Optimal Time-sampling Problem in a Statistical Control with a Quadratic Cost Functional
Analytical and Numerical Approaches

Valery Y. Glizer and Vladimir Turetsky
Department of Applied Mathematics, ORT Braude College of Engineering, P.O.B. 78, Karmiel 2161002, Israel

Keywords: Statistical Control, Statistical Information, Quadratic Cost Functional, Optimal Time-sampling, Pontryagin's Maximum Principle, Quadratic Optimization.

Abstract: We consider the problem of constructing an optimal time-sampling for a Statistical Process Control (or, briefly, Statistical Control (SC)). The aim of this time-sampling is to minimize the expected loss, caused by a delay in the detection of an undesirable process change. We study the case where this loss is a quadratic functional of the sampling time-interval. This problem is modeled by a nonstandard calculus of variations problem. We propose two approaches to the solution of this calculus of variations problem. The first approach is based on its equivalent transformation to an optimal control problem. The latter is solved by application of the Pontryagin's Maximum Principle, yielding an analytical expression for the optimal time-sampling in the SC. The second approach uses a discretization of the calculus of variations problem, resulting in a finite dimensional quadratic optimization problem. Solution of the latter provides a suboptimal time-sampling in the SC. The time-samplings, obtained by these two approaches, are compared to each other in numerical examples.

1 INTRODUCTION

The Statistical Control is a quality control method (see, e.g., (Qiu, 2013) and references therein). It consists in a monitoring of a process state using a statistical information on samples of its characteristic index in some time-intervals. The SC is applied in industry, medicine, veterinary, environment control, etc. Its objective is to minimize losses which can be caused by delay in the detection of undesirable process changes, subject to reasonable inspection expenses. For many years, the traditional SC practice was to take samples of the process characteristic index with a fixed time-sampling. The idea of using a variable time-sampling (VTS) in the SC was suggested for the first time in the work (Reynolds et al., 1988). Then, this idea was developed in a number of works (see, e.g., (Amin and Hemasinha, 1993); (Amin and Miller, 1993); (Bashkansky and Glizer, 2012); (Chew et al., 2015); (Costa, 1994); (Costa, 1997); (Costa, 1998); (Costa, 1999b); (Costa, 1999a); (Costa and Magalhaes, 2007); (Glizer et al., 2015); (Hatjimihail, 2009); (Sultana et al., 2014); (Li and Qiu, 2014); (Prabhu et al., 1994); (Reynolds, 1995)).

In (Reynolds et al., 1988), the delay in the detection of a process change was considered as a criterion for optimality of the variable sampling time-interval in the SC. Another possible criterion, proposed in (Taguchi et al., 2007). is the expected loss due to such a delay. The latter criterion is more general and, therefore, more suitable for various applications.

In many processes, the relation between the expected loss and the delay in the detection of a process change is non-linear. Among such processes, we can mention: (i) fires propagation (Babrauskas, 2008), (ii) oil spills spreading (Sebastião and Soares, 1995), (iii) cholesterol plaque growth (Bulelzai and Dubbeldam, 2012), (iv) epidemics propagation (Carpenter et al., 2011), (v) fatigue crack growth in ship hull structures (Kim and Frangopol, 2011).

Genichi Taguchi (see e.g. (Taguchi et al., 2007)) proposed a quadratic dependence of the expected loss on some critical performance parameter of a process. In modern industry, medicine, veterinary, natural environment protection, etc, the statistical control of a process becomes its indispensable part. Therefore, the delay in the detection of a process change can be considered as a critical performance parameter of the process. This observation yields a quadratic dependence of the expected loss on the detection delay.

Glizer, V. and Turetsky, V.
Optimal Time-sampling Problem in a Statistical Control with a Quadratic Cost Functional - Analytical and Numerical Approaches.
DOI: 10.5220/0006829000210032
In *Proceedings of the 15th International Conference on Informatics in Control, Automation and Robotics (ICINCO 2018) - Volume 1*, pages 21-32
ISBN: 978-989-758-321-6

We model the problem of the SC time-sampling optimization by some calculus of variations problem. This problem consists of a cost functional (the expected loss) and two types of constraints (geometric and integral inequality constraints). The geometric constraint gives the lower and upper bounds of each sampling time-interval. The integral inequality constraint means that the average of the sampling time-interval is not prescribed but it belongs to a given interval. This model for the SC time-sampling optimization is more general than those studied in ((Bashkansky and Glizer, 2012); (Glizer et al., 2015)). Moreover, both types of the constraints are not studied in the classical calculus of variations theory. Thus, the considered extremal problem is nonstandard. We propose two methods of its solution, which are not based on a preliminary approximate decomposition of this problem. The first method converts equivalently the original extremal problem into an optimal control problem. This optimal control problem is solved using the Pontryagin's Maximum Principle (PMP), which yields an *exact analytical* solution to the original calculus of variations problem. This solution constitutes the *optimal time-sampling of the SC*. In the second method, the original calculus of variations problem is replaced approximately with a finite-dimensional optimization problem. The latter is solved using corresponding mathematical programming tools, which yields an approximate solution of the original extremal problem. This solution constitutes the *suboptimal time-sampling of the SC*.

It is important to note, that the SC time-sampling, designed in this paper, depends on the current state of the process. It is not designed in advance for an entire period of the process control. Applying the terminology of control engineering, this SC time-sampling can be called a *state-feedback* time-sampling.

Also, it should be noted that in most VTS schemes, described in the literature, the sampling time-interval of only two different lengths is considered. In the present paper, as well as in (Li and Qiu, 2014) and (Glizer et al., 2015), more than two different lengths of the sampling time-interval are proposed for the SC. In (Li and Qiu, 2014), the multiple lengths sampling time-interval is related to the p-value of the charting statistic, while in (Glizer et al., 2015) and the present paper, such sampling time-intervals are derived from solutions (exact and approximate) of the optimization problems.

2 PROBLEM STATEMENT

We analyze the SC case where the monitoring of a characteristic index x of the process state is carried out based on the information about its sample mean. Namely, at some prescribed/precalculated time instance t a batch of n observations x_j, $(j=\overline{1,n})$ of the value x is obtained, and the sample mean $\bar{x}$ of these observations is derived. We assume that the sample size n is independent of t. Let μ and σ be the mean value and the standard deviation of the random value x. Then, the mean value and the standard deviation of the random value $\bar{x}$ are μ and $\sigma/\sqrt{n}$. In this paper, we deal with the case where the random value $\bar{x}$ is normally distributed, i.e., $\bar{x}\sim N(\mu,\sigma/\sqrt{n})$. This occurs when either the random value x is normally distributed, or the sample size n is considerably large ($n\geq 30$). In the latter case, by virtue of the Central Limit Theorem, the normal distribution $N(\mu,\sigma/\sqrt{n})$ provides a good approximation of $\bar{x}$ even if x does not strictly fit a normal distribution (see, e.g., (Qiu, 2013) and references therein). Thus, the normalized sample mean, called the *standard score*, is $z=\big(\bar{x}-\mu\big)/\big(\sigma/\sqrt{n}\big)\sim N(0,1)$. The upper and lower limits of the standard Shewhart control chart for z are $z_{\min}=-3$ and $z_{\max}=3$, respectively, (see (Qiu, 2013)). Therefore, the false alarm probability α (type I error), i.e., the probability of the event $z\notin[-3,3]$, is $\alpha\approx 0.0027$.

Let the mean value of the index x is shifted by Δ, i.e., a new mean value is $\mu'=\mu+\Delta$, while the standard deviation σ remains unchanged. Then the distribution of z becomes $z\sim N(\delta,1)$, where the normalized shift $\delta=\Delta/\big(\sigma/\sqrt{n}\big)$ is the so-called *signal-to-noise ratio*. The probability of discovering the shift (receiving the signal) by a single sample is the probability of the event $z\notin[-3,3]$ for $z\sim N(\delta,1)$:

$$1-\beta=1-\frac{1}{\sqrt{2\pi}}\int_{-3}^{3}\exp\big(-(z-\delta)^2/2\big)dz \\ =1-[\Phi(3-\delta)-\Phi(-3-\delta)], \qquad (1)$$

where $\beta=\beta(\delta)$ is the probability of a type II error (not discovering the shift).

Consider the SC with a variable sampling time-interval $u(z)$, depending on the standard score $z=\big(\bar{x}-\mu\big)/\big(\sigma/\sqrt{n}\big)$. Since the value of the sampling time-interval should depend only on $|z|$, the function $u(z)$ is even $\big(u(-z)=u(z)\big)$. Therefore, in what follows, we consider the function $u(z)$ in the interval $[0,3]$. Also, for the sake of simplicity, we assume that $\delta\geq 0$. The case $\delta\leq 0$ is treated similarly. Further, we assume that the function $u(z)$ is bounded as:

$$0<u_{\min}\leq u(z)\leq u_{\max},\quad z\in[0,3]. \qquad (2)$$

The inequality (2) is a geometric constraint, imposed on the function $u(z)$. Now, let us consider the following integral constraint, imposed on $u(z)$:

$$aT_{\min} \le \int_0^3 \exp(-z^2/2)u(z)dz \le aT_{\max}, \qquad (3)$$

where $0 < T_{\min} < T_{\max}$, and

$$a \stackrel{\triangle}{=} \int_0^3 \exp(-z^2/2)dz > 0. \qquad (4)$$

Remark 1. *The inequality (3) means that the expected sampling time-interval in the case of unshifted z ($\delta = 0$) belongs to a prescribed nominal interval $[T_{\min}, T_{\max}]$. We assume that*

$$T_{\min} > u_{\min}, \quad T_{\max} < u_{\max}. \qquad (5)$$

If the shift of the mean in the process characteristic index remains constant, the time t_d, required for discovering this shift (so-called *time to signal*), is the sum of a random amount K_d of random independent and identically distributed sampling time-intervals u_i, conditionally independent of K_d: $t_d = \sum_{i=1}^{K_d} u_i$. The value K_d is distributed geometrically with the success probability $1-\beta$, given by (1). Its mathematical expectation and variance are (Ross, 2009): $\mathbf{E}(K_d) = 1/(1-\beta)$, $\mathbf{Var}(K_d) = \beta/(1-\beta)^2$.

The cost functional, to be minimized by a proper choice of the sampling time-interval $u(z)$, is the mathematical expectation $\mathbf{E}(L)$ of the loss L, caused by the delay in the detection of the shift. Here, we consider the loss L as a quadratic function of the delay t_d, i.e., $L = kt_d^2$, where $k = k(\delta) \ge 0$ is an increasing function of $\delta \ge 0$ and $k(0) = 0$. Thus, the expected loss is

$$\mathbf{E}(L) = k(\delta)\mathbf{E}(t_d^2). \qquad (6)$$

By routine calculations, we have

$$\mathbf{E}(t_d^2) = A(\delta)\left[\int_0^3 \psi(z,\delta)u^2(z)dz + B(\delta)\left(\int_0^3 \psi(z,\delta)u(z)dz\right)^2\right], \qquad (7)$$

where

$$A(\delta) \stackrel{\triangle}{=} \frac{\exp(-\delta^2/2)}{(1-\beta)\beta\sqrt{2\pi}} > 0,\ B(\delta) \stackrel{\triangle}{=} \frac{2\exp(-\delta^2/2)}{(1-\beta)\sqrt{2\pi}} > 0, \qquad (8)$$

$$\psi(z,\delta) \stackrel{\triangle}{=} 2\exp\left(-z^2/2\right)\cosh(\delta z) > 0,\ z \in [0,3]. \qquad (9)$$

Since $k(\delta) > 0$ for all $\delta > 0$, then due to (7) – (8), the minimization of the cost functional (6) for any given $\delta > 0$ is equivalent to the minimization of the following cost functional:

$$J\big(u(z)\big) \stackrel{\triangle}{=} \int_0^3 \psi(z,\delta)u^2(z)dz + B(\delta)\left(\int_0^3 \psi(z,\delta)u(z)dz\right)^2. \qquad (10)$$

Remark 2. *Two cases can be distinguished with respect to the information on the value of δ: (i) the value of δ is known; (ii) the value of δ is unknown. In the first case, one should minimize with respect to $u(z)$ the cost functional (10), calculated for the known δ. In the second case, one should minimize with respect to $u(z)$ the cost functional (6) – (7) robustly in $\delta \ge 0$, i.e., one should minimize with respect to $u(z)$ the new cost functional $J_{\text{new}}\big(u(z)\big) = \max_{\delta\ge 0} k(\delta)\mathbf{E}(t_d^2)$. In this paper, we restrict our analysis with the first case.*

Thus, we can formulate the following extremal problem.

Extremal Problem (EP): for a known $\delta \ge 0$, to find the function $u(z)$, $z \in [0,3]$, which minimizes the cost functional (10) subject to the constraints (2), (3) and the inequality (5).

In subsequent sections, we solve the EP in a closed analytical form and numerically, thus designing the optimal and suboptimal SC time-sampling.

3 ANALYTICAL SOLUTION OF THE EP

The EP is a nonstandard calculus of variations problem with two types of constraints, the geometric constraint (2) and the integral inequality constraint (3), imposed on the minimizing function. The classical calculus of variations theory does not study extremal problems with such types of constraints (see, e.g., (Gelfand and Fomin, 1963)). We propose another approach to the solution of this problem, which consists in an equivalent transformation of the EP into an optimal control problem. The latter is analyzed by application of the control optimality necessary condition – the Pontryagin's Maximum Principle (PMP) (Pontryagin et al., 1962).

3.1 Transformation of the EP

Let us introduce the auxiliary vector-valued function $w(z) = \big(w_1(z), w_2(z), w_3(z)\big)^T$, $z \in [0,3]$, where

$$w_1(z) = \int_0^z \psi(\zeta,\delta)u^2(\zeta)d\zeta, \qquad (11)$$

$$w_2(z) = \int_0^z \psi(\zeta,\delta)u(\zeta)d\zeta, \qquad (12)$$

$$w_3(z) = \int_0^z \exp\left(-\zeta^2/2\right)u(\zeta)d\zeta. \qquad (13)$$

The functions $w_i(z)$, $(i = 1,2,3)$, satisfy the differential equations

$$dw_1/dz = \psi(z,\delta)u^2(z), \qquad (14)$$

$$dw_2/dz = \psi(z,\delta)u(z), \tag{15}$$

$$dw_3/dz = \exp\left(-z^2/2\right)u(z), \tag{16}$$

and the initial conditions

$$w_1(0) = 0, \quad w_2(0) = 0, \quad w_3(0) = 0. \tag{17}$$

Based on (13), the integral inequality (3) of the EP becomes

$$aT_{\min} \le w_3(3) \le aT_{\max}. \tag{18}$$

This inequality can be rewritten equivalently as the set of two inequalities

$$g_1\big(w(3)\big) \stackrel{\triangle}{=} -w_3(3) + aT_{\min} \le 0, \tag{19}$$

$$g_2\big(w(3)\big) \stackrel{\triangle}{=} w_3(3) - aT_{\max} \le 0. \tag{20}$$

Using (11) – (12), the cost functional (10) becomes

$$J\big(u(z)\big) = w_1(3) + B(\delta)\big(w_2(3)\big)^2. \tag{21}$$

Thus, we have transformed the EP into the equivalent optimal control problem: to find the control function $u(z)$, transferring the system (14) – (16) from the initial position (17) to the set of terminal positions (19) – (20) and minimizing the cost functional (21), subject to the geometric constraint (2) and the inequality (5). This optimal control problem is non-linear with respect to $u(z)$, and in what follows, it is called the Non-linear Optimal Control Problem (NOCP). Due to (Ioffe and Tihomirov, 1979) (see Section 9.2, Theorem 3), the NOCP has a solution (optimal control).

3.2 Solution of the NOCP by Application of the PMP

The Variational Hamiltonian of the NOCP is

$$H = H(w,u,\lambda,z)$$
$$= \lambda_1\psi(\delta,z)u^2 + \lambda_2\psi(\delta,z)u + \lambda_3\exp\left(-z^2/2\right)u, \tag{22}$$

where $\lambda = \lambda(z) = \big(\lambda_1(z),\lambda_2(z),\lambda_3(z)\big)^T$, and $\lambda_i = \lambda_i(z)$, $(i = 1,2,3)$ are the costate variables. These costate variables satisfy the differential equations

$$d\lambda_1/dz = -\partial H/\partial w_1 = 0, \quad z \in [0,3], \tag{23}$$

$$d\lambda_2/dz = -\partial H/\partial w_2 = 0, \quad z \in [0,3], \tag{24}$$

$$d\lambda_3/dz = -\partial H/\partial w_3 = 0, \quad z \in [0,3], \tag{25}$$

and the terminal conditions

$$\lambda_1(3) = -C_0\big(\partial J/\partial w_1(3)\big) = -C_0, \tag{26}$$

$$\lambda_2(3) = -C_0\big(\partial J/\partial w_2(3)\big) = -2C_0B(\delta)\gamma,\ \gamma \stackrel{\triangle}{=} w_2(3), \tag{27}$$

$$\lambda_3(3) = -C_0\big(\partial J/\partial w_3(3)\big)$$
$$-C_1\Big(\partial g_1\big(w(3)\big)/\partial w_3(3)\Big)$$
$$-C_2\Big(\partial g_1\big(w(3)\big)/\partial w_3(3)\Big) = C_1 - C_2. \tag{28}$$

In these terminal conditions,

$$C_0 \ge 0, \quad C_1 \ge 0, \quad C_2 \ge 0 \tag{29}$$

are some constants, such that

$$C_0 + C_1 + C_2 > 0, \tag{30}$$

and

$$C_1g_1\big(w(3)\big) = 0, \quad C_2g_2\big(w(3)\big) = 0. \tag{31}$$

Denote $I_u \stackrel{\triangle}{=} \big\{u : u_{\min} \le u \le u_{\max}\big\}$. Due to the PMP, an optimal control $u^*(z)$ of the NOCP necessarily satisfies the following condition for all $z \in [0,3]$:

$$\max_{u(z)\in I_u} H\big(w(z),u(z),\lambda(z),z\big) =$$
$$H\big(w(z),u^*(z),\lambda(z),z\big). \tag{32}$$

Thus, any control $u(z)$, satisfying the equations (32), (14) – (17) and (23) – (28), the conditions (19) – (20) and (29) – (31), is an optimal control candidate in the NOCP. To obtain such a control, first, we solve the equations (23) – (28). These equations yield the following solution for $z \in [0,3]$: $\lambda_1(z) = -C_0$, $\lambda_2(z) = -2C_0B\gamma$, $\lambda_3(z) = C_1 - C_2$. By substitution of this solution into (22) and using (9), the Variational Hamiltonian of the NOCP becomes

$$H = -\exp\left(-z^2/2\right)G\big(u,z,\gamma,C_0,C_1,C_2\big), \tag{33}$$

where the function $G\big(u,z,\gamma,C_0,C_1,C_2\big)$ has the form

$$G\big(u,z,\gamma,C_0,C_1,C_2\big) =$$
$$\Big(4C_0B(\delta)\gamma\cosh(\delta z) - C_1 + C_2\Big)u +$$
$$2C_0\cosh(\delta z)u^2. \tag{34}$$

Let us show that $C_0 > 0$. For this purpose, we assume the opposite which, due to (29), is $C_0 = 0$. In this case, the use of (33) – (34) and (30) yields

$$H = \exp(-z^2/2)(C_1 - C_2)u \tag{35}$$

and $C_1 + C_2 > 0$. Note that $C_1 \ne C_2$. Indeed, if $C_1 = C_2$, then both constants are non-zero. In such a case, by virtue of (31), $g_1\big(w(3)\big) = 0$ and $g_2\big(w(3)\big) = 0$. The latter, along with (19) – (20) and the inequality $T_{\min} < T_{\max}$ (see (5)), yields a contradiction. Thus, $C_1 - C_2 \ne 0$. The unique control, satisfying (32) with the Variational Hamiltonian of the form (35) is

$$u^*(z) = \begin{cases} u_{\min}, & \text{if } C_1 - C_2 < 0, \\ u_{\max}, & \text{if } C_1 - C_2 > 0. \end{cases} \tag{36}$$

Now, substituting (36) into (16) instead of $u(z)$ and solving the resulting equation subject to the initial condition from (17), we obtain

$$w_3(3) = \begin{cases} au_{\min} & \text{if } C_1 - C_2 < 0, \\ au_{\max}, & \text{if } C_1 - C_2 > 0. \end{cases} \tag{37}$$

The latter, along with the inequality (5), means that $w_3(3)$ does not belong to the set of terminal positions (19) – (20). Therefore, the control (36), obtained from (32), (35) under the assumption $C_0 = 0$, is not admissible. This means that the assumption $C_0 = 0$ is wrong, i.e., $C_0 > 0$.

Due to the PMP, we can set $C_0 = 1$ and rewrite the equations (33) – (34) as:

$$H = -\exp\big(-z^2/2\big)G_1\big(u, z, \gamma, C_1, C_2\big), \tag{38}$$

$$G_1\big(u, z, \gamma, C_1, C_2\big) = \Big(4B(\delta)\gamma\cosh(\delta z) - C_1 + C_2\Big)u + 2\cosh(\delta z)u^2. \tag{39}$$

Thus, applying (32) to (38) – (39), we obtain the optimal control of the NOCP in the form

$$u^*(z) = u^*(z, \gamma, C_1, C_2) = \begin{cases} u_{\min}, & \tilde{u}(z, \gamma, C_1, C_2) \le u_{\min}, \\ \tilde{u}(z, \gamma, C_1, C_2), & \tilde{u}(z, \gamma, C_1, C_2) \in (u_{\min}, u_{\max}], \\ u_{\max}, & \tilde{u}(z, \gamma, C_1, C_2) > u_{\max}, \end{cases} \tag{40}$$

where

$$\tilde{u}(z, \gamma, C_1, C_2) = \big(C_1 - C_2\big)/\big(4\cosh(\delta z)\big) - B\gamma \tag{41}$$

is the unique solution of the following equation with respect to u: $\partial G_1(u, z, \gamma, C_1, C_2)/\partial u = 0$.

In order to use the equation (40), we need to know the constants γ, C_1 and C_2. These constants should be chosen in such a way that the resulting control (40) will transfer the system (14) – (16) from the initial position (17) to the intersection of the set of terminal positions (19) – (20) and the plane $w_2(3) - \gamma = 0$ in the 3D-space $\big(w_1(3), w_2(3), w_3(3)\big)$. Substituting (40) into the system (14) – (16) instead of $u(z)$, solving the resulting system subject to the initial conditions (17), and using the above mentioned requirement yield the following set of the inequality and the algebraic equation with respect to γ, C_1 and C_2:

$$aT_{\min} \le \Phi_1(\gamma, C_1, C_2) \le aT_{\max}, \tag{42}$$

$$\Phi_2(\gamma, C_1, C_2) - \gamma = 0, \tag{43}$$

where

$$\Phi_1(\gamma, C_1, C_2) \triangleq \int_0^3 \exp(-z^2/2)u^*(z, \gamma, C_1, C_2)dz, \tag{44}$$

$$\Phi_2(\gamma, C_1, C_2) \triangleq \int_0^3 \psi(z, \delta)u^*(z, \gamma, C_1, C_2)dz. \tag{45}$$

Note that $\Phi_1(\gamma, C_1, C_2)$ and $\Phi_2(\gamma, C_1, C_2)$ are the values $w_3(3)$ and $w_2(3)$, generated by the control (40).

Further, due to (19) – (20) and (31), the constants γ, C_1 and C_2 should satisfy the algebraic equations

$$C_1\big(aT_{\min} - \Phi_1(\gamma, C_1, C_2)\big) = 0, \tag{46}$$

$$C_2\big(\Phi_1(\gamma, C_1, C_2) - aT_{\max}\big) = 0. \tag{47}$$

Moreover, by (29),

$$C_1 \ge 0, \quad C_2 \ge 0. \tag{48}$$

Remark 3. *Since the NOCP has the solution, the set (42) – (48) has a solution. If this set has more than one solution, we choose the solution* $\big(\gamma = \gamma^*, C_1 = C_1^*, C_2 = C_2^*\big)$, *which provides the minimum value of the NOCP cost functional (21) in comparison with the other solutions.*

3.3 Analysis of the Set (42) – (48)

Let (γ, C_1, C_2) be a solution of this set. Since the control (40) satisfies the inequality $u^*(z, \gamma, C_1, C_2) \ge u_{\min} > 0$ for all $z \in [0, 3]$, then due to (9), (43) and (45), $\gamma > 0$. For the further analysis, we distinguish the following cases with respect to C_1 and C_2: (I) $C_1 > 0$, $C_2 > 0$; (II) $C_1 = C_2 = 0$; (III) $C_1 = 0$, $C_2 > 0$; (IV) $C_1 > 0$, $C_2 = 0$.

We start with the first case. Due to (46)-(47), this case yields the contradictory equality $aT_{\min} = aT_{\max}$, meaning that the first case is impossible.

Proceed to the second case. In this case, the function $\tilde{u}(z, \gamma, C_1, C_2)$ (see (41)) becomes $\tilde{u}(z, \gamma, C_1, C_2) = \tilde{u}(z, \gamma, 0, 0) = -B(\delta)\gamma$, $z \in [0, 3]$. Since $B(\delta)$ and γ are positive, then $\tilde{u}(z, \gamma, C_1, C_2) < 0$, $z \in [0, 3]$. Therefore, due to the equation (40) and the inequality $u_{\min} > 0$, we have $u^*(z, \gamma, C_1, C_2) = u_{\min}$, $z \in [0, 3]$. Now, substitution of this control into (44), and using (42) and the positiveness of a yield the inequality $T_{\min} \le u_{\min}$, which contradicts the inequality in (5). Thus, the second case also is impossible.

Now, let us treat the third case. In this case, we have $\tilde{u}(z, \gamma, C_1, C_2) = \tilde{u}(z, \gamma, 0, C_2) = -\big[C_2/\big(4\cosh(\delta z)\big) + B(\delta)\gamma\big] < 0$, $z \in [0, 3]$. Thus, by the same arguments as in the case (II), the third case is impossible.

Finally, let us consider the case (IV). In this case, by denoting $C \triangleq C_1$, the optimal control (40) becomes

$$u^*(z) = u^*(z, \gamma, C) = \begin{cases} u_{\min}, & \bar{u}(z, \gamma, C) \le u_{\min}, \\ \bar{u}(z, \gamma, C), & u_{\min} < \bar{u}(z, \gamma, C) \le u_{\max}, \\ u_{\max}, & \bar{u}(z, \gamma, C) > u_{\max}, \end{cases} \tag{49}$$

where

$$\bar{u}(z, \gamma, C) = C/\big(4\cosh(\delta z)\big) - B(\delta)\gamma \tag{50}$$

is the unique solution of the following equation with respect to u: $\partial G_2(u,z,\gamma,C)/\partial u = 0$, $G_2\big(u,z,\gamma,C\big) = \big(4B(\delta)\gamma\cosh(\delta z) - C\big)u + 2\cosh(\delta z)u^2$. Moreover, using the equation (46) and denoting $T \triangleq T_{\min}$, the set (42) – (43) becomes

$$\Lambda_1(\gamma,C) \triangleq \int_0^3 \exp(-z^2/2)u^*(z,\gamma,C)dz - aT = 0, \tag{51}$$

$$\Lambda_2(\gamma,C) \triangleq \int_0^3 \psi(z,\delta)u^*(z,\gamma,C)dz - \gamma = 0. \tag{52}$$

Thus, to construct the optimal control $u^*(z,\gamma,C)$ and to design the optimal SC sampling time-interval, one should solve the system (51) – (52) with respect to (γ,C) and substitute its solution into the equation (49). Due to Remark 3, the system (51) – (52) has a solution. Some properties of this solution, helpful for its numerical obtaining, are presented in the next subsection. Examples of such an obtaining are presented in Section 5.

Remark 4. *It is important to note that in the optimal control $u^*(z,\gamma,C)$ of the NOCP both bounds $u_{\min}$ and $u_{\max}$ of the geometric constraint (2) are used. At the same time, only the lower bound $T_{\min}$ of the terminal state inequality (18), equivalent to the integral inequality (3), is used in the optimal control of the NOCP. Thus, the optimal SC time-sampling depends on both bounds of the geometric constraint (2) and only on the lower bound of the integral inequality constraint (3).*

3.4 Properties of the Solution to the System (51) – (52)

Let us introduce into the consideration the following values:

$$C_{\min}(\gamma) \triangleq 4\big(B(\delta)\gamma + u_{\min}\big),$$
$$C_{\max}(\gamma) \triangleq 4\cosh(3\delta)\big(B(\delta)\gamma + u_{\max}\big),$$

$$\gamma_{\min} \triangleq u_{\min}\int_0^3 \psi(z,\delta)dz, \quad \gamma_{\max} \triangleq u_{\max}\int_0^3 \psi(z,\delta)dz,$$

$$\Gamma_{\min} \triangleq \max\{\gamma_{\min}\ ,\ 2aT\},$$
$$\Gamma_{\max} \triangleq \min\{\gamma_{\max}\ ,\ 2aT\cosh(3\delta)\}.$$

Also, in the plane (γ,C), we consider the non-empty domain

$$\Omega \triangleq \Big\{\gamma \in \big(\Gamma_{\min},\Gamma_{\max}\big), C \in \big(C_{\min}(\gamma),C_{\max}(\gamma)\big)\Big\}.$$

Based on the above introduced values and domain, we obtain the following assertions.

Assertion 1. *Let $\delta > 0$. Let (γ,C) be a solution of the system (51) – (52). Then, $(\gamma,C) \in \Omega$.*

Remark 5. *Due to Assertion 1, in the case $\delta > 0$, we can look for the solution of the system (51) – (52) not in the entire plane (γ,C), but in the bounded domain Ω, which decreases considerably the computational effort. In the particular case of $\delta = 0$, we can solve the system (51) – (52) analytically.*

Assertion 2. *Let $\delta = 0$. Then, the system (51) – (52) has the unique solution $\big(\gamma = 2aT, C = 8B(0)aT + 4T\big)$.*

Assertion 3. *For any given $\delta > 0$ and $\gamma \in \big(\Gamma_{\min},\Gamma_{\max}\big)$, the equation (51) has the unique solution $C = \widetilde{C}(\gamma)$, and*

$$\widetilde{C}(\gamma) \in \big(C_{\min}(\gamma),C_{\max}(\gamma)\big). \tag{53}$$

Moreover, $\widetilde{C}(\gamma)$ is a monotonically increasing function of $\gamma \in \big(\Gamma_{\min},\Gamma_{\max}\big)$.

Assertion 4. *For any given $\delta > 0$ and $\gamma \in \big(\Gamma_{\min},\Gamma_{\max}\big)$, the equation (52) has the unique solution $C = \bar{C}(\gamma)$, and*

$$\bar{C}(\gamma) \in \big(C_{\min}(\gamma),C_{\max}(\gamma)\big). \tag{54}$$

Moreover, $\bar{C}(\gamma)$ is a monotonically increasing function of $\gamma \in \big(\Gamma_{\min},\Gamma_{\max}\big)$.

Remark 6. *Based on Assertions 3 and 4, the γ-component of the solution to the system (51) – (52) can be obtained by solving with respect to $\gamma \in \big(\Gamma_{\min},\Gamma_{\max}\big)$ either the equation $\Lambda_2\big(\gamma,\widetilde{C}(\gamma)\big) = 0$, or the equation $\Lambda_1\big(\gamma,\bar{C}(\gamma)\big) = 0$, or the equation $\widetilde{C}(\gamma) = \bar{C}(\gamma)$.*

The proofs of Assertions 1 – 4 are presented in Appendix.

4 APPROXIMATE NUMERICAL SOLUTION OF EP

Let us divide the interval $[0,3]$ into N equal subintervals by the collocation points

$$z_i = i\Delta z, \quad i = 0,1,\ldots,N, \ \ \Delta z = 3/N. \tag{55}$$

Then, based on (55) and using the left rectangles formula (Davis and Rabinowitz, 2007), we approximate the integrals in the cost functional (10) and in the integral constraint (3).

Thus, the cost functional is approximated as:

$$J\big(u(z)\big) \approx \tilde{J}^N(U) \triangleq \Delta z\sum_{i=0}^{N-1}\psi(z_i,\delta)U_i^2 + B\left(\Delta z\sum_{i=0}^{N-1}\psi(z_i,\delta)U_i\right)^2, \tag{56}$$

where the vector $U \in E^N$ is

$$U = \left(U_0, U_1, \dots, U_{N-1}\right)^T = \left(u(z_0), u(z_1), \dots, u(z_{N-1})\right)^T. \quad (57)$$

The constraint (3) is approximated as:

$$\Delta z T_{\min} \sum_{i=0}^{N-1} \exp(-z_i^2/2) \le \Delta z \sum_{i=0}^{N-1} \exp(-z_i^2/2) U_i \le \Delta z T_{\max} \sum_{i=0}^{N-1} \exp(-z_i^2/2). \quad (58)$$

The approximation (58) of the constraint (3) is derived using the following approximation of the value a:

$$a \approx a_N \Delta z, \quad a_N \stackrel{\triangle}{=} \sum_{i=0}^{N-1} \exp(-z_i^2/2).$$

The geometric constraint (2), appearing in the EP, is approximated as:

$$u_{\min} \le U_i \le u_{\max}, \quad i = 0, 1, \dots, N-1. \quad (59)$$

Dividing the expression in the right-hand side of (56) and the inequality (58) by Δz, we obtain the following finite-dimensional cost functional and the constraint:

$$J^N(U) \stackrel{\triangle}{=} \sum_{i=0}^{N-1} \psi(z_i, \delta) U_i^2 + B\Delta z \left(\sum_{i=0}^{N-1} \psi(z_i, \delta) U_i\right)^2, \quad (60)$$

$$a_N T_{\min} \le \sum_{i=0}^{N-1} \exp(-z_i^2/2) U_i \le a_N T_{\max}. \quad (61)$$

Now, based on (60) – (61), we can formulate the following quadratic programming problem.

Quadratic Programming Problem (QPP): for a known $\delta \ge 0$, to find the vector $U = \left(U_0, U_1, \dots, U_{N-1}\right)^T$ which minimizes the cost functional (60) subject to the constraints (59), (61) and the inequality (5).

The QPP can be solved using standard optimization tools, for example, the MATLAB function "quadprog". It is reasonable to expect that for a sufficiently large N, the components U_i, $(i = 0, 1, \dots, N-1)$ of the QPP solution will be close to the corresponding values $u^*(z_i, \gamma^*, C^*)$ of the optimal control in the NOCP. In such a case, the optimal value of the cost functional (60) in the QPP multiplied by Δz will be close to the optimal value of the cost functional (21) in the NOCP.

5 NUMERICAL EVALUATION OF THE OPTIMAL AND SUBOPTIMAL SAMPLING TIME-INTERVALS

For the numerical evaluation, the following two sets of parameters are chosen:

$$\text{(I)} \quad u_{\min} = 0.5, \; u_{\max} = 3.5;$$

$$\text{(II)} \quad u_{\min} = 0.1, \; u_{\max} = 2.5.$$

5.1 Numerical Solution of the System (51) – (52)

To obtain the sampling time-interval $u^*(z, \gamma, C)$, the system (51) – (52) was solved numerically. The value of γ was calculated by application of the bisection algorithm to the equation $\widetilde{C}(\gamma) = \bar{C}(\gamma)$ for $\gamma \in (\Gamma_{\min}, \Gamma_{\max})$.

Using Assertions 3 and 4, the functions $\widetilde{C}(\gamma)$ and $\bar{C}(\gamma)$ also were derived by the bisection method for $(\gamma, C) \in \Omega$.

In Figs. 1 – 2, the functions $\widetilde{C}(\gamma)$ and $\bar{C}(\gamma)$, along with the functions $C_{\min}(\gamma)$ and $C_{\max}(\gamma)$, are depicted in the logarithmic scale for the set (I) with $\delta = 2.5$ (Fig. 1) and for the set (II) with $\delta = 2.9$ (Fig. 2). The value of $T = T_{\min} = 1$. It is seen that the functions $\widetilde{C}(\gamma)$ and $\bar{C}(\gamma)$ are monotonically increasing, which corresponds to the claims of Assertions 3 and 4. The solution (γ, C) of the set (51) – (52), depicted by the circle, belongs to the set Ω as it is stated in Assertion 1. Moreover, since the derivative of $\bar{C}(\gamma)$ with respect to γ is larger than the derivative of $\widetilde{C}(\gamma)$, this solution is unique.

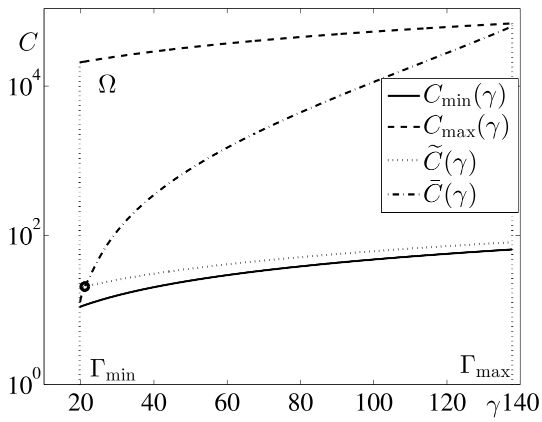

Figure 1: $\widetilde{C}(\gamma)$, $\bar{C}(\gamma)$, $C_{\min}(\gamma)$ and $C_{\max}(\gamma)$: set (I).

The numerical solution (γ, C) of the set (51) – (52), and the absolute values of the functions $\Lambda_1(\gamma, C)$

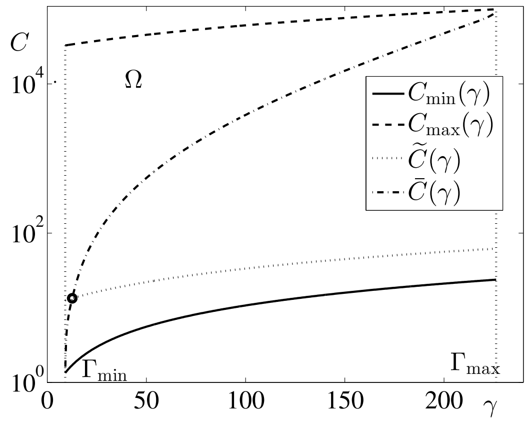

Figure 2: $\widetilde{C}(\gamma)$, $\bar{C}(\gamma)$, $C_{\min}(\gamma)$ and $C_{\max}(\gamma)$: set (II).

and $\Lambda_2(\gamma, C)$ are presented in Table 1. It is seen that the obtained numerical solution provides the deviations of $\Lambda_1(\gamma, C)$ and $\Lambda_2(\gamma, C)$ from zero smaller than 10^{-4}.

Table 1: Numerical solution of (51) – (52).

Set	δ	γ	C	$\lvert\Lambda_1(\gamma,C)\rvert$	$\lvert\Lambda_2(\gamma,C)\rvert$
(I)	2.5	21.1	20.6	$2.9 \cdot 10^{-5}$	$7.0 \cdot 10^{-5}$
(II)	2.9	12.6	13.4	$7.1 \cdot 10^{-5}$	$7.6 \cdot 10^{-5}$

5.2 Optimal Sampling Time-Interval (EP Solution) vs. Approximate Sampling Time-Interval (QPP Solution)

In Figs. 3 – 4, the optimal sampling time-interval $u^*(z, \gamma, C)$ (the EP solution), given by the analytical expression (49), is compared to the approximate sampling time-interval U^* (the QPP solution) for the set (I) with $\delta = 2.5$ (Fig. 3) and for the set (II) with $\delta = 2.9$ (Fig. 4). In both cases, $T_{\min} = 1$, $T_{\max} = 2$. It is seen that the approximation, obtained for $N = 100$, and the optimal sampling time-interval match well.

Note that in both cases, for the approximate solution, the left-hand side inequality in the constraint (61) is satisfied as the equality

$$\sum_{i=0}^{N-1} \exp(-z_i^2/2) U_i = a_N T_{\min},$$

thus mimicking the corresponding property of the EP analytical solution.

Based on the equation (49) and the above presented numerical calculations, the optimal sampling time-interval for the set (I), depicted in Fig. 3, can be rewritten as:

$$u^*(z, \gamma, C) = \begin{cases} \bar{u}(z, \gamma, C), & z \in [0, 0.47), \\ u_{\min}, & z \in [0.47, 3]. \end{cases}$$

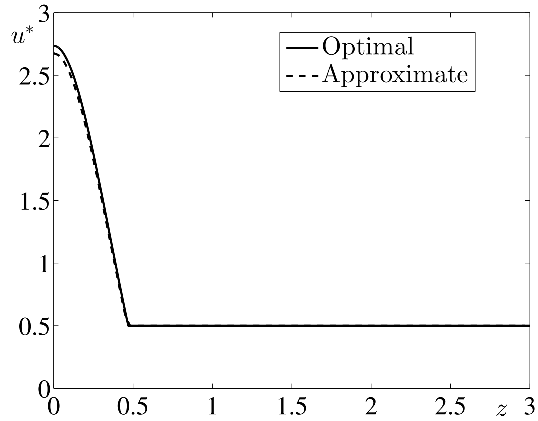

Figure 3: Optimal vs. approximate sampling time-interval: set (I).

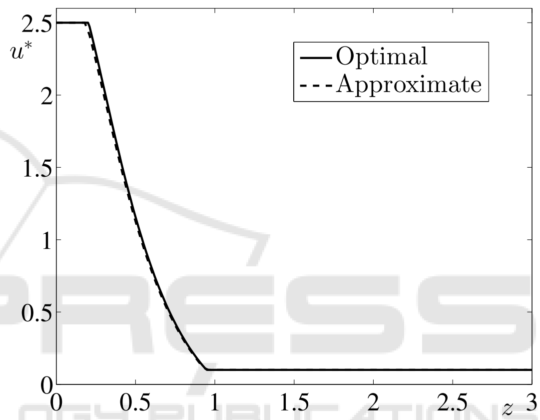

Figure 4: Optimal vs. approximate sampling time-interval: set (II).

The optimal sampling time-interval for the set (II), depicted in Fig. 4, can be rewritten in the form

$$u^*(z, \gamma, C) = \begin{cases} u_{\max}, & z \in [0, 0.21), \\ \bar{u}(z, \gamma, C), & z \in [0.21, 0.95), \\ u_{\min}, & z \in [0.95, 3]. \end{cases}$$

In Fig. 5, the optimal (minimum) value of the expected loss, given by (6)–(9), is depicted as a function of δ for the set (I), $T_{\min} = 1$, and different coefficient functions $k = k(\delta)$. It is seen that for each considered $k(\delta)$, the optimal expected loss admits the maximum for some value of δ, belonging to the interval $(0, 1)$.

6 CONCLUSIONS

In this paper, the problem of constructing an optimal state-feedback sampling time-interval for the statistical control was considered. The expected loss, quadratically dependent on the delay in the detection of a process change, was chosen as the criterion of

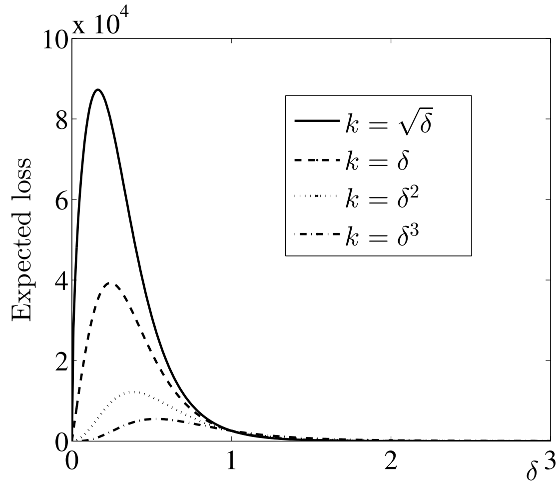

Figure 5: Optimal expected loss as a function of δ: set (I).

the optimization (minimization). This expected loss also depends on the numerical parameter δ, characterising a magnitude of the process change. The problem of the expected loss minimization was reduced to the extremal problem in the form of a nonstandard calculus of variations problem where the sampling time-interval of the statistical control is a minimizing function. This minimizing function depends of the process state. Two methods of the solution of this extremal problem were proposed. The first method transforms the original extremal problem to an equivalent optimal control problem. Then, the latter was solved using the Pontryagin's Maximium Principle, which yields the explicit analytical expression for the optimal sampling time-interval. This expression contains two parameters. For obtaining these parameters, the set of two algebraic equations was derived and analyzed. Based on this analysis, the method of solution of this set was proposed. The second method of the solution of the original extremal problem uses its discretization. This leads to a finite-dimensional extremal problem (the quadratic programming problem), approximating the original one. This quadratic programming problem was solved using the MATLAB function quadprog, providing the suboptimal sampling time-interval for the statistical control. The optimal and suboptimal sampling time-intervals were evaluated by numerical examples. This evaluation has shown a good match of the optimal analytical sampling time-interval and the suboptimal numerical sampling time-interval. The optimal (minimum) value of the expected loss was constructed numerically as a function of the parameter δ. It was shown that this function has a single maximum.

It should be noted that the results, obtained in this paper, are rather theoretical, and these results are strongly based on two assumptions: (a) the sample mean $\bar{x}$ of the characteristic index x in the statistical control is normally distributed; (b) the value of the parameter δ, characterising a magnitude of the process change, is known. Therefore, one can ask the following: "What will happen if at least one of these assumptions is violated?" To answer this question, the following issues will be studied in a future:

(I) an evaluation (by extensive computer simulations) of the sampling time-interval, obtained in this paper, in the cases where either a distribution of the sample mean $\bar{x}$ differs from the normal one, or the value of δ is unknown;

(II) a design of optimal sampling time-interval in the case where a distribution of the sample mean $\bar{x}$ is not normal;

(III) a design of optimal sampling time-interval, robust with respect to δ, in the case where the value of this parameter is unknown.

Results of these studies will be presented in forthcoming papers.

REFERENCES

Amin, R. W. and Hemasinha, R. (1993). The switching behavior of $\bar{X}$ charts with variable sampling intervals. *Communication in Statistics - Theory and Methods*, 22(7):2081–2102.

Amin, R. W. and Miller, R. W. (1993). A robustness study of charts with variable sampling intervals. *Journal of Quality Technology*, 25(1):36–44.

Babrauskas, V. (2008). Heat release rates. In *SFPE Handbook of Fire Protection Engineering*, pages 1–59. National Fire Protection Association.

Bashkansky, E. and Glizer, V. Y. (2012). Novel approach to adaptive statistical process control optimization with variable sampling interval and minimum expected loss. *International Journal of Quality Engineering and Technology*, 3(2):91–107.

Bulelzai, M. A. K. and Dubbeldam, J. L. A. (2012). Long time evolution of atherosclerotic plaques. *Journal of Theoretical Biology*, 297:1–10.

Carpenter, T. E., OBrien, J. M., Hagerman, A., and McCarl, B. (2011). Epidemic and economic impacts of delayed detection of foot-and-mouth disease: a case study of a simulated outbreak in california. *Journal of Veterinary Diagnostic Investigation*, 23(1):26–33.

Chew, X. Y., Khoo, M. B., Teh, S. Y., and Castagliola, P. (2015). The variable sampling interval run sum $\bar{X}$ control chart. *Computers & Industrial Engineering*, 90:25–38.

Costa, A. F. B. (1994). X charts with variable sample size. *Journal of Quality Technology*, 26(3):155–163.

Costa, A. F. B. (1997). $\bar{X}$ charts with variable sample size and sampling intervals. *Journal of Quality Technology*, 29(2):197–204.

Costa, A. F. B. (1998). Joint $\bar{X}$ and r control charts with variable parameters. *IIE Transactions*, 30(6):505–514.

Costa, A. F. B. (1999a). $\bar{X}$ charts with variable parameters. *Journal of Quality Technology*, 31(4):408–416.

Costa, A. F. B. (1999b). Joint $\bar{X}$ and r charts with variable sample sizes and sampling intervals. *Journal of Quality Technology*, 31(4):387–397.

Costa, A. F. B. and Magalhaes, M. S. D. (2007). An adaptive chart for monitoring the process mean and variance. *Quality and Reliability Engineering International*, 23(7):821–831.

Davis, P. J. and Rabinowitz, P. (2007). *Methods of Numerical Integration*. Dover Publications Inc., Mineola, NY.

Gelfand, I. M. and Fomin, S. V. (1963). *Calculus of Variations*. Prentice-Hall, Englewood Cliffs, NJ.

Glizer, V. Y., Turetsky, V., and Bashkansky, E. (2015). Statistical process control optimization with variable sampling interval and nonlinear expected loss. *Journal of Industrial and Management Optimization*, 11(1):105–133.

Hatjimihail, A. T. (2009). Estimation of the optimal statistical quality control sampling time intervals using a residual risk measure. *PLoS ONE*, 4(6):e5770.

Ioffe, A. D. and Tihomirov, V. M. (1979). *Theory of Extremal Problems*. North-Holland Pub. Co., Amsterdam, Netherlands.

Kim, S. and Frangopol, D. M. (2011). Optimum inspection planning for minimizing fatigue damage detection delay of ship hull structures. *International Journal of Fatigue*, 33(3):448–459.

Li, Z. and Qiu, P. (2014). Statistical process control using dynamic sampling. *Technometrics*, 56(3):325–335.

Pontryagin, L. S., Boltyanskii, V. G., Gamkrelidze, R. V., and Mishchenko, E. F. (1962). *The Mathematical Theory of Optimal Processes*. Interscience, New York, NY.

Prabhu, S. S., Montgomery, D. C., and Runger, G. C. (1994). A combined adaptive sample size and sampling interval x control scheme. *Journal of Quality Technology*, 26(3):164–176.

Qiu, P. (2013). *Introduction to Statistical Process Control*. Chapman and Hall/CRC, Boca Raton, FL.

Reynolds, M. R. (1995). Evaluating properties of variable sampling interval control charts. *Sequential Analysis*, 14(1):59–97.

Reynolds, M. R., Amin, R. W., Arnold, J. C., and Nachlas, J. (1988). $\bar{X}$ charts with variable sampling intervals. *Technometrics*, 30(2):181–192.

Ross, S. (2009). *A First Course in Probability*. Prentice Hall, Upper Saddle River, NJ.

Sebastião, P. and Soares, C. G. (1995). Modeling the fate of oil spills at sea. *Spill Science and Technology Bulletin*, 2(2-3):121–131.

Sultana, I., Ahmed, I., Chowdhury, A. H., and Paul, S. K. (2014). Economic design of $\bar{X}$ control chart using genetic algorithm and simulated annealing algorithm. *International Journal of Productivity and Quality Management*, 14(3):352–372.

Taguchi, G., Chowdhury, S., and Wu, Y. (2007). *Taguchi's Quality Engineering Handbook*. John Wiley and Sons Inc., Hoboken, NJ.

APPENDIX

The proofs of Assertions 1-4 are based on the following auxiliary propositions.

Auxiliary Propositions

Proposition 1. *For any given $\delta \geq 0$, there are no solutions of the set (51) – (52) in the half-planes*

$$C \leq C_{\min}(\gamma) \tag{62}$$

and

$$C \geq C_{\max}(\gamma) \tag{63}$$

of the plane (γ, C).

Proof. First, let us prove the claim of the proposition with respect to the half-plane (62), i.e., in the case where the pair (γ, C) satisfies this inequality. In this case, since $\cosh(\delta z) \geq 1$ for $z \geq 0$, then $\bar{u}(z, \gamma, C) \leq u_{\min}$ for $z \geq 0$. By virtue of (49), the latter means that $u^*(z, \gamma, C) = u_{\min}$. Thus, due to (4), (5), (51), and the notation $T = T_{\min}$,

$$\Lambda_1(\gamma, C) = a\big(u_{\min} - T\big) < 0, \tag{64}$$

meaning that the above mentioned pair (γ, C) does not satisfy the equation (51). The claim of the assertion with respect to the half-plane (63) is proven similarly. □

Proposition 2. *For any given $\delta \geq 0$, there are no solutions of the set (51) – (52) in the half-planes*

$$\gamma \leq \gamma_{\min} \tag{65}$$

and

$$\gamma \geq \gamma_{\max} \tag{66}$$

of the plane (γ, C).

Proof. First of all, let us note the following. Since $0 < u_{\min} < u_{\max}$ and $\psi(z, \delta) > 0$, $z \in [0, 3]$, then

$$0 < \gamma_{\min} < \gamma_{\max}. \tag{67}$$

Consider the case where the pair (γ, C) satisfies the strict inequality in (65). In this case, due to (9), (49), (52) and (65), we have

$$\Lambda_2(\gamma, C) > \int_0^3 \psi(z, \delta)\Big(u^*(z, \gamma, C) - u_{\min}\Big)dz \geq 0, \tag{68}$$

meaning that (γ, C) does not satisfy the equation (52). Now, let us consider the case $\gamma = \gamma_{\min}$. In this case, using (49) and (52), we have

$$\Lambda_2(\gamma_{\min}, C) =$$

$$\int_0^3 \psi(z,\delta)\big(u^*(z,\gamma_{\min},C) - u_{\min}\big)dz \geq 0. \qquad (69)$$

By virtue of (49), the inequality in (69) becomes equality only if $u^*(z,\gamma_{\min},C) \equiv u_{\min}$ for all $z \in [0,3]$, yielding $\bar{u}(z,\gamma_{\min},C) \leq u_{\min}$ for all $z \in [0,3]$. Using (50), one directly obtains that the latter inequality for $z = 0$ is equivalent to the inequality (62) with $\gamma = \gamma_{\min}$. However, in this case by virtue of Proposition 1, the equation (51) is not satisfied. Thus, the claim of the assertion with respect to the half-plane (65) has been proven. The claim of the proposition with respect to the half-plane (66) is proven similarly. $\square$

Proposition 3. *Let $\delta > 0$. Let (γ, C) be a solution of the set (51) – (52). Then, the component γ of this solution satisfies the inequality*

$$2aT < \gamma < 2aT\cosh(3\delta). \qquad (70)$$

Proof. Substitution of (9) into (52) yields after a simple rearrangement

$$\gamma = 2\int_0^3 \exp\big(-z^2/2\big)\cosh(\delta z)u^*(z,\gamma,C)dz. \qquad (71)$$

Applying the Mean Value Theorem to the integral in the right-hand side of (71) and taking into account the fact that $\cosh(\delta z)$ monotonically increases with respect to $z \in [0,3]$ for $\delta > 0$, we obtain

$$\gamma = 2\cosh(\delta\bar{z})\int_0^3 \exp\big(-z^2/2\big)u^*(z,\gamma,C)dz. \qquad (72)$$

where $\bar{z}$ is some value from the interval $(0,3)$.

Due to (51), we can replace the integral in (72) with aT, which leads to the equality $\gamma = 2aT\cosh(\delta\bar{z})$. The latter, along with the inequality $1 < \cosh(\delta\bar{z}) < \cosh(3\delta)$, directly implies the statement of the proposition. $\square$

Proposition 4. *For $\delta > 0$, the following inequality holds:*

$$0 < \Gamma_{\min} < \Gamma_{\max}. \qquad (73)$$

Proof. Using the definition of $\gamma_{\max}$ (see Subsection 3.4), the equation (4), the inequality (5), the notation $T = T_{\min}$ and the same arguments as in the proof of Proposition 2, we obtain

$$\gamma_{\max} = 2au_{\max}\cosh(\delta\tilde{z}) > 2aT, \qquad (74)$$

where $\tilde{z}$ is some value from the interval $(0,3)$.

Similarly, using the definition of $\gamma_{\min}$ (see Subsection 3.4), the equation (4) and the inequality (5), we have

$$\gamma_{\min} = 2au_{\min}\cosh(\delta\tilde{z}) < 2aT\cosh(3\delta). \qquad (75)$$

Now, the inequalities (67), (74) and (75), along with the inequality $2aT < 2aT\cosh(3\delta)$, yield immediately the statement of the proposition. $\square$

Proposition 5. *For any $\gamma \geq 0$, the following inequality is valid:*

$$0 < C_{\min}(\gamma) < C_{\max}(\gamma). \qquad (76)$$

Proof. The assertion directly follows from the definitions of $C_{\min}(\gamma)$ and $C_{\max}(\gamma)$ (see Subsection 3.4). $\square$

Proof of Assertion 1

First of all let us note that, due to Propositions 4 and 5, the domain Ω indeed is non-empty. Now, the statement of the assertion directly follows from Propositions 1-3 and the definitions of $\Gamma_{\min}$, $\Gamma_{\max}$ and Ω.

Proof of Assertion 2

For $\delta = 0$, the function $\bar{u}(z,\gamma,C)$, given by (50), becomes a constant, i.e.,

$$\bar{u}(z,\gamma,C) \equiv \frac{C}{4} - B(0)\gamma, \quad z \in [0,3]. \qquad (77)$$

Moreover, due to Proposition 1, in order to be a solution of the set (51) – (52), the pair (γ, C) should satisfy the inequality

$$u_{\min} < \frac{C}{4} - B(0)\gamma < u_{\max}. \qquad (78)$$

The latter, along with (49) and (77), means that

$$u^*(z,\gamma,C) \equiv \frac{C}{4} - B(0)\gamma, \quad z \in [0,3]. \qquad (79)$$

Substituting (79) into the set (51) – (52) and using (9) and the fact that $\delta = 0$ directly yield the unique solution of (51) – (52) in the form

$$\frac{C}{4} - B(0)\gamma = T, \quad \gamma = 2aT. \qquad (80)$$

The latter yields the unique $C = 8B(0)aT + 4T$, which completes the proof of the assertion.

Proof of Assertion 3

The existence and uniqueness of $\widetilde{C}(\gamma)$ is proven similarly to the work (Glizer et al., 2015) (see Lemma 5.1 and its proof where $\gamma = 0$). The inclusion (53) follows from the proof of Proposition 1. Let us prove the monotonicity of $\widetilde{C}(\gamma)$, $\gamma \in (\Gamma_{\min}, \Gamma_{\max})$. We prove this feature of $\widetilde{C}(\gamma)$ by contradiction. Namely, we assume that the statement on the monotonic increasing of this function is wrong. This means the existence of $\gamma_1 \in (\Gamma_{\min}, \Gamma_{\max})$ and $\gamma_2 \in (\Gamma_{\min}, \Gamma_{\max})$ such that

$$\gamma_1 < \gamma_2, \tag{81}$$

while

$$\widetilde{C}(\gamma_1) \geq \widetilde{C}(\gamma_2). \tag{82}$$

The equation (50), and the inequalities (81) and (82) directly yield

$$\bar{u}\big(z, \gamma_1, \widetilde{C}(\gamma_1)\big) > \bar{u}\big(z, \gamma_2, \widetilde{C}(\gamma_2)\big), \quad z \in [0,3]. \tag{83}$$

Due to the equation (49) and the inequality (83), we immediately have

$$u^*\big(z, \gamma_1, \widetilde{C}(\gamma_1)\big) \geq u^*\big(z, \gamma_2, \widetilde{C}(\gamma_2)\big) \quad \forall z \in [0,3]. \tag{84}$$

By virtue of the inclusion (53), we obtain that for any $\gamma \in (\Gamma_{\min}, \Gamma_{\max})$ and all $z \in [0,3]$ the function $u^*\big(z, \gamma, \widetilde{C}(\gamma)\big)$ is neither identical $u_{\min}$, nor identical $u_{\max}$. This fact, along with (49) and (83), yields the existence of a point $\tilde{z} \in [0,3]$, such that

$$u^*\big(\tilde{z}, \gamma_1, \widetilde{C}(\gamma_1)\big) > u^*\big(\tilde{z}, \gamma_2, \widetilde{C}(\gamma_2)\big). \tag{85}$$

Further, from (49) and (50), one directly concludes that for any $\gamma \in (\Gamma_{\min}, \Gamma_{\max})$ the control $u^*\big(z, \gamma, \widetilde{C}(\gamma)\big)$ is continuous function of $z \in [0,3]$. This observation, along with the inequality (85), yields the existence of the interval $[z_1, z_2]$, $(z_1 < z_2)$, such that $[z_1, z_2] \subset [0,3]$ and

$$u^*\big(\tilde{z}, \gamma_1, \widetilde{C}(\gamma_1)\big) > u^*\big(\tilde{z}, \gamma_2, \widetilde{C}(\gamma_2)\big) \quad \forall z \in [z_1, z_2]. \tag{86}$$

Now, the definition of $\Lambda_1(\gamma, C)$ (see (51)), along with the inequalities (84) and (86), yields

$$\Lambda_1\big(\gamma_1, \widetilde{C}(\gamma_1)\big) > \Lambda_1\big(\gamma_2, \widetilde{C}(\gamma_2)\big). \tag{87}$$

The latter contradicts the fact that $\widetilde{C}(\gamma_1)$ and $\widetilde{C}(\gamma_2)$ are solutions of the equation (51) with respect to C for $\gamma = \gamma_1$ and $\gamma = \gamma_2$. This contradiction implies that the function $\widetilde{C}(\gamma)$ monotonically increases for $\gamma \in (\Gamma_{\min}, \Gamma_{\max})$, which completes the proof of the assertion.

Proof of Assertion 4

We start the proof with the first two statements of the assertion. Let the pair (γ, C) be any fixed satisfying the inequalities $\Gamma_{\min} < \gamma < \Gamma_{\max}$, $C \leq C_{\min}(\gamma)$.

Using the equation (50), the definition of $C_{\min}(\gamma)$ (see Subsection 3.4), the inequality (73), as well as the positiveness of $B(\delta)$ and the inequality $\cosh(\delta z) \geq 1$, $z \in [0,3]$, one directly has the following chain of the inequalities:

$$\bar{u}(z, \gamma, C) \leq \frac{u_{\min}}{\cosh(\delta z)}$$

$$+B(\delta)\gamma\left(\frac{1}{\cosh(\delta z)} - 1\right) \leq u_{\min}, \quad z \in [0,3]. \tag{88}$$

Hence, due to (49), $u^*(z, \gamma, C) \equiv u_{\min}$, $z \in [0,3]$. Using the latter and (52), (65) – (66) yield

$$\Lambda_2(\gamma, C) < 0, \quad \gamma \in (\Gamma_{\min}, \Gamma_{\max}), \quad C \leq C_{\min}(\gamma). \tag{89}$$

It is shown similarly, that

$$\Lambda_2(\gamma, C) > 0, \quad \gamma \in (\Gamma_{\min}, \Gamma_{\max}), \quad C \geq C_{\max}(\gamma). \tag{90}$$

Further, from (49) and (50), one directly concludes that for any $z \in [0,3]$ and $\gamma \in (\Gamma_{\min}, \Gamma_{\max})$ the control $u^*(z, \gamma, C)$ is continuous and monotonically increasing function of $C \in (-\infty, +\infty)$. Therefore, $\Lambda_2(\gamma, C)$ (see (52)) is a continuous and monotonically increasing function of $C \in (-\infty, +\infty)$ for any $\gamma \in (\Gamma_{\min}, \Gamma_{\max})$. The latter, along with (89) – (90), implies immediately the existence of the unique solution $C = \bar{C}(\gamma)$ of the equation (52) for any $\gamma \in (\Gamma_{\min}, \Gamma_{\max})$. Moreover, the inclusion (54) is satisfied for any $\gamma \in (\Gamma_{\min}, \Gamma_{\max})$.

The monotonicity of $\bar{C}(\gamma)$, $\gamma \in (\Gamma_{\min}, \Gamma_{\max})$ is proven similarly to the same feature of the function $\widetilde{C}(\gamma)$ (see the proof of Assertion 3). This completes the proof of the assertion.