IIE Transactions (1999) **31**, 517–527

Integrated design of run-to-run PID controller and SPC monitoring for process disturbance rejection

FUGEE TSUNG[1] and JIANJUN SHI[2]

[1]*Department of Industrial Engineering and Engineering Management, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong*
E-mail: season@ust.hk
[2]*Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor, MI 48109-2117, USA*
E-mail: shihang@engin.umich.edu

Received July 1997 and accepted June 1998

An integrated design methodology has been developed for a run-to-run PID controller and SPC monitoring for the purpose of process disturbance rejection. In the paper, the process disturbance is assumed to be an ARMA (1,1) process. A detailed procedure is developed to design a PID controller which minimizes process variability. The performance of the PID controller is also discussed. A joint monitoring of input and output, using Bonferroni's approach, is then designed for the controlled process. The ARL performance is studied. One major contribution of the paper is to develop a complete procedure and design plots, which serve as tools to conduct all the aforementioned tasks. An example is provided to illustrate the integrated design approach.

1. Introduction

The concept of integrating the Automatic Process Control (APC) and Statistical Process Control (SPC) techniques was introduced approximately ten years ago [1–3], and since these initial papers a considerable amount of progress has been repored in the literature.

MacGregor [1] and also Box and Kramer [2] have presented overview descriptions of the concepts associated with APC/SPC integration by assuming a first-order Integrated Moving Average (IMA(0, 1, 1)) disturbance process. Both of these papers have suggested the use of a minimum-cost strategy. They used SPC charts as dead bands for feedback-controlled processes, and adjusted the processes at the signal of Exponentially Weighted Moving Average (EWMA) charts. Del Castillo [4] has extended the dead-band idea to multivariate situations. English and Case [5] used SPC charts as filters for feedback controllers. They also considered the effect of a trend or ramp disturbance on a feedback-controlled process [6].

In addition, Run-To-Run (RTR) process control techniques that combine APC and SPC concepts have been developed and applied to the disturbance rejection of semiconductor manufacturing processes [7–9]. An RTR process refers to a process, such as a wafer-etching process or an auto-body stamping process, in which a control action, e.g., a change in a process parameter, can only be implemented between runs instead of during a run. The objective of RTR control is to reject various disturbances frequently found in RTR processes, such as shifts and trends, as well as autocorrelated disturbances. SPC acts as a supervisor indicating the need for RTR control action. A recent review of RTR control can be found in Del Castillo and Hurwitz [10].

This paper builds on previous work [11–13] and places a different emphasis on APC/SPC integration by using APC and SPC to perform the separate functions of control and monitoring respectively. Van der Wiel *et al.* [11] have proposed an algorithmic statistical process control (ASPC) scheme whose goal is to reduce predictable quality variations using APC and then monitor the system to detect and remove unpredictable variations using SPC. A batch polymerization example was used to illustrate the ASPC scheme, where an Auto Regressive Moving Average (ARMA(1,1)) disturbance process was assumed and a Minimum-Mean-Squared-Error (MMSE) control scheme as well as a CUSUM chart monitoring the process output were applied.

Although the concept of integrating APC and SPC has been proposed and discussed for years, little literature can be found to demonstrate the integrated design procedures of APC and SPC for process disturbance rejection. This paper proposes a framework and methodology to develop an integrated design of a Proportional-Integral-Derivative (PID) controller and its associated SPC charts for process disturbance rejection. In this paper, a PID controller, rather than MMSE, is used as the basic automatic

process control strategy to minimize process variability. This effort is motivated by the fact that in industrial practice, the PID controller is by far the most common controller [14], and many process control devices in industry are only equipped with a PID controller rather than an MMSE controller. Also, a PID controller is more robust with regard to process models [15] and can handle non-minimum phase processes. A non-minimum phase process is a process that has zeros outside the unit disc [16]. The application of control theory to this situation suggests that the PID controller is superior to the MMSE controller since the MMSE controller will have unstable modes and therefore can not be used in practice [16]. Therefore, we focus on the problem of minimizing process variability under the constraint of using a PID control strategy.

It is reported in the SPC literature that an in-control process is stable and predictable over time, whereas an out-of-control process has an unpredictable amount of variation [17]. In this research, an in-control process is defined as being stable and predictable by a stationary model, such as an ARMA model, of which an i.i.d. model is a special case. A process that is shown to be unstable and unpredictable by the model is an out-of-control process. Thus, type I error (concluding the process is out of control when it is really in control) and type II error (concluding the process is in control when it is really out of control) are still useful under our generalised definitions.

In this paper, the process disturbance is modeled as an ARMA (1,1) process. A PID controller is designed to minimize the process variation. Then, based on the model parameters, an SPC control chart is designed to monitor the PID-controlled process. Performance measures of both the PID controller and the SPC charts, using the Absolute Efficiency (AE), Relative Efficiency (RE) and the Average Run Length (ARL), are presented. A general framework of the proposed integrated design approach is shown in Fig. 1.

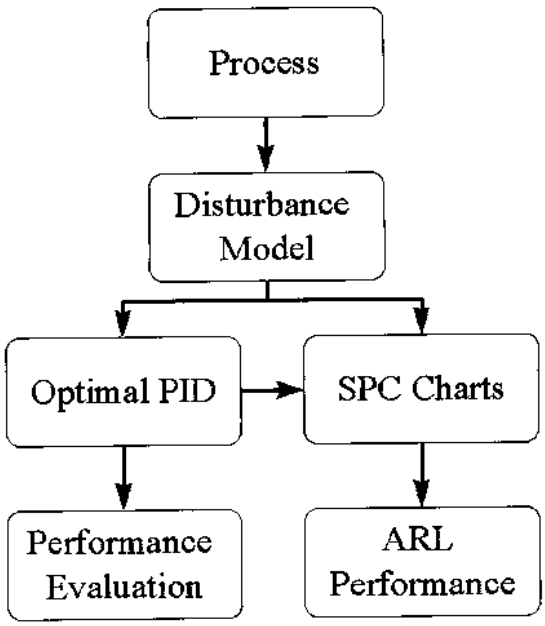

Fig. 1. The framework of the proposed integrated design of PID and SPC strategies.

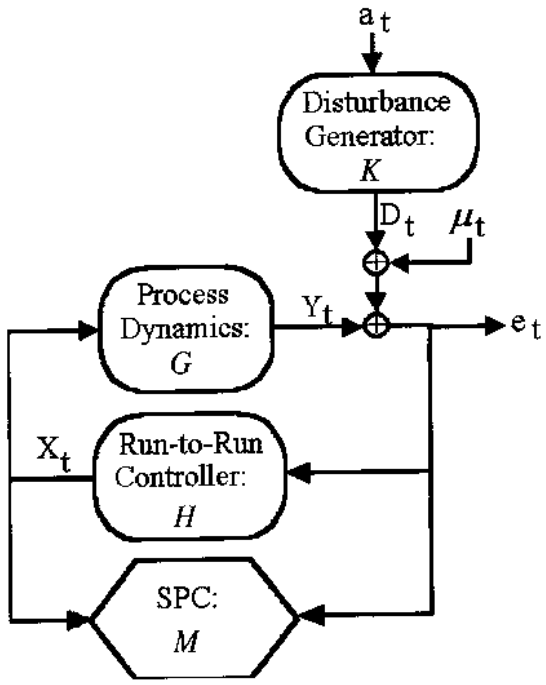

Fig. 2. The process.

2. Process model

The process is shown in Fig. 2. In the figure, G is the process dynamics, H is the RTR controller, K is the disturbance generator, and M is the SPC monitoring scheme.

RTR processes differ from conventional control schemes used within runs. The RTR control responds to post-process measurements of certain quality characteristics by modeling the disturbance process between runs, and then providing new set-points for use in the next run of the process. Consider a process under RTR feedback control, and assume without loss of generality that the target value is zero. The process output e_t, which can then be viewed as the deviation from the target, is the sum of three components: the output from the process dynamics Y_t, the process disturbance D_t, and the mean shift μ_t. Let X_t denote the control action, that is, the run-to-run adjustment of the process set-point, with the initial input assumed to be zero. Also, let B be the usual backward shift operator, i.e., $Ba_t = a_{t-1}$. Throughout the paper, a dynamic model $G = B$ for process output will be considered: $Y_t = X_{t-1}$, which means that the output at run t

depends on the input variables at the end of run $t-1$, or equivalently, at the start of run t. This implies that the control action immediately has its full effect on the process output in one run or one batch, which is reasonable in discrete part manufacturing and is widely used in the RTR control literature [10].

The output e_t can be written as:

$$e_t = Y_t + D_t + \mu_t = X_{t-1} + D_t + \mu_t, \qquad (1)$$

where μ_t can be any type of mean shift such as a step, ramping, or cyclical change.

It is known that most of the process disturbances can be modeled as a stationary time-series process, which is stable and invertible in general. In this situation, both an AR model and an ARMA model can be used to describe the process. In many cases, an ARMA model is more efficient (needs less parameters) compared with an AR model in describing a stationery process [18]. Also an ARMA (1, 1) model can be used to approximate a broad class of disturbances with different ϕ and θ. In particular, MacGregor [19] has pointed out that an AR(1) disturbance process with white noise measurement error is equivalent to an ARMA(1,1) process. In this paper, the disturbance process assumes D_t to be a stationary ARMA(1,1) process, i.e., $K = (1-\theta B)/(1-\phi B)$:

$$D_t = \frac{1-\theta B}{1-\phi B} a_t, \qquad (2)$$

where a_t represents white noise. It is assumed that the parameters ϕ and θ of the ARMA(1,1) processes satisfy the conditions $|\phi| < 1$ and $|\theta| < 1$. Although some of the following results can be extended to other situations, for simplicity, attention in this paper will be restricted to the ARMA(1,1) stationary disturbance process.

Note that in this paper, the disturbance process is modeled as a stationary stochastic process. This disturbance is part of the inherent process, which is the cumulative effect of many small, essentially unavoidable causes called "common causes". Here the "in-control" condition of a run-to-run process is defined as having no mean shift or variance changes in the disturbance model. The impacts of those disturbances can be reduced by using the run-to-run control strategy, which is why a PID control is needed.

Other kinds of variability, such as the mean shifts μ_t, may also occasionally be present. These are called "special causes". Thus, the "out-of-control" condition is defined as having a mean shift or variance change in the disturbance model. Although in this situation the run-to-run feedback control can partially compensate for the mean shift on the output, SPC monitoring is desired to detect this mean shift, and eventually eliminate the root cause of this mean shift. Here the definitions of "in-control" and "out-of-control" conditions for a run-to-run process are of more practical use than conventional SPC definitions, where an "in-control" process is usually assumed to be an independent and identically distributed (i.i.d.) process [20].

The purpose of the following research is to develop an integrated design approach to: (1) design a PID controller (H) to minimize the process variation due to the process disturbance; and (2) design a SPC chart (M) to monitor the PID-controlled process.

3. Design of a PID controller

The PID control schemes studied in this paper are of the form:

$$X_t = -k_P e_t - k_I \frac{1}{1-B} e_t - k_D(1-B)e_t, \qquad (3)$$

where k_P, k_I, and k_D are constants. Here k_P determines the amount of proportional adjustment, k_I determines the amount of integral adjustment, and k_D determines the amount of derivative adjustment [21].

The purpose of PID control is to manipulate the process input so as to minimize the process variation due to process disturbance. Thus, the concept of the PID controller design used in this paper is different to conventional PID controller design [14,22], in which the performance of the servo system is the primary focus and the disturbance is usually considered as white noise.

3.1. *PID controller design*

The primary objective in quality improvement is to reduce process variation and the Mean-Squared-Error (MSE) of the process outputs is used as an optimization index to serve this objective. Also, optimizing the process MSE is equivalent to minimizing the quality loss of this process since the MSE criterion can meaningfully approximate the quality loss of most industrial processes [23]. A study has been performed by Tsung [24] in which he derived the relationship between the PID control parameters and the process MSE for an ARMA (1,1) process. The relationship can be described as:

$$MSE(k_P, k_D, k_I, \phi, \theta) \equiv (I + 2\rho_1(II + III))\sigma_D^2, \qquad (4)$$

where I, II, and III are functions of the PID parameters and disturbance parameters and can be calculated by substituting into either Equation (A10) or (A14), that are discussed in Appendix A. Also, ρ_1 is the first-order autocorrelation of $\{D_t\}$ and σ_D^2 is the variance of $\{D_t\}$, both of which can be easily obtained during disturbance process modeling [21]. Note that both disturbance process modeling and parameter estimation are required for this design and they can be obtained by open loop data collection and time series model fitting. The model fitting of an ARMA process by a three-stage iterative procedure based on identification, estimation, and diagnostic checking is described extensively in Box *et al.* [21].

Based on the optimization index, PID controllers have been designed for operation across the entire parameter domain of the ARMA (1, 1) model. The design of the PID parameters k_P, k_I, and k_D are obtained by minimizing *MSE* within the stability region:

$$(k_P, k_I, k_D) = \begin{cases} k_I \geq 0, \\ k_P + k_I/2 + 2k_D < 1, \\ -1 < k_D < 1, \\ -k_D(1 + k_P + k_I) - k_P < 1. \end{cases} \quad (5)$$

where, to ensure process stationarity, the stability region is based on the results in Box *et al.* [21] and Tsung [24]. The MSE in (4) is a complicated function of the parameters, so numerical methods are used for its optimization. Because the regions of stability for k_P, k_I, and k_D are reasonably small, a grid search method is used to determine the optimum values.

To present the choices of k_P, k_I, and k_D in a two-dimensional parameter space, contour plots are used to show the computational results. The contour plots are graphic representations with lines (as on a map) connecting the points on a response surface that have the same elevation or response value. Here, the vertical axis is the disturbance parameter ϕ, and the horizontal axis is the disturbance parameter θ. The response value is the corresponding control parameter, which is labeled on each contour line.

The PID parameters and their relationships are summarized in the PID design maps shown in Fig. 3(a–c). As an illustration, if a process disturbance model is given as $\phi = 0.86$ and $\theta = 0.16$, the PID parameters can be obtained from the design maps as $k_P = 0.24$, $k_I = 0.58$, and $k_D = -0.1$. This process model will be further used to demonstrate the integrated design procedure. Note from Fig. 3b that, if $\phi \leq 0.5$ or if $\phi \leq \theta$, the value of k_I is essentially zero for the PID schemes. In this region, PID control coincides with PD control. Only in the complementary region are the PID and PD schemes different. In control theory, when $\phi < \theta$ the transfer function of a control system has its pole closer to the image axis than zero. In this situation, Integral (I) control action is not needed [16]. Near the region $\phi = \theta$, the choices of k_P, k_I and k_D are close to zero. This is because when $\phi = \theta$, the ARMA(1,1) model in (2) becomes a white noise process, and under this situation the best control strategy is to not adjust the process. This is consistent with Deming's philosophy [25].

3.2. *PID controller performance evaluation*

The traditional control settings can be selected with different emphasis, such as stability margin, robustness to model error, smaller transient time period, or high gain for disturbance rejection, etc. It is hard to make a comprehensive performance comparison with traditional control settings. This research emphasizes the minimization of process variation using a PID controller. Thus, we compare the proposed PID control scheme with no-control and MMSE control, as measures of lower and upper bounds. Thus, two criteria are used to assess the efficiency of the proposed control scheme. The first is the Relative Efficiency (RE) defined by

$$RE \equiv \frac{MSE_{\text{No-Control}}}{MSE_{\text{PID}}}, \quad (6)$$

a comparison of the variability of the control scheme of interest with the variability of the no-control strategy. Because the no-control strategy is a special case of the PID scheme with $k_P = k_I = k_D = 0$, it is clear that $RE \geq 1$. The values of RE give us a measure of the improvement over the no-control strategy, i.e., the lower bound of the control performance.

The second criterion is the Absolute Efficiency (AE) measured by

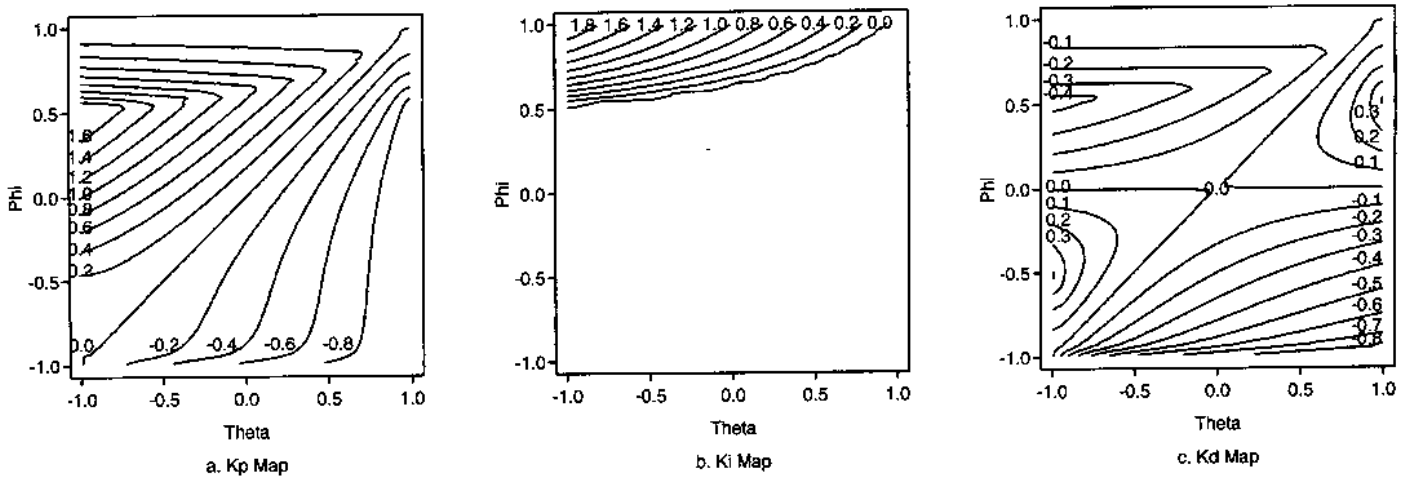

Fig. 3. PID design maps for ARMA(1, 1) disturbance processes.

$$AE \equiv \frac{MSE_{\text{MMSE}}}{MSE_{\text{PID}}}, \tag{7}$$

which compares the performance of a given control scheme with the MMSE scheme. Harris [26] was the first to suggest the use of this ratio to assess control performance. The use of this criterion has subsequently attracted considerable attention [27]. Note that $AE \leq 1$ as the MMSE scheme minimizes the mean squared error (under the assumed model and when the model parameters are known). Although this may not be achievable in practice, AE still provides a benchmark in terms of how close one can come to the best possible performance in terms of variation reduction, i.e., the upper bound of the control performance.

Here the contour plots are used again to present the AE and RE in a two-dimensional parameter space. The response value is the corresponding AE or RE, which is labeled on each contour line. Figure 4(a and b) contains the AE and RE contour plots for different values of the ARMA(1,1) parameters.

As seen from Fig. 4(a and b), the PID schemes perform well. The absolute efficiency is greater than 90% for most of the ARMA(1,1) parameter space. Only when ϕ is close to -1 is the AE not as good, but the RE is still much larger than 1.

4. Design of SPC charts for the associated PID-controlled process

In general, data from a PID-controlled process are autocorrelated. Various research has been performed on the topic of autocorrelated SPC [28–31]. However, a PID-controlled process has its unique features, especially when the integral (I) control mode takes action. It is known from the control theory literature that if there is an integral control involved in the controlled process, a steady-state mean shift of the process outputs will be eliminated immediately after a transient time period occurring immediately after a process change. Thus, there is only a limited "window of opportunity" during which the process change must be detected [32]. All conventional SPC control charts suffer from this problem. One possible solution for this is to monitor the inputs along with the outputs as the mean shifts may lead to some unusually large inputs created by feedback.

Therefore, we propose to jointly monitor both the output and input using bivariate SPC to improve the efficiency of detection. Note that the proposed joint monitoring strategy can be implemented using any bivariate SPC scheme [33]. However, for the purpose of demonstration, Bonferroni's approach is investigated.

Bonferroni's approach is known to be a conservative procedure, which has been recommended by Alt [34] because of its simplicity in determining which variables are responsible for an out-of-control condition. This approach is applied to monitor multiple characteristics simultaneously, while using Bonferroni's inequity to control the overall error probabilities. This approach, unlike most of the multivariate SPC techniques, does not need to assume a known covariance structure among process variables. Thus, it is a popular and simple charting method for use in multivariate processes where covariance information is either missing or is uncertain. As opposed to conventional SPC chart design, where the

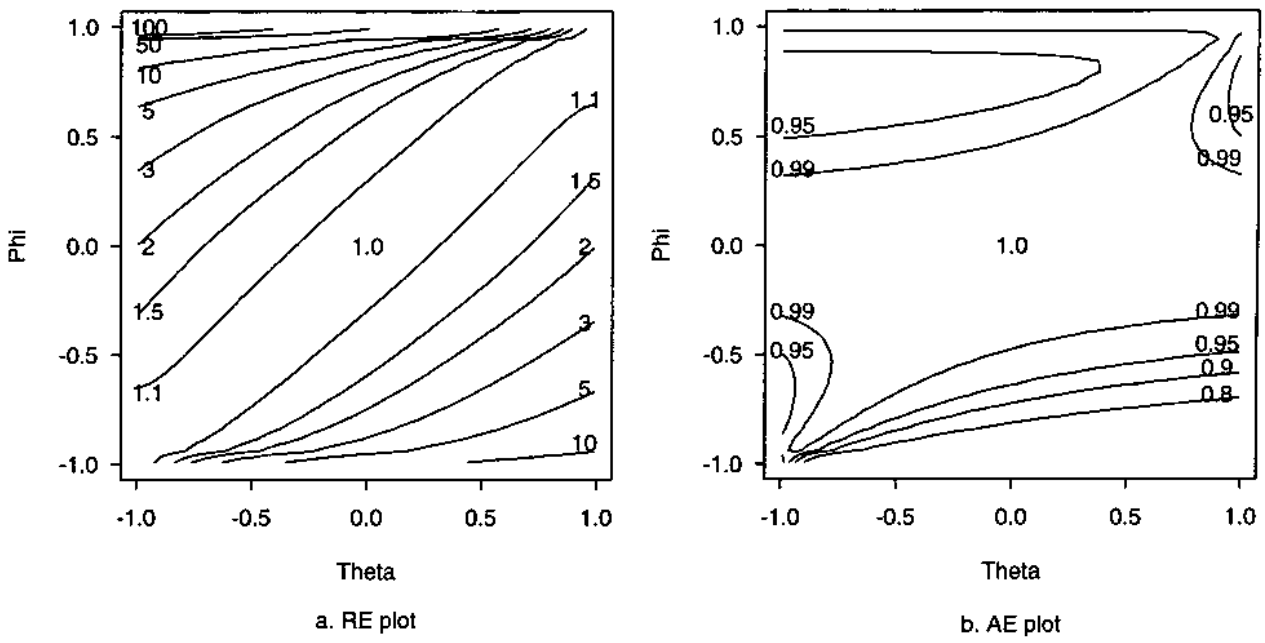

Fig. 4. PID performance evaluation plots.

process mean and standard deviation are obtained from process observations, the proposed SPC chart design is directly based on the PID-controlled process model.

4.1. *Joint monitoring of the output and input of a PID-controlled process*

From control theory and selected PID control parameters, the standard deviation of the output after process control can be described as

$$\sigma_e = \sqrt{I + 2\rho_1(II + III)}\sigma_D. \tag{8}$$

Similarly, the standard deviation of the manipulated input can be described as

$$\sigma_X = \sqrt{I' + 2\rho_1(II' + III')}\sigma_D, \tag{9}$$

where I', II', and III' are also functions of PID parameters and disturbance parameters, and can be calculated by substituting into Equation (A20) or (A23), provided in Appendix B.

Based on (8) and (9), joint charts can be designed using Bonferroni's approach. Note that the center lines of the joint charts are the means of e_t and X_t, which are assumed to be zero. The control limits are set at the value of the constants L_e and L_X time standard deviations (σ_e and σ_X) above and below the center lines. Here the constants are chosen as:

$$L_e = L_X = z_{(1-\alpha/4)}. \tag{10}$$

Note that the process data is assumed to be normally distributed, and the constants are modified from $z_{(1-\alpha/2)}$ to $z_{(1-\alpha/4)}$ using Bonferroni's inequity to control the overall type I error probabilities [34]. The control limits of the joint charts (CL_e and CL_X) can then be written as

$$\begin{aligned} CL_e &= \pm L_e\sigma_e, \\ CL_X &= \pm L_X\sigma_X. \end{aligned} \tag{11}$$

The joint decision rule suggests that the controlled process is out of control when either the controlled output or the manipulated input is outside the control limits.

From (8), (9), and (11), we can see that the control limits are a function of PID control parameters and thus the original ARMA (1,1) parameters. As a result, for a given ARMA (1,1) model, PID control parameters can be selected using methods proposed in the last section. Associated SPC chart parameters can then be determined using (11).

4.2. *Performance analysis of the developed SPC charts*

A performance measure study using the ARL is presented in this section. In order to more effectively compare the performance of the monitoring strategy, the control limits of the control charts are manipulated so that the ARL, when there is no shift in the mean, is the same for all charts (for example, an in-control ARL = 370 is used in this study). Thus the chart with the lowest out-of-control ARL when a shift in the mean has occurred is considered superior. This is analogous to matching the type I errors so that the type II errors can be compared in a more meaningful way. Also, the ARL is determined via Monte Carlo simulation. The simulation procedure is similar to that reported by Wardell *et al.* [31]: disturbance processes are generated according to (2), with a step change in the mean introduced at time zero. The observed control inputs and outputs are calculated by (3) and (1). The number of time periods is then measured until the first out-of-control condition is signaled for each chart. Our simulation data shows that the geometric assumption of the run-length distribution is acceptable. If the run length is geometrically distributed with a mean and a standard deviation equal to the ARL, then 10000 replications produce an estimated ARL with a standard error within 1%, which is reasonably precise. Thus, for each disturbance model parameter set, the process is simulated 10000 times in order to obtain the ARL value.

SPC performance plots in Fig. 5(a–c) contain the contour plots of the ARL of joint monitoring. Bonferroni's approach was used for the PID-controlled processes. The plots lead to the following conclusions: (1) for large mean shifts (e.g., $\mu_t = 2\sigma_D$), Figure 5a shows that Bonferroni's approach for the PID-controlled processes performs quite well (ARL < 10) for most of the parameter space. Especially when $\phi < \theta$, most of the ARL are less than 3; (2) for medium mean shifts (e.g., $\mu_t = 1\sigma_D$). Figure 5b indicates that Bonferroni's approach performs well when $\phi < \theta$, but it does not perform well (ARL > 50) when $\phi > \theta$; (3) for small mean shifts (e.g., $\mu_t = 0.5\sigma_D$) as shown in Fig. 5c, it performs well only when near the region of ϕ close to -1 and θ close to 1. It is known that Shewhart-type charts including Bonferroni's approach are not sensitive to small mean shifts. However, the power of small-shift detection of the proposed SPC monitoring strategy can be improved by replacing Bonferroni's approach with a more advanced multivariate SPC scheme such as multivariate CUSUM charts [35] and multivariate EWMA charts [36].

Note that Bonferroni's approach should not be used for the pure Proportional (P) controlled processes. For a pure P-controlled process, the manipulated inputs are proportional to the controlled outputs, so the simultaneous monitoring of the inputs and outputs are redundant. In this situation, the performance of joint monitoring degrades as the control limits and the corresponding type I error are wrongly designed by Bonferroni's inequity. Thus, when ϕ is close to zero and where the best choices of k_I and k_D are close to zero (see Fig. 3(a–c)), the joint monitoring approach is not recommended.

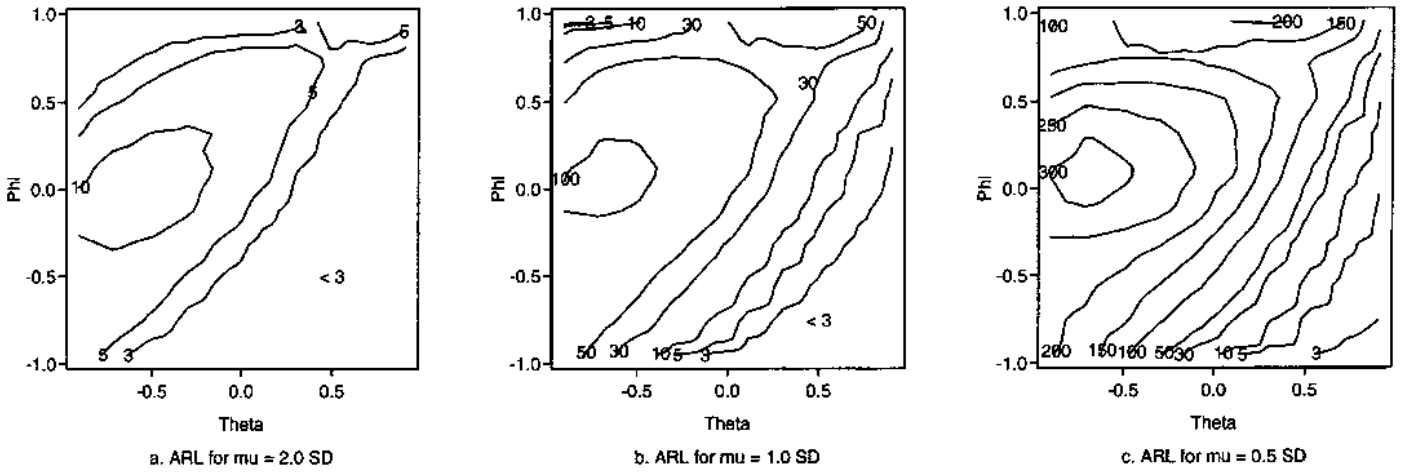

Fig. 5. SPC performance evaluation plots.

5. Integrated design procedure and an example

In this section, an example is used to demonstrate the applicability and efficiency of the integrated design. Van der Wiel *et al.* [11] have proposed an ASPC scheme and used a batch polymerization process as an example. In their example, an MMSE feedback control scheme was used for an ARMA(1,1) disturbance process with $\phi = 0.86$ and $\theta = 0.16$. Along with that, a conventional SPC was put in place to monitor the process output. We suggest using the PID control scheme for RTR feedback control on the same disturbance model with simulated data. Along with PID, joint monitoring using bivariate SPC is implemented. The step-by-step integrated design procedure is illustrated as follows:

(i) An ARMA(1,1) model with $\phi = 0.86$ and $\theta = 0.16$ is simulated for the disturbance process. In practice, an adequate disturbance model can be obtained by open-loop data collection and time-series model fitting.

(ii) Based on the ARMA(1,1) model and using the PID design maps in Fig. 3(a–c), a PID scheme is designed with $k_P = 0.24$, $k_I = 0.58$, and $k_D = -0.08$.

(iii) The corresponding control performance can be obtained from the PID performance evaluation plots in Fig. 4(a and b). Its *AE* is 0.95, very close to 1, which means it is very close to the best possible performance. Also, its *RE* is 3.0, which means it is three times better than a no-control strategy in terms of variation reduction.

(iv) Based on the model and control parameters, the joint monitoring design using Bonferroni's approach is determined by (11). The control limits of the joint monitoring charts are $CL_e = \pm 3.21$ and $CL_X = \pm 4.83$.

(v) The corresponding ARL performances can be obtained from the SPC performance evaluation plots in Fig. 5(a–c). The designed in-control ARL is about 370. For a larger mean shift ($\mu = 2\sigma_D$), the *ARL* is 3. For a smaller mean shift ($\mu = 1\sigma_D$), however, the performance of the Shewhart-type joint monitoring is not very good (*ARL* = 50). If the detection of a smaller mean shift is critical, some advanced joint monitoring schemes such as multivariate CUSUM charts [35] should be used.

The integrated design of the PID controller and SPC charts and its performance is summarized in Table 1. Overall, the proposed integrated design indicates that the PID controller is highly efficient and very close to the MMSE schemes. Furthermore, the associated joint monitoring scheme is efficient when the process mean shift is large.

6. Conclusions

Integration of the APC and SPC techniques is an emerging area which has attracted attention from both academia and industry. Successful integration of the APC and SPC approaches will provide better quality control

Table 1. The integrated design and its performance for an ARMA(1,1) disturbance process with $\phi = 0.859$ and $\theta = 0.164$

APC					*SPC*			
PID controller			*Control performance*		*Joint monitoring*		*ARL performance*	
k_P	k_I	k_D	*AE*	*RE*	CL_e	CL_X	$\mu = 1\sigma_D$	$\mu = 2\sigma_D$
0.24	0.58	−0.08	0.95	3.0	±3.21	±4.83	50	3

and process improvements in manufacturing. However, lack of research on the integrated design of those tools has proven to be a barrier to the implementation of the concept. In this paper, an integrated design procedure is developed to design a PID controller and its associated SPC monitoring algorithms for process disturbance rejection. The paper provides complete tools for the integrated design: for a given process disturbance model, a set of PID control parameters can be selected from the design maps, which minimize the process variability. PID evaluation plots can then be used to provide an assessment of the PID controller performance. After that, a joint SPC monitoring algorithm, based on Bonferroni's approach, can be obtained using provided equations. Finally, the obtained SPC performance can be evaluated according to its ARL using the provided SPC evaluation plots. A heuristic approach, with step-by-step procedures and examples, has been presented in Section 5. Although the research results are presented with the assumption that the process disturbance is an ARMA(1,1) model, other disturbance models can be studied following the same procedure. Thus, the approach used in this study is generic. The research presented in this paper is an initial attempt to provide tools for the integrated design of APC and SPC. There are many challenges and opportunities in this field. More effort is needed toward eventually integrating the APC and SPC approaches to produce process improvements in manufacturing.

Acknowledgments

The authors are grateful to the referees for their helpful comments. J. Shi was supported by NSF CAREER award: DMI 9624402.

References

[1] MacGregor, J.F. (1988) On-line statistical process control. *Chemical Engineering Process*, **10**, 21–31.

[2] Box, G.E.P. and Kramer, T. (1992) Statistical process monitoring and feedback adjustment—a discussion. *Technometrics*, **34**, 251–285.

[3] Montgomery, D.C., Keats, J.B., Runger, G.C. and Messina, W.S. (1994) Integrating statistical process control and engineering process control. *Journal of Quality Technology*, **26**, 79–87.

[4] Del Castillo, E. (1996) A multivariate self-tuning controller for run-to-run process control under shift and trend disturbances. *IIE Transactions*, **28**, 1011–1021.

[5] English, J.R. and Case, K.E. (1990) Control charts applied as filtering devices within a feedback loop. *IIE Transactions*, **22**, 255–269.

[6] English, J.R. and Case, K.E. (1994) Statistical process control in continuous flow processes—ramp disturbances. *IIE Transactions*, **26**, 22–28.

[7] Sachs, E., Hu, A. and Ingolfsson, A. (1995) Run-by-run process control: combining SPC and feedback control. *IEEE Transactions on Semiconductor Manufacturing*, **8**, 26–43.

[8] Butler, S.W. and Stefani, J.A. (1994) Supervisory run-to-run control of polysilicon gate etch using *in situ* ellipsometry. *IEEE Transactions on Semiconductor Manufacturing*, **7**, 193–201.

[9] Mozumder, P.K., Saxena, S. and Collins, D.J. (1994) A monitor wafer based controller for semiconductor processes. *IEEE Transactions on Semiconductor Manufacturing*, **7**, 400–410.

[10] Del Castillo, E. and Hurwitz, A.M. (1997) Run to run process control: review and extensions. *Journal of Quality Technology*, **29**, 184–196.

[11] Van der Wiel, S.A., Tucker, W.T., Faltin, F.W. and Doganaksoy, N. (1992) Algorithmic statistical process control: concepts and an application. *Technometrics*, **34**, 286–297.

[12] Tucker, W.T., Faltin, F.W. and Van der Wiel, S.A. (1993) Algorithmic statistical process control: an elaboration. *Technometrics*, **35**, 363–375.

[13] Faltin, F.W., Hahn, G.J., Tucker, W.T. and Van der Wiel, S.A. (1993) Algorithmic statistical process control: some practical observations. *International Statistical Review*, **61**, 67–80.

[14] Astrom, K.J. (1988) *Automatic Tuning of PID Controllers*, Instrument Society of America, Research Triangle Park, NC.

[15] Tsung, F., Wu, H. and Nair, V.N. (1998) On the efficiency and robustness of discrete proportional-integral control schemes. *Technometrics*, **40**, 214–222.

[16] Astrom, K.J. and Wittenmark, B. (1990) *Computer Controlled System: Theory and Design*, 2nd edn., Prentice Hall, Englewood Cliffs, NJ.

[17] Leitnaker, M.G., Sanders, R.D. and Hild, C. (1996) *The Power of Statistical Thinking*, Addison-Wesley, Reading, MA.

[18] Pandit, S.M. and Wu, S.M. (1983) *Time Series and System Analysis with Applications*, John Wiley, New York.

[19] MacGregor, J.F. (1990) A different view of the funnel experiment. *Journal of Quality Technology*, **22**, 255–259.

[20] Montgomery, D.C. (1997) *Introduction to Statistical Quality Control*, 3rd edn., John Wiley, New York.

[21] Box, G.E.P., Jenkins, G.M. and Reinsel, G.C. (1994) *Time Series Analysis Forecasting and Control*, 3rd edn., Prentice-Hall, Englewood Cliffs, NJ.

[22] Ziegler, J.G. and Nichols, N.B. (1942) Optimum settings for automatic controllers. *Transactions of ASME*, **64**, 759–768.

[23] Taguchi, G. (1986) *Introduction to Quality Engineering*, Asian Productivity Organization, Tokyo, Japan.

[24] Tsung, F. (1997) Run-to-run proportional-integral-derivative process control and monitoring schemes. Ph.D. dissertation, University of Michigan, Ann Arbor, MI 48109, USA.

[25] Deming, W.E. (1986) *Out of the Crisis*, MIT Center for Advanced Engineering Study, Cambridge, MA.

[26] Harris, T.J. (1989) Assessment of control loop performance. *Canadian Journal of Chemical Engineering*, **67**, 856–861.

[27] Desborough, L. and Harris, T. (1993) Performance assessment measures for univariate feedforward/feedback control. *The Canadian Journal of Chemical Engineering*, **71**, 605–616.

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

[29] Alwan, L.C. and Roberts, H.V. (1988) Time-series Modeling for statistical process control. *Journal of Business and Economic Statistics*, **6**, 87–95.

[30] Montgomery, D.C. and Mastrangelo, C.M. (1991) Some statistical process control methods for autocorrelated data. *Journal of Quality Technology*, **23**, 179–204.

[31] Wardell, D.G., Moskowitz, H. and Plante, R.D. (1994) Run-length distributions of special-cause control charts for correlated observations. *Technometrics*, **36**, 3–17.

[32] Van der Wiel, S.A. (1996) Monitoring processes that wander using integrated moving average models. *Technometrics*, **38**, 139–151.

[33] Tsung, F., Shi, J. and Wu, C.F.J. (1999) Joint monitoring of PID controlled processes. *Journal of Quality Technology*, in press.
[34] Alt, F.B. (1985) Multivariate Quality Control. *Encyclopedia of the Statistical Sciences*, Johnson, N.L., Kotz, S. and Read, C.R. (eds), John Wiley, New York, pp. 111–122.
[35] Woodall, W.H. and Ncube, M.M. (1985) Multivariate CUSUM quality control procedures. *Technometrics*, **27**, 285–292.
[36] Lowry, C.A. and Woodall, W.H. (1992) A multivariate exponentially weighted moving average control chart. *Technometrics*, **34**, 46–53.

Appendix A

The variance of the controlled outputs

By applying the PID scheme in (3) to the output from model (1), we have

$$e_t = \varphi(B)^{-1}(1-B)D_t \equiv \sum_{j=0}^{\infty} \psi_j D_{t-j}\,, \tag{A1}$$

where

$$\varphi(B) = 1 - \alpha B - \beta B^2 - \gamma B^3\,, \tag{A2}$$

with $\alpha = 1 - k_P - k_I - k_D$, $\beta = k_P + 2k_D$, and $\gamma = -k_D$. Hence the variance of the output error is

$$\begin{aligned}\sigma_e^2 &= \left(\sum_{j=0}^{\infty}\sum_{l=0}^{\infty} \psi_j\psi_l\rho_{|j-l|}\right)\sigma_D^2\,,\\ &= \left(\sum_{j=0}^{\infty}\psi_j^2 + 2\psi_0\sum_{l=1}^{\infty}\psi_l\rho_l + 2\sum_{j=1}^{\infty}\sum_{l=1}^{\infty}\psi_j\psi_{j+l}\rho_l\right)\sigma_D^2.\end{aligned} \tag{A3}$$

When the disturbance $\{D_t\}$ is an ARMA(1,1) model,

$$\begin{aligned}\rho_1 &= (1-\phi\theta)(\phi-\theta)/(1+\theta^2-2\phi\theta)\,,\\ \rho_l &= \phi^{l-1}\rho_1 \quad (l \geq 1).\end{aligned} \tag{A4}$$

By substituting (A4) into (A3),

$$\begin{aligned}\sigma_e^2 &= \left(\sum_{j=0}^{\infty}\psi_j^2 + 2\rho_1\left(\psi_0\sum_{l=1}^{\infty}\psi_l\phi^{l-1} + \sum_{j=1}^{\infty}\sum_{l=1}^{\infty}\psi_j\psi_{j+l}\phi^{l-1}\right)\right)\sigma_D^2\,,\\ &\equiv (I + 2\rho_1(II+III))\sigma_D^2,\end{aligned} \tag{A5}$$

where the terms I, II, and III will be derived in (A10) and (A14). Let Δ be the discriminant of $\varphi(B)$

$$\Delta = -18\alpha\beta\gamma + \alpha^2\beta^2 - 4\alpha^3\gamma + 4\beta^3 - 27\gamma^2. \tag{A6}$$

Two cases are considered:
(i) when $\Delta \neq 0$, $\varphi(B) = 0$ has the three distinct roots

$$\begin{aligned}\epsilon &= U - U/V + \alpha/3,\\ \zeta &= -U/2 + V/(2U) + \alpha/3 + i\sqrt{3}(U + V/U)/2,\\ \eta &= -U/2 + V/(2U) + \alpha/3 - i\sqrt{3}(U + V/U)/2,\end{aligned} \tag{A7}$$

where

$$\begin{aligned}U &= (\alpha\beta/6 + \Gamma/2 + \alpha^3/27 + \sqrt{-3\Delta}/18)^{1/3}\,,\\ V &= -\beta/3 - \alpha^2/9.\end{aligned}$$

When $\Delta < 0$, ζ and η are complex roots. This leads to

$$\begin{aligned}\psi_0 &= 1\,,\\ \psi_j &= \Gamma\epsilon^{j-1} + \Theta\zeta^{j-1} + \Lambda\eta^{j-1} \quad (j \geq 1)\,,\end{aligned} \tag{A8}$$

where

$$\begin{aligned}\Gamma &= \frac{\epsilon^2(-1+\epsilon)}{(\epsilon-\zeta)(\epsilon-\eta)},\ \Theta = \frac{\zeta^2(-1+\zeta)}{(\zeta-\epsilon)(\zeta-\eta)},\\ \Lambda &= \frac{\eta^2(-1+\eta)}{(\eta-\epsilon)(\eta-\zeta)},\end{aligned} \tag{A9}$$

and the three terms in (A5) are obtained:

$$\begin{aligned}I &= \frac{\Gamma^2}{1-\epsilon^2} + \frac{\Theta^2}{1-\zeta^2} + \frac{\Lambda^2}{1-\eta^2} + \frac{2\Gamma\Theta}{1-\epsilon\zeta}\\ &\quad + \frac{2\Gamma\Lambda}{1-\epsilon\eta} + \frac{2\Theta\Lambda}{1-\zeta\eta} + 1,\\ II &= \frac{\Gamma}{1-\phi\epsilon} + \frac{\Theta}{1-\phi\zeta} + \frac{\Lambda}{1-\phi\eta},\\ III &= \frac{\Gamma\epsilon}{1-\epsilon\phi}\left(\frac{\Gamma}{1-\epsilon^2} + \frac{\Theta}{1-\epsilon\zeta} + \frac{\Lambda}{1-\epsilon\eta}\right)\\ &\quad + \frac{\Theta\zeta}{1-\zeta\phi}\left(\frac{\Gamma}{1-\epsilon\zeta} + \frac{\Theta}{1-\zeta^2} + \frac{\Lambda}{1-\zeta\eta}\right)\\ &\quad + \frac{\Lambda\eta}{1-\eta\phi}\left(\frac{\Gamma}{1-\epsilon\eta} + \frac{\Theta}{1-\zeta\eta} + \frac{\Lambda}{1-\eta^2}\right).\end{aligned} \tag{A10}$$

(ii) When $\Delta = 0$, it has roots with multiplicity,

$$\begin{aligned}\epsilon &= 2U + \alpha/3,\\ \zeta &= \eta = -U + \alpha/3,\end{aligned} \tag{A11}$$

with $U = (\alpha\beta/6 + \gamma/2 + \alpha^3/27)^{1/3}$, which leads to

$$\begin{aligned}\psi_0 &= 1,\\ \psi_j &= \Gamma\epsilon^{j-1} + \Theta\zeta^{j-1} + \Lambda(j-1)\zeta^{j-1} \quad (j \geq 1),\end{aligned} \tag{A12}$$

where

$$\begin{aligned}\Gamma &= \frac{\epsilon^2(-1+\epsilon)}{(\epsilon-\zeta)^2},\ \Theta = \frac{\zeta(2\epsilon - 3\epsilon\zeta - \zeta + 2\zeta^2)}{(\epsilon-\zeta)^2},\\ \Lambda &= \frac{\zeta(1-\zeta)}{(\epsilon-\zeta)}.\end{aligned} \tag{A13}$$

It follows that the three terms in (A5) are

$$I = \frac{\Gamma^2}{1-\epsilon^2} + \frac{\Theta^2}{1-\zeta^2} + \frac{\Lambda^2\zeta^2(1+\zeta^2)}{(1-\zeta^2)^3} + \frac{2\Gamma\Theta}{1-\epsilon\zeta} + \frac{2\Gamma\Lambda\epsilon\zeta}{(1-\epsilon\zeta)^2} + \frac{2\Theta\Lambda\zeta^2}{(1-\zeta^2)^2} + 1,$$

$$II = \frac{\Gamma}{1-\phi\epsilon} + \frac{\Theta}{1-\phi\zeta} + \frac{\Lambda\phi\zeta}{(1-\phi\zeta)^2},$$

$$III = \frac{\Gamma\epsilon}{1-\epsilon\phi}\left(\frac{\Gamma}{1-\epsilon^2} + \frac{\Theta}{1-\epsilon\zeta} + \frac{\Lambda\epsilon\zeta}{(1-\epsilon\zeta)^2}\right) + \left(\frac{(\Theta+\Lambda)\zeta}{1-\zeta\phi} + \frac{\Lambda\zeta^2\phi}{(1-\zeta\phi)^2}\right)\left(\frac{\Gamma}{1-\epsilon\zeta} + \frac{\Theta}{1-\zeta^2} + \frac{\Lambda\zeta^2}{(1-\zeta^2)^2}\right) + \frac{\Lambda\zeta}{1-\zeta\phi}\left(\frac{\Gamma\epsilon\zeta}{(1-\epsilon\zeta)^2} + \frac{\Theta\zeta^2}{(1-\zeta^2)^2} + \frac{\Lambda\zeta^2(1+\zeta^2)}{(1-\zeta^2)^3}\right). \quad \text{(A14)}$$

Appendix B

The variance of the manipulated inputs

From (1), (3), and (A1) we have

$$X_t = \left(-k_P - k_I\frac{1}{1-B} - k_D(1-B)\right)\varphi(B)^{-1}(1-B)D_t \equiv \sum_{j=0}^{\infty}\vartheta_j D_{t-j}, \quad \text{(A15)}$$

where $\varphi(B)$ is the same as in (A2), and ϑ_j will be derived in (A18) and (A21). Thus

$$\sigma_X^2 = \left(\sum_{j=0}^{\infty}\sum_{l=0}^{\infty}\vartheta_j\vartheta_l\rho_{|j-l|}\right)\sigma_D^2. \quad \text{(A16)}$$

Let the disturbance $\{D_t\}$ be an ARMA(1,1) model. Then σ_X^2 has the same form as in (A5), i.e.,

$$\sigma_X^2 = (I' + 2\rho_1(II' + III'))\sigma_D^2, \quad \text{(A17)}$$

where the terms I', II', and III' will be derived in (A20) and (A23). Two cases are considered:

(i) When the discriminant Δ in (A6) is non-zero:

$$\vartheta_0 = k_P + k_I + k_D,$$
$$\vartheta_j = \Gamma'\epsilon^{j-1} + \Theta'\zeta^{j-1} + \Lambda'\eta^{j-1} \quad (j \geq 1), \quad \text{(A18)}$$

where the three roots of $\varphi(B)$: ϵ, ζ and η are the same as in (A7), but Γ', Θ', and Λ' are different from Γ, Θ, and Λ in (A9):

$$\Gamma' = \frac{-\epsilon^3(k_P+k_I+k_D) + \epsilon^2(k_P+2k_D) - \epsilon k_D}{(\epsilon-\zeta)(\epsilon-\eta)},$$

$$\Theta' = \frac{-\zeta^3(k_P+k_I+k_D) + \zeta^2(k_P+2k_D) - \zeta k_D}{(\zeta-\epsilon)(\zeta-\eta)},$$

$$\Lambda' = \frac{-\eta^3(k_P+k_I+k_D) + \eta^2(k_P+2k_D) - \eta k_D}{(\eta-\epsilon)(\eta-\zeta)}. \quad \text{(A19)}$$

Therefore, the three terms of σ_X^2 are obtained:

$$I' = \frac{\Gamma'^2}{1-\epsilon^2} + \frac{\Theta'^2}{1-\zeta^2} + \frac{\Lambda'^2}{1-\eta^2} + \frac{2\Gamma'\Theta'}{1-\epsilon\zeta} + \frac{2\Gamma'\Lambda'}{1-\epsilon\eta} + \frac{2\Theta'\Lambda'}{1-\zeta\eta} + (k_P+k_I+k_D)^2,$$

$$II' = \left(\frac{\Gamma'}{1-\phi\epsilon} + \frac{\Theta'}{1-\phi\zeta} + \frac{\Lambda'}{1-\phi\eta}\right)(k_P+k_I+k_D),$$

$$III' = \frac{\Gamma'\epsilon}{1-\epsilon\phi}\left(\frac{\Gamma'}{1-\epsilon^2} + \frac{\Theta'}{1-\epsilon\zeta} + \frac{\Lambda'}{1-\epsilon\eta}\right) + \frac{\Theta'\zeta}{1-\zeta\phi}\left(\frac{\Gamma'}{1-\epsilon\zeta} + \frac{\Theta'}{1-\zeta^2} + \frac{\Lambda'}{1-\zeta\eta}\right) + \frac{\Lambda'\eta}{1-\eta\phi}\left(\frac{\Gamma'}{1-\epsilon\eta} + \frac{\Theta'}{1-\zeta\eta} + \frac{\Lambda'}{1-\eta^2}\right). \quad \text{(A20)}$$

(ii) When the discriminant Δ is zero:

$$\vartheta_0 = k_P + k_I + k_D,$$
$$\vartheta_j = \Gamma'\epsilon^{j-1} + \Theta'\zeta^{j-1} + \Lambda'(j-1)\zeta^{j-1} \quad (j \geq 1). \quad \text{(A21)}$$

Thus, σ_X^2 also has a similar form to (A5) except that

$$\Gamma' = \frac{-\epsilon^3(k_P+k_I+k_D) + \epsilon^2(k_P+2k_D) - \epsilon k_D}{(\epsilon-\zeta)^2},$$

$$\Theta' = \frac{-2\zeta^3(k_P+k_I+k_D) + \zeta^2(3\epsilon(k_P+k_I+k_D) + k_P + 2k_D)}{(\epsilon-\zeta)^2} + \frac{-2\zeta\epsilon(k_P+2k_D) + \epsilon k_D}{(\epsilon-\zeta)^2},$$

$$\Lambda' = \frac{\zeta^2(k_P+k_I+k_D) - \zeta(k_P+2k_D) + k_D}{(\epsilon-\zeta)}. \quad \text{(A22)}$$

It follows that

$$I' = \frac{\Gamma'^2}{1-\epsilon^2} + \frac{\Theta'^2}{1-\zeta^2} + \frac{\Lambda'^2\zeta^2(1+\zeta^2)}{(1-\zeta^2)^3} + \frac{2\Gamma'\Theta'}{1-\epsilon\zeta} + \frac{2\Gamma'\Lambda'\epsilon\zeta}{(1-\epsilon\zeta)^2} + \frac{2\Theta'\Lambda'\zeta^2}{(1-\zeta^2)^2} + (k_P+k_I+k_D)^2,$$

$$II' = \left(\frac{\Gamma'}{1-\phi\epsilon} + \frac{\Theta'}{1-\phi\zeta} + \frac{\Lambda'\phi\zeta}{(1-\phi\zeta)^2}\right)(k_P+k_I+k_D),$$

$$III' = \frac{\Gamma'\epsilon}{1-\epsilon\phi}\left(\frac{\Gamma'}{1-\epsilon^2} + \frac{\Theta'}{1-\epsilon\zeta} + \frac{\Lambda'\epsilon\zeta}{(1-\epsilon\zeta)^2}\right) + \frac{\Theta'\zeta}{1-\zeta\phi}\left(\frac{\Gamma'}{1-\epsilon\zeta} + \frac{\Theta'}{1-\zeta^2} + \frac{\Lambda'\zeta^2}{(1-\zeta^2)^2}\right) + \frac{\Lambda'\phi\zeta^2}{(1-\phi\zeta)^2}\left(\frac{\Gamma'}{1-\epsilon\zeta} + \frac{\Theta'}{1-\zeta^2} + \frac{\Lambda'\zeta^2}{(1-\zeta^2)^2}\right). \quad \text{(A23)}$$