estimations of dynamics nonzero costs. Bull. Internat. Statist. Inst. 34, 943–974.

Box, G.E.P., Jenkins, G.M., Reinsel, G., 1994. Time Series Analysis, Forecasting, and Control, third ed. Prentice-Hall, Englewood Cliffs.

Crowder, S.V., 1992. An SPC model for short production runs: minimizing expected cost. Technometrics 34, 64–73.

Doucet, A., de Freitas, N., Gordon, N., 2001. Sequential Monte Carlo Methods in Practice. Springer, New York.

Duncan, D.B., Horn, S.D., 1972. Linear dynamic recursive estimation from the viewpoint of regression analysis. J. Amer. Statist. Assoc. 67 (340), 815–821.

Jensen, K.L., Vardeman, S.B., 1993. Optimal adjustment in the presence of deterministic process drift and random adjustment error. Technometrics 35, 376–389.

Lian, Z., Colosimo, B.M., del Castillo, E., 2006. Setup adjustment of multiple lots using a sequential Monte Carlo method. Technometrics 48(3), 373–385.

Lian, Z., del Castillo, E., 2006. Setup adjustment under unknown process parameters and fixed adjustment cost. J. Statist. Plann. Inference 136, 1039–1060.

Spiegelhalter, D.J., Abrams, K.R., Myles, J.P., 2004. Bayesian Approaches to Clinical Trials and Health-Care Evaluation. Wiley, New York.