

SMALL SAMPLE PERFORMANCE OF SOME STATISTICAL SETUP ADJUSTMENT METHODS.

Enrique del Castillo and Rong Pan

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

Bianca M. Colosimo

Dipartimento di Meccanica, Politecnico di Milano, Via Bonardi 9, 20133, Milan, Italy

Key Words: Process Adjustment, Stochastic Approximation, Integral Control.

ABSTRACT

The setup adjustment problem occurs when a machine experiences an upset at setup that needs to be compensated for. In this paper, feedback methods for the setup adjustment problem are studied from a small-sample point of view, relevant in modern manufacturing. Sequential adjustment rules due to Grubbs (1983) and an integral controller are considered. The performance criteria is the quadratic off-target cost incurred over a small number of parts produced. Analytical formulae are presented and numerically illustrated. Two cases are considered, the first one where the setup error is a constant but unknown offset and the second one where the setup error is a random variable with unknown first two moments. It is shown how Grubbs' harmonic rule and a simple integral controller provide a robust adjustment strategy in a variety of circumstances. As a by-product, the formulae presented in this paper allow to compute the expected off-target quadratic cost when a sudden shift occurs during production (not necessarily at setup) and the adjustment scheme compensates immediately after its occurrence.

1. INTRODUCTION.

Suppose that due to a defective setup machine operation, the quality characteristic generated by a production process is in a state of statistical control but the process starts off-target. Adjusting the process is justified if the only relevant cost is the cost of running the process off-target. Sequential adjustment rules for this problem were proposed by Grubbs (1983). These rules have been shown to derive from a much broader class of setup adjustment problems based on stochastic control methods by Del Castillo, Pan and Colosimo (2002). In this paper, the small sample performance of Grubbs' adjustment rules are studied and contrasted with other feedback adjustment methods.

To introduce some basic notation that will be used in what follows, let Y_t denote the observed deviation from target of some quality characteristic of interest. Following Grubbs (1983), a simple

but useful model for the setup adjustment problem is to assume

$$Y_t = d + U_{t-1} + v_t = \mu_t + v_t \quad (1)$$

where d is the setup error, U_t is the level of the controllable factor set after producing part t (this will have an immediate effect on part $t + 1$), μ_t is the mean deviation from target for part t , and $v_t \sim N(0, \sigma_v^2)$ models both the part-to-part variability and the measurement error.

Two different control rules were derived by Grubbs depending on two sets of assumptions made on the setup error d :

1. If d is an unknown constant, minimization of $\text{Var}(\mu_{n+1})$ subject to $E[\mu_{n+1}] = 0$ results in Grubbs' "harmonic rule":

$$U_t = -\hat{d}_t, \quad \hat{d}_t = \hat{d}_{t-1} + K_t Y_t, \quad K_t = 1/t \quad (2)$$

or

$$U_t - U_{t-1} = \nabla U_t = -\frac{Y_t}{t}$$

thus the weights K_t follow a harmonic series. An initial (a priori) estimate \hat{d}_0 is required to set the first setting of the controllable factor at $U_0 = -\hat{d}_0$. Del Castillo, Pan and Colosimo (2002) point out how adjustment rule (2) is a particular case of Robbins and Monro's (1951) celebrated stochastic approximation method. Therefore, this paper contains (indirectly) a small-sample performance study of stochastic approximation methods applied to the simple case of estimation of an unknown constant.

2. If $d \sim N(\bar{d}, \sigma_d^2)$ with both \bar{d} and σ_d^2 known, minimization of $E[\sum_{t=1}^n \mu_t^2]$ is achieved by Grubbs' second adjustment rule. In this case,

$$\hat{d}_t = \hat{d}_{t-1} + K_t Y_t, \quad K_t = \frac{1}{t + \frac{\sigma_v^2}{\sigma_d^2}} \quad \text{and} \quad \nabla U_t = -K_t Y_t. \quad (3)$$

This rule is called Grubbs' "extended" rule by Trietsch (1998).

Del Castillo, Pan and Colosimo (2002) provide a Bayesian formulation based on a Kalman filter to the solution of the second setup adjustment problem above (random d) that yields Grubbs extended rule as a solution. Define the posterior variance of μ_t as $P_t = \text{Var}(\mu_t | Y_t, Y_{t-1}, \dots, Y_1)$. The Kalman filter formulation yields in this case (Del Castillo, Pan, and Colosimo, 2002):

$$P_t = \frac{P_{t-1} \sigma_v^2}{\sigma_v^2 + P_{t-1}} = \frac{\sigma_v^2 P_0^2}{\sigma_v^2 + t P_0} \quad \hat{d}_t = \hat{d}_{t-1} + K_t Y_t, \\ K_t = \frac{1}{t + \frac{\sigma_v^2}{P_0}}, \quad \text{and} \quad \nabla U_t = -K_t Y_t. \quad (4)$$

The interpretation is that, a priori, $d \sim (\hat{d}_0, P_0)$. Thus for Grubbs' extended rule to be optimal with respect to $E[\sum_{t=1}^n \mu_t^2]$ we need to know \bar{d} (so we can set $\hat{d}_0 = \bar{d}$) and σ_d^2 and σ_v^2 must be known for us to use (3). Evidently, if $P_0 = \sigma_d^2$, (3) and (4) are identical. Note how under this interpretation, if there is no prior information on the offset ($P_0 \rightarrow \infty$) (4) is equivalent to (2), Grubbs' simpler harmonic rule.

A pure Bayesian point of view will stop after stating the solution (4). It is still of practical interest, however, to study how this rule behaves if the prior distribution is not exactly equal to the true setup distribution. In other words, what if \bar{d} and σ_v^2 are not known but someone still applies Grubbs' extended rule with $P_0 \neq \sigma_d^2$ and/or $\hat{d}_0 \neq \bar{d}$? The opposite is also of interest: to study the performance of Grubbs' simpler harmonic rule under the assumption of a random setup error.

An additional adjustment rule that will be contrasted is an integral controller (Box and Luceño, 1997);

$$\nabla U_t = -\lambda Y_t \quad (5)$$

which, contrary to Grubb's rules, does not converge to zero since λ is a constant.

The remainder of this paper is organized as follows. Section 2 presents the small sample performance indices which will be used to compare the three adjustment rules given by equations (2), (4) and (5). Given the emphasis in modern manufacturing for short production runs, the performance analysis in this paper focuses on the performance for a small number of parts. Section 3 shows the numerical results for the case the setup error is an unknown constant. Section 4 presents the corresponding results for the random setup error case. All the numerical results in sections 3 and 4 were obtained analytically and are not based on simulation. The paper concludes with a summary of results.

2. PERFORMANCE INDICES FOR SMALL SAMPLES.

The performance indices that will be used in the remainder of this paper are presented in this section for the two cases considered by Grubbs.

Consider first the case where the setup error d is an unknown constant or "offset". For this case, the performance index considered is the scaled Average Integrated Square Deviation (AISD) incurred over m time instants or parts. This is defined for integer $m > 0$ as:

$$\text{AISD}(m) = \frac{1}{m\sigma_v^2} \sum_{i=1}^m E[Y_i^2] = \frac{1}{m\sigma_v^2} \sum_{i=1}^m \left(V[Y_i] + E[Y_i]^2 \right). \quad (6)$$

The AISD is a common performance index in the control engineering literature. Since Y_i models deviations from target, the AISD index is like an average "variance plus squared bias" calculation, and is a surrogate of a quadratic off-target "loss" function. We avoid dependency on σ_v^2 by dividing by this quantity.

One important byproduct of the AISD formulae (presented later in the paper) is that they provide a measure of the quadratic cost incurred by a process after a shift of size d occurs *at any point in time* (not only at startup) assuming the process is adjusted after the occurrence of the shift with one of the adjustment policies here discussed.

Consider now the case where the setup error d is a random variable. The performance measure to be used when d is random is once again the AISD but we need to account for the additional variability in the setup error, so we define

$$\text{AISD}_d(m) = \frac{1}{m\sigma_v^2} E_d \left[\sum_{i=1}^m E[Y_i^2] \right] = \frac{1}{m\sigma_v^2} \int_{-\infty}^{\infty} \sum_{i=1}^m E[Y_i^2] f_d(x) dx \quad (7)$$

where the outer expectation is taken over the distribution of d . Note that if d is a non-random constant, then $\text{AISD}(m) = \text{AISD}_d(m)$. The case when d is normal with known mean and known variance was discussed by Trietsch (1998). Under such conditions, Grubbs' extended rule is optimal for the AISD_d criterion.

3. PERFORMANCE FOR AN UNKNOWN CONSTANT SETUP ERROR.

Suppose d is an unknown constant but unaware of this fact a user applies Grubbs' extended rule (i.e., the Kalman filter adjustment scheme given by (4)) to the process. It is shown in Appendix A that this rule applied to the process $Y_t = d + U_{t-1} + v_t$ (with d constant) results in

$$\frac{E[Y_t]}{\sigma_v} = \frac{A}{B_1(t-1) + 1} \quad (8)$$

and

$$\frac{V[Y_t]}{\sigma_v^2} = 1 + \frac{t-1}{(1/B_1 + t-1)^2} \quad (9)$$

where $A = (d - \hat{d}_0)/\sigma_v$ measures how far off the initial estimate of the offset was. The quantity $B_1 = P_0/\sigma_v^2$ is a measure of the "confidence" on the initial offset estimate.

To study the performance of Grubbs' adjustment rules, equations (8) and (9) can be substituted into equation (6) and the sum computed for given values of A, B_1 and m . Note that our analytic expressions are exact and avoid use of simulation to estimate the $\text{AISD}(m)$. Alternatively, an expression for the sum in $\text{AISD}(m)$ is given by formula (17) in Appendix A which can be easier to use if a software that computes the polygamma function is available (e.g., Mathematica or Maple). We note that the corresponding expressions for Grubbs' harmonic rule are obtained from (8) and (9) by letting $B_1 \rightarrow \infty$.

For a discrete integral controller (or EWMA controller), it can be shown that

$$\frac{E[Y_t]}{\sigma_v} = (1 - \lambda)^{t-1} A \quad (10)$$

and

$$\frac{V[Y_t]}{\sigma_v^2} = \frac{2 - \lambda(1 - \lambda)^{2(t-1)}}{2 - \lambda} \quad (11)$$

from where $\text{AISD}(m) = \frac{1}{m\sigma_v^2} \sum_{i=1}^m (E[Y_i]^2 + V[Y_i])$ can be computed or one can use the closed-form expression (eq. 18) in Appendix A. The $\text{AISD}(m)$ expressions allow to study the trade-offs between the sum of the variances and the sum of squared expected deviations (squared bias). For the Kalman filter scheme, as $B_1 = P_0/\sigma_v^2 \rightarrow 0$, implying increasingly higher confidence in the a priori offset estimate, then $m^{-1} \sum_{i=1}^m V[Y_i]/\sigma_v^2 \rightarrow 1$ (i.e., we get lower variance), but $m^{-1} \sum_{i=1}^m E[Y_i]^2/\sigma_v^2 \rightarrow A^2$ (i.e., we get larger bias). Similarly, for the EWMA or integral controller, as $\lambda \rightarrow 0$, implying less weight given to the last observation we have that $m^{-1} \sum_{i=1}^m V[Y_i]/\sigma_v^2 \rightarrow 1$ (lower variance), but $m^{-1} \sum_{i=1}^m E[Y_i]^2/\sigma_v^2 \rightarrow A^2$ (larger bias).

The performance of the following adjustment rules was evaluated based on the AISD criterion (in all cases, adjustments are given by $\nabla U_t = -\nabla \hat{d}_t$):

1. Grubbs harmonic rule, where $\hat{d}_t = \hat{d}_{t-1} + Y_t/t$;
2. Kalman filter rule 1 (assumes σ_v^2 is known), where $\hat{d}_t = \hat{d}_{t-1} + \frac{Y_t}{\sigma_v^2/P_0 + t}$. This is equivalent to Grubbs' extended rule;
3. Kalman filter rule 2 which is same as above but σ_v^2 is estimated on-line from $z_t = Y_t - U_{t-1}$ using only the data available at time t ;
4. Discrete integral controller (EWMA controller), where $\hat{d}_t = \hat{d}_{t-1} + \lambda Y_t$.

There are two parameters that can be modified in Grubbs' adjustment rules: \hat{d}_0 and P_0 . The effect of these parameters can be studied from looking at the effect of changes in A and B_1 , as previously defined. Therefore, the four scenarios presented in Table 1 were investigated.

	B_1 small	B_1 large
$ A $ small	good choice (case 1)	bad choice (case 2)
$ A $ large	bad choice (case 3)	good choice (4)

Table 1: Scenarios of interest, adjusting schemes, $A = (d - \hat{d}_0)/\sigma_v$, $B_1 = P_0/\sigma_v^2$.

In the table, if the initial prior variance P_0 is large relative to σ_v^2 (i.e., if B_1 is large), the weights K_t will be close to $1/t$ (Grubbs' harmonic rule), i.e., the initial estimate \hat{d}_0 will be discounted faster. This turns out to be a good decision if the initial offset estimate is far from d , where the distance between d and \hat{d}_0 is measured relative to σ_v . A similar good decision is when P_0 is low and \hat{d}_0 is a good estimate of the offset (B_1 small, $|A|$ small). In such case, $K_t < 1/t$, so there will be a slower discounting of the initial estimate \hat{d}_0 . Cases (2) and (3) on the table indicate bad decisions, when

the value of P_0 does not reflect how good the initial offset estimate really is. Since in the absence of historical information it is difficult to know a priori the value of d it is of practical interest to study the four cases on the Table.

Table 2 contrasts the AISD performance of Grubbs' harmonic rule, the discrete integral controller (EWMA controller) and the Kalman filter adjusting scheme (σ_v^2 known). The table shows the values of AISD(m) for $m = 5, 10$ and 20 . As can be seen from the Table, the "gap" between the column minimum and the AISD provided by Grubbs rule shrinks as the offset d gets much larger than σ_v (i.e., as $|A|$ increases). This gap, however, is quite moderate except in the case where one is very confident ($B_1 = P_0/\sigma_v^2$ small) of our a priori offset estimate and the a priori offset estimate turns out to be quite accurate (i.e., $A = 0$). This is not a practical case because it implies we practically know the value of the offset d .

If $A = 0$, it can be seen from (8) and (10) that the AISD indices equal the average scaled variance since the deviations from target will always equal zero on average. If $d = \hat{d}_0 = 0$, the AISD quantifies the average *inflation in variance* we will observe for adjusting a process when there was no need to do so. Note that for $A = 0$ (no offset), one can get an inflation in variance equal to zero if $B_1 = 0$ in the Kalman filter scheme or if $\lambda = 0$ in the integral control scheme. This inflation in variance has been studied, for discrete integral controllers, by Box and Luceño (1997) and Del Castillo (2001), although these authors looked at asymptotic variances, and not at small-sample variances as we do here.

Perhaps it should be pointed out that if one were extremely confident on the estimate of the offset of the machine ($B_1 \rightarrow 0$), simply setting $U_t = -\hat{d}_0$ for $t = 0, 1, \dots$ will result in an on-target process assuming we indeed have $\hat{d}_0 = d$. Thus, for most practical cases where a sequential adjustment rule is needed, the Kalman Filter rule (and Grubbs' extended rule) does not perform significantly better than Grubbs' harmonic scheme in the case of a constant unknown setup error.

Intuitively, if the variance σ_v^2 is unknown the performance of the Kalman filter scheme (Grubbs' extended rule) can only worsen. This was confirmed by estimating AISD using simulation. Thus Grubbs harmonic rule is also superior, in the single realization case, to the Kalman Filter scheme with variance unknown.

Turning to the discrete integral controller, it can be seen that it also provides a very competitive scheme compared to the Kalman filter (Grubbs' extended rule) scheme. The parameter λ has the effect of bringing the process back to target more rapidly the larger λ is. The trade-off is that there is an increase, for small A , of the AISD index as λ is increased. That is, the inflation in variance due to unnecessarily adjusting an on-target process ($d = 0, A = 0$) increases as λ increases. From

$m = 5$				
$B_1 :$	$ A = 0$	$ A = 1$	$ A = 2$	$ A = 3$
1/90	1.00023	1.95791	4.83092	9.61929
0.5	1.09344	1.48656	2.66589	4.63144
1	1.16394	1.45667	2.33483	3.79844
2	1.24138	1.47815	2.18847	3.37233
90	1.41043	1.61046	2.21056	3.21074
Grubbs	1.41667	1.61667	2.21667	3.21667
I controller ($\lambda = 0.1$)	1.01655	1.70215	3.75895	7.18696
I controller ($\lambda = 0.2$)	1.05601	1.55191	3.03962	5.51914
I controller ($\lambda = 0.3$)	1.10922	1.49030	2.63354	4.53894
$m = 10$				
$B_1 :$	$ A = 0$	$ A = 1$	$ A = 2$	$ A = 3$
1/90	1.00049	1.91004	4.63869	9.18644
0.5	1.09038	1.31359	1.98323	3.09930
1	1.13792	1.29290	1.75783	2.53271
2	1.18491	1.30578	1.66840	2.27276
90	1.27952	1.37954	1.67959	2.17969
Grubbs	1.28290	1.38290	1.68290	2.18290
I controller ($\lambda = 0.1$)	1.02830	1.49063	2.87761	5.18925
I controller ($\lambda = 0.2$)	1.08060	1.35518	2.17890	3.55178
I controller ($\lambda = 0.3$)	1.14190	1.33782	1.92558	2.90518
$m = 20$				
$B_1 :$	$ A = 0$	$ A = 1$	$ A = 2$	$ A = 3$
1/90	1.00090	1.82739	4.30685	8.43929
0.5	1.07243	1.19211	1.55117	2.14960
1	1.10008	1.17989	1.41931	1.81835
2	1.12585	1.18691	1.37009	1.67539
90	1.17564	1.22565	1.37568	1.62573
Grubbs	1.17739	1.22739	1.37739	1.62739
I controller ($\lambda = 0.1$)	1.03899	1.29825	2.07606	3.37240
I controller ($\lambda = 0.2$)	1.09568	1.23455	1.65116	2.34552
I controller ($\lambda = 0.3$)	1.15917	1.25721	1.55133	2.04152

Table 2: Kalman filter adjusting scheme (σ_v^2 known), Grubbs' harmonic rule and Integral controller AISD performance. $A = (d - \hat{d}_0)/\sigma_v$, $B_1 = P_0/\sigma_v^2$. Bold numbers are minimums by column.

$m = 10, N = 100$				
$B_1 :$	$A = 0$	$A = 1$	$A = 2$	$A = 3$
1/90	1.0009	1.8208	4.2807	8.3805
0.5	1.0760	1.1233	1.2653	1.5019
1	1.0948	1.1177	1.1865	1.3012
2	1.1082	1.1224	1.1648	1.2354
90	1.1277	1.1377	1.1677	1.2177
Grubbs	1.1182	1.1287	1.1583	1.2083
I controller ($\lambda = 0.1$)	1.0493	1.2050	1.6719	2.4502
I controller ($\lambda = 0.2$)	1.1130	1.1508	1.2643	1.4535
I controller ($\lambda = 0.3$)	1.1798	1.2001	1.2610	1.3626

Table 3: Kalman filter adjusting scheme (σ_v^2 known), Grubbs' harmonic rule and I-controller AISD(10,100) performance. Ten adjustments were made after which 90 additional parts were produced. $A = (d - \hat{d}_0)/\sigma_v$, $B_1 = P_0/\sigma_v^2$. Bold numbers are minimums by column.

the Table, it appears the value $\lambda = 0.2$ provides a relatively good trade-off between fast return to target and inflation of variance if the process is really on-target (no offset).

Adjusting only the first few times

It could be argued that in practice, only the first few adjustments will be implemented after which no further adjustments are made to the machine. The machine then runs at the final setting that resulted at the end of the adjustments until completion of the batch of N parts. The performance of the adjustment rules should be investigated on this assumption instead. For this reason, let m denote the number of adjustments implemented in a batch of size N parts produced. For discrete m such that $1 \leq m \leq N$, the average integrated squared deviation index is defined as:

$$\text{AISD}(m, N) = \frac{m \text{AISD}(m)}{N} + \frac{(N - m)(V[Y_{m+1}] + E[Y_{m+1}]^2)}{N\sigma_v^2}.$$

Closed-form expressions for $\text{AISD}(m, N)$ for the Kalman filter, Grubbs harmonic, and discrete integral control rules can be found in Appendix A. They were used to produce Table 3 where the AISD(10, 100) performance indices were computed. Clearly, if $m = N$, then $\text{AISD}(m, N) = \text{AISD}(m)$.

From Table 3 and similar computations for other values of m and N , it was observed that

the conclusions expressed before based on the AISD(m) index, which measures off-target cost only while the adjustments take place, are practically unchanged if we consider in addition the cost incurred after adjustments stop and the process keeps operating. Only when $A \rightarrow 0$ (no offset or perfect offset estimate) and $B_1 = P_0/\sigma_v^2$ is small, the Kalman filter rule outperforms the harmonic rule. The discrete integral controller with $\lambda = 0.2$ is a good intermediate value that balances a rapid return to target for large offsets with a low inflation in variance in case the process was really on-target but we nevertheless adjust.

4. PERFORMANCE WHEN THE SETUP ERROR IS A RANDOM VARIABLE.

Suppose now that the offset d is a random variable such that $d \sim (\bar{d}_0, \sigma_d^2)$. Note that no assumption on the distribution of d is made. We wish to evaluate the performance of the different adjustment methods by averaging over the possible realizations of the random offset d .

As mentioned earlier, if the mean and variance of d are known, then the Kalman Filter scheme, and hence, Grubbs extended rule are optimal for a quadratic loss function such as our AISD $_d$ criterion. This was the case discussed by Grubbs (1983) and Trietsch (1998). In this section we consider the more general case when the mean and variance of d are both unknown.

When d is random, we need to use a prior estimate \hat{d}_0 with associated variance P_0 to start the Kalman filter scheme (4). The situation is depicted in Figure 1. Using (7) as our performance index, it is shown in Appendix B that for the Kalman Filter (KF) scheme:

$$\text{AISD}_d(m)_{KF} = C_1(B_2 + A_2^2) + C_2 \quad (12)$$

where $A_2 = (\bar{d}_0 - \hat{d}_0)/\sigma_v$ is a measure of the *average* error in the offset estimate, $B_2 = \sigma_d^2/\sigma_v^2$ is a measure of the variability of the setup, and C_1 , and C_2 (and C_3 used next) are functions of $B_1 = P_0/\sigma_v^2$ and m shown in Appendix A. For Grubbs' (G) harmonic rule this reduces to:

$$\text{AISD}_d(m)_G = \frac{B_2 + A_2^2}{m} + C_3. \quad (13)$$

Recall that B_1 is a measure of confidence in \hat{d}_0 , therefore, since σ_d^2 is not known, we have that in general $B_1 \neq B_2$. Closed-form expressions can be obtained for AISD $_d(m)_{\text{EWMA}}$, see Appendix B.

The AISD $_d$ performance of the Kalman filter approach, Grubbs basic harmonic rule, and that of an integral controller were evaluated using equations (12-13) and (22). Figure 2 shows cases when the Kalman filter approach is better than Grubbs' harmonic rule for different values of B_1 , B_2 , A_2 , and m . The shaded regions correspond to cases where $\text{AISD}_d(m)_{KF} < \text{AISD}_d(m)_G$. As it can be seen, for large average offsets (A_2 large) and/or large setup noise (B_2 large), Grubbs harmonic rule is better. Here "large" and "small" are terms relative to the process variance σ_v^2 . The advantage of the harmonic rule over the Kalman filter scheme decreases with increasing value of $B_1 = P_0/\sigma_v^2$.

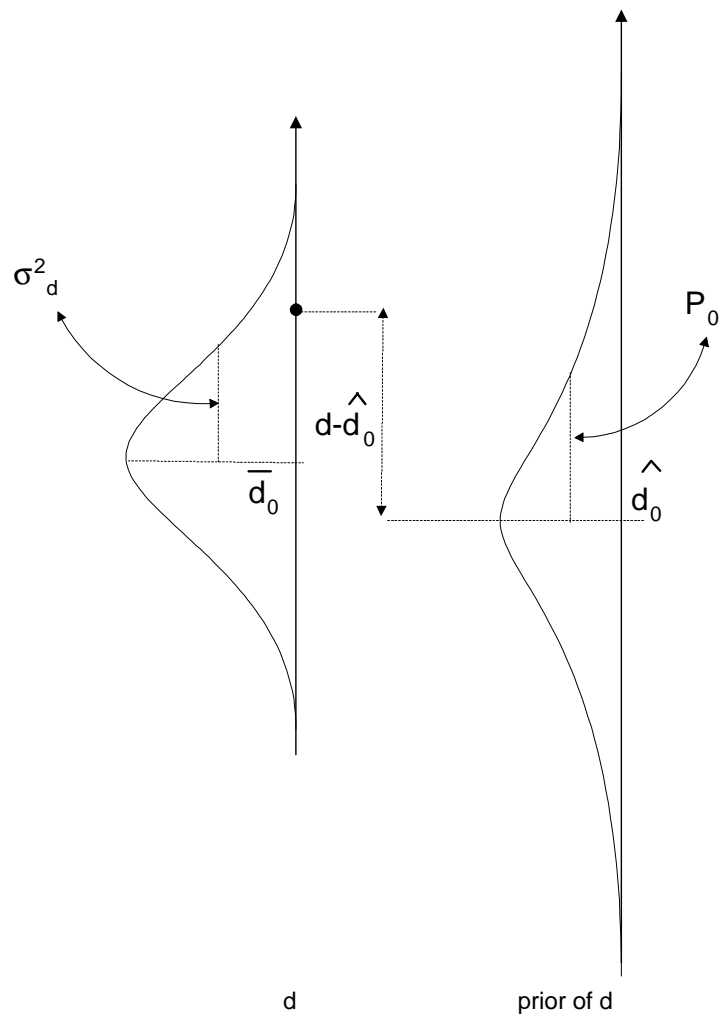


Figure 1. Starting the Kalman filter scheme, \bar{d}_0 and/or σ_d^2 unknown.

Note that under the assumptions in Case 1 above (when \bar{d}_0 and σ_v^2 are known), i.e., when we have that $B_1 = B_2$ and $A_2 = 0$, the Kalman filter method always dominates Grubbs rule. This agrees with our earlier comment which indicated that the Kalman filter scheme (and Grubbs extended rule) is optimal for the AISD_d criterion if the parameters are known.

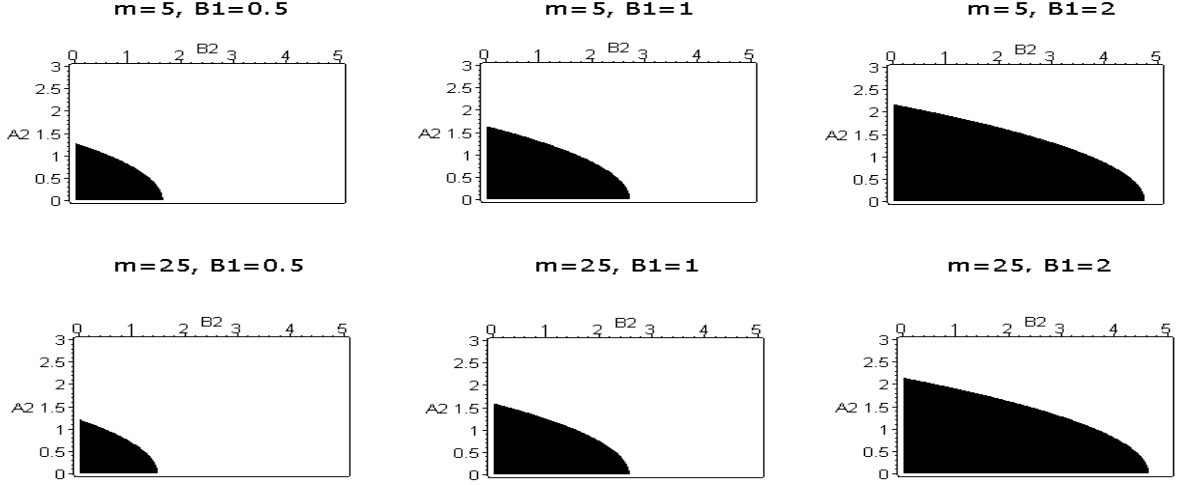


Figure 2. Kalman filter and Grubbs rule performance, random setup error. Shaded regions indicate cases for which $\text{AISD}_d(m)_{KF} < \text{AISD}_d(m)_G$. $B_1 = P_0/\sigma_v^2$, $B_2 = \sigma_d^2/\sigma_v^2$, $A_2 = |d_0 - \hat{d}_0|/\sigma_v$.

Figures 3 and 4 compare the AISD_d performance of the Kalman filter approach with that of an Integral controller with $\lambda = 0.2$ and $\lambda = 0.1$, respectively. As it can be seen, the Kalman filter scheme is to be preferred over more cases as the number of observations m increases. The integral controller should be preferred when the average offset is large (A_2 large) and/or the setup is very variable (large B_2). This is even more true as the confidence in the initial offset mean decreases (i.e., the larger B_1). Observe how for cases where the average offset is very small the integral controller also dominates the Kalman filter approach.

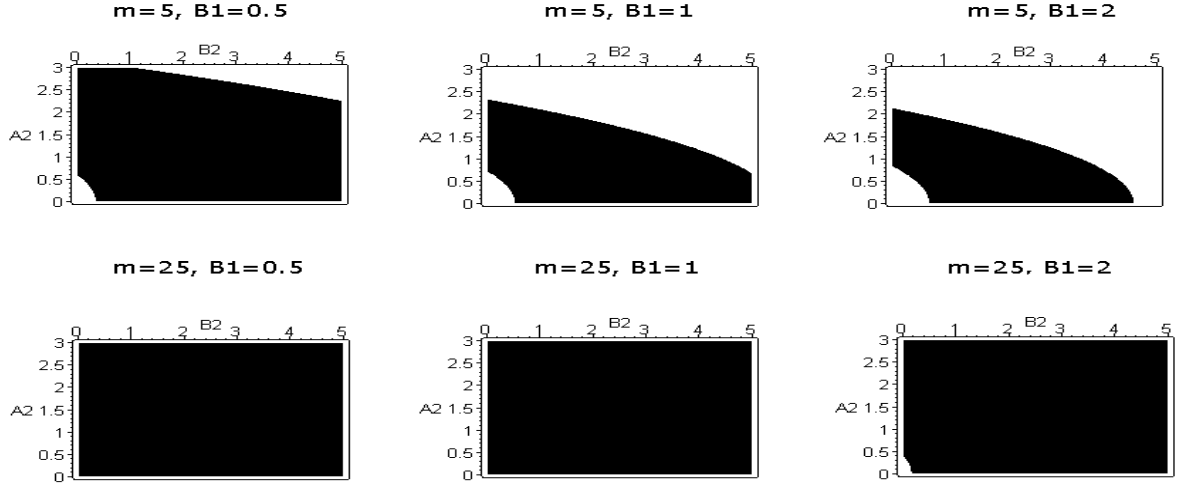


Figure 3. Kalman filter and discrete integral (EWMA) controller (with $\lambda = 0.2$) performance, random setup error. Shaded regions indicate cases for which $\text{AISD}_d(m)_{KF} < \text{AISD}_d(m)_{EWMA}$.

$$B_1 = P_0/\sigma_v^2, B_2 = \sigma_d^2/\sigma_v^2, A_2 = |d_0 - \hat{d}_0|/\sigma_v.$$

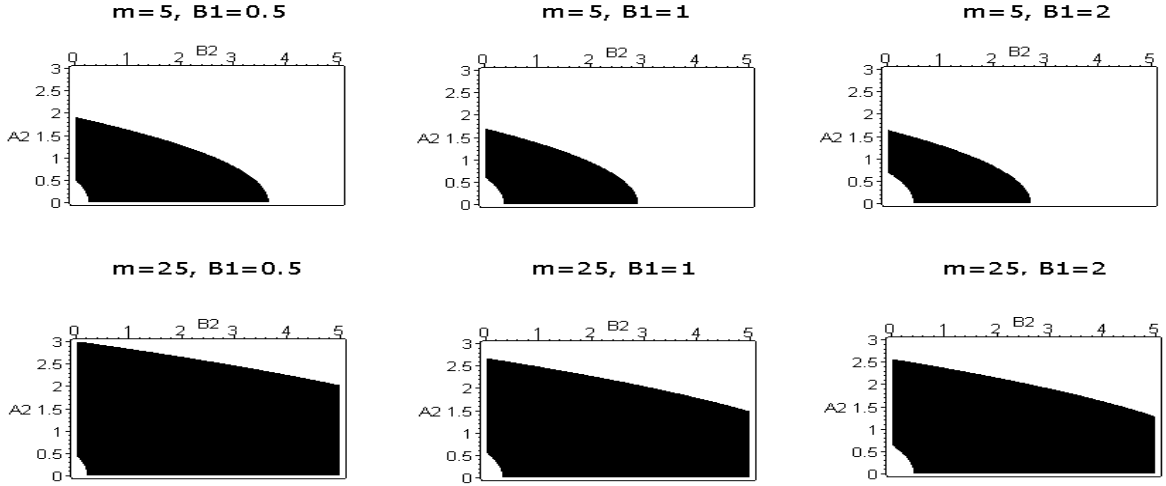


Figure 4. Kalman filter and discrete integral (EWMA) controller (with $\lambda = 0.1$) performance, random setup error. Shaded regions indicate cases for which $\text{AISD}_d(m)_{KF} < \text{AISD}_d(m)_{EWMA}$.

$$B_1 = P_0/\sigma_v^2, B_2 = \sigma_d^2/\sigma_v^2, A_2 = |d_0 - \hat{d}_0|/\sigma_v.$$

Finally, Figure 5 shows the AISD_d comparison between Grubbs' harmonic rule and an integral controller. The integral controller outperforms the harmonic rule for cases near the origin, when A_2 is small (small average error in offset estimate) and B_2 is small (low setup variance). As the sample size increases, Grubbs' harmonic rule dominates the integral controller scheme.

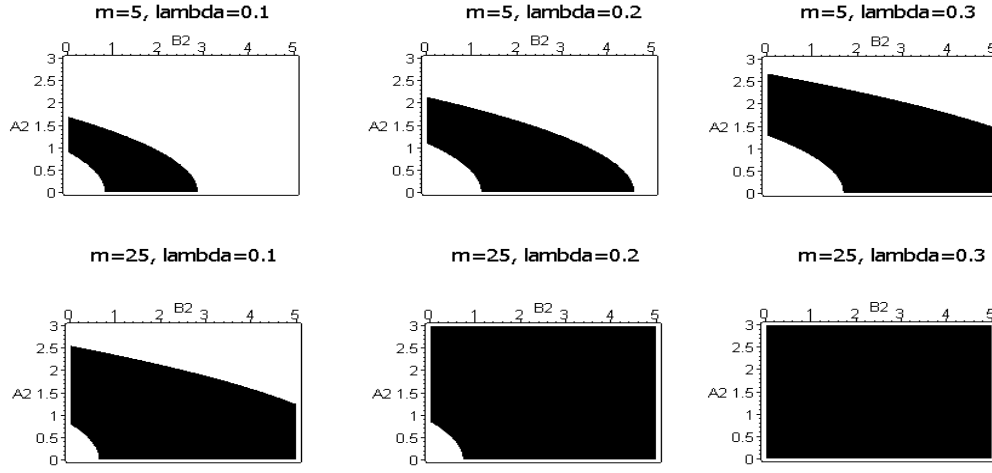


Figure 5. Grubbs harmonic rule and discrete integral (EWMA) controller performance, random setup error. Shaded regions indicate cases for which $\text{AISD}_d(m)_G < \text{AISD}_d(m)_{\text{EWMA}}$.

$$B_2 = \sigma_d^2 / \sigma_v^2, A_2 = |d_0 - \hat{d}_0| / \sigma_v.$$

5. CONCLUSIONS AND FURTHER RESEARCH.

The small sample properties of Grubbs (1983) adjustment schemes and that of an integral controller were analyzed for the case a setup error is systematic (non-random) and when it is a random variable with unknown mean and variance. The performance metric used was a quadratic off-target cost.

In the setup error is an unknown constant it was shown that for most practical cases when sequential adjustment is necessary, Grubbs (1983) harmonic rule represents a better strategy than the Kalman filter scheme (and therefore it performs better than Grubbs' extended rule). The even simpler integral or EWMA controller with weight $\lambda = 0.2$ provides a competitive alternative to the harmonic rule for cases when the offset is small (in the order of less than one standard deviation of the process). It was shown that these conclusions remain essentially unchanged if the performance is evaluated based not only while adjustments take place but also by considering additional runs in which no further adjustments are made to the process.

If the setup error is instead a random variable, an integral controller performs better than the Kalman filter scheme when the setup noise is relatively high and the offset is very large on average. When the offset is large and/or the setup noise is large, Grubbs harmonic rule outperforms the Kalman filter scheme. The analytic formulae in Appendix B allow to obtain similar results for other values of the process and controller parameters without recourse to simulation. Further recommendations about when to use each method in the random setup error case can be reached by looking at Figures 2-5.

The main advantage of the discrete integral controller over the other methods considered in this paper is that it stays alert for compensating for further shifts in the process that occur while in manufacturing (i.e., not at setup). The weights in Grubbs rule, for example, would have to be reset every time a shift is detected. This points toward integration of the adjustment schemes with a detection mechanism that will trigger the adaptation of the weights, along the lines recently followed by Guo et al. (2000). Such integrated SPC/EPC approaches require further research. The important connection between Grubbs harmonic rule and Robbins and Monro's Stochastic Approximation method (1951) point out to a very large body of literature and theoretical results some of which may prove useful in process adjustment applications.

Acknowledgements.- This paper was partially funded by NSF grant DMI 9996031.

APPENDIX A: MEAN AND VARIANCE OF THE QUALITY CHARACTERISTIC AND AISD FORMULAE FOR AN UNKNOWN CONSTANT SETUP ERROR.

The set point at time t is given by $U_t = -\sum_{i=1}^t K_i Y_i - \hat{d}_0$. Thus, $Y_t = d - \hat{d}_0 - \sum_{i=1}^{t-1} K_i Y_i + v_t$. So, for $t = 1$ we have that

$$Y_1 = d - \hat{d}_0 + v_1 \quad (14)$$

from which $E[Y_1]/\sigma_v = (d - \hat{d}_0)/\sigma_v = A$ and $V[Y_1]/\sigma_v^2 = 1$ (compare with eqs. 8 and 9). Given that

$$K_t = \frac{1}{\sigma_v^2/P_0 + t},$$

we have that

$$Y_t = \frac{(d - \hat{d}_0) \frac{\sigma_v^2}{P_0}}{\frac{\sigma_v^2}{P_0} + 1} \prod_{i=2}^{t-1} \left(1 - \frac{1}{\frac{\sigma_v^2}{P_0} + i} \right) + f_1 v_1 + f_2 v_2 \dots + f_{t-1} v_{t-1} + v_t \quad (15)$$

where the quantities $\{f_k\}_{k=1}^{t-1}$ are functions of $\frac{\sigma_v^2}{P_0}$ but not of d or \hat{d}_0 . After some algebra, equation (15) simplifies to:

$$Y_t = \frac{(d - \hat{d}_0) \frac{\sigma_v^2}{P_0}}{\frac{\sigma_v^2}{P_0} + 1} \left(\frac{\frac{P_0}{\sigma_v^2} + 1}{t \frac{P_0}{\sigma_v^2} - \frac{P_0}{\sigma_v^2} + 1} \right) + f_1 v_1 + f_2 v_2 + \dots + v_t.$$

Taking expected value and variance in the previous expression results in equations (8) and (9).

An expression for

$$\text{AISD}(m) = \frac{1}{m\sigma_v^2} \sum_{i=1}^m (V[Y_i] + E[Y_i]^2) \quad (16)$$

can be obtained after some algebra from properties of the sums involved. Define

$$C_1 = \frac{\Psi'(1/B_1) - \Psi'(m + 1/B_1)}{B_1^2 m}$$

and

$$C_2 = \frac{B_1[\Psi(m + \frac{1}{B_1}) - \Psi(\frac{1}{B_1})] + \Psi'(m + \frac{1}{B_1}) - \Psi'(\frac{1}{B_1})}{B_1 m} + 1$$

where $\Psi(x) = d \ln \Gamma(x)/dx$ (Psi or digamma function), and $\Psi'(x) = d\Psi(x)/dx$ (trigamma function). If a software is available that computes these functions, finding the AISD is facilitated by the formula. Then, equation (16) can be written as:

$$\text{AISD}(m) = C_1 A^2 + C_2. \quad (17)$$

Taking limit as $B_1 \rightarrow \infty$, we get a simple formula for the AISD given by Grubbs harmonic rule:

$$\text{AISD}_G(m) = 1 + \frac{\Psi(m) + \gamma + A^2}{m} \equiv A^2/m + C_3$$

where $\gamma \approx 0.5772156$ is Euler's constant and

$$C_3 = 1 + \frac{\Psi(m) + \gamma}{m}.$$

For the discrete integral (or EWMA) controller, the corresponding expression for the AISD is given directly from equations (10) and (11):

$$\text{AISD}_{\text{EWMA}}(m) = \frac{2}{2 - \lambda} + \left(\frac{1 - (1 - \lambda)^{2m}}{m(2 - \lambda)} \right) \left(\frac{A^2}{\lambda} - \frac{1}{2 - \lambda} \right). \quad (18)$$

Adjusting only the first few items.

Closed-form expressions for $\text{AISD}(m, N)$ can be obtained by direct substitution of the $\text{AISD}(m)$ formulae presented above into:

$$\text{AISD}(m, N) = \frac{m \text{AISD}(m)}{N} + \frac{(N - m)(V[Y_{m+1}] + E[Y_{m+1}]^2)}{N \sigma_v^2}.$$

After some algebra, for the Kalman filter scheme this turns out to be:

$$\text{AISD}_{\text{KF}}(m, N) = \frac{m}{N} (A^2 C_1 + C_2) + \frac{N - m}{N} \left(1 + \frac{m B_1^2 + A^2}{(1 + m B_1)^2} \right) \quad (19)$$

where C_1 , and C_2 are the functions of m and B_1 defined before. If $B_1 \rightarrow \infty$, one gets the corresponding expression for Grubbs harmonic rule:

$$\text{AISD}_G(m, N) = 1 + \frac{A^2 + \gamma + \Psi(m) - 1}{N} + \frac{1}{m}. \quad (20)$$

Finally, for the discrete integral (EWMA) controller this turns out to be

$$\begin{aligned} \text{AISD}_{\text{EWMA}}(m, N) &= \frac{1}{N} \left\{ \frac{2m}{2-\lambda} + \frac{1-(1-\lambda)^{2m}}{2-\lambda} \left(\frac{A^2}{\lambda} + \frac{1}{2-\lambda} \right) \right\} \\ &+ \frac{1}{N} \left\{ (N-m) \left[\frac{2-\lambda(1-\lambda)^{2m}}{2-\lambda} + (1-\lambda)^{2m} A^2 \right] \right\}. \end{aligned} \quad (21)$$

APPENDIX B. AISD_d (RANDOM SETUP ERROR) FORMULAE.

Assume that \bar{d}_0 and/or σ_d^2 are unknown. We have that $\text{AISD}_{KF}(m) = C_1 A^2 + C_2$. The expression for $\text{AISD}_d(m)$ is obtained from its definition as follows

$$\begin{aligned} \text{AISD}_d(m)_{KF} &= \int_{-\infty}^{\infty} \text{AISD}_{KF}(m) f_d(x) dx \\ &= C_1 \int_{-\infty}^{\infty} \frac{(d - \hat{d}_0)^2}{\sigma_v^2} f_d(x) dx + C_2 \int_{-\infty}^{\infty} f_d(x) dx \\ &= \frac{C_1}{\sigma_v^2} \int_{-\infty}^{\infty} (d^2 - 2d\hat{d}_0 + \hat{d}_0^2) f_d(x) dx + C_2 \end{aligned}$$

Since $\int_{-\infty}^{\infty} d^2 f_d(x) dx = \sigma_d^2 + \bar{d}_0^2$ and $\int_{-\infty}^{\infty} d f_d(x) dx = \bar{d}_0$, then

$$\begin{aligned} \text{AISD}_d(m)_{KF} &= \frac{C_1}{\sigma_v^2} (\sigma_d^2 + (\bar{d}_0 - \hat{d}_0)^2) + C_2 \\ &= C_1 (B_2 + A_2^2) + C_2. \end{aligned}$$

The $\text{AISD}_d(m)$ formula for Grubbs' harmonic rule is obtained in a similar way. Note that the AISD_d depends only on the mean and variance of d . For the integral (or EWMA) controller, the corresponding expression is

$$\text{AISD}_d(m)_{\text{EWMA}} = \frac{2}{2-\lambda} + \frac{1-(1-\lambda)^{2m}}{(2-\lambda)m} \left[\frac{B_2 + A_2^2}{\lambda} - \frac{1}{2-\lambda} \right]. \quad (22)$$

It is possible to obtain $\text{AISD}_d(m, N)$ formulae for when only the first few adjustments are implemented using each method, and this can be useful to determine the optimal number of adjustments when considering the performance of the adjustment scheme under random setup errors.

REFERENCES.

Box, G.E.P., Luceño, A. (1997). *Statistical Control by Monitoring and Feedback Adjustment*. John Wiley & Sons, New York, NY.

Del Castillo, E. (2001), "Some properties of EWMA feedback quality adjustment schemes for drifting disturbances", *Journal of Quality Technology*, 33(2), 153-166.

Del Castillo, E., Pan, Rong, and Colosimo, B.M. (2002). “An Unifying View of Some Process Adjustment Methods”. Submitted for publication. Technical Paper, Applied Statistics Laboratory, Department of Industrial Engineering, Penn State University.

Grubbs, F.E., (1983). “An Optimum Procedure for Setting Machines or Adjusting Processes,” *Journal of Quality Technology*, 15, 4, pp. 186-189.

Guo, R.S., Chen, A., and Chen, J.J., (2000). “An Enhanced EWMA Controller for Processes Subject to Random Disturbances”, in Moyne, J., Del Castillo, E., and Hurwitz, A., *Run to run control in semiconductor manufacturing*, CRC press.

Robbins, H. and Monro, S. (1951). “A Stochastic Approximation Method”. *Annals of Mathematical Statistics*, 22, pp. 400-407.

Trietsch, D., (1998). “The Harmonic Rule for Process Setup Adjustment with Quadratic Loss”, *Journal of Quality Technology*, 30, 1, pp. 75–84.