ISSN 2782-2427

CONTROL SCIENCES

3/2021

[1]Russian Academy of Sciences.

CONTROL SCIENCES
Scientific Technical Journal

6 issues per year
ISSN 2782-2427
Open access

Published since 2021

Original Russian Edition
Problemy Upravleniya
Published since 2003

FOUNDER AND PUBLISHER
V.A. Trapeznikov
Institute of Control Sciences
of Russian Academy of Sciences

Editorial address
65 Profsoyuznaya st., office 410,
Moscow 117997, Russia

☎/🖹 +7(495) 198-17-20, ext. 1410

✉ pu@ipu.ru

URL: http://controlsciences.org

Published: August 3, 2021

Registration certificate of
Эл № ФС 77-80482
of 17 February 2021
issued by the Federal Service
for Supervision of Communications,
Information Technology, and Mass
Media

CONTROL SCIENCES

3.2021

CONTENTS

DOI: http://doi.org/10.25728/cs.2021.3.1

INVESTIGATION OF MULTIVARIABLE AUTOMATIC CONTROL SYSTEMS FOR COMPLEX DYNAMIC OBJECTS BASED ON PETROV'S PARADIGM[1]

B.G. Il'yasov and G.A. Saitova

Ufa State Aviation Technical University, Ufa, Russia

✉ ilyasov@ugatu.su, ✉ saitova@bk.ru

Abstract. This paper considers some approaches to studying the properties of multivariable automatic control systems (MACSs), particularly their stability, based on different descriptive models. The theory presented below extends the previously known ideas of Academician B.N. Petrov, which are fundamental in the classical theory of automatic control. Petrov's theory is based on the structural and functional decomposition of MACSs into separate real subsystems and multiple connections between them, described by a new model, and the study of system properties using frequency-domain methods. Therefore, this theory is related to the physical (engineering) approach to dynamic systems analysis. A method for describing MACSs by the individual characteristics of subsystems and multiple connections is suggested. Stability criteria for linear MACSs with identical subsystems and a stability criterion for the system's equilibrium are established. A technology for finding the parameters of periodic motions and assessing their stability for nonlinear MACSs is introduced. Some numerical examples with technical objects illustrate this technology for studying the properties of MACSs.

Keywords: multivariable system, decomposition, frequency-domain methods, linear system, nonlinear system.

Dedicated to the blessed memory of B.N. Petrov, Academician of the USSR Academy of Sciences

INTRODUCTION

In the first half of the twentieth century, complex dynamic objects (CDOs) (aircrafts, power and propulsion systems, electrical installations, complex technological processes in the petrochemical, engineering, and other industries) appeared in operation. Multiple output variables of such objects needed automatic control via appropriate actions applied to some of their control variables. As a result, a new class of controlled systems was created and called multivariable automatic control systems (MACSs) for CDOs. (Note that the terms "multiply connected systems" and "multi-connected systems" are also used in the literature.)

This class of systems has the following peculiarity: when maintaining a given value of its output variable, each subsystem of the system inevitably affects the operation of other subsystems due to the physical processes (aerodynamic, gas-dynamic, electrical, chemical, thermal, etc.) occurring in the controlled object and connecting the subsystems. Practice demanded studying the new class of systems, and new complex problems arose in automatic control theory.

This paper reveals the content of Petrov's paradigm (a model of problems and their solutions) subject to the investigation of MACSs. Petrov's idea is based on the structural and functional decomposition of MACSs for CDOs into physical subsystems and multiple connections between them and analysis of their properties using frequency-domain methods. It extends his earlier idea [1] formulated in 1945, which underlies classical control theory.

We have a modest intention: to show the possibilities of Petrov's paradigm by his students' publications, thus approving it as a new technology for examining the properties of MACSs equally with the technologies

[1]This work was supported by the Russian Foundation for Basic Research, project no. 20-08-01172 A.

based on other paradigms. Therefore, when speaking about the advantages of the classical paradigms for studying the stability of MACSs, we would like to emphasize the practical benefits of Petrov's paradigm in solving this problem.

1. MODELS OF MULTIVARIABLE AUTOMATIC CONTROL SYSTEMS FOR COMPLEX DYNAMIC OBJECTS

We separate three types of models of the MACSs for CDOs, writing them in the vector-matrix form:

$$X(s)=W(s)U(s)+Q(s)F(s),$$
$$U(s)=R(s)\left[X^o(s)-X(s)\right],$$

where $X(s), X^o(s), U(s)$, and $F(s)$ are the vectors of controlled variables, reference signals, control variables, and disturbances, respectively; $W(s)=\left\|W_{ij}(s)\right\|_{n\times n}$ and $R(s)=\left\|R_{ij}(s)\right\|_{n\times n}$ are the matrix transfer functions (MTFs) of the object and controller, respectively (the controller includes the actuator); finally, $Q(s)=\left\|Q_{ij}(s)\right\|_{n\times n}$ is the MTF of the disturbance.

Let **the first model** reflect only natural connections between the subsystems through a multivariable controlled object. Then the MTF $R(s)$ is a diagonal matrix, $R(s)=\left\|R_{ij}(s)\right\|_{n\times n}$, in which the transfer function $R_i(s)$ of the control device and actuator of the corresponding subsystem stands on the diagonal.

Consider **the second model** of MACSs, in which a multivariable object represents a set of autonomously operating objects (power units, robots, electric motors, etc.). Then its MTF will be a diagonal matrix $W(s)=\left\|W_i(s)\right\|_{n\times n}$. In the system, the set of objects is complexly controlled by a multivariable controller with the MTF $R(s)=\left\|R_{ij}(s)\right\|_{n\times n}$. The design problems remain the same as in the first model.

Finally, in **the third model**, connections between subsystems are through a multivariable controlled object and a multivariable controller. The possibilities of such MACSs are still underinvestigated.

The structural diagram of MACSs is shown in Fig. 1.

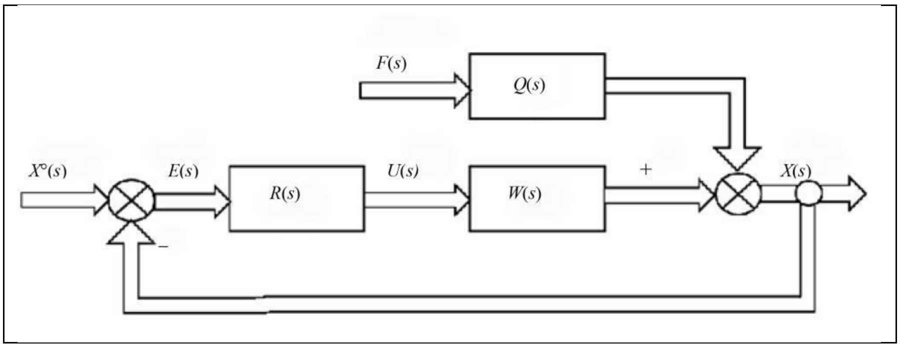


Fig. 1. The structural diagram of MACSs for CDOs.

For MACSs for CDOs, the first and main problem is stability. Suppose that the stability of each subsystem is established using a well-known classical criterion (Lyapunov, Routh, Hurwitz, Stodola, Nyquist, or Hermite–Mikhailov). In this case, the characteristic equation of the multivariable system is the product of the characteristic equations of the subsystems considering their interconnections. As a result, we obtain a characteristic equation of the form

$$D(s)=a_0s^n+a_1s^{n-1}+\ldots+a_{n-1}s+a_n=0, \qquad (1)$$

but with a very high degree n > (25–40). Nowadays, the system's stability can be assessed by equation (1) without finding the eigenvalues. However, apart from stability analysis, equation (1) yields no constructive conclusions about the reasons of stability of the entire multivariable system using the above criteria. Modern algorithms and programs allow determining the stability of MACSs but not the potential effect on their degree of stability exerted, e.g., by a simultaneous change in the set of physical parameters of the subsystems and their interconnection coefficients: the relationship between the coefficients a_i of equation (1) and the physical parameters of the system is often implicit.

2. VOZNESENSKII'S AUTONOMY PARADIGM

In 1938, I.N. Voznesenskii formulated the autonomy principle of MACSs with respect to free motions; see the paper [2].

To implement this principle, we should consider the third model of MACSs and design artificial connections between the subsystems through a multivariable controller to compensate for the natural connections through the multivariable object. Then the entire MACS will be decomposed into separate stable subsystems. However, the complete compensation of the natural connections is often impossible due to the inertia of the system's elements. In this case, the matter concerns compensating the connections only for a particular operating mode of the system. With a change in the operating modes, the entire MACS needs to be retuned.

Certain success can be achieved in the case of an optimal tuning of MACSs [3].

In the paper [4], the autonomy problem was solved using the structures of the subsystems with infinitely increasing gains without violating the stability of the entire MACS.

When a high performance of the subsystems is achieved, their mutual effect on each other becomes minimal: due to the rapid execution of their control task, the subsystems have no time to respond to the effects of other subsystems fully. As noted by the author of [4], with such properties of the system, it can be made invariant with respect to the load. At the same time, for several important objects, autonomy contradicts technological processes.

As was demonstrated later, multivariable control is vital for complex dynamic objects: it guarantees higher efficiency than the transition to an autonomous structure. Despite this fact, many industrial systems are designed within the autonomy paradigm: metallurgy, power engineering, steam boilers control, and heating turbines control in stable conditions, to name a few.

3. A PARADIGM BASED ON THE STATE-SPACE DESCRIPTION OF MACS

This paradigm implements a purely mathematical approach: an MACS described by an n-order differential equation is reduced to a system of first-order differential equations (the Cauchy form). This approach gave a huge impetus to studying the properties of dynamic systems with feedback control.

Let a dynamic system be described by a system of equations in the vector-matrix form:

$$\dot{X} = AX + BU, \quad Y = CX, \qquad (2)$$

where $A = \|a_{ij}\|_{n\times n}, B = \|b_{ik}\|_{n\times m}, C = \|c_{\lambda i}\|_{p\times n}$, are numerical matrices with numerical entries; X is the vector of state variables; Y is the vector of output variables; U is the vector of input variables (control).

Note that such MACSs should be designed in the class of fully controllable, fully observable, and structurally stable systems. The MACS structure completely merges with the mathematical structure (2).

By assigning different values to the numerical parameters of the matrices A, B, and C, we determine the system's state X (the motion trajectory, or the solution of system (2)) for any time instant. This process can be repeated. The wide variety of research methods obtained by different modifications of the model (2) does not change the essence of studying dynamic systems in the state space. However, after transforming the characteristic equation of the MACS,

$$D(s) = \det[Is - A] = 0,$$

where I denotes an identity matrix of compatible dimensions, stability analysis reduces to the characteristic equation (1). In other words, the genesis (origin) of the stability of MACSs remains an open problem.

Another drawback of this paradigm, from the engineering viewpoint, is the difficulty of establishing a direct relationship between the state variables and the physical parameters of a real system. For this reason, implementing this paradigm in the design of real MACSs encounters definite difficulties and limitations. Using this paradigm, the design engineer has no clear answer to the question: how will the properties of an MACS change if the characteristics of certain elements of the real system's structure are simultaneously modified? Will the real system remain stable in this case?

Despite the applied problems mentioned above, this paradigm is effective when studying MACSs at the level of their mathematical description (2). Some problems can be eliminated if the model (2) is constructed not top-to-bottom but bottom-to-top: the models of elements are gradually combined into the system (2).

Even if the state matrix A of a multivariable system is the multi-connected composition of the submatrices of the corresponding physical subsystems and their interconnection coefficients, the investigation can be more efficient, e.g., when analyzing the stability margins of the MACS under variations of the individual parameters of the subsystems. However, the problem becomes more complicated if the system characteristics of the subsystems are simultaneously changed. In addition, the stability of an MACS is determined by complexes, i.e., the combinations of two, three, …, n subsystems related by multi-connection characteristics without an explicit matrix form. As we therefore believe, studying the genesis of the stability of MACSs by varying the system characteristics of subsystems will require additional research, causing definite difficulties and limitations in the design of real MACSs.

The state-space paradigm can be successfully applied for assessing the stability of a given MACS by its mathematical model. This paradigm underlies the modern control of dynamic feedback systems and is considered by many researchers [5–15].

4. A PARADIGM BASED ON THE MATRIX TRANSFER FUNCTION DESCRIPTION OF MACS

The matrix description of the MACS dynamics is quite convenient and makes the results visual. However, the MACS description using the MTF (Fig. 1) is somewhat incomplete and reflects the entire system's behavior to some extent only: it provides no information about the behavior of the uncontrolled and unobserved parts of the system. For a complete description of the system, we should pass to the state-space representation (2), whose dimension equals the number of the system's degrees of freedom. The vector of controlled outputs of the MACS has a linear relationship with the vector of state variables.

This paradigm, to some extent, repeats the shortcomings of the state-space paradigm. Performing matrix transformations, the design engineer obtains the final result under the given system parameters and input actions $U(t)$, and the matrix transformations themselves do not reflect physical processes.

The question remains: how to establish a relationship between the physical parameters of a real system and the parameters of matrix transformations? As in the previous paradigm, determining the effect of a change in any physical parameter, e.g., on the system's stability, requires repeating the numerical experiment many times. The problem becomes even more complicated if a group of physical parameters is varied simultaneously.

Unlike classical control theory, e.g., stability criteria cannot be derived in an analytical form through matrix transformations. The stability analysis of MACSs still involves the characteristic equation (1), obtained by transforming the equation

$$D(s) = \det\left[I + W(s)R(s)\right] = 0. \qquad (3)$$

The stability analysis by equations (1) or (3) does not explain how the properties of individual subsystems and the properties of interconnections between the subsystems affect the stability of the entire MACS.

The properties of MACSs were examined using matrix methods by Krasovskii [16], Meerov [17], Chinaev [18], Morozovskii [19], and Sobolev [20]. The applications of the theory of multivariable control systems were described by Bodner [21, 22], Shevyakov [23], Yanushevskii [24], Ray [25], and other researchers.

5. PETROV'S PARADIGM BASED ON THE STRUCTURAL AND FUNCTIONAL DECOMPOSITION OF MACS AND FREQUENCY-DOMAIN METHODS

In the late 1970s, Academician B.N. Petrov posed the following problem: to describe MACSs by larger (than the elements of subsystems) physical blocks and multiple connections between them. Petrov is known in automatic control theory for the paradigm of passing from a system of differential equations to its structural representation by functional blocks with operators and connections between them [1]. The paradigm gave a new and huge impetus to developing and creating the classical automatic control theory. The researchers of MACSs faced a similar problem. The main goal was to preserve the physicality of the structure and all the transformations so that the design engineer knew exactly (without solving the system of differential equations) what changes would contribute to improving the dynamic properties of the MACS.

The solution of this problem—a new description of MACSs through the physical characteristics of the subsystems and multi-connection characteristics—was presented in [26]. This description was used later in [27–29].

According to Fig. 1, an MACS consists of a set of interconnected and closed separate subsystems, each controlling a particular output of the object. We will consider the first model of MACSs, the most common in engineering practice, in which different subsystems are connected through a multivariable controlled object. In this case, as noted above, the controller's MTF $R(s)$ is diagonal, and the transfer functions $R_i(s)$ are located on the diagonal.

The main requirement to describe the dynamic characteristics of MACSs is to make one subsystem (connection) distinguishable from another. In other words, each characteristic must have its "individuality." Therefore, the characteristics of both subsystems and connections are labeled in a systematic way, and their dimensions are indicated.

As a separate ith subsystem, consider a closed loop system with its internal structure, which controls the ith output of the multivariable controlled object. As an individual characteristic (IC) of the ith subsystem, consider the one that fully reflects the investigated properties of the subsystem and expresses these properties and distinctive features. For example, these requirements are satisfied by the individual transfer function $\Phi_i(s)$ in the control mode, when the ith subsystem operates in an isolated (autonomous) mode independently of the other subsystems:

$$\forall i : \Phi_i(s) = \frac{X_i(s)}{X_i^o(s)} = \frac{R_i(s)W_{ii}(s)}{1 + R_i(s)W_{ii}(s)}.$$

Note that its gain-phase response (GPR) $\Phi_i(j)$ and error transfer function $\Phi_\varepsilon(s) = 1 - \Phi_i(s)$ can also be considered as the IC of the ith subsystem. In this case, the ith subsystem corresponds to a real physical system with an independent design value and an individual dynamic characteristic (model), widely used in the classical automatic control theory with its well-developed methods for studying closed loop single-input single-output (SISO) systems.

A multi-characteristic of interconnections (MCI) is introduced to concretize the relationships between the subsystems and express their features. This characteristic (model) reflects the existing relationships between the subsystems, and their mathematical model is built from typical dynamic links of the classical automatic control theory. For the class of MACSs under consideration, cross-connections between the subsystems are determined only by off-diagonal elements $W_{ij}(s)(i \neq j)$ of the MTF $W(s)$ of the multivariable controlled object. They form the connection matrix $\|W_{ij}(s)\gamma_{ij}\|$, where $\gamma_{ij} = \begin{cases} 1, & i \neq j, \\ 0, & i = j, \end{cases}$ $i, j = \overline{1,n}$. This matrix reflects the individual relationships between the pairs, triples, quads, etc. of the subsystems.

For this class of MACSs, we should identify the absolute effect of cross-connections and, most importantly, their effect relative to direct connections through the controlled object. The latter connections are characterized by the diagonal matrix $\|W_{ij}(s)\delta_{ij}\|$, where $\delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases}$ $i, j = \overline{1,n}$.

This relative connection between the subsystems is considered as the MCI in MACSs. The mathematical model of the MCI between k subsystems is given by

$$H_k(s) = \frac{\det\|W_{ij}(s)\gamma_{ij}\|_{k\times k}}{\det\|W_{ij}(s)\delta_{ij}\|_{k\times k}}. \qquad (4)$$

The characteristic $H_k(s)$ can be real, complex, or imaginary.

By the nature of its effect, the MCI can be flexible or rigid; stabilizing or destabilizing; forcing, inertial, or lagging. In the general case, it characterizes the sign, magnitude (“strength”), and the character of connections in a group of k subsystems combined into a single whole through this multiple connection.

Thus, the model (4) concretizes the nature of the connections between different subsystems. Changing the sign, parameters, and structure of the model, we can design connections in an MACS, ensuring the required properties of the entire system.

Among the various types of MACSs, a class of homogeneous (identical, single-type) MACSs is often distinguished: the ICs $\Phi_i(s)$ of their subsystems are identical and equal. For this class of MACSs, it is reasonable to introduce the concept of a generalized characteristic of connections (GCC) as the sum of the connection characteristics for the subsystems of one equivalence class. For example, for all interconnected pairs of the subsystems, the GCC will have the form $H_2(s) = \sum_{i,j=1}^{C_n^2} H_{ij}(s)$; for all interconnected triplets in an n-dimensional system, the form $H_3(s) = \sum_{i,j,k=1}^{C_n^3} H_{ijk}(s)$, where $C_n^k = \frac{n!}{k!(n-k)!}$ is the number of k-combinations from n elements; and so on.

The GCC expresses the total connection $H_k(s)$ created by a group of C_n^k identical subsystems. The terms in $H_k(s)$ can be of different signs: the connections between the subsystems within one equivalence class may compensate each other (in this case, $H_k(s) = 0$). The same occurs if there are no connections between the subsystems within the kth equivalence class.

The ICs of subsystems and the MCIs introduced above allow passing from the MACS description at the level of elementary dynamic links to that at the level of subsystems and MCIs formed from these elements.

The new description of MACSs will also have new structural diagrams, e.g., in the form of a loop labeled digraph (Fig. 2).

For the MACSs shown in Fig. 2, the characteristic equation can be expressed through the ICs of the subsystems and their MCIs [27–29]:

$$D(\Phi, H) = 1 + \sum_{i,j=1}^{C_n^2} \Phi_i \Phi_j H_{ij} + \sum_{i,j,k=1}^{C_n^3} \Phi_i \Phi_j \Phi_k H_{ijk} + \ldots + H_{1n} \prod_{i=1}^{n} \Phi_i = 0. \qquad (5)$$

Here the ICs $\Phi_i(s)$ and the MCIs $H_k(s)$ are functions of the complex variable s. Hence, frequency-domain methods can be applied to study the characteristic equation (5).

Consider the characteristic equation (5) for the class of MACSs with homogeneous subsystems. Since

$$\Phi_1(s) = \Phi_2(s) = \cdots = \Phi_n(s) = \Phi(s),$$

we obtain

$$D(\Phi, H) = 1 + H_2(s)\Phi^2(s) + H_3(s)\Phi^3(s) + \cdots + = H_k(s)\Phi^k(s) + \cdots + H_n(s)\Phi^n(s) = 0,$$

where $H_k(s)$ is the GCC of the subsystem of dimension k given by (4).

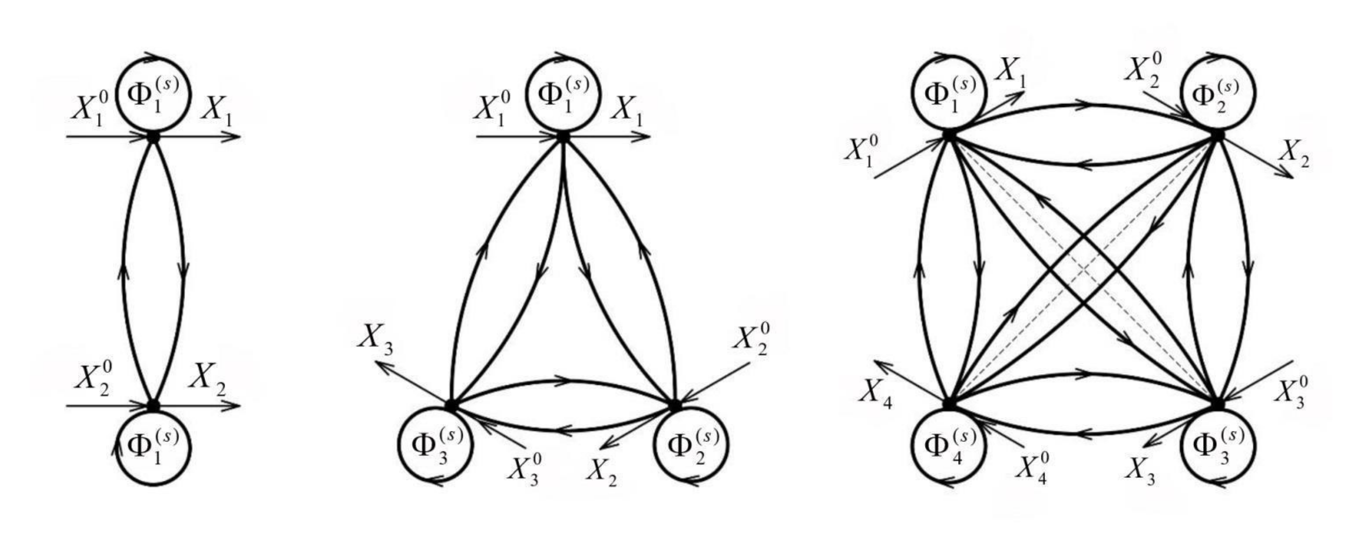


Fig. 2. MACS as a loop labeled digraph

This characteristic equation for one variable Φ(*s*) has complex-valued coefficients.

Now let the connection be either through the object or the controller, expressed by numerical coefficients k_{ij}. Then the MCI has the form

$$h_k = \frac{\det\left\|K_{ij}\gamma_{ij}\right\|_{k\times k}}{\det\left\|K_{ij}\delta_{ij}\right\|_{k\times k}}. \quad (6)$$

As a result, the characteristic equation for the homogeneous systems [30] reduces to

$$D(\Phi,h) = 1 + h_2\Phi^2(s) + h_3\Phi^3(s) + \cdots + h_n\Phi^n(s) = 0 \quad (7)$$

The characteristic equation (7) can be written in another form. For $\Phi(s) = \frac{1}{M(s)}$, where *M*(*s*) is the characteristic polynomial of the subsystem, the characteristic equation of the MACS [31] becomes

$$D(M,h) = M^n(s) + h_2M^{n-2}(s) + h_3M^{n-3}(s) + \cdots + h_{n-1}M(s) + h_n = 0. \quad (8)$$

Here the Hermite–Mikhailov characteristic polynomial *M*(*s*) acts as an IC of the subsystems.

Writing the characteristic equations in the form (7) and (8) opens up new possibilities in the investigation of multivariable systems.

Thus, Petrov's paradigm allows studying separately the individual characteristics $\Phi_i(s)$ and $M_i(s)$ of the subsystems and their GCC $H_k(s)$ and integrating them into a single characteristic of a real MACS to study its system properties.

6. STABILITY ANALYSIS OF LINEAR MACS BASED ON PETROV'S PARADIGM

Since the 1980s, the scientific direction based on Petrov's paradigm was developed at Ufa Aviation Institute; since 1992, at Ufa State Aviation Technical University in the scientific school headed by Prof. B.G. Il'yasov.

In the first stages, the characteristic equation (5) was solved using frequency-domain and numerical methods. They allowed assessing the stability of MACSs for gas turbine engines of supersonic aircrafts in various flight conditions.

The early results were presented in the monographs [27, 28] jointly with researchers from Trapeznikov Institute of Control Sciences and the Central Institute of Aviation Motors.

In that time, the conditions of static stability (the positivity of the free term of the characteristic equation) were derived for the MACS consisting of identical astatic subsystems with Φ(0) = 1,

$$D(h,\Phi) = 1 + h_2 + h_3 + \cdots + h_n > 0, \quad (9)$$

and the MACS consisting of identical static subsystems,

$$D(h,\Phi) = 1 + h_2\Phi^2(0) + h_3\Phi^3(0) + \cdots + h_n\Phi^n(0) > 0, \quad (10)$$

where $\Phi(0) = \frac{k}{1+k}$, and *k* denotes the gain of the open loop subsystem.

In contrast to the matrix form, this form allows easily analyzing the effect of connections between subsystems on the static stability of MACS, i.e., easily assessing the violation of structural stability due to numerical changes in the connections between the subsystems.

For an MACS with different static subsystems, the static stability condition has the form

$$D(h,\,\Phi)=1+h_2\Phi_1(0)\Phi_2(0)+h_3\prod_{i=1}^{3}\Phi_i(0)+$$
$$+\ldots+h_n\prod_{i=1}^{n}\Phi_i(0)>0$$

For assessing dynamic stability, the stability problem was analytically solved [28, 29] by the *D*-partition method for a three-variable system consisting of three identical second-order subsystems interconnected through the outputs by numerical coefficients. Its characteristic equation

$$D(h,\Phi)=1+h_2\Phi^2(s)+h_3\Phi^3(s)=0 \qquad (11)$$

was represented on the plane of the interconnection coefficients (h_2, h_3); see Fig. 3. The function

$$\Phi(s)=1/(\tau^2s^2+2\xi\tau s+1),$$

where $\tau=0.5$ s and $\xi\in(0,1\ldots1)$, was taken as an individual characteristic of the subsystem.

According to Fig. 3, the less the subsystems are damped, the smaller the stability domain of the entire MACS will be.

Using numerical and frequency-domain methods, we can solve equation (11) for a more complex form and a higher order of the function $\Phi(s)$. Note that this approach was applied for assessing the stability of the designed three-variable ACSs for gas turbine engines of supersonic aircrafts based on their mathematical models [27, 28], for the first time in practice.

The results were used to formulate the fundamental postulates (regularities) for the MACSs consisting of stable identical subsystems interconnected through output variables.

Postulate 1. For this class of MACSs, the static stability conditions (9) and (10) are simultaneously a structural stability condition: in the case of their violation, the stability of MACSs cannot be achieved due to changes in the subsystems' parameters. This conclusion also applies to the MACSs with stable subsystems having different structures and individual characteristics.

Postulate 2. If a structurally unstable subsystem appears in a stable MACS, this will be a sufficient condition for the structural instability of the entire MACS in which all subsystems are interconnected through the output variables by numerical interconnection coefficients: a change in the interconnection coefficients h_2 or h_2 prevents from restoring the structural stability of the entire MACS.

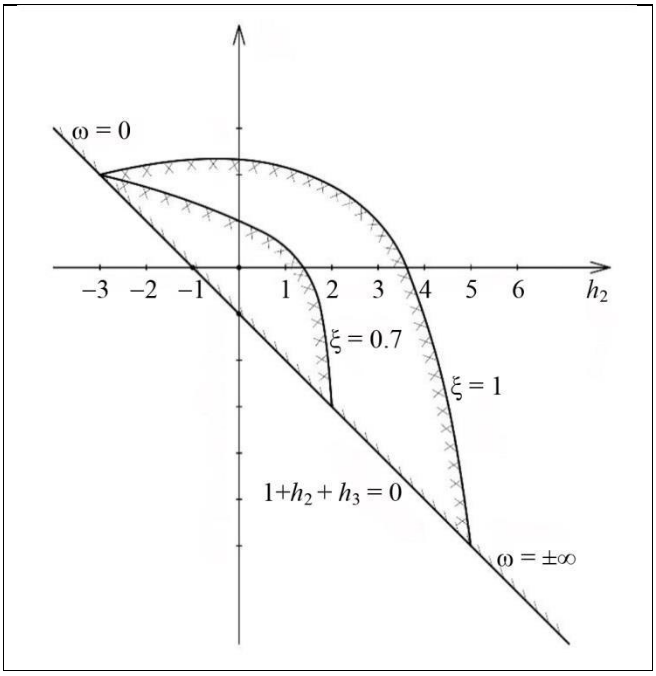


Fig. 3. Stability domains of three-variable ACS under different values of ξ.

Postulate 3. Consider an MACS in which the numerical connections between *n* subsystems are implemented either through a multivariable controlled object or a multivariable controller. For this MACS, there additionally exist *n* critical points located on its stability boundary and determined by the multiconnection equations.

Postulate 4. The critical points are determined by the roots of the mult-connection equation that are obtained either from the characteristic equation (7) by replacing $\Phi(s)$ with *x*,

$$D(h,x)=1+h_2x^2+h_3x^3+\cdots+h_nx^n=0, \qquad (12)$$

or from the characteristic equation (8) by replacing $M(s)$ with *z*,

$$D(h,z)=z^n+h_2z^{n-2}+h_3z^{n-3}+\cdots+h_n=0, \qquad (13)$$

where the equation order *n* is given by the number of interconnected subsystems.

The equation of multiple connections through the interconnection coefficients h_i in Petrov's paradigm was introduced for the first time.

Postulate 5. Consider an MACS in which identical subsystems are rigidly interconnected through the output variables. This MACS is dynamically stable if and only if the corresponding subsystem's GPR $\Phi(j\omega)$ neither hits nor encircles any critical point of the multi-connection equation (12) as the frequency ω varies from 0 to + ∞, and the Hermite–Mikhailov characteristic polynomial $M(j\omega)$ does not hit but encircles all critical points of the multi-connection equation (13).

Postulate 6. For an MACS with identical subsystems, stability margins (gain and phase margins) are defined as the distance of the corresponding frequency response of a subsystem on the complex plane to the nearest critical point of the multi-connection equations (12) or (13); see [32]. This postulate also holds if identical subsystems have elements with pure delay [33].

Postulate 7. A linear MACS with n identical subsystems interconnected through the output variables lies on the boundary of oscillatory stability: steady oscillations (periodic motions with frequency ω_n and amplitude α_n) occur in the system if one of the characteristics of an identical subsystem passes through the nearest critical point of the multi-connection equations (12) or (13). In this case, the amplitude and frequency of oscillations are determined from the corresponding individual characteristics of the subsystem and the multi-connection characteristics [30, 31].

Postulate 3 was used to formulate two frequency-domain stability criteria.

Criterion 1. A linear MACS with identical subsystems and numerical interconnection coefficients is stable if and only if the subsystem's GPR $\Phi(j\omega)$ neither hits nor encircles any critical point given by the roots of the multi-connection equation (12) as the frequency ω varies from 0 to $+\infty$ [30].

This criterion was confirmed by numerical examples in [31, 32]; also, see Fig. 4.

Example 1. The transfer function of the closed loop stable and separate subsystem of a three-variable system has the form $\Phi(s) = 1/(s^3 + 3s^2 + 2s + 1)$.

The subsystems have the multiple connection given by

$$h = \begin{pmatrix} 0 & 0.2 & 0.4 \\ 0.2 & 0 & 0.5 \\ 0.2 & 0.5 & 0 \end{pmatrix}.$$

The characteristic equation of the system with numerical coefficients has the form

$$D(\Phi) = 1 + h_2\Phi^2 + h_3\Phi^3 = 0, \qquad (14)$$

where h_3 = –0.6 and h_2 = –0.37 according to formula (6). Replacing the function Φ in (14) with the complex variable x, we obtain the multi-connection equation

$$D(x) = 1 + h_2x^2 + h_3x^3 = 0. \qquad (15)$$

Its roots (the critical points) are x_1 = –5.64, x_2 = –2.00, and x_3 = 1.48.

We construct on the complex plane the hodograph of the function $\Phi(j\omega)$ for $\omega \in (0, +\infty)$. On the same plane, we arrange the critical points x_i, $i = \overline{1, 3}$. By Criterion 1, the multivariable system will be stable since the GPR $\Phi(j\omega)$ of the autonomous subsystem does not encircle any critical point of equation (15) as the frequency ω varies from 0 to $+\infty$; see Fig. 4. ♦

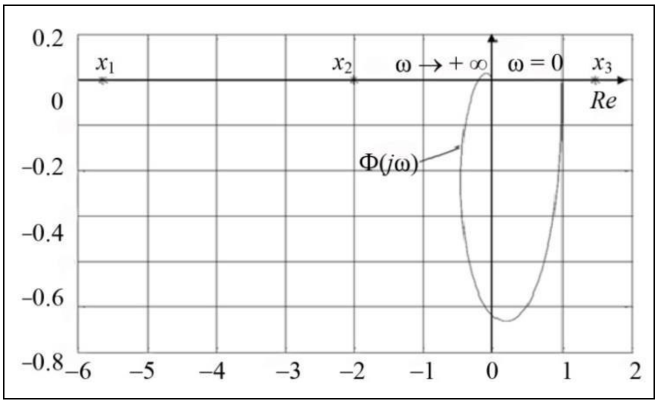


Fig. 4. The hodograph of Φ(jω) and critical points x_i, $i = \overline{1, 3}$.

Criterion 2. A linear MACS with identical subsystems and numerical interconnection coefficients is stable if and only if the subsystem's characteristic hodograph (the Hermite–Mikhailov curve) neither hits nor encircles any critical point given by the roots of the multi-connection equation (13) as the frequency ω varies from 0 to $+\infty$ [31].

A numerical example of calculating a three-variable system confirms this criterion; see Fig. 5.

Example 2. We write the characteristic equation of Example 1 according to formula (8):

$$D(M) = M^3 + h_2M + h_3 = 0, \qquad (16)$$

where the coefficients h_3 = –0.06 and h_2 = –0.37 are given by (6).

The critical points satisfy the equation $z^3 + h_2z + h_3 = 0$. The closed separate subsystems have the characteristic equation corresponding to a stable subsystem: $M(s) = s^3 + 3s^2 + 2s + 1 = 0$. The roots of the critical points equation are z_1 = 0.68, z_2 = –0.5, and z_3 = –0.18.

We construct on the complex plane the Hermite–Mikhailov curve $M(j\omega)$ for $\omega \in (0, +\infty)$. On the same plane, we arrange the roots z_i. According to Fig. 5, the Hermite–Mikhailov curve $M(j\omega)$ encircles all the roots. By Criterion 2, the multivariable system is stable. This conclusion is confirmed by the transient processes of the system. ♦

Postulate 8. According to the studies presented above, the postulates also hold for an MACS containing identical subsystems with a digital or discrete-time control part [31].

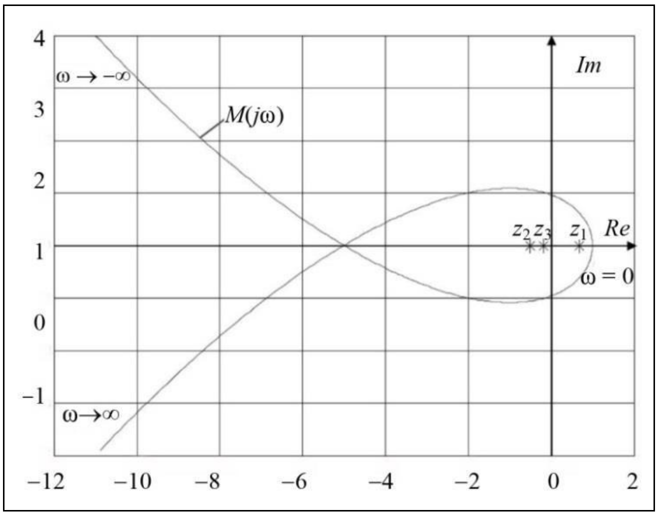


Fig. 5. The Hermite–Mikhailov curve *M(jω)* and critical points z_i, $i=\overline{1,3}$, in Example 2.

7. STUDYING THE PROPERTIES OF NONLINEAR MACS BASED ON PETROV'S PARADIGM

This approach can also be used to study the properties of nonlinear MACSs. For example, consider a class of nonlinear MACSs with identical subsystems containing elements with nonlinear static characteristics. The subsystems are connected through a multivariable controlled object.

Let the harmonic linearization method be applied to this class of MACSs. An additional strict requirement is that the characteristics of all the subsystems and resulting closed loops satisfy the filtering condition.

We represent the nonlinear MACS as the interconnection of a nonlinear element (NE) and the linear part W_{lin} of the system; see Fig. 6.

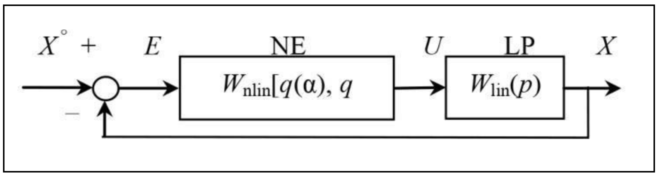


Fig. 6. The structural diagram of nonlinear MACS.

Here p denotes the differentiation operator. The functions $W_{\text{lin}}(p)$ and $W_{\text{nlin}}\left[q(\alpha),\, q'(\alpha)\right]$ form the operators of the system's linear part (LP) and nonlinear element (NE), respectively. The latter element is subjected to harmonic linearization.

The individual characteristic of the harmonically linearized identical subsystem has the form

$$\Phi(p,a)=\frac{W_{\text{nlin}}\left[q(\alpha),q'(\alpha)\right]W_{\text{lin}}(p)}{1+W_{\text{nlin}}\left[q(\alpha),q'(\alpha)\right]W_{\text{lin}}(p)},$$

where $q(\alpha)$ and $q'(\alpha)$ are the harmonic linearization coefficients; α is the input signal amplitude.

We write the characteristic equation for the entire MACS with identical nonlinear subsystems:

$$D(p,\alpha)=1+h_2\Phi^2(p,\alpha)+h_3\Phi^3(p,\alpha)+ +\dots+h_n\Phi^n(p,\alpha)=0.$$

Passing from the function $\Phi(p,\alpha)$ to $M(p,\alpha)=1/\Phi(p,\alpha)$, we obtain

$$D(p,\alpha)=M^n(p,\alpha)+h_2M^{n-2}(p,\alpha)+ +h_3M^{n-3}(p,\alpha)+\dots+h_n=0.$$

For these two equations, the multi-connection equations (12) and (13), respectively, hold as well.

We introduce the frequency-domain characteristics with the change $p=j\omega$. Then each characteristic equation is a function of ω and α: $D(\omega,\ \alpha)=0$. Postulates 1–8 are true for the harmonically linearized MACS. Hence, we may formulate another postulate for it.

Postulate 9. Consider a nonlinear MACS with harmonically linearized identical subsystems. This MACS is stable if and only if the characteristics $\Phi(j\omega,\alpha)$ do not encircle any critical point as the frequency ω varies from 0 to $\pm\infty$, and the curves $M(j\omega,\alpha)$, where the amplitude α belongs to some range, encircle all critical points of the multi-connection equations without hitting them.

Postulate 10. Periodic motions occur in the nonlinear MACS if either the characteristic $\Phi(j\omega,\alpha)$ or the Hermite–Mikhailov curve (the subsystem's characteristic polynomial) $M(j\omega,\alpha)$ hit a critical point of the corresponding multi-connection equation (12) or (13). The frequency ω_{per} and amplitude α_{per} of periodic motions are determined using classical control theory methods. The amplitude α_{per} is calculated by the corresponding characteristic; the frequency ω_{per}, by the multi-connection equation.

Postulate 11. Like in the classical control theory, the stability of periodic motions is assessed by the direction of deformation of the curves $\Phi(j\omega,\alpha)$ or $M(j\omega,\alpha)$ under increasing the amplitude α.

The periodic motions in a linear homogeneous MACS with fuzzy controllers in separate subsystems were analyzed using the same technique.

Example 3. It is required to investigate a nonlinear three-variable system for the presence of self-oscillations. The multivariable system consists of identical nonlinear subsystems interconnected through the output variables Y by numerical coefficients. The nonlinear subsystem is a standard structure consisting of a nonlinear element (NE) and a linear part (LP); see Fig. 7.

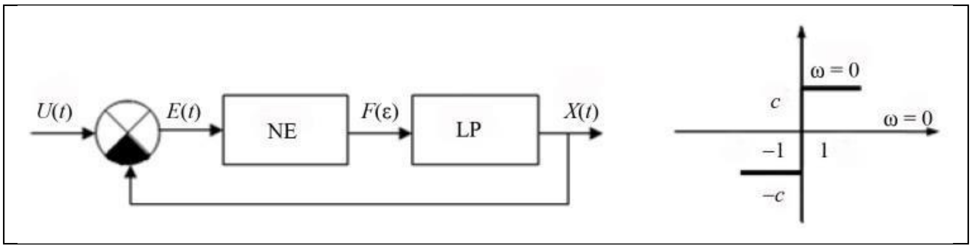


Fig. 7. The structural diagram of nonlinear MACS.

The nonlinear element is a relay (Fig. 7) with the harmonic linearization coefficients $q(\alpha) = 4c/\pi\alpha$, where $c = \pi$, and $q'(\alpha) = 0$.

Let the linear part have the transfer function $W_{\text{lp}} = \dfrac{2}{p(p^2+p+1)}$.

Then the characteristic polynomial $M(p)$ of the subsystem's harmonically linearized equation is written as

$$M(p) = \alpha(p^3 + p^2 + p) + 8. \qquad (17)$$

The interconnection coefficients of the subsystems are given by the matrix

$$h = \begin{bmatrix} 0 & 0.75 & -1.45 \\ 0.18 & 0 & 0.75 \\ 0.75 & 0.18 & 0 \end{bmatrix}.$$

The characteristic equation of the three-variable system has the form

$$D(M) = M^3 + h_2 M + h_3 = 0. \qquad (18)$$

For the autonomous subsystems, the self-oscillation parameters are $\omega_{\text{per}} = 1$ and $\alpha_{\text{per}} = 8$. Replacing M with z, from formula (17) and equation (13) we obtain the critical point equation

$$D(z) = z^3 + h_2 z + h_3 = 0. \qquad (19)$$

We calculate the coefficients h_2 and h_3 of the characteristic equation (18) by formulas (6). For the given numerical values of the interconnection coefficients, we obtain $h_2 = 0.8175$ and $h_3 = 1.728$. Then the roots of (19) are $z_1 = -0.976$ and $z_{2,3} = 0.488 \pm 1.238\,j$.

This problem can be solved graphically. We construct the hodograph of the function $M(j\omega)$ (17) for $\omega \in (0, +\infty)$ and $\alpha \in (0, +\infty)$. On the same plane, we arrange the eigenvalues of the critical point equation, z_i (see Fig. 8). Of all the roots, the nearest critical one is $z_1 = -0.976$: the curve $M(j\omega, \alpha)$ hits the remaining roots, and the nonlinear MACS is unstable according to the above criterion.

Hence, there are stable periodic motions with the parameters $\omega_{\text{per}} = 1$ and $\alpha_{\text{per}} = 8.976$ in this multivariable system. Note that under numerical connections in multivariable systems, the frequency ω of the subsystems remains the same, whereas the amplitude α_{per} of oscillations changes compared to the autonomous subsystem. ♦

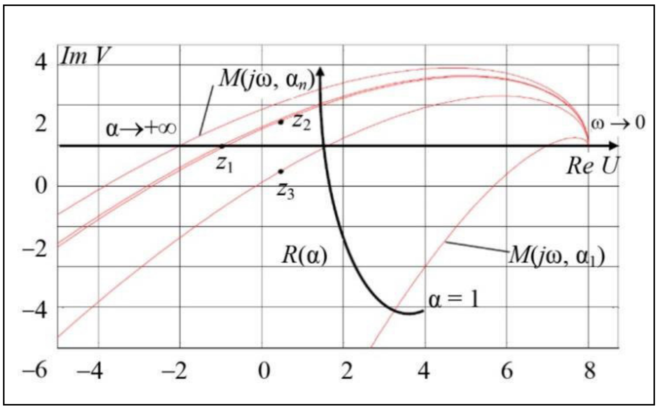


Fig. 8. The Hermite–Mikhailov curves $M(j\omega, \alpha)$ and critical points z_i, $i = \overline{1, 3}$, in Example 3.

Thus, Petrov's paradigm allows extending traditional control theory methods to the class of linear and nonlinear MACSs, including those with logical elements [32], artificial intelligence elements [35], adaptive systems [36], and variable structure systems [37]. What is important, the physical meaning of the effect of system elements on the properties of the entire MACS is preserved.

CONCLUSIONS

Academician B.N. Petrov and his students proposed a new paradigm for studying MACSs based on the description of the characteristic equation of a linear MACS through the individual characteristics of its subsystems and multiple connections between them. Within this paradigm, the system properties are investigated in the frequency domain. Such an approach was pioneering in the theory of multivariable systems.

Forming the multi-connection equation from the interconnection coefficients and finding new critical points for the subsystems to assess the stability of the entire MACS was novel in the theory of multivariable systems. As a result, new criteria for the stability of MACSs were established.

Petrov's paradigm involves the structural and functional decomposition of MACSs and frequency-domain methods to investigate the properties of MACSs. This paradigm fundamentally differs from the existing approaches: it preserves the physical meaning of each element of the subsystem and each connection and their role in the properties of MACSs.

Over the past 40 years, the theory of MACSs based on Petrov's paradigm was developed in the works of his students: new theoretical results were obtained and used in applications. As was shown, the theoretical and practical results established for the linear and nonlinear MACSs based on Petrov's paradigm keep the "spirit" of the classical control theory and the physical (engineering) sense of the ongoing research of complex systems.

Petrov's paradigm is a significant contribution of Russian scientists to global research as a new approach to studying various classes of MACSs for CDOs and revealing their unique properties.

Investigations of MACSs based on Petrov's paradigm open up new opportunities for studying various classes of MACSs for complex dynamic objects.

REFERENCES

1. Petrov, B.N., On Construction and Transformation of Block Diagrams, *Izv. Akad. Nauk SSSR. Otd. Tekh. Nauk*, 1945, no. 12, pp. 1146–1162. (In Russian.)
2. Voznesenskii, I.N., On Control of Machines with a Large Number of Controlled Parameters, *Avtomatika i Telemekhanika*, 1938, no. 4-5, pp. 65–78. (In Russian.)
3. Kornilov, Yu.G., Autonomous Control as an Extremum Problem, *Izv. Akad. Nauk SSSR. Otd. Tekh. Nauk*, 1954, no. 4. (In Russian.)
4. Meerov, M.V., On the Autonomy of Multivariable Systems Stable under an Infinite Increase of the Steady-State Accuracy, *Avtomatika i Telemekhanika*, 1956, no. 5, pp. 410–424. (In Russian.)
5. *Spravochnik po teorii avtomaticheskogo upravleniya* (Handbook of Automatic Control Theory), Krasovskii, A.A., Ed., Moscow: Nauka, 1987. (In Russian.)
6. Andreev, Yu.N., *Upravlenie konechnomernymi lineinymi ob"ektami* (Control of Finite-Dimensional Linear Plants), Moscow: Nauka, 1976. (In Russian.)
7. Voronov, A.A., *Ustoichivost', upravlyaemost', nablyudaemost'* (Stability, Controllability, Observability), Moscow: Nauka, 1979. (In Russian.)
8. Derusso, P.M., Roy, R.J., and Close, C.M., *State Variables for Engineers*, New York: John Wiley, 1967.
9. Kwakernaak, H. and Sivan, R., *Linear Optimal Control Systems*, 1st ed., Hoboken: Wiley-Interscience, 1972.
10. Hsu, J.C. and Meyer, A.U., *Modern Control Principles and Applications*, McGraw-Hill, 1968.
11. Kalman, R.E., Falb, P.L., and Arbib, M.A., *Topics in Mathematical System Theory*, New York: McGraw-Hill, 1969.
12. Tou, J.T., *Modern Control Theory* (*Electrical & Electronic Engineering*), 1st ed., McGraw-Hill, 1964.
13. Bukov, V.N., *Vlozhenie sistem. Analiticheskii podkhod k analizu i sintezu matrichnykh sistem* (Nesting of Systems. An Analytical Approach to Analysis and Design of Matrix Systems), Kaluga: Izd. Nauchn. Lit. N.F. Bochkarevoi, 2006. (In Russian.)
14. Bazhenov, S.G., Kozyaychev, A.N., and Korolyov, V.S., Stability Analysis of Airplane with MIMO Control System Based on Frequency Methods, *Control Sciences*, 2020, no. 2, pp. 20–27. (In Russian.)
15. Parsheva, E.A., A Modified High Order Adaptation Algorithm for Decentralized Control of Multivariable Plants with State Delay, *Control Sciences*, 2008, no. 3, pp. 37–43. (In Russian.)
16. Krasovskii, A.A., On Two-Channels Control Systems Described by Equations with Complex Parameters, *Avtomatika i Telemekhanika*, 1957, no. 2, pp. 123–136. (In Russian.)
17. Meerov, M.V., *Sistemy mnogosvyaznogo regulirovaniya* (Multivariable Control Systems), Moscow: Nauka, 1965. (In Russian.)
18. Chinaev, P.I., *Metody analiza i sinteza mnogomernykh avtomaticheskikh system* (Analysis and Design Methods for Multi-Dimensional Automatic Systems), Kiev: Tekhnika, 1969. (In Russian.)
19. Morozovskii, V.T., *Mnogosvyaznye sistemy avtomaticheskogo regulirovaniya* (Multivariable Automatic Control Systems), Moscow: Energiya, 1970. (In Russian.)
20. Sobolev, O.S., *Metody issledovaniya lineinykh mnogosvyaznykh system* (Methods for Studying Linear Multivariable Systems), Moscow: Energoizdat, 1985. (In Russian.)
21. Bodner, V.A., *Teoriya avtomaticheskogo upravleniya poletom* (Automatic Flight Control Theory), Moscow: Nauka, 1961. (In Russian.)
22. Bodner, V.A., Ryazanov, Yu.A., and Shaimardanov, F.A., *Sistemy avtomaticheskogo upravleniya dvigatelyami letatel'nykh apparatov* (Automatic Control Systems for Aircraft Engines), Moscow: Mashinostroenie, 1973. (In Russian.)
23. Shevyakov, A.A., *Avtomatika aviatsionnykh i raketnykh silovykh ustanovok* (Automation of Aircraft and Missile Propulsion Systems), Moscow: Mashinostroenie, 1970. (In Russian.)
24. Yanushevskii, R.T., *Teoriya lineinykh optimal'nykh mnogosvyaznykh sistem upravleniya* (Theory of Linear Optimal Multivariable Control Systems), Moscow: Nauka, 1973. (In Russian.)
25. Ray, W.H., *Advanced Process Control*, McGraw-Hill, 1981.
26. Petrov, B.N., Cherkasov, B.A., Il'yasov, B.G., and Kulikov, G.G., A Frequency Method for Analyzing and Designing Multivariable Automatic Control Systems, *Doklady Akad. Nauk SSSR*, 1979, vol. 247, no. 2, pp. 304–307. (In Russian.)
27. *Proektirovanie sistem avtomaticheskogo upravleniya gazoturbinnykh dvigatelei. Normal'nye i neshtatnye rezhimy* (Designing Automatic Control Systems for Gas-Turbine Engines. Normal and Emergency Modes), Petrov, B.N., Ed., Moscow: Mashinostroenie, 1981. (In Russian.)
28. Shevyakov, A.A., Mart'yanova, T.S., Rutkovskii, V.Yu., Il'yasov, B.G., et al., *Optimizatsiya mnogomernykh sistem upravleniya gazoturbinnykh dvigatelei letatel'nykh apparatov* (Optimization of Multivariable Control Systems for Aircraft Gas-Turbine Engines), Moscow: Mashinostroenie, 1989. (In Russian.)
29. Babak, S.F., Vasil'ev, V.I., Il'yasov, B.G., et al., *Osnovy teorii mnogosvyaznykh sistem avtomaticheskogo upravleniya letatel'nymi apparatami* (Foundations of the Theory of Multivariable Automatic Control Systems for Aircrafts), Moscow. Moscow Aviation Institute, 1995. (In Russian.)
30. Il'yasov, B.G., and Kabal'nov, Yu.S., An Investigation into the Stability of Single-type Multiply Connected Automatic Control Systems with Holonomic Ties between Subsystems, *Autom. Remote Control*, 1995, vol. 56, no. 8, pp. 1120–1125.
31. Il'yasov, B.G. and Saitova, G.A., Stability Analysis of Dynamic Systems in the Polynomial Vector-Matrix Representation, *J. Comput. Syst. Sci. Int.*, 2018, vol. 57, pp. 171–178.

32. Il'yasov, B.G. and Saitova, G.A., A Systems Approach to Studying Multiconnected Automated Control Systems Based on Frequency Methods, *Autom. Remote Control*, 2013, vol. 74, no. 3, pp. 456–470.
33. Il'yasov, B.G., Saitova, G.A., and Elizarova, A.V., Investigation of Multi-connected System of Automatic Control with Delay Decomposition Method, *Modern High Technologies*, 2019, no. 3, part 2, pp. 177–181. (In Russian.)
34. Il'yasov, B.G., and Saitova, G.A., A Study of Periodic Motions in Homogeneous Nonlinear Multivariable Systems Written in the Polynomial Vector-Matrix Representation, *J. Comput. Syst. Sci. Int.*, 2020, vol. 59, pp. 1–7.
35. Il'yasov, B.G., Munasypov, R.A., Saitova, G.A., et al., Analyzing Periodic Motions in Multivariable Systems with Fuzzy Controllers in Separate Subsystems, *Mekhatronika, Avtomatiz., Upravlen.*, 2004, no. 8, pp. 24–29. (In Russian.)
36. Il'yasov, B.G., Saitova, G.A., and Nazarov, A.Sh., An Approach to Designing Adaptive Multivariable Automatic Control Systems for an Complex Dynamic Plant, *Mekhatronika, Avtomatiz., Upravlen.*, 2010, no. 8, pp. 13–20. (In Russian.)
37. Il'yasov, B.G., Saitova, G.A., and Nazarov, A.Sh., An Structural Reconfiguration Algorithm for a Multivariable Automatic Control System Using a Stability Condition Based on Frequency Methods, *Vestn. Ufimsk. Gos. Aviats. Tekh. Univ.*, 2012, vol. 16, no. 3(48), pp. 3–10. (In Russian.)

This paper was recommended for publication by N.V. Kuznetsov, a member of the Editorial Board.

Received November 27, 2020, and revised February 15, 2021.
Accepted February 24, 2021.

Author information

Il'yasov, Baryi Galeevich. Dr. Sci. (Eng.), Ufa State Aviation Technical University, Ufa, Russia,
✉ ilyasov@ugatu.su
Saitova, Guzel' Askhatovna. Cand. Sci. (Eng.), Ufa State Aviation Technical University, Ufa, Russia,
✉ saitova@bk.ru

Cite this article

Il'yasov, B.G., Saitova, G.A. Investigation of Multivariable Automatic Control Systems for Complex Dynamic Objects Based on Petrov's Paradigm. *Control Sciences* **3,** 2–13 (2021). http://doi.org/10.25728/cs.2021.3.1

DOI: http://doi.org/10.25728/cs.2021.3.2

INFORMATION COMMUNITIES IN SOCIAL NETWORKS. PART III: APPLIED ASPECTS OF DETECTION AND ANALYSIS[1]

L.M. Boiko[2], D.A. Gubanov[3], and I.V. Petrov[4]

Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia

[2]✉ boiko.lilia@gmail.com, [3]✉ dmitry.a.g@gmail.com, [4]✉ zyxzy@protonmail.ch

Abstract. This paper overviews the empirical studies of the formation and detection of information communities in social networks. In parts I and II of the survey, we outlined the concept of an information community and considered the relevant mathematical models describing the formation of beliefs. Model identification, data gathering, and data analysis become highlighted areas of current research due to the uncertainty about social learning mechanisms and networked interaction structure. To solve the identification problem, researchers carry out behavioral experiments and field investigations. In practice, researchers analyze communities on available real-world data, applying methods based on the structural properties of the network of information interactions between agents, the individual characteristics of agents, and a combination of structural and individual characteristics. Part III of the survey presents studies on identifying belief formation models and discusses some practical aspects of analyzing information communities in social networks.

Keywords: social networks, information community, formation of information communities, belief formation, detection of information communities.

INTRODUCTION

In parts I and II of the survey (see [1, 2]), the problems of identifying (detecting) and studying information communities in social networks were introduced. In addition, mathematical models of belief dynamics and the formation of information communities in social networks were presented, and the factors and conditions for the formation of information communities were considered. In practice, the identification of such models is nontrivial: many parameters are exogenous, and a significant aspect of the learning process remains unobservable in applied research. In many situations, people demonstrate neither their true beliefs nor information available for decision-making (the mechanisms for processing this information).

During social interaction, people receive incomplete information from their opponents, e.g., information about the results of actions (activity) of other people but not why and how these decisions were made. Many factors can cause this limitation, e.g., the nature of social interaction means or the high costs of receiving and (or) transmitting complete information. Researchers conduct field investigations and behavioral experiments to identify the real-life mechanisms of information processing by people despite the arising difficulties. Numerous methods for analyzing information communities were proposed using examples of publicly available data.

Part III of the survey is organized as follows. Section 1 discusses publications on the identification of belief formation models in networks. Section 2 considers applied research of information communities in social networks.

1. IDENTIFICATION OF BELIEF FORMATION MODELS IN NETWORKS

Depending on information processing methods, two types of agents can be distinguished (see parts I and II of the survey [1, 2]):

[1]This research was supported by the Russian Foundation for Basic Research, projects nos. 19-17-50225 and 20-07-00190 (D.A. Gubanov), and 20-31-90133 (I.V. Petrov).

- rational agents (e.g., within the concept of Bayesian rationality), which can further be divided into *myopic* agents (choosing the best response in the short term) and strategic agents (choosing an optimal response based on some game-theoretic concept, e.g., Nash equilibrium);
- *naive* agents, often described by DeGroot's rule (following it, the agents form their beliefs by averaging the observed opinions of other agents).

Early research works were devoted to identifying the types of agents in the laboratory and field investigations [3–5]. The existence of agents with different learning mechanisms within the same group was not assumed, which is a drawback of these studies. In the recent paper [6], an attempt was made to identify the types of agents on several sets of real data: it was shown that social groups consist of a mixture of rational (Bayesian) and naive (acting according to DeGroot's rule) agents, and the relationship between the types varies for different data sets. For example, in a series of behavioral experiments involving residents of 19 Indian villages, the identification procedure yielded the following results: 10% of the population behave in accordance with Bayesian rationality, and the rest of the agents prefer to average the responses of their neighbors in the social relations network. In the same experiment involving the students of the National Technological Institute of Mexico, the share of Bayesian agents reached 50%. The number of experiments for each group was 95 and 50, respectively, and the number of participants was 665 and 350, respectively.

The authors [6] identified four learning patterns to distinguish the agents with Bayesian and DeGroot's rule-based learning in the model with incomplete information. Moreover, they identified a key network characteristic separating learning types, called a *clan* (a strongly connected component of the graph):

1) If the clan consists entirely of agents following DeGroot's rule, who reach consensus on the state of the world at some time instant, then they will not change their ideas at the subsequent time instants (even if they are false).

2) In the model with complete information, Bayesian agent i, whose neighbors belong to the set of neighbors of Bayesian agent j, copies the estimate of the state of the world of agent j.

3) Regardless of the type of agent i, Bayesian agent j never considers his estimate of the state of the world (complement of pattern 2).

4) Even in the case of incomplete information, a Bayesian agent identifying the simple majority of the estimates of the state of the world of his neighbors will never change his estimate under any changes in the estimates of his particular neighbors.

In [7], a smaller-scale experimental study with similar results was carried out: the authors discovered that the agents' decisions agree with DeGroot's rule in 80–98% of cases, and the forecast errors depend on the agent's position in the network. However, the central observation was as follows: the real learning process of agents matches the naive DeGroot rule only in comparative statics, and the dynamics of reaching consensus in the laboratory experiments have more complex rules for information processing. Moreover, the authors identified these heuristics, tested a wide class of other learning rules, and modified the classical DeGroot model by allowing the agents to adjust the weight of their previous states.

In addition to the complexity of identifying information processing mechanisms by agents, the structure of social interaction is often difficult (or even impossible) to detect for an external observer. At the same time, this structure may have a crucial effect on agents' learning [8–10]. From this point of view, modern technologies (e.g., implemented via online social platforms) have radically changed the way people interact and consume information. Nevertheless, some phenomena preventing the identification of social relations and sources of information arise here as well. A key aspect in this area is the policy of processing personal data by online platforms [11], when a user has to choose between privacy and the disclosure of various personal information (biography, geolocation data, or the so-called digital traces—the history of activity on the Internet) to other users, owners of the platform, or third-party applications. Therefore, a user has to decide on the availability of information about his social relations with other network members. In addition, despite the increased efficiency of information transmission, users still have cognitive and temporal limitations. As a result, recommender systems have been developed, and there is an increasing interest in algorithmic personalization. The effect of algorithmic filtering on social learning is still underinvestigated, but several models (for example, see [12]) showed that the order of information messages received can significantly influence the effectiveness of learning and consensus reaching. All these factors play a decisive role in identifying the structures of information interaction and complicate the observability of social relations.

Thus, uncertainty arises both about the mechanisms of information processing by individuals and the structure of interactions within which agents exchange their information. These features motivate further research in the area.

2. STUDIES OF INFORMATION COMMUNITIES IN SOCIAL NETWORKS

2.1. Detection of Information Communities

There is no consensus in the literature regarding the formal definition of information communities. In applied research, the authors choose fairly general definitions that reasonably reflect the essence of the phenomena occurring in information interaction networks. Several such phenomena indicate the presence of information communities, together usually characterized as *controversy*:

– *echo chamber*, a socio-psychological phenomenon when opinions or beliefs are supported in communities of like-minded people approving and strengthening each other's opinions;

– *filter bubble*, a phenomenon when the algorithms of personalized recommender systems offer content consistent with the information earlier received by a user, thereby excluding his opportunity to get acquainted with alternative or new information.

The overwhelming majority of research into information communities is associated with significant restrictions. Such publications consider public opinions about political issues, focusing on large-scale and long-term events (e.g., elections). In many countries, citizens actively discuss socially significant issues on online social networks (Twitter, Facebook, etc.). As a result, huge thematic data sets containing information about users and their actions become available for analysis. Therefore, many works can be characterized as case studies, in which information communities are investigated on a specific data set related to a particular social phenomenon.

In these studies of information communities, as a rule, the processes of information propagation and their properties are considered; the formation mechanisms of network participants' beliefs are not identified. (For these problems, see the corresponding models in parts I and II of the survey [1, 2].) In addition to the complex identification procedure for learning rules, the reason is that in most theoretical models, the network structure is determined exogenously and does not depend on learning results: learning does not change the mutual influence of participants of the information process. However, empirical studies of the phenomena characterizing information communities reveal evidence of a relationship between learning and the structure of interactions. Identifying and formalizing these phenomena in theoretical models could significantly reduce the gap between theory and practice. Modern applied research is limited to the development of identification methods for the state of individuals (assessment of private beliefs based on observed information) and the analysis of comparative statics.

Attempts to identify the states of information interaction participants often rest on the following observation. Generally, the formation of an information community can be represented as a diffusion process on a network (known as the diffusion of innovations, ideas, or information) in which joining a new community is analogous to the acceptance of ideas or beliefs. The converse is also true: any propagation process on a network can be viewed as the formation of a network community in which the elements are grouped by their states. One example of such processes is information propagation called information cascade; see Fig. 1.

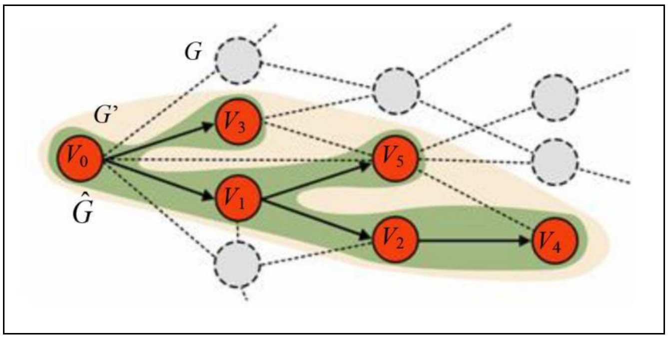


Fig. 1. Community formation process interpreted as a diffusion process on the network [13]. The thick directed edges in the graph G show the information propagation process starting in vertex V_0 and covering the vertices of the subgraph $\hat{G}$. The undirected edges show the relations of social interactions between network nodes (e.g., friendship ties). All together, these relations induce a friends subgraph G'.

According to a natural assumption, such information processes (cascades) should correlate with the beliefs of the individuals involved and affect their beliefs. This analogy with diffusion processes often becomes a starting point when studying the formation of information communities: the authors apply methods based on the structural properties of the information interaction network, the properties of the network elements, or a combination of structural and individual characteristics.

Thus, the problem is identifying two main characteristics of information interaction: the communication structure of the participants and their individual characteristics. The table below presents the most cited papers on identifying information communities: a brief description of the data used, the proposed measures, and the methods adopted by the authors. As mentioned above, both identification problems are complex, and the choice of an appropriate method for identifying information communities largely depends on the set of real data available to the researchers.

Most cited papers on identifying information communities

Paper title	Concept of community	Graph type	Measure (characteristic) type	Data type	Data source
Testing Models of Social Learning on Networks: Evidence from Two Experiments [6]	A set of nodes more connected to each other than to those outside the group	An artificially created network of relations between the experiment participants	Structural (clan)	Offline	Laboratory experiments
Ideological Segregation Online and Offline [14]	A community with equivalent characteristics of members	Relations between interaction participants are not considered	Individual characteristics (isolation index)	Online, offline	Internet news, offline media, personal interaction
Quantifying Controversy in Social Media – [15]	Opinions or beliefs are supported in the communities created by like-minded people, who strengthen and approve of each other's opinions	A dialog graph: a graph corresponding to thematic discussions, where relations between participants are formed in the case of users' responses to each other's messages	Structural (Random Walk Controversy, Betweenness Centrality Controversy, Embedding Controversy)	Online	Twitter
Political Discourse on Social Media: Echo Chambers, Gatekeepers, and the Price of Bipartisanship [16]	Preferences for the content received by network users match the preferences for the content they distribute	A subscriber graph: a directed edge (link) between participants arises if one participant monitors information updates from another participant	Individual characteristics (production polarity, consumption polarity)	Online	Twitter
Community Interaction and Conflict on the Web [17]	Community members interact primarily with other members of their community	A bipartite multigraph between users and communities. Relations arise in the case of communication between users within a given community	Mixed	Online, time series	Reddit
Quantifying Echo Chamber Effects in Information Spreading over Political Communication Networks [18]	Beliefs are strengthened through repeated interactions with people sharing the same viewpoints	A subscriber graph: a directed edge (link) between participants arises if one participant monitors information updates from another participant	Mixed	Online, time series	Twitter
An Empirical Examination of Echo Chambers in US Climate Policy Networks [19]	A community is characterized by two attributes: information coinciding with established beliefs and a clustered structure of interaction	A network of interaction between experts	Mixed	Online	Survey
Echo Chambers: Emotional Contagion and Group Polarization on Facebook [20]	Groups of like-minded people with extreme-value beliefs	Relations between interaction participants are not considered	Individual characteristics (user sentiment polarization)	Online	Facebook
Exposure to Ideologically Diverse News and Opinion on Facebook [21]	Two types of communities: – a set of participants exposed only to information from like-minded people; – the information offered by the algorithms matches the history of user actions	A graph of friendship ties between social network participants	Mixed (Alignment score)	Online	Facebook
Filter Bubbles, Echo Chambers, and Online News Consumption [22]	Two types of communities: – a set of participants exposed only to information from like-minded people; – the information offered by the algorithms matches the history of user actions	Relations between interaction participants are not considered	Individual characteristics (audience-based measure of outlet slant)	Online	Net-surfing history

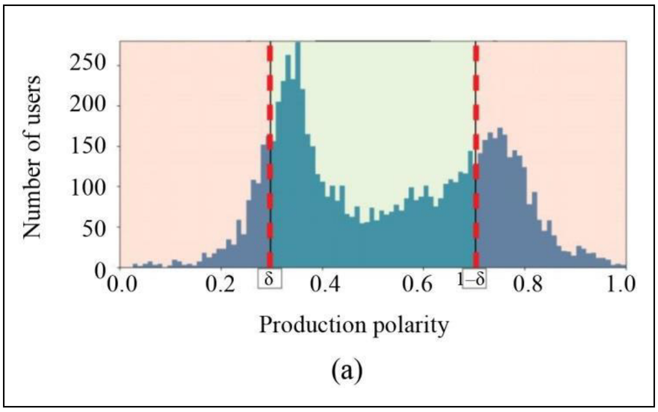
Subsection 2.2 provides some of the most common methods and measures for identifying information communities.

2.2. Identification of Information Communities: Methods Based on the Properties of Network Structure Elements

In applied research, content is crucial for measuring the effects that characterize the presence of information communities. The essential characteristic of an information community is the degree of correspondence between the content consumed and produced by the network participants. In this regard, the authors [16] divided the general process of information interaction into information consumption and information production processes. According to their approach, each message t of a social network belongs to one of two subclasses: $l(t)=l_n\in\{0,1\}$. (The study involved data from Twitter with the following notations: P_u is information published by the user on his page; C_u is publications that the user receives from other users. The information was classified by the users' political views, where l_n = 1 for conservative views and l_n = 0 for the liberal ones.) Based on the set of all information produced (P_u) and consumed (C_u) by the user u, the degree of diversity of the content produced and consumed by the users (production and consumption polarity, respectively) was determined as the amount of information from one class divided by the total amount of information produced and consumed by the user:

$$p(u)=\frac{\sum_{t\in P_u} l(t)}{|P_u|}\ ,\ c(u)=\frac{\sum_{t\in C_u} l(t)}{|C_u|}.$$

There are variations of these measures (where the variability is due to the specifics of real data), yielding practically interpretable conclusions about information propagation among the users. In particular, by calculating statistical characteristics (variance, correlation, or measures of difference between probability distributions) for the diversity of consumed and produced content, the researchers demonstrated the presence of information communities; see Fig. 2.



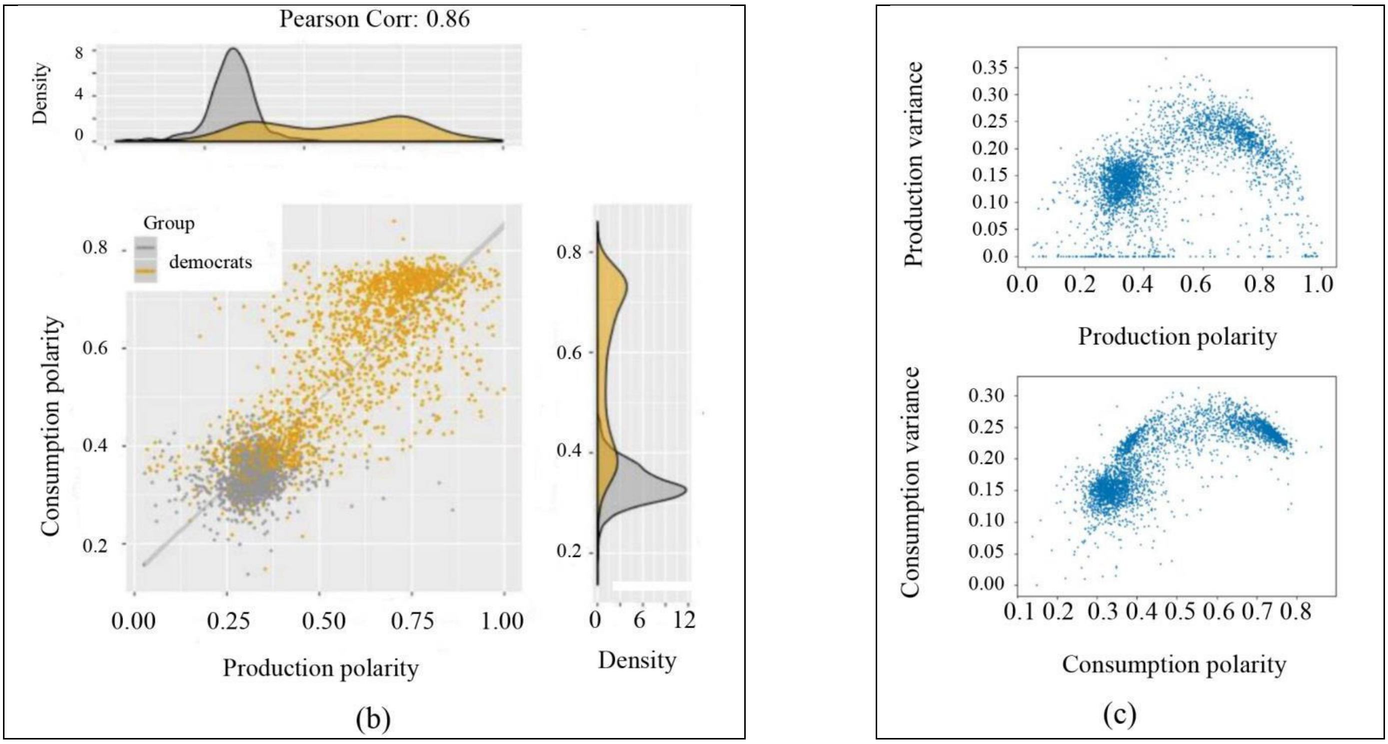


Fig. 2. Estimated statistical characteristics of information interaction between Twitter users on the legislative regulation of arms traffic [16].

Figure 2*a* presents the distribution of users by production polarity: the double-peaked property of this distribution indicates, among other signs, the presence of information communities (echo chambers). The graph of the relationship between consumption and production polarity (Fig. 2*b*) demonstrates the high-degree clustering of the values for representatives of different user groups. The graphs on the right are intended to assess the relationship between production (consumption) polarity and its variance.

2.3. Analysis of Information Communities: Methods Based on the Structural Properties of Information Interaction

Analyzing the structural properties of information communities, researchers focus on comparing interaction processes between the nodes of different communities. The analysis tool is often the properties of random processes on graphs or centrality measures reflecting the effectiveness of nodes during information propagation.

Random Walk Controversy (RWC [19]) is a random walk-based measure defined as follows. Let the graph be partitioned by some criterion into two subgraphs, *X* and *Y*, with nonintersecting sets of vertices. Consider two random walks, one ending in the subgraph *X* and the other in the subgraph *Y*. RWC is the difference between the probabilities of two events: (1) both random walks started in the same subgraph where they ended and (2) both random walks started in a subgraph differing from the one where they ended). That is,

$$\mathrm{RWC} = P_{XX} P_{YY} - P_{YX} P_{XY},$$

where

$P_{AB} = P(\text{the process starts in } A \mid \text{the process ends in } B)$

denotes the corresponding conditional probability with $A, B \in \{X, Y\}$.

Another random walk-based method for identifying information communities is a personalized version of the PageRank algorithm [23], in which the damping factor changes depending on the group of the graph vertex in which the random walk process starts [17]. In the classical version of the algorithm, transitions occur either to neighbors or any other node selected equiprobably. (A practical interpretation is the end of the link click process and the beginning of a new one). In the personalized version, the probability distribution on the set of vertices is different for vertices from different communities. Thus, the method allows assessing the controversy of communities by comparing the probabilities of interaction between members of different communities [17].

A measure based on betweenness centrality [15]. The betweenness centrality $bc(e)$ of a network edge *e* is defined as

$$bc(e) = \sum_{s \neq t \in V} \frac{\sigma_{s,t}(e)}{\sigma_{s,t}},$$

where $\sigma_{s,t}$ is the total number of shortest paths between vertices *s* and *t* in the graph, and $\sigma_{s,t}(e)$ is the number of shortest paths passing through the edge *e*. The authors [15] proposed to analyze the differences in the centralities of the vertices from two sets forming a graph partition. (In the original work, they used the METIS algorithm [24].) The idea is to compare the centralities of the edges included in the graph cut-set (i.e., the edges connecting the vertices from different subsets of the graph vertices) and the centralities of the edges in the rest of the graph. If a "good" graph cut-set is obtained, most of the shortest paths from one graph part to another will pass through the cut-set edges, and the centrality of these edges will have higher values compared to the centrality of the edges in the rest of the graph. Comparing two distributions of centralities––inside the cut-set and outside it—for example, using the KL-divergence d_{KL} and performing normalization, we obtain the following expression for *Betweenness Centrality Controversy* (BCC):

$$\mathrm{BCC} = 1 - e^{-d_{KL}}.$$

In addition to these methods, classical clustering techniques without considering diffusion processes on networks or calculating paths between vertices are used. Researchers associate the resulting structural characteristics with the individual characteristics of separate graph nodes, thereby combining the methods demonstrated above.

2.4. Analysis of Information Communities: Methods Based on the Combination of Structural and Individual Characteristics

Combining the individual characteristics of the participants and the structural characteristics of information interaction is a nontrivial problem underinvestigated in the literature. One solution is to employ machine learning methods: all characteristics of the information interaction process available to researchers (the set of all produced or consumed information, information content, structural characteristics of the network and individual participants, etc.) are considered and placed in a single feature space. Here, classification methods are used to detect information communities.

When combining individual and structural characteristics, a promising line is to adopt various transformations of the initial data: *embedding* and, particularly, node/edge/graph embedding [25].

This operation generally transforms the original feature space into another space, often of a lower dimension. From this viewpoint, all the methods mentioned above can be understood as special cases of such transformations. The clustering problem can be solved by classical methods in a new space [26, 27]. For example, the *Embedding Controversy* (EC) measure

$$\mathrm{EC} = 1 - \frac{d_X + d_Y}{2d_{XY}},$$

where d_X (d_Y) is the average distance between the pairs of elements from the set X (Y, respectively), and d_{XY} is the average distance between the pairs of elements from different sets, yields another method for identifying information communities [15]. An EC value close to 1 indicates the presence of information communities and a high degree of graph clustering; an EC value close to 0 indicates the opposite.

The graph embedding method is more complicated; however, it allows analyzing not individual nodes but entire graphs. The method involves graph kernels—transformations for the pairwise comparison of structures with each other—and can be used both for comparative analysis of individual groups of graph vertices [28] and information processes occurring on networks [29–31]. This approach allows studying the sequences of information flows and comparing and predicting the characteristics of information cascades (such as size, speed, etc.) in information communities.

CONCLUSIONS

This paper has overviewed studies of information communities in complementary areas: the formation models of information communities in social networks (with microeconomic, cognitive, and socio-psychological foundations), identification methods for information communities, and applied research into information communities in social networks.

Parts I and II of the survey have outlined the concept of an information community and considered belief formation models for individuals seeking to eliminate uncertainty about (a) given question(s), eventually forming information communities. Approaches to model the belief updating process of individuals and the effect of various factors on reaching true beliefs and forming different (or identical) stable beliefs in the network, leading to the emergence of information communities, have been described.

This part of the survey has presented empirical studies of the existence of information communities in real social networks. Uncertain mechanisms of information processing by individuals, an uncertain structure of interaction, and abundant real data sets (mainly from online social networks) cause a wide variety of empirical methods for identifying information communities and research focusing on real data sources. Due to the specifics of the available data, the considered methods characterize the information produced and consumed by social network users rather than their beliefs. The absence of any prerequisites for belief formation mechanisms is a significant drawback of these methods: only indirect conclusions can be drawn both about the true beliefs of the participants in information interactions and the formation dynamics of information communities. The transition from the analysis of individual node interactions to the analysis of higher-order structure interactions characterizing the evolution of the information process seems promising for identifying communities in information interaction structures. Investigations in this area can significantly expand the understanding of the relationship between information processes and the formation of information communities.

REFERENCES

1. Gubanov, D.A. and Petrov, I.V., Information Communities in Social Networks. Part I: From Concept to Mathematical Models, *Control Sciences*, 2021, no. 1, pp. 13–20.
2. Gubanov, D.A. and Petrov, I.V., Information Communities in Social Networks. Part II: Networked Models of Formation, *Control Sciences*, 2021, no. 2, pp. 16–28.
3. Hung, A.A. and Plott, C.R., Information Cascades: Replication and an Extension to Majority Rule and Conformity-Rewarding Institutions, *American Economic Review*, 2001, vol. 91, no. 5, pp. 1508–1520.
4. Kübler, D. and Weizsäcker, G., Limited Depth of Reasoning and Failure of Cascade Formation in the Laboratory, *The Review of Economic Studies*, 2004, vol. 71, no. 2, pp. 425–441.
5. Choi, S., Gale, D., and Kariv, S., *Behavioral Aspects of Learning in Social Networks: An Experimental Study*, Emerald Group Publishing Limited, 2005.
6. Chandrasekhar, A.G., Larreguy, H., and Xandri, J.P., Testing Models of Social Learning on Networks: Evidence From Two Experiments, *Econometrica*, 2020, no. 1 (88), pp. 1–32.
7. Grimm, V. and Mengel, F., Experiments on Belief Formation in Networks, *Journal of the European Economic Association*, 2020, vol. 18, no. 1, pp. 49–82.
8. Acemoglu, D. and Ozdaglar, A., Opinion Dynamics and Learning in Social Networks, *Dynamic Games and Applications*, 2011, no. 1 (1), pp. 3–49.
9. Golub, B. and Sadler, E., Learning in Social Networks, in *The Oxford Handbook of the Economics of Networks*, Bramoullé, Y., Galeotti, A., and Rogers, B.W., Eds., 2016. DOI: 10.1093/oxfordhb/9780199948277.013.12.
10. Young, J.G., Cantwell, G.T., and Newman, M.E. J., Robust Bayesian Inference of Network Structure from Unreliable Data, *arXiv preprint arXiv:2008.03334*, 2020.

11. Lovato, J., Allard, A., Harp, R., and Hébert-Dufresne, L., Distributed Consent and Its Impact on Privacy and Observability in Social Networks, *arXiv preprint arXiv:2006.16140*, 2020.
12. Perra, N. and Rocha, L.E.C., Modelling Opinion Dynamics in the Age of Algorithmic Personalisation, *Scientific Reports*, 2019, vol. 9, no. 1, pp. 1–11.
13. Cheng, J., Adamic, L., Dow, P.A., et al., Can Cascades Be Predicted? *Proceedings of the 23rd International Conference on World Wide Web*, 2014, pp. 925–936.
14. Gentzkow, M. and Shapiro, J.M., Ideological Segregation Online and Offline, *The Quarterly Journal of Economics*, 2011, vol. 126, no. 4, pp. 1799–1839.
15. Garimella, K., Morales, G.D.F., Gionis, A., and Mathioudakis, M., Quantifying Controversy on Social Media, *ACM Transactions on Social Computing*, 2018, vol. 1, no. 1, pp. 1–27.
16. Garimella, K., Morales, G.D.F., Gionis, A., and Mathioudakis, M., Political Discourse on Social Media: Echo Chambers, Gatekeepers, and the Price of Bipartisanship, *ArXiv180101665 Cs*, 2018.
17. Kumar, S., Hamilton, W.L., Leskovec, J., and Jurafsky, D., Community Interaction and Conflict on the Web, *Proceedings of the 2018 World Wide Web Conference*, 2018, pp. 933–943.
18. Cota, W., Ferreira, S.C., Pastor-Satorras, R., and Starnini, M., Quantifying Echo Chamber Effects in Information Spreading over Political Communication Networks, *EPJ Data Science*, 2019, vol. 8, no. 1, p. 35.
19. Jasny, L., Waggle, J., and Fisher, D.R., An Empirical Examination of Echo Chambers in US Climate Policy Networks, *Nature Climate Change*, 2015, vol. 5, no. 8, pp. 782–786.
20. Del Vicario, M., Vivaldo, G., Bessi, A., et al., Echo Chambers: Emotional Contagion and Group Polarization on Facebook, *Scientific Reports*, 2016, vol. 6, p. 37825.
21. Bakshy, E., Messing, S., and Adamic, L.A., Exposure to Ideologically Diverse News and Opinion on Facebook, *Science*, 2015, vol. 348, no. 6239, pp. 1130–1132.
22. Flaxman, S., Goel, S., and Rao, J.M., Filter Bubbles, Echo Chambers, and Online News Consumption, *Public Opinion Quarterly*, 2016, vol. 80, no. S1, pp. 298–320.
23. Page, L., Brin, S., Motwani, R., and Winograd, T., The PageRank Citation Ranking: Bringing Order to the Web, Stanford InfoLab, 1999.
24. Karypis, G., METIS: Unstructured Graph Partitioning and Sparse Matrix Ordering System, Technical Report, 1997.
25. Rossi, R.A., Jin, D., Kim, S., et al., On Proximity and Structural Role-Based Embeddings in Networks: Misconceptions, Techniques, and Applications, *ACM Transactions on Knowledge Discovery from Data* (*TKDD*), 2020, vol. 14, no. 5, pp. 1–37.
26. Chattopadhyay, S. and Ganguly, D., Community Structure Aware Embedding of Nodes in a Network, *arXiv preprint arXiv:2006.15313*, 2020.
27. Billings, J.C.W., Hu, M., Lerda, G., et al., Simplex2Vec Embeddings for Community Detection in Simplicial Complexes, *arXiv preprint arXiv:1906.09068*, 2019.
28. Rossi, R.A., Jin, D., Kim, S., et al., From Community to Role-Based Graph Embeddings, *arXiv preprint arXiv:1908.08572*, 2019.
29. Zhou, X. and Zafarani, R., A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities, *arXiv preprint arXiv:1812.00315*, 2018.
30. Ma, J., Gao, W., and Wong, K.F., Detect Rumors in Microblog Posts Using Propagation Structure via Kernel Learning, Association for Computational Linguistics, 2017.
31. Wu, K., Yang, S., and Zhu, K.Q., False Rumors Detection on Sina Weibo by Propagation Structures, *2015 IEEE 31st International Conference on Data Engineering*, 2015, pp. 651–662.

This paper was recommended for publication by D.A. Novikov, a member of the Editorial Board.

Received August 26, 2020, and revised November 9, 2020.
Accepted November 24, 2020.

Author information

Boiko, Liliya Mikhailovna. Junior researcher, Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia, ✉ boiko.lilia@gmail.com

Gubanov, Dmitry Alekseevich. Cand. Sci. (Eng.), Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia, ✉ dmitry.a.g@gmail.com

Petrov, Il'ya Vladimirovich. Postgraduate, Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia, ✉ zyxzy@protonmail.ch

Cite this article

Boiko, L.M., Gubanov, D.A., Petrov, I.V. Information Communities in Social Networks. Part III: Applied Aspects of Detection and Analysis. *Control Sciences* **3,** 14–21 (2021).
http://doi.org/10.25728/cs.2021.3.2

DOI: http://doi.org/10.25728/cs.2021.3.3

CONSTRUCTING POWER-EXPONENTIAL AND LINEAR-LOGARITHMIC REGRESSION MODELS

M.P. Bazilevskiy

Irkutsk State Transport University, Irkutsk, Russia

✉ mik2178@yandex.ru

Abstract. When using nonlinear regression models, the estimates of the resulting dependence are often difficult or even impossible to interpret. This paper develops nonlinear regression specifications in which any estimated parameter, except the free term, can always be given some practical interpretation. A multiplicative power-exponential regression generalizing the Cobb–Douglas production function and an additive linear-logarithmic regression are constructed. Three construction strategies are formulated for each of them, and the issues of interpreting their estimates are considered in detail. The construction strategies based on the least absolute deviations method are formalized as linear and partially Boolean linear programming problems. The mathematical apparatus developed in this paper is illustrated by modeling rail freight traffic in Irkutsk oblast.

Keywords: regression model, interpretation, multiplicative power-exponential regression, linear-logarithmic regression, feature selection, least absolute deviations, rail freight traffic.

INTRODUCTION

Regression analysis is a worldwide recognized tool for mathematical modeling based on statistics [1, 2]. One of the first (and, perhaps, most important) stages in constructing a regression model is specification, i.e., choosing an appropriate composition of the variables and a mathematical relationship among them. A significant number of such specifications have been developed to date, and most of them can be found in [3–6]. The simplest specification is the multiple linear regression model:

$$y_i = \alpha_0 + \sum_{j=1}^{l} \alpha_j x_{ij} + \varepsilon_i, \quad i = \overline{1,n}, \tag{1}$$

where y_i, $i = \overline{1,n}$, are the observed values of the independent (output) variable y; x_{ij}, $i = \overline{1,n}$, $j = \overline{1,l}$, are the observed values of the explanatory (input) variables x_1, x_2, …, x_l; ε_i, $i = \overline{1,n}$, are approximation errors; α_0, α_1, α_2, …, α_l are unknown parameters.

The linear regression (1) is easily estimated, e.g., using the least squares method (LSM). Let the estimated equation have the form

$$\tilde{y} = \tilde{\alpha}_0 + \sum_{j=1}^{l} \tilde{\alpha}_j x_j, \tag{2}$$

where $\tilde{y}$ is the model value of the independent variable; $\tilde{\alpha}_0, \tilde{\alpha}_1, \tilde{\alpha}_2, ..., \tilde{\alpha}_l$ are the estimates of the unknown parameters.

The coefficient $\tilde{\alpha}_s$ at the explanatory variable x_s in equation (2) is interpreted in the following way: if the value of the explanatory variable x_s varies by 1, then the value of the independent variable y varies by $\tilde{\alpha}_s$ on average.

Note that the development of new specifications for regression models continues to the present time. For example, a linear multiplicative regression (LMR) and a regression contrary to the Leontief production function were proposed in [7] and [8], respectively. Later on, they were combined in [9]. Another specification is an index regression introduced in [10].

For solving the specification problem, a technology to organize a "competition" of regression models was developed; for details, see the monograph [6]. The competition is intended to form a set of alternative regressions and select the best one among them.

The following algorithm for forming alternatives was considered in [6]. First, the set of original explanatory variables is enlarged using some transformations, e.g., the elementary functions $\ln x$, e^x, x^{-1}, x^2, x^3, $\sqrt{x}$, etc. Then, by a complete enumeration of all combinations, m features are selected [11]. Unfortunately, the resulting regression equation often turns out to be significantly nonlinear, making it difficult (or even impossible) to interpret the estimates found.

This paper develops nonlinear regression specifications in which any estimated parameter, except the free term, can always be given some practical interpretation during the competition of regression models.

1. MULTIPLICATIVE POWER-EXPONENTIAL REGRESSION

The exponential regression with one explanatory variable [12, 13] has the form

$$y_i = \alpha_0 \cdot e^{\alpha_1 x_i} \varepsilon_i, \ i = \overline{1, n}. \tag{3}$$

The model (3) is nonlinear in the estimated parameters but can be linearized by taking the logarithm:

$$\ln y_i = c_0 + \alpha_1 x_i + u_i, \ \ i = \overline{1,n}, \tag{4}$$

where $c_0 = \ln \alpha_0$ and $u_i = \ln \varepsilon_i$.

The linear model (4) is called the semi-log (left-log, or log-linear) regression [13].

The book [13] suggested the following interpretation of the estimated coefficient $\tilde{\alpha}_1$ of the models (3) and (4): if the explanatory variable x changes by 1, then the independent variable y changes by $100\tilde{\alpha}_1$% on average.

Unfortunately, as noted in [13], this interpretation of the coefficient $\tilde{\alpha}_1$ of the models (3) and (4) applies to small $\tilde{\alpha}_1$ only.

Consider a generalization of the model (3): the additive multiple exponential regression

$$y_i = \alpha_0 + \sum_{j=1}^{l} \alpha_j e^{\beta_j x_{ij}} + \varepsilon_i, \ \ i = \overline{1,n}, \tag{5}$$

where β_j, $j = \overline{1, l}$, are unknown parameters.

It seems impossible to linearize the model (5). Even if its estimates were found, it would be difficult to give them any practical interpretation. Therefore, consider the multiplicative multiple exponential regression

$$y_i = \alpha_0 \prod_{j=1}^{l} e^{\alpha_j x_{ij}} \varepsilon_i, \ i = \overline{1,n}. \tag{6}$$

The model (6) is linearized by taking the logarithm, and all its coefficients have the interpretation described above.

The regression (6) resembles by properties the Cobb–Douglas production function (the power regression)

$$y_i = \alpha_0 \prod_{j=1}^{l} x_{ij}^{\alpha_j} \ \varepsilon_i, \ i = \overline{1,n}. \tag{7}$$

The model (7) is also linearized by taking the logarithm, and the estimated coefficient $\tilde{\alpha}_s$ at the explanatory variable x_s is interpreted in the following way: if the explanatory variable x_s changes by 1%, then the independent variable y changes by $\tilde{\alpha}_s$% on average. In other words, $\tilde{\alpha}_s$ gives the elasticity of the variable y in x_s.

We construct a multiplicative combination of the models (6) and (7):

$$y_i = \alpha_0 \prod_{j=1}^{l} x_{ij}^{\alpha_j} \prod_{j=1}^{l} e^{\beta_j x_{ij}} \varepsilon_i, \ i = \overline{1,n}. \tag{8}$$

The expression (8) will be called the multiplicative power-exponential regression (MPER).

Note that the power and exponential regressions were also combined previously. For example, a modification of the Cobb–Douglas production function was considered in the paper [14]: labor and capital were included as power functions and scientific and technical information as an exponential function. In addition, we mention Tinbergen's production function [6], representing the product of the power regression (7) and a factor $e^{\gamma t}$ describing the "neutral" technical progress effect. However, the MPER generalizes all these known modifications.

The logarithmized MPER (8) has the form

$$\ln y_i = c_0 + \sum_{j=1}^{l} \alpha_j \ln x_{ij} + \sum_{j=1}^{l} \beta_j x_{ij} + u_i, \ i = \overline{1,n}. \tag{9}$$

Clearly, the MPER is easily estimated. However, a problem arises with a practical interpretation of its coefficients: each explanatory variable enters into the model (9) both linearly and logarithmically. Therefore, for interpreting any coefficient of the MPER, we should perform feature selection in modeling.

For the further presentation, we introduce the following Boolean variables:

$$\sigma_j^{\text{pow}} = \begin{cases} 1 \text{ if } x_j \text{ enters into the MPER} \\ \quad \text{via the power function,} \\ 0 \text{ otherwise,} \end{cases}$$

$$\sigma_j^{\text{exp}} = \begin{cases} 1 \text{ if } x_j \text{ enters into the MPER exponentially,} \\ 0 \text{ otherwise.} \end{cases}$$

Then linear constraints can be imposed on the coefficients of the models (8) and (9):

$$-M\sigma_j^{\text{pow}} \le \alpha_j \le M\sigma_j^{\text{pow}}, \ \ j = \overline{1, l}, \tag{10}$$

$$-M\sigma_j^{\exp} \le \beta_j \le M\sigma_j^{\exp},\ j=\overline{1,l}, \quad (11)$$

where M is a large positive number.

If $\sigma_j^{\text{pow}}=1$ and $\sigma_j^{\exp}=0$, $j=\overline{1,l}$, then the MPER (8) is transformed to the power regression (7); if $\sigma_j^{\exp}=0$ and $\sigma_j^{\exp}=1$, $j=\overline{1,l}$, to the exponential regression (6).

We formulate three strategies to construct the MPER:

- Strategy 1. There are no restrictions on how the variables enter into the model. In this case, we need to estimate the linear regression (9) with $(2l+1)$ parameters and pass to the MPER (8). The estimated equation can be used for prediction, but the coefficients cannot be interpreted.
- Strategy 2. Each explanatory variable enters into the model either via the power function or exponentially. This strategy is formally described by

$$\sigma_j^{\text{pow}}+\sigma_j^{\exp}=1,\ \ j=\overline{1,l}. \quad (12)$$

In this case, we need to estimate 2^l linear regressions (9) with $(l+1)$ parameters, select the best one, and pass to the MPER (8). In the resulting equation, any coefficient (possibly except the free term) can always be given a practical interpretation if its sign corresponds to the problem's sense. Also, the resulting equation can be used for prediction. But if the number of variables l is large, then the problem arises with selecting a given number of the most informative ones.

- Strategy 3. Each explanatory variable enters into the model either via the power function or exponentially, and the total number of linear features is *m*. This strategy is formally described by

$$\sigma_j^{\text{pow}}+\sigma_j^{\exp}\le 1,\ j=\overline{1,l}, \quad (13)$$

$$\sum_{j=1}^{l}\left(\sigma_j^{\text{pow}}+\sigma_j^{\exp}\right)=m. \quad (14)$$

In this case, we need to estimate $C_l^m\cdot 2^m$ linear regressions of the form (9) with $(m+1)$ parameters, select the best one, and pass to the MPER (8). The resulting equation can be used for prediction and interpretation as well.

2. LINEAR-LOGARITHMIC REGRESSION

The logarithmic [12] (right-log, log-linear) regression with one explanatory variable has the form

$$y_i=\alpha_0+\alpha_1\ln x_i+\varepsilon_i,\ \ i=\overline{1,n}. \quad (15)$$

According to [15], the estimated coefficient $\tilde{\alpha}_1$ of the model (15) is interpreted in the following way: if the explanatory variable x changes by 1%, then the independent variable y changes by $\tilde{\alpha}_1/100$ on average.

As we believe, the estimate $\tilde{\alpha}_1$ of the logarithmic model (15) can be also given another interpretation: if the explanatory variable x changes by e times, then the independent variable y changes by $\tilde{\alpha}_1$ on average.

Consider a generalization of the model (15): the additive multiple logarithmic regression

$$y_i=\alpha_0+\sum_{j=1}^{l}\alpha_j\ln x_{ij}+\varepsilon_i,\ i=\overline{1,n}. \quad (16)$$

The model (16) is linear in the parameters, and any estimated coefficient at the logarithm of an explanatory variable can be interpreted as mentioned above.

Note that there is no sense to use logarithms with different bases in (16). For example, consider the model with two explanatory variables

$$y_i=\alpha_0+\alpha_1\log_2 x_1+\alpha_2\log_3 x_2+\varepsilon_i,\ \ i=\overline{1,n}.$$

With the well-known logarithmic relation $\log_a x=\frac{\log_c x}{\log_c a}$, this model is written as

$$y_i=\alpha_0+\alpha_1\frac{\ln x_1}{\ln 2}+\alpha_2\frac{\ln x_2}{\ln 3}+\varepsilon_i,\ i=\overline{1,n}.$$

Redenoting $\alpha_1=\frac{\alpha_1}{\ln 2}$ and $\alpha_2=\frac{\alpha_2}{\ln 3}$, we obtain a particular case of the regression (16).

Now we construct an additive combination of the models (1) and (16):

$$y_i=\gamma_0+\sum_{j=1}^{l}\gamma_j x_{ij}+\sum_{j=1}^{l}\delta_j\ln x_{ij}+\varepsilon_i,\ i=\overline{1,n}, \quad (17)$$

The expression (17) will be called the linear-logarithmic regression (LLR).

Trying to interpret the LLR, we face the same problem as for the MPER: each explanatory variable enters into equation (17) both linearly and logarithmically.

Let us introduce the following Boolean variables:

$$\sigma_j^{\text{lin}}=\begin{cases}1 \text{ if } x_j \text{ enters into the LLR linearly,}\\ 0 \text{ otherwise,}\end{cases}$$

$$\sigma_j^{\log}=\begin{cases}1 \text{ if } x_j \text{ enters into the LLR logarithmically,}\\ 0 \text{ otherwise.}\end{cases}$$

Then linear constraints can be imposed on the coefficients of the model (17):

$$-M\sigma_j^{\text{lin}}\le\gamma_j\le M\sigma_j^{\text{lin}},\ \ j=\overline{1,l}, \quad (18)$$

$$-M\sigma_j^{\log}\le\delta_j\le M\sigma_j^{\log},\ \ j=\overline{1,l}. \quad (19)$$

If $\sigma_j^{\text{lin}} = 1$ and $\sigma_j^{\log} = 0$, $j = \overline{1, l}$, then the LLR (17) is transformed to the linear regression (1); if $\sigma_j^{\text{lin}} = 0$ and $\sigma_j^{\log} = 1$, $j = \overline{1, l}$, to the logarithmic regression (16).

By analogy with the MPER, we formulate three strategies to construct the LLR:

• Strategy 1. There are no restrictions on how the variables enter into the model. In this case, we need to estimate the linear regression (17) with $(2l+1)$ parameters. The estimated equation can be used for prediction, but the coefficients cannot be interpreted.

• Strategy 2. Each explanatory variable enters into the model either linearly or logarithmically. This strategy is formally described by

$$\sigma_j^{\text{lin}} + \sigma_j^{\log} = 1, \; j = \overline{1, l}. \tag{20}$$

In this case, we need to estimate 2^l linear regressions of the form (17) with $(l+1)$ parameters and select the best one. It can be used for prediction and interpretation.

• Strategy 3. Each explanatory variable enters into the model either linearly or logarithmically, and the total number of features is *m*. This strategy is formally described by

$$\sigma_j^{\text{lin}} + \sigma_j^{\log} \le 1, \;\; j = \overline{1, l}, \tag{21}$$

$$\sum_{j=1}^{l} \left(\sigma_j^{\text{lin}} + \sigma_j^{\log} \right) = m. \tag{22}$$

In this case, we need to estimate $C_l^m \cdot 2^m$ linear regressions of the form (17) with $(m+1)$ parameters and select the best one.

3. CONSTRUCTION OF MPER AND LLR USING PARTIALLY BOOLEAN LINEAR PROGRAMMING

Mathematical programming is widely used in regression analysis; for example, see [16–18].

Let the logarithmized MPER (9) be estimated using the least absolute deviations (LAD) method. As shown in the monograph [6], the LAD estimates of this regression can be obtained by solving the linear programming (LP) problem

$$\lambda_i^+ + \lambda_i^- \to \min, \tag{23}$$

$$v_i = c_0 + \sum_{j=1}^{l} \alpha_j z_{ij} + \sum_{j=1}^{l} \beta_j x_{ij} + \lambda_i^+ - \lambda_i^-, \; i = \overline{1, n}, \tag{24}$$

$$\lambda_i^+, \lambda_i^- \ge 0, \tag{25}$$

where $v_i = \ln y_i$, $z_{ij} = \ln x_{ij}$,

$$\lambda_i^+ = \begin{cases} v_i - c_0 - \sum_{j=1}^{l} \alpha_j z_{ij} - \sum_{j=1}^{l} \beta_j x_{ij} \\ \text{if } v_i - c_0 - \sum_{j=1}^{l} \alpha_j z_{ij} - \sum_{j=1}^{l} \beta_j x_{ij} > 0, \\ 0 \text{ otherwise}, \end{cases}$$

$$\lambda_i^- = \begin{cases} c_0 + \sum_{j=1}^{l} \alpha_j z_{ij} + \sum_{j=1}^{l} \beta_j x_{ij} - v_i \\ \text{if } c_0 + \sum_{j=1}^{l} \alpha_j z_{ij} + \sum_{j=1}^{l} \beta_j x_{ij} - v_i > 0, \\ 0 \text{ otherwise}. \end{cases}$$

Then one of the following problems should be solved depending on the strategy to construct the MPER:

• for strategy 1, the LP problem with the objective function (23) and the linear constraints (24) and (25);

• for strategy 2, the partially Boolean linear programming (PBLP) problem with the objective function (23) and the linear constraints (24), (25), and (10)–(12);

• for strategy 3, the PBLP problem with the objective function (23) and the linear constraints (24), (25), (10), (11), (13), and (14).

The problem of constructing the LLR is formalized by analogy. The LAD estimates of the LLR (17) are found by solving the LP problem

$$\theta_i^+ + \theta_i^- \to \min, \tag{26}$$

$$y_i = \gamma_0 + \sum_{j=1}^{l} \gamma_j x_{ij} + \sum_{j=1}^{l} \delta_j z_{ij} + \theta_i^+ - \theta_i^-, \; i = \overline{1, n}, \tag{27}$$

$$\theta_i^+, \; \theta_i^- \ge 0, \tag{28}$$

where $z_{ij} = \ln x_{ij}$,

$$\theta_i^+ = \begin{cases} y_i - \gamma_0 - \sum_{j=1}^{l} \gamma_j x_{ij} - \sum_{j=1}^{l} \delta_j z_{ij} \\ \text{if } y_i - \gamma_0 - \sum_{j=1}^{l} \gamma_j x_{ij} - \sum_{j=1}^{l} \delta_j z_{ij} > 0, \\ 0 \text{ otherwise}, \end{cases}$$

$$\theta_i^- = \begin{cases} \gamma_0 + \sum_{j=1}^{l} \gamma_j x_{ij} + \sum_{j=1}^{l} \delta_j z_{ij} - y_i \\ \text{if } \gamma_0 + \sum_{j=1}^{l} \gamma_j x_{ij} + \sum_{j=1}^{l} \delta_j z_{ij} - y_i > 0, \\ 0 \text{ otherwise}. \end{cases}$$

Then one of the following problems should be solved depending on the strategy to construct the LLR:

- for strategy 1, the LP problem with the objective function (26) and the linear constraints (27) and (28);
- for strategy 2, the PBLP problem with the objective function (26) and the linear constraints (27), (28), and (18)–(20);
- for strategy 3, the PBLP problem with the objective function (26) and the linear constraints (27), (28), (18), (19), (21), and (22).

4. MODELING RAIL FREIGHT TRAFFIC IN IRKUTSK OBLAST

Nowadays, a topical problem is to model rail freight traffic; for example, see [19, 20]. To demonstrate the mathematical apparatus proposed above, we considered this problem for Irkutsk oblast. Models were constructed based on the annual data of the Federal State Statistics Service for 2000–2018, available at the official website, with the following indicators:

– freight forward by public railway transport, y (million tons);

– labor force, x_3 (thousand people);

– gross regional product, x_{14} (million rubles);

– the number of enterprises and organizations, x_{18};

– industrial output (million rubles);

– electricity production, x_{22} (billion kWh);

– the average annual nominal wage in the mining industry, x_{23} (rubles);

– the average annual nominal wage in the manufacturing industry, x_{24} (rubles);

– agricultural output, x_{25} (million rubles);

– the average annual nominal wage in agriculture, hunting, and forestry (rubles);

– the number of active construction organizations;

– retail trade turnover, x_{31} (million rubles).

A special script was written in the hansl language of Gretl (an open-source statistical package for econometrics) to construct the MPER and LLR.

First, the MPER was constructed based on the initial data using strategy 3. The problem was solved by enumeration, and the estimates were obtained by the least squares method with m = 3 features. The complete enumeration of $C_{11}^3 \cdot 2^3 = 1320$ alternatives yielded the best one in terms of the coefficient of determination R^2. The resulting regression has the prologarithmic form

$$\ln \tilde{y} = -1.2502 + \underset{(2.972)}{8.431 \cdot 10^{-6}}\, x_{23} - \underset{(-6.098)}{3.388 \cdot 10^{-5}}\, x_{25} + \underset{(11.13)}{0.5176} \ln x_{31}, \quad (29)$$

and the coefficient of determination is $R^2 = 0.9334$. In equation (29), the values of the Student's t-test are indicated under the coefficients of the explanatory variables. According to these values, all coefficients are significant for the significance level $\alpha = 0.05$.

Unfortunately, due to the multicollinearity effect, the coefficient at the variable x_{25} changed its sign. Therefore, an attempt to interpret equation (29) leads to an absurd conclusion: we should reduce agricultural output for increasing rail freight traffic. Hence, when enumerating the models, we should check whether the signs of the regression equation coefficients agree with the practical interpretation of the variables. If at least one coefficient does not agree with its interpretation, then such a model is eliminated from further consideration. This recommendation can be found in the monograph [6]. Therefore, the MPER was rebuilt: an expert group determined that all explanatory variables should affect y with the "+" sign. The script was modified and launched with the same settings. As it turned out, among the 1320 alternatives, only 64 ones match the practical interpretation. The best of them is the logarithmic form model

$$\ln \tilde{y} = -6.4889 + \underset{(2.812)}{0.00127}\, x_3 + \underset{(3.555)}{0.533} \ln x_{18} + \underset{(3.447)}{0.754} \ln x_{22}, \quad (30)$$

where all coefficients of the explanatory variables are significant and $R^2 = 0.7437$.

The MPER corresponding to equation (30) is

$$\tilde{y} = 0.00152 \cdot e^{0.00127 x_3} \cdot x_{18}^{0.533} \cdot x_{22}^{0.754}. \quad (31)$$

The sum of the squared residuals for the model (31) is 229.598.

The model (31) is interpreted in the following way: with an increase in the labor force x_3 by 1 thousand people, the freight forward y raises by 0.127% on average; with an increase in the number of enterprises and organizations x_{18} by 1%, the freight forward y raises by 0.533% on average; with an increase in the electricity production x_{22} by 1%, the freight forward y raises by 0.754% on average.

Then, the LLR was constructed based on the initial data using strategy 3. The script settings were the same as for the MPER. The enumeration of the 1320 alternatives yielded the model

$$\tilde{y} = -267.173 + \underset{(2.55)}{0.000639}\, x_{24} - \underset{(-6.185)}{0.00193}\, x_{25} + \underset{(12.87)}{29.124} \ln x_{14}, \quad (32)$$

where all coefficients at the features are significant and $R^2 = 0.9328$.

In the model (32), the coefficient at the variable x_{25} again does not match the practical interpretation of the problem. Therefore, this model was rebuilt by checking the signs of the coefficients. As it turned out, among the 1320 alternatives, only 64 ones match the practical interpretation. The best of them is the regression

$$\tilde{y} = -552.38 + \underset{(2.764)}{0.0746}\, x_3 + \\ + \underset{(3.489)}{31.1352} \ln x_{18} + \underset{(3.282)}{42.7013} \ln x_{22}, \quad (33)$$

where all coefficients of the explanatory variables are significant, $R^2 = 0.7312$, and the sum of the squared residuals is 233.236.

Clearly, in terms of the sum of the squared residuals, the LLR (33) is somewhat worse than the MPER (31). Note that the LLR (33) includes the same features as the MPER (31).

The model (33) is interpreted in the following way: with an increase in the labor force x_3 by 1 thousand people, the freight forward y raises by 0.0746 million tons on average; with an increase in the number of enterprises and organizations x_{18} by 1%, the freight forward y raises by 0.3113 million tons on average; with an increase in the electricity production x_{22} by 1%, the freight forward y raises by 0.427 million tons on average. In addition: if the number of enterprises and organizations x_{18} increases by e times, the freight forward y will grow by 31.1352 million tons on average; if the electricity production x_{22} increases by e times, the freight forward y will grow by 42.7013 million tons on average.

Thus, if the researcher needs to predict the freight forward y, he should apply the models (29) and (32). If the researcher is also interested in interpreting the effect of different features on y, he should choose the MPER (31) and LLR (33), approximately of the same quality but with different meanings.

CONCLUSIONS

This paper has introduced two new specifications for regression models: the multiplicative power-exponential regression (MPER) and the linear logarithmic regression (LLR). The issues of their estimation and practical interpretation have been considered. The main advantage of these specifications is that each regression coefficient, except the free term, can always be given some practical interpretation. The MPER and LLR specifications allow identifying and studying new nonlinear regularities of processes or objects. Generally speaking, these specifications increase the usefulness of regression analysis.

REFERENCES

1. Pardoe, I., *Applied Regression Modeling*, Wiley, 2020.
2. Westfall, P.H., and Arias, A.L., *Understanding Regression Analysis: a Conditional Distribution Approach*, Chapman and Hall/CRC, 2020.
3. Kleiner, G.B., *Proizvodstvennye funktsii: teoriya, metody, primenenie* (Production Functions: Theory, Methods, Application), Moscow: Finansy i Statistika, 1986. (In Russian.)
4. Kleiner, G.B., *Ekonomika. Modelirovanie. Matematika. Izbrannye trudy* (Economics. Modeling. Mathematics. Selected Works), Moscow: Central Economics and Mathematics Institute, 2016. (In Russian.)
5. Khatskevich, G.A., Pronevich, A.F., and Chaikovskii, M.V., Two-Feature Production Functions with a Given Marginal Rate of Substitution, in *Ekonomicheskaya Nauka Segodnya*, 2019, no. 10, pp. 169–181. (In Russian.)
6. Noskov, S.I., *Tekhnologiya modelirovaniya ob''ektov s nestabil'nym funktsionirovaniem i neopredelennost'yu v dannykh* (A Technology to Model Objects with Unstable Operation and Data Uncertainty), Irkutsk: Oblinformpechat', 1996. (In Russian.)
7. Bazilevsky, M.P. and Noskov, S.I., Formalization of the Problem of Construction of Linear Multiplicative Regressions in the Form of a Partial-Boolean Linear Programming Problem, *Modern Technologies. System Analysis. Modeling*, 2017, vol. 55(3), pp. 101–105. (In Russian.)
8. Ivanova, N.K., Lebedeva, S.A., and Noskov, S.I., Parameter Identification for Some Nonsmooth Regressions, *Informatsionnye Tekhnologii i Problemy Matematicheskogo Modelirovaniya Slozhnykh Sistem*, 2016, no. 17, pp. 107–110. (In Russian.)
9. Noskov, S.I. and Khonyakov, A.A., Software Complex for Building Some Types of Piecewise Linear Regressions, *Information Technology and Mathematical Modeling in the Management of Complex Systems*, 2019, vol. 4, no. 3, pp. 47–55. (In Russian.)
10. Bazilevskii, M.P. and Noskov, S.I., Estimation of Index Regression Models Using the Least Absolute Deviations, *Vestnik of Russian New University. Complex Systems: Models, Analysis, Management*, 2020, no. 1, pp. 17–23. (In Russian.)
11. Bazilevskii, M.P., Vergasov, A.S., and Noskov, S.I., Group Selection of Informative Variables in Regression Models, *South-Siberian Scientific Bulletin*, 2019, vol. 28, no. 4(1), pp. 36–39. (In Russian.)
12. Eliseeva, I.I., Kurysheva, S.V., Kosteeva, T.V., et al., *Ekonometrika* (Econometrics), Moscow: Finansy i Statistika, 2007. (In Russian.)
13. Dougherty, C., *Introduction to Econometrics*, Oxford: Oxford University Press, 2011.
14. Gorid'ko, N.P. and Nizhegorodtsev, R.M. Elaborating of Long-Run Lag Regressional Models of Cobb–Douglas Type, *Control Sciences*, 2012, no. 3, pp. 55–63. (In Russian.)

15. URL: https://ru.coursera.org/lecture/ekonometrika/3-1-3-intierprietatsiia-koeffitsiienta-pri-logharifmirovanii-DdROH.
16. Park, Y.W. and Klabjan, Y.W., Subset Selection for Multiple Linear Regression via Optimization, *Journal of Global Optimization*, 2020, vol. 77, pp. 543–574.
17. Chung, S., Park, Y.W., and Cheong, T., A Mathematical Programming Approach for Integrated Multiple Linear Regression Subset Selection and Validation, *Pattern Recognition*, 2020, vol. 108, article no. 107565.
18. Bertsimas, D. and Li, M.L., Scalable Holistic Linear Regression, *Operations Research Letters*, 2020, vol. 48, no. 3, pp. 203–208.
19. Bazilevskiy, M.P., Vrublevskiy, I.P., Noskov, S.I., et al. Medium-Term Forecasting of Performance Indicators of Functioning of Krasnoyarsk Railway, *Fundamental Research*, 2016, no. 10(3), pp. 471–476. (In Russian.)
20. Noskov, S.I. and Vrublevskii, I.P., A Regression Model for the Dynamics of Railway Transport Performance Indicators, *Modern Technologies. System Analysis. Modeling*, 2016, vol. 50(2), pp. 192–197. (In Russian.)

This paper was recommended for publication by N.N. Bakhtadze, a member of the Editorial Board.

Received January 11, 2021, and revised March 4, 2021.
Accepted March 10, 2021.

Author information

Bazilevskiy, Mikhail Pavlovich. Cand. Sci. (Eng.), Irkutsk State Transport University, Irkutsk, Russia, ✉ mik2178@yandex.ru

Cite this article

Bazilevskiy, M.P. Constructing Power-Exponential and Linear-Logarithmic Regression Models. *Control Sciences* **3,** 22–28 (2021). http://doi.org/10.25728/cs.2021.3.3

DOI: http://doi.org/10.25728/cs.2021.3.4

DESIGN OF MULTIVARIABLE TRACKING SYSTEMS VIA ENGINEERING PERFORMANCE INDICES BASED ON H_∞ APPROACH

V.N. Chestnov and D.V. Shatov

Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia

✉ vnchest@yandex.ru, ✉ dvshatov@gmail.com

Abstract. This paper proposes an algorithm for designing a measured output-feedback controller with given or achievable engineering performance indices for linear multivariable systems. The plant is subjected to bounded exogenous disturbances from the class of polyharmonic functions with an infinite number of harmonics and a bounded sum of their amplitudes for each disturbance component. As a result, additional tracking errors appear in controlled variables. The problem is to design a multivariable output-feedback controller ensuring given or achievable tracking errors, the settling time determined by a given or achievable degree of stability of the closed loop system, and a set of the oscillation indices M_i for the *i*th closed loop relating the *i*th reference signal g_i to the *i*th controlled variable z_i. In addition, the controller should ensure the conditions $M_i \le \gamma$, where γ is a given number or the minimand. As shown below, H_∞ control methods are quite convenient for solving such problems. An illustrative example of designing an interconnected electric drive is presented.

Keywords: linear multivariable systems, bounded exogenous disturbances, tracking errors, settling time, degree of stability, oscillation index of the *i*th loop.

INTRODUCTION

The classical theory of automatic control of minimum-phase neutral single-input single-output (SISO) plants has shown high practical efficiency due to the physical clarity of the engineering performance indices underlying it: the tracking error, the settling time, and the oscillation index [1].

For multivariable (multi-input multi-output) plants, such an approach to controller design has not yet been formed, although the obvious first step is to solve the problem of autonomous control [2–4] going back to Voznesenskii [5].

The authors' recent works [6, 7] were devoted to single-variable plants (both minimum-phase and non-minimum-phase, stable and unstable), and controllers were constructed based on the H_∞ theory. This paper deals with multivariable plants and can be considered an extension of the approach [8] to the class of tracking systems: the robustness of a closed loop system is assessed not using the radius of stability margins but the oscillation index, a more natural and generally accepted performance index in the theory and applications of tracking systems.

Let us clarify the concept of the oscillation index of a multivariable system: it means a set of the oscillation indices M_i for the *i*th closed loop relating the *i*th reference signal g_i to the *i*th controlled variable z_i. The controller should ensure the conditions $M_i \le \gamma$, where γ is a given number or the minimand. As shown below, H_∞ control methods are quite convenient for solving such problems.

In practice, automatic systems are subjected to bounded exogenous disturbances causing additional tracking errors. This paper considers polyharmonic exogenous disturbances with an infinite number of unknown harmonics and a bounded sum of their (unknown) amplitudes for each disturbance component. The controller should ensure given (or achievable) tracking errors in the presence of such disturbances.

Note that they cover an applications-relevant class of continuous disturbances with piecewise continuous time derivatives [8, 9].

Another engineering performance index employed in control design is the settling time, characterizing the response speed of the closed loop system under nonzero initial conditions and (or) a stepwise change in the reference signal or disturbance. Below, the settling time is indirectly taken into account by ensuring a given degree of stability of the closed loop system. Although this index estimates the rate of transient processes very approximately (especially when the roots of the characteristic polynomial of the closed loop system are close to each other), it has proven itself well in applications with an initial estimate of the settling time. As demonstrated in [6, 10], an excessive increase in the degree of stability (a smaller distance from the plant's zero to the imaginary axis) catastrophically reduces the radius of stability margins (raises the oscillation index) and the phase and modulus margins even in the minimum-phase case with an output-feedback controller. Such results are unacceptable in practice due to large overshoots in the step response of the closed loop system. This phenomenon is analogous to the burst effect [11] for the output-feedback controllers. Therefore, the controller design algorithm suggested below, like the one developed in [8], provides for a gradual increase in the degree of stability.

To the authors' knowledge, Aleksandrov [12, 13] and his students are among the leading researchers stating and solving particular problems of this very difficult class within the theory of LQ control and H_∞ optimization. A detailed survey of the corresponding results was provided in the paper [8]. In the Western literature, with the appearance of the H_∞ theory in the early 1980s, stability margins were given much attention [10, 14]. At the same time, the issues of accuracy, response speed, and stability margins, combined in a unified output-feedback controller design method for multivariable tracking systems, have not received proper coverage.

Let us mention some important publications on different aspects of this range of problems. For example, the issues of nonsmooth H_∞ optimization were considered in [15], and the results were later used in [16] to design controllers of a given structure and order (particularly PID controllers). The matter is that controllers based on modern design techniques have a high order. Hence, they are "fragile": lose stability under small deviations of their parameters from the calculated ones [17]. This property is usually expressed in small phase and modulus margins of control loops. Also, a still-unsolved problem is choosing weight functions in the design of multivariable H_∞ controllers, noted in [18]. (For scalar systems, some rules were suggested in the monograph [14].) Below, we introduce a strict mathematical rule for choosing a weight for a given tracking accuracy.

Besides a different measure of robustness for the closed loop system, this paper involves a fundamentally novel approach to accuracy compared to [8]: a new vector of weighted controlled variables is not introduced, but the vector of exogenous disturbances is weighted. With this approach, the degree of sufficiency of the results is considerably decreased.

As shown below, the problem to ensure the engineering performance indices reduces to a special H_∞ optimization problem [19, 20]. A numerical solution of such a degenerate problem can be conveniently obtained using the technique of Linear Matrix Inequalities (LMIs) [21, 22], e.g., in MatLab's Robust Control Toolbox [23]. Finally, an illustrative example of designing an interconnected electric drive [8] is presented.

1. PROBLEM STATEMENT

Consider a plant described by the state-space equations

$$\dot{x} = Ax + B_1 f + B_2 u, \quad z = Cx, \qquad (1)$$

where $x \in R^n$, $u \in R^m$, and $z \in R^{m_1}$ denote the plant's state vector, the control vector, and the vector of controlled variables, respectively; $f \in R^{m_3}$ is the vector of unmeasured exogenous disturbances.

Let the plant (1) be looped by a stabilizing output-feedback controller

$$\dot{x}_c = A_c x_c + B_c \varepsilon, \quad u = C_c x_c + D_c \varepsilon, \quad \varepsilon = g - z, \qquad (2)$$

where $x_c \in R^{n_c}$ ($n_c \le n$), $g \in R^{m_1}$, and $\varepsilon \in R^{m_1}$ denote the controller's state vector, the vector of reference signals, and the vector of measured tracking errors; A_c, B_c, C_c, and D_c are numerical matrices.

The exogenous disturbance vector has bounded components of the form

$$f_i(t) = \sum_{k=1}^{\infty} f_{ik} \sin(\omega_k t + \psi_{ik}), \quad i = \overline{1, m_3}, \qquad (3)$$

where the amplitudes $f_{ik} \ge 0$, the initial phases ψ_{ik}, and the frequencies ω_k ($i = \overline{1, m_3}$, $k = \overline{1, \infty}$) are unknown, and the number of harmonics is infinite.

Assume that the exogenous disturbance is bounded:

$$\sum_{k=1}^{\infty} f_{ik} \le f_i^*, \quad i = \overline{1, m_3}, \qquad (4)$$

where $f_i^* > 0$, $i = \overline{1, m_3}$, are given numbers.

Conditions (3) and (4) mean the inequalities $|f_i(t)| \le f_i^*$, ($i = \overline{1, m_3}$). The model (3) and (4) covers a wide applications-relevant class of continuous exogenous disturbances with piecewise continuous time derivatives [8]. Therefore, such disturbances can be expanded into an absolutely convergent Fourier series [9], representing a special case of (3) with multiple frequencies. In addition, the series (3) is not necessarily a periodic function of time. For example, choosing all frequencies in (3) equal to 0 and the initial phases equal to $(2k+1)\pi/2$, where $k=\overline{0,\infty}$, we arrive at a step function.

The tracking errors caused by the exogenous disturbance (3), (4) are defined as

$$\varepsilon_{i,st}=\sup_{t\geq t_{\text{set}}}|\varepsilon_i(t)|,\ \ i=\overline{1,m_1},$$

where t_{set} denotes the settling time. A requirement common in practice is

$$\varepsilon_{i,st}\leq\varepsilon_i^*,\ \ i=\overline{1,m_1}, \tag{5}$$

where $\varepsilon_i^*>0$ are given numbers (the desired tracking errors).

The settling time in the closed loop system (1) and (2) can be approximately estimated as $t_{\text{set}}\approx 3/\beta$, where β specifies the system's degree of stability (the minimum distance from the eigenvalues of the closed loop system matrix

$$A_{cl}=\begin{bmatrix}A-B_2D_cC & B_2C_c\\ -B_cC & A_c\end{bmatrix}$$

to the imaginary axis).

Problem. Find a stabilizing controller (2) under which:

The system's accuracy requirements

$$\varepsilon_{i,st}\leq\gamma_1\varepsilon_i^*,\ \ i=\overline{1,m_1}, \tag{6}$$

hold, where γ_1 is a given number or the minimand.

The oscillation indices do not exceed a given number (or the minimand) γ_2,

$$M_i=\|t_i\|_\infty\leq\gamma_2,\ \ i=\overline{1,m_1}, \tag{7}$$

where $t_i(s)$ is the transfer function of the closed loop system relating the ith reference signal g_i to the ith controlled variable z_i, and $\|t_i\|_\infty$ denotes its H_∞ norm.

The eigenvalues of the matrix A_{cl} of the closed loop system (1) and (2) satisfy the condition

$$\text{Re}\ \lambda_i(A_{cl})\leq-\beta,\ \ i=\overline{1,n+n_c}, \tag{8}$$

where $\beta\geq 0$ is a given number.

Let us comment on this problem, further referred to as the original problem.

If the plant (1) is non-minimum-phase in the control variable (has zeros in the right half-plane), then the initial accuracy requirement (5) cannot be satisfied for any ε_i^*; therefore, the requirements (6) should be considered instead. Moreover, if the plant is also unstable [14], then the value γ_2 on the right-hand side of inequality (7) always exceeds 1, and it has a lower bound on M_i that cannot be overcome by any linear controller. And finally, the degree of stability β cannot be made greater than the plant's zero closest to the imaginary axis, which sharply decreases the stability margin: the high accuracy requirement contradicts the requirement of low oscillation indices, and high performance (a large value of β) contradicts the requirement for stability margins (small values of M_i).

This paper seeks a reasonable compromise between the mutually contradictory engineering performance indices based on the H_∞ optimization technique, which became a very convenient tool for designing applications-relevant controllers.

2. SOLUTION BASED ON H_∞ APPROACH

For solving the original problem, we first establish a connection between the oscillation indices M_i, $i=\overline{1,m_1}$, and the H_∞ norm of the transfer function $T(s)$ of the closed loop system relating the reference signal vector g to the controlled variable vector z. The following result is true.

Lemma. If

$$\|T\|_\infty\leq\gamma, \tag{9}$$

then

$$M_i\leq\gamma,\ \ i=\overline{1,m_1}. \tag{10}$$

Note that due to (9), a similar inequality will hold for any element of the matrix $T(s)$ [24], particularly for any diagonal element $t_i(s)$: $\|t_i\|_\infty\leq\gamma$. Since the transfer function $t_i(s)$ relates the ith reference signal g_i to the ith controlled variable z_i, by definition we obtain $M_i=\|t_i\|_\infty$ and consequently (10).

At its input, the closed loop system (1) and (2) receives two exogenous signals, g and f. We form the augmented vector $w^{\text{T}}=\begin{bmatrix}g^{\text{T}} & f^{\text{T}}\end{bmatrix}$ and choose the vector z as the controlled output. In the closed loop system, these vectors are connected via the transfer function $T_{zw}(s)$:

$$z=T_{zw}(s)w=\begin{bmatrix}T(s) & T_f(s)\end{bmatrix}w, \tag{11}$$

where $T(s)$ denotes the transfer function of the closed loop system relating the vector g to the vector z; $T_f(s)$ denotes the transfer function of the closed loop system relating the vector f to the vector z.

Consider an auxiliary H_∞ optimization problem of the form

$$\|T_{zw}\|_\infty \le \gamma\,, \qquad (12)$$

where γ is a given number or the minimand.

Due to the transfer function structure (11), condition (12) can be written in the equivalent frequency representation

$$T^{\mathrm{T}}(-j\omega)T(j\omega) + T_f^{\mathrm{T}}(-j\omega)T_f(j\omega) \le \gamma^2 I, \quad \omega \in [0,\infty), \qquad (13)$$

where I denotes an identity matrix of compatible dimensions. Hence,

$$T^{\mathrm{T}}(-j\omega)T(j\omega) \le \gamma^2 I \quad T_f^{\mathrm{T}}(-j\omega)T_f(j\omega) \le \gamma^2 I, \quad \omega \in [0,\infty), \qquad (14)$$

where the former condition is equivalent to (9), and the latter one means that $\|T_f\|_\infty \le \gamma$.

Thus, with the controller (2) obtained by solving problem (12) numerically, we satisfy the target requirement (7) for $\gamma_2 = \gamma$, where γ is the value realized during the calculations.

Next, consider the accuracy requirements (6). To account for them when solving problem (12), we replace the matrix B_1 of the plant (1) by $B_1 \cdot Q^{1/2}$, where $Q^{1/2}$ is a scalar weight specified below. Then the second condition of (14) takes the form

$$T_f^{\mathrm{T}}(-j\omega)QT_f(j\omega) \le \gamma^2 I, \quad \omega \in [0,\infty)\,.$$

Using the lemma on the working process from [8], for the steady-state values of the controlled variables $z_{i,st} = \sup_{t \ge t_p} |z_i(t)|, \; i = \overline{1,m_1}$, we get

$$Qz_{i,st}^2 \le \gamma^2 \left(\sum_{j=1}^{m_3} f_j^* \right)^2, \; i = \overline{1,m_1}\,, \qquad (15)$$

where f_j^* are known bounds on the exogenous disturbance components from (4). In contrast to the paper [8], formula (15) involves the common weight Q for all variables $z_{i,st}$. Therefore, we choose it based on the least error $z_{\min} = \varepsilon_{\min} = \min\left\{\varepsilon_1^*, \varepsilon_2^*, \ldots, \varepsilon_{m_1}^*\right\}$:

$$Q = \left(\sum_{j=1}^{m_3} f_j^* \right)^2 \Big/ \left(\varepsilon_{\min}^*\right)^2. \qquad (16)$$

In this case, the tracking error due to the exogenous disturbance f satisfies the relation $\varepsilon_{i,st} = z_{i,st}$. It follows from inequality (15) that

$$\varepsilon_{i,st} \le \gamma \varepsilon_{\min}^*\,, \; i = \overline{1,m_1}\,,$$

and the accuracy requirements (6) are satisfied.

Now, we account for the stability requirements (8) to the closed loop system, which determine the settling time. Following the paper [8], when solving problem (12), we replace the matrix A of the plant (1) by $\tilde{A} = A + \beta I$, where β is the desired degree of stability. Then the solution of the shifted problem (12) with $\tilde{A}$ yields a shifted controller with matrices $\tilde{A}_c$, B_c, C_c, and D_c. According to [8], the desired controller (2) solving the original problem has the matrices

$$A_c = \tilde{A}_c - \beta I,\; B_c,\; C_c,\; D_c\,. \qquad (17)$$

Summarizing the considerations above, we formulate the following result.

Theorem. The controller (2) and (17) solves the original problem if the weight Q in the shifted H_∞ problem (12) satisfies condition (16). In this case, the values of γ_1 and γ_2 (see the target requirements (6) and (7)) coincide with the value of γ realized when solving problem (12) and (13) numerically.

Note that transition from inequality (12) to inequalities (14) makes this result sufficient.

3. NUMERICAL SOLUTION

Since the controlled variable vector of the system (11) contains no controls, problem (12) is singular and cannot be solved numerically using the 2-Riccati method [19]. A preferable approach is based on the LMI technique [21, 22] and calculations in MatLab [23]; see the details below. As noted earlier, an excessive increase in the degree of stability β sharply raises the oscillation indices, causing large overshoots in the step response of the closed loop system. Therefore, the design algorithm presented below involves the principle of gradually increasing the response speed or the value of β. This algorithm includes the following steps.

1. Replace the plant's matrix A by the matrix $\tilde{A} = A + \beta I$, first letting $\beta = 0$.

2. Choose a weight from equality (16) and construct the four matrices A_{gen}, B_{gen}, C_{gen}, and D_{gen} of the state-space equations of the generalized plant: $A_{\text{gen}} = \tilde{A}$, $B_{\text{gen}} = \begin{bmatrix} 0 & B_1 \cdot Q^{1/2} & B_2 \end{bmatrix}$, $C_{\text{gen}} = \begin{bmatrix} C \\ -C \end{bmatrix}$, and $D_{\text{gen}} = \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix}$, where $D_{11} = \begin{bmatrix} 0 & 0 \end{bmatrix}$, $D_{21} = \begin{bmatrix} I & 0 \end{bmatrix}$, $D_{22} = 0$, $D_{12} = 0$, and all matrices have compatible dimensions.

3. Form the generalized plant's system matrix using the procedure

$$P = \text{ltisys}(A_{\text{gen}}, B_{\text{gen}}, C_{\text{gen}}, D_{\text{gen}})\,.$$

4. Find the optimal value $\gamma_0 = \mathrm{hinflmi}(P, [m_2, m])$ in problem (12), where m_2 and m are the numbers of the controller's inputs and outputs.

5. Choose $\gamma > \gamma_0$ and construct a controller's system matrix K that solves problem (12) using the procedure $[\gamma, K] = \mathrm{hinflmi}(P, [m_2, m], \gamma, \varepsilon)$, where ε is the accuracy of calculating γ.

6. Using the procedure $[\tilde{A}_c, B_c, C_c, D_c] = \mathrm{ltiss}(K)$, extract the state-space matrices of the shifted controller from the system matrix K.

7. Find the matrices $A_c = \tilde{A}_c - \beta I$, B_c, C_c, and D_c of the desired controller (17).

8. Construct the step response of the closed loop system under the exogenous disturbances (3) and (4), and find the tracking errors, the settling time, and the oscillation indices M_i. If the accuracy requirements (6) and (or) the oscillation requirements (7) do not hold, the problem is unsolvable by the proposed approach. Otherwise, proceed to Step 9.

9. If the response speed requirements do not hold, increase the value of β and get back to Step 1. Otherwise, the problem is solved.

For the first iteration of the algorithm, a natural choice is $\gamma = \gamma_0$. If the requirements (6) and (or) (7) do not hold, consider separate controller design problems with the oscillation index ($T_{zw}(s) = T(s)$) or with a given accuracy ($T_{zw}(s) = T_f(s)$ with the weight from formula (16)). These problems have a necessary and sufficient character, yielding γ_0 that determines the achievable accuracy (6) or oscillation indices $M_i = \gamma_0$.

4. CONTROLLER DESIGN FOR AN INTERCONNECTED ELECTRIC DRIVE

Consider an interconnected electric drive model described in the paper [8]. In [25], it was classified as a parallel system. The structural diagram of the model is shown in Fig. 1.

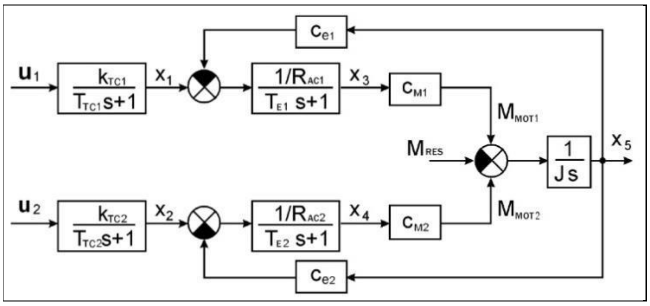


Fig. 1. The structural diagram of plant.

This diagram has the following notations: x_1 and x_2 are the deviations of the output voltages of the thyristor converters from the rated ones supplied to the armature circuits of the motors; x_3 and x_4 are the deviations of the armature currents of the drive motors; x_5 is the deviation of the angular rate of rotation of the motor shaft; u_1 and u_2 are the deviations of the control voltages supplied to the thyristor converters from the drive control system; M_{MOT1} and M_{MOT2} are the deviations of the electromagnetic moments developed by the motors from the rated values; M_{RES} is the deviation of the moment of resistance (load); T_{TC1} and T_{TC2} are the time constants of the thyristor converters; k_{TC1} and k_{TC2} are the gains of the thyristor converters; c_{M1}, c_{M2}, c_{e1}, and c_{e2} are the design constants of the motors; R_{AC1} and R_{AC2} are the active resistances of the armature circuits of the motors; T_{E1} and T_{E2} are the electromagnetic constants of the armature circuits of the motors; J is the total moment of inertia reduced to one of the motor shafts.

The model parameters in this diagram have the following numerical values: $c_{M1} = 8.1 \frac{\mathrm{N \cdot m}}{\mathrm{A}}$, $c_{M2} = 8.262 \frac{\mathrm{N \cdot m}}{\mathrm{A}}$, $c_{e1} = 8.15 \frac{\mathrm{V \cdot s}}{\mathrm{rad}}$, $c_{e2} = 8.313 \frac{\mathrm{V \cdot s}}{\mathrm{rad}}$, $T_{E1} = 0.0886\,\mathrm{s}$, $T_{E2} = 0.090372\,\mathrm{s}$, $T_{TC1} = 0.01\,\mathrm{s}$, $T_{TC2} = 0.012\,\mathrm{s}$, $R_{AC1} = 0.0819\ \mathrm{Ohm}$, $R_{AC2} = 0.08358\ \mathrm{Ohm}$, $k_{TC1} = 161.2$, $k_{TC2} = 164.424$, and $J = 32.5\ \mathrm{kg \cdot m^2}$.

The exogenous disturbance $f = M_c$ is the deviation of the moment of resistance (load) from the rated value. It does not exceed $f^* = 600$ Nm (20% of the rated motor moment). The measured variables for this plant are related to the physical variables: $y_1 = x_3$, $y_2 = x_4$, and $y_3 = x_5$. The exogenous disturbance f and the controls u_1 and u_2 are applied at different points. The main controlled variable of the plant is the angular rate of rotation of the motors: $z_3 = y_3 = x_5$. In addition, an important practical requirement to parallel systems is an equal load of the motors (close values of their armature currents) when operating on a common load. This requirement is often not satisfied when using standard PI controllers [25]. According to experimental evidence, if the angular rate of rotation of the motors is chosen as the only controlled variable ($z_3 = x_5$), then the equal load requirement may not hold, and most importantly, the stability margins for the measured variables $y_1 = x_3$ and $y_2 = x_4$ (the motor currents) at the plant's output may be very small, which is unacceptable in applications. Therefore, we will consider all measured variables of the plant as the controlled variables: $z_1 = y_1 = x_3$, $z_2 = y_2 = x_4$, and $z_3 = y_3 = x_5$ (the currents and the angular rate of rotation of the motors). Thus, the reference signals g_1 and g_2 will be fictitious and used for controller design only. Note that the behavior of the motor currents in the rated mode is completely determined by the variations in the load moment (disturbance).

The plant's matrices (1) have the following form [8]:

$$A = \begin{bmatrix} -100 & 0 & 0 & 0 & 0 \\ 0 & -83.333 & 0 & 0 & 0 \\ 137.811 & 0 & -11.287 & 0 & -1123.155 \\ 0 & 132.459 & 0 & -11.065 & -1101.133 \\ 0 & 0 & 0.2487 & 0.254 & 0 \end{bmatrix},$$

$$B_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ -0.031 \end{bmatrix}, \; B_2 = \begin{bmatrix} 16120 & 0 \\ 0 & 13702 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix},$$

$$C = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

The system requirements are:

• The tracking error in the angular rate of rotation should be $\varepsilon_{3,st} = z_{3,st} \le z_3^* = 1$ rad/s, and the current deviations in the transient modes should be $|z_1| \le 375$ A and $|z_2| \le 375$ A.

• The oscillation indices for the plant's measured outputs (y_1, y_2, , and y_3) should not exceed 1.

• The settling time should be $t_{\text{set}} = 0.25$ s.

We will design an appropriate controller using the algorithm from Section 3. For this purpose, we find the weight $Q^{1/2} = f^* / z_3^* = 600/1 = 600$ from formula (16) and let $\beta = 0$. The resulting controller matrices (2) and the realized value of γ are:

$$A_c = \begin{bmatrix} -387.822 & -91.669 & 421.791 & 902.372 & -4003.386 \\ -18.692 & -371.046 & -82.211 & 1772.776 & -7073.098 \\ -74.743 & -12.302 & -65.282 & 170.106 & -613.257 \\ -6.386 & -24.3 & -3.326 & -10.924 & 135.406 \\ 0.0525 & 0.0697 & 0.066 & -484.352 & -484.698 \end{bmatrix},$$

$$B_c = \begin{bmatrix} -0.00119 & 0.00132 & 0.000937 \\ 0.000693 & -0.000701 & 0.00003 \\ -0.00223 & 0.00252 & 0.000215 \\ 0.552 & 0.579 & 4.817 \\ 63.954 & 62.713 & -13.418 \end{bmatrix},$$

$$C_c = \begin{bmatrix} -0.00639 & 0.0118 & 0.0162 & -0.0697 & 0.265 \\ 0.0197 & 0.0161 & -0.0237 & -0.12 & 0.506 \end{bmatrix},$$

$$D_c = 0_{2\text{x}3}, \; \gamma = 51.86.$$

Note that the response of the closed loop system (1) and (2) ($g_3 \to z_3$) to the reference signal $g_3 \to z_3$ gives a large static error in the angular rate of rotation. To eliminate it, we rescale the reference signal by a value inverse to this error (the closed loop system gain, easily found from the open loop Nyquist contour for the angular rate of rotation; see below). After such scaling, the amplitude-frequency response of the closed loop system for the reference signal is presented in Fig. 2*b*, and the system's response in the angular rate-of-rotation channel is shown in Fig. 3*b*. The amplitude-frequency response of the closed loop system in the disturbance channel ($f \to z_3$) scaled by 600 Nm is demonstrated in Fig. 2*a*. This response is monotonically decreasing: for the closed loop system, the worst disturbance from the class (3) and (4) is the step function. In the angular rate-of-rotation channel, the corresponding response to such a disturbance of 600 Nm is shown in Fig. 3*a*. Clearly, the accuracy requirements are satisfied, like the response speed requirements to the closed loop system ($t_{\text{set}} \le 0.25$ s). The motor currents in the transient modes are very close

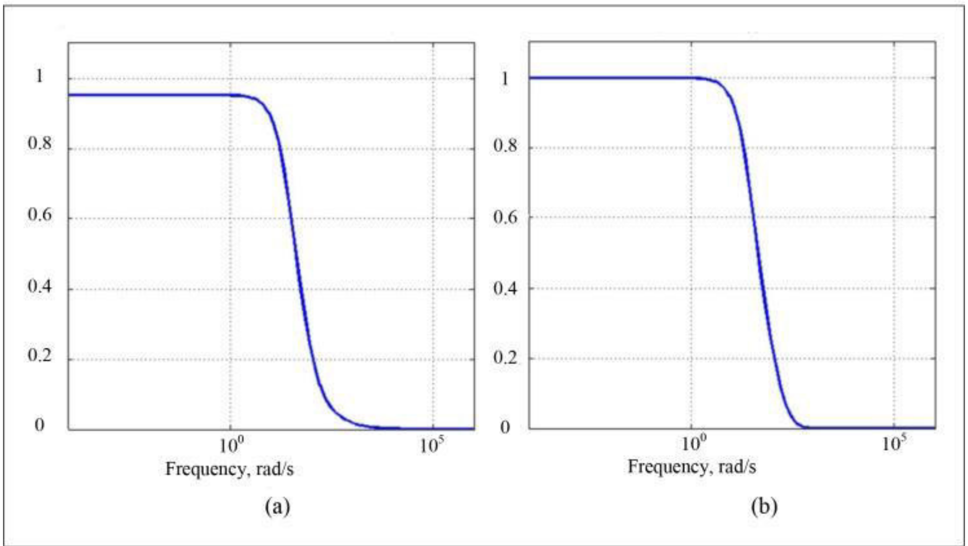


Fig. 2. The amplitude-frequency response of the closed loop system:
(a) disturbance and (b) reference signal

(the motors are equally loaded), and their deviations from the rated value are much less than the admissible value of 375 A.

Figure 4 provides the Nyquist contour of the open loop system in the corresponding measured variables: the motor currents (on the left) and the angular rate of rotation (on the right). Obviously, the system has an infinite phase margin for the measured variables - since the curves $w_i(j\omega)$ are entirely inside the unit circle. The modulus margins for the first and second measured variables (the currents of the first and second motors) are 1.6 and 1.4, respectively. The modulus margin for the main controlled variable (the angular rate of rotation of the motors) is 833. Thus, the system has significant stability margins for the measured variables. Moreover, the contours do not encircle the critical point $(-1, j0)$: the open loop systems are stable, which is important from a practical viewpoint. Checking the stability margins at the plant's physical input, we establish that the stability margins for the control variables u_1 and u_2 have radii 0.46 and 0.42, respectively, and open loop systems are stable in the control variable u_1 or u_2.

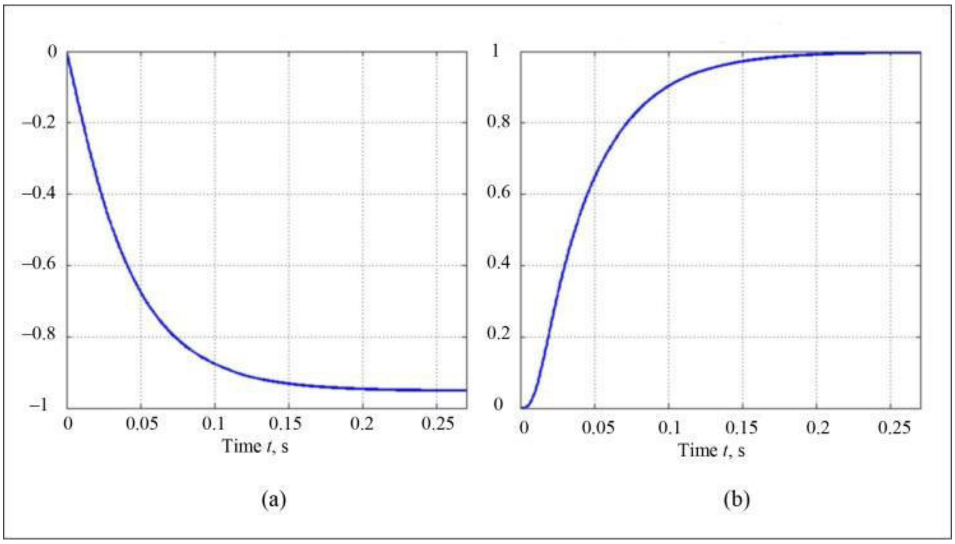


Fig. 3. The system's response to step disturbance and reference signal:
(a) output under $f(t) = 600$ and (b) output under $g_3(t) = 85.47$.

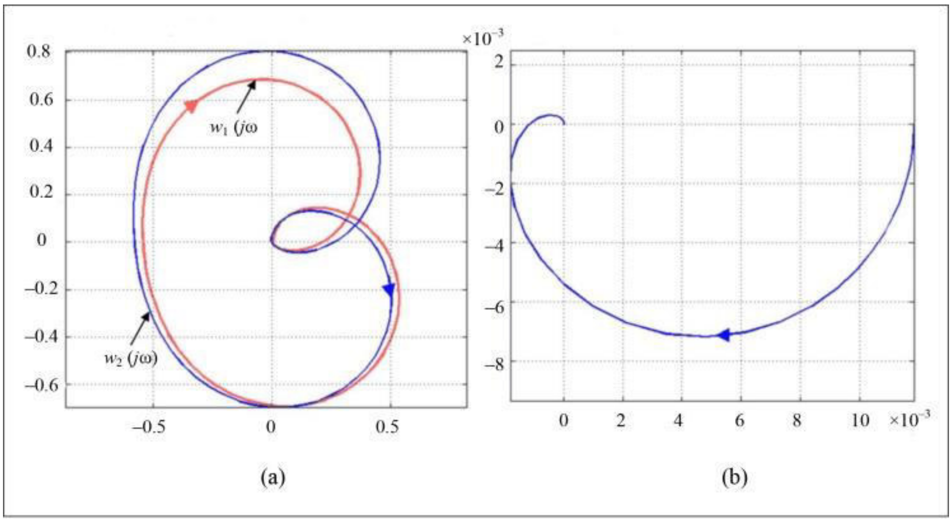


Fig. 4. The Nyquist contour of the closed loop system for different outputs:
(a) $w_1(j\omega)$ and $w_2(j\omega)$, (b) $w_3(j\omega)$.

CONCLUSIONS

This paper has presented an approach to design multivariable tracking systems based on engineering performance indices: the tracking errors due to unmeasured exogenous disturbances, a set of oscillation indices M_i for the ith closed loop relating the ith reference signal g_i to the ith controlled variable z_i, and the settling time. Note some features of the approach that are attractive from an engineering viewpoint:

- Clear engineering performance indices are applied.
- The controller design procedure reduces to the standard H_∞ optimization problem, solved by powerful software tools.
- The controller's order does not exceed that of the original physical object.
- An applications-relevant class of continuous disturbances with piecewise continuous time derivatives is considered.

If the oscillation indices should be provided at the physical input of a plant, then extra (fictitious) exogenous disturbances have to be introduced additively with the controls, and the controls have to be treated as the controlled variables.

REFERENCES

1. Besekerskii, V.A. and Popov, E.P., *Teoriya sistem avtomaticheskogo regulirovaniya* (Theory of Automatic Control Systems), Moscow: Nauka, 1975. (In Russian.)
2. Wang, Q.-G., *Decoupling Control*, Berlin: Springer, 2003.
3. Aleksandrov, A.G., Design of Controllers by Indices of Precision and Speed. III. Control-stable Multidimensional Plants, *Automation and Remote Control*, 2018, vol. 79, no. 2, pp. 241–257.
4. Aleksandrov, A.G. Design of Controllers by the Indices of Precision and Speed. IV. Method of Diagonal Domination, *Automation and Remote Control*, 2018, vol. 79, no. 4, pp. 581–593.
5. Voznesenskii, I.N., On Regulation of Machines with a Large Number of Parameters Regulated, *Automation and Remote Control*, 1938, vol. 4-5, pp. 65–78.
6. Chestnov, V.N. and Shatov, D.V., Design of Given Oscillation Index Scalar Controllers: Modal and H_∞-approaches, *Control Sciences*, 2019, no. 2, pp. 2–8. (In Russian.)
7. Chestnov, V.N. and Shatov, D.V, Design of H_∞ Controllers for One-Dimensional Tracking Systems via Engineering Performance Indices, *Trudy 13-go Vserossiiskogo soveshchaniya po problemam upravleniya (VSPU XIII)* (Proceedings of the 13th All-Russian Meeting on Control Problems (AMCP-13)), Moscow: Trapeznikov Institute of Control Sciences, 2019, pp. 88–93. (In Russian.)
8. Chestnov, V.N., Synthesis of Multivariable Systems According to Engineering Quality Criteria Based on H_∞-optimization, *Automation and Remote Control*, 2019, vol. 80, no. 10, pp. 241–257.
9. Bari, N.K., *Trigonometric Series*, New York: Holt, Rinehart and Winston, 1967.
10. Åström, K.J. and Murray, R.M., *Feedback Systems: an Introduction for Scientists and Engineers*, New Jersey: Princeton University Press, 2008.
11. Polyak, B.T., Khlebnikov, M.V., and Shcherbakov, P.S., *Upravlenie lineinymi sistemami pri vneshnikh vozmushcheniyakh. Tekhnika lineinykh matrichnykh neravenstv* (Control of Linear Systems under Exogenous Disturbances), Moscow: LENAND, 2014. (In Russian.)
12. Aleksandrov, A.G., *Sintez regulyatorov mnogomernykh system* (Controller Design for Multidimensional Systems), Moscow: Mashinostroenie, 1986. (In Russian.)
13. Aleksandrov, A.G., *Metody postroeniya sistem avtomaticheskogo upravleniya* (Methods to Construct Automatic Control Systems), Moscow: Fizmatlit, 2008. (In Russian.)
14. Skogestad, S. and Postlethwaite, I., *Multivariable Feedback Control. Analysis and Design*, New York: John Wiley and Sons, 2006.
15. Apkarian, P. and Noll, D., Nonsmooth H_∞ Synthesis, *IEEE Trans. Autom. Contr.*, 2006, vol. 51, no. 1, pp. 71–86.
16. Gahinet, P. and Apkarian, P., Structured H_∞-Synthesis in MATLAB, *Preprints of the 18th IFAC World Congress*, Milano, Italy, 2011, pp. 1435–1440.
17. Keel, L.H. and Bhattacharyya, S.P., Robustness and Fragility of High Order Controllers: A Tutorial, *Proceedings of IEEE Conference on Control Applications (CCA)*, Buenos Aires, Argentina, 2016, pp. 191–202.
18. Pittet, C., Mignot, J., and Viaud, F., Attitude and Active Payload Control: the H_∞-revolution, *Preprints of the 20th IFAC World Congress*, Toulouse, France, 2017, pp. 6619–6624.
19. Doyle, J.C., Glover, K., Khargonekar, P.P., and Francis, B.A., State-Space Solution to Standard H_2 and H_∞-control Problem, *IEEE Trans. Autom. Contr.*, 1989, vol. 34, no. 8, pp. 831–846.
20. Zhou, K., Doyle, J.C., and Glover, K., *Robust and Optimal Control*, New Jersey: Prentice-Hall, 1996.
21. Boyd, S., El Ghaoui, L., Feron, E., and Balakrishnan, V., *Linear Matrix Inequalities in System and Control Theory*, Philadelphia: SIAM, 1994.
22. Gahinet, P. and Apkarian, P., A Linear Matrix Inequality Approach to H_∞ Control, *Int. J. Robust and Nonlinear Control*, 1994, vol. 4, pp. 421–448.
23. Balas, G.J., Chiang, R.Y., Packard, A., and Safonov, M.G., *Robust Control Toolbox 3. User's Guide*, Natick: The MathWorks, 2010.
24. *The Control Handbook*, Levine, W.S., Ed., IEEE Press, 1996.
25. Åström, K. J. and Hägglund, T., *Advanced PID Control*, NC: ISA, 2006.

This paper was recommended for publication by S.A. Krasnova, a member of the Editorial Board.

Received February 13, 2020, and revised May 13, 2021. Accepted May 13, 2021.

Author information

Chestnov, Vladimir Nikolaevich. Dr. Sci. (Eng.), Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, ✉ vnchest@yandex.ru

Shatov, Dmitry Vladimirovich. Cand. Sci. (Eng.), Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, ✉ dvshatov@gmail.com

Cite this article

Chestnov, V.N, Shatov, D.V. Design of Multivariable Tracking Systems via Engineering Performance Indices Based on H_∞ Approach. *Control Sciences* **3,** 29–36 (2021). http://doi.org/10.25728/cs.2021.3.4

DOI: http://doi.org/10.25728/cs.2021.3.5

PRICING MECHANISMS FOR COST REDUCTION UNDER BUDGET CONSTRAINTS

V.N. Burkov and A.V. Shchepkin

Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia

✉ vlab17@bk.ru, ✉ av_shch@mail.ru

Abstract. The problem of evaluating the prices (cost) of individual projects of a megaproject or program is considered. The megaproject manager evaluates the cost of each project based on its planned cost reported by the project executors under the budget constraint on the total cost of the program. The executor of each project is a monopolist in the relevant area and cannot be replaced by another executor. In the deterministic case, the executors know the exact actual cost of their project; the manipulability of the mechanism for forming the cost of projects is investigated. In the stochastic case, the executors do not know the actual cost of their projects; when evaluating the planned cost, they estimate a probable value of the actual cost. For this estimate, the distribution function of the project's actual cost is used. The paper proposes a pricing mechanism for cost reduction under the budget constraint on the total cost of the program and a given probability distribution of the project's actual cost.

Keywords: cost price, limit price, profit, mathematical expectation, manipulability, cost reduction.

INTRODUCTION

Expectations that a market economy would maximize production efficiency did not realize. In such an economy, competition is a prerequisite for increasing efficiency, and no competition occurs in the presence of monopolists.

According to the monograph [1], the mechanism of cost reduction is a mechanism that encourages every employee to increase production efficiency and manufacture better quality products at lower costs and lower prices. The main results on control mechanisms of cost reduction were obtained for deterministic models [1–4]. Determinism was primarily understood as the absence of random disturbances. Note that cost-reducing control mechanisms were developed within the theory of active systems [5] to fight monopoly effects such as eliminating or preventing competition [6]. (Note that they are also called counter-expensive mechanisms.) At the same time, most studies of Russian researchers were aimed at cost-reducing measures without considering the monopoly effect [7–12]. The main focus in Western literature was on applying antitrust laws [13–16], including the ones to split monopolies. In addition to these papers, cost reduction and management were widely studied [17–19].

Russian researchers associate the solution of stochastic problems in the theory of active systems primarily with the analysis of incentive mechanisms; for details, see [20–22]. Similar problems in Western literature are considered within contract theory [23–28] and risk analysis [29, 30]. This paper deals with the following case: the project manager does not know the project cost price in advance, but he knows the cost price distribution.

1. PROBLEM STATEMENT

Consider a two-level system composed of a Principal (the upper level allocating budget funds for program execution) and agents (the lower level represented by program executors). The program consists of n projects executed by n organizations (agents), each being a monopolist in the corresponding area. Each agent knows the project's cost and limit prices. The Principal has budget funds in an amount R, restricting the entire program's cost, and knows the limit price (cost) of each project. Let c_i and l_i denote the cost and limit prices of project i, respectively, i = 1,…, n. The problem is to determine the cost of each project.

Consider a game-theoretic statement of this problem.

1. Each agent reports an estimate of the project's cost (his strategy).

2. The Principal determines the cost (prices) of all projects based on the information received.

3. The agents and the Principal determine their payoffs. The agent's payoff is the expected profit. The Principal's payoff function can be different, but this is not important: this aims to analyze the agents' strategies.

In the sequel, we will suppose that $C < R$, where $C = \sum_{i=1}^{n} c_i$. The cost C_i of project i , i = 1,…, n, will be determined using a pricing mechanism for cost reduction [1–4]:

$$C_i = s_i + k(l_i - s_i),\ i = 1,\dots, n,$$

where s_i denotes the cost price estimate reported by agent i. A natural assumption is $s_i < l_i$: the Principal will not consider the cost price estimates higher or equal to the limit price. Another natural assumption has the form $\sum_{i=1}^{n} l_i = L > R$.

The value k is obtained from the condition

$$R = \sum_{i=1}^{n} C_i = S + k\left(L - S\right), \qquad (1)$$

where $S = \sum_{i=1}^{n} s_i$.

From condition (1) we find

$$k = \frac{R - S}{L - S}. \qquad (2)$$

Hence, $k < 0$ if $S > R$: the project's price established by the Principal is smaller than the cost price estimate reported by the agent. In the case $S < R$, we have $1 > k > 0$: the project's price exceeds its cost price estimate.

The agent's profit is given by

$$P_i = C_i - c_i = s_i + k\left(l_i - s_i\right) - c_i,\ i = 1,\dots, n. \qquad (3)$$

We write the expression (3) as

$$P_i = k\left(l_i - s_i\right) + s_i - c_i,\ i = 1,\dots, n. \qquad (4)$$

Let $P_i^{(\text{pl})} = k\left(l_i - s_i\right)$ be the agent's planned profit, and $P_i^{(\text{sp})} = s_i - c_i$ be the super-planned profit. Naturally, the matter concerns the super-planned profit if $s_i > c_i$. In this paper, the profit of agent i is calculated as

$$P_i = P_i^{(\text{pl})} + qP_i^{(\text{sp})} = k\left(l_i - s_i\right) + q\left(s_i - c_i\right),\quad i = 1, \dots, n, \qquad (5)$$

where $q \le 1$. If $q \in (0, 1]$, then the Principal allocates a share of the super-planned profit to the agent, and q is a normative value of the agent's super-planned profit. Finally, if $q \le 0$, then q specifies a penalty coefficient for any project's cost price distortions.

In the case $s_i < c_i$, the agent's profit will be written in the form (4). Clearly, for $s_i = c_i$, the expression (4) coincides with (5).

2. STUDY OF MANIPULABILITY

Let the agents act under the hypothesis of weak contagion [5]. In this case, agent i neglects the effect of his estimate s_i on the value k. The agents will not benefit by overestimating the cost prices of their projects under the condition

$$\frac{\partial P_i}{\partial s_i} = -k + q < 0,\ i = 1,\dots, n. \qquad (6)$$

Inequality (6) will hold if $q < k$. Condition (6) being valid, each agent benefits by reporting the true estimate of the cost price c_i, i = 1,…, n, and the value k is given by

$$k = \frac{R - C}{L - C}.$$

Hence, we arrive in the following conclusion: even if $0 < q < k$ and the Principal allocates part of the super-planned profit to the agents, they will benefit not by overestimating the cost prices of their projects (to gain the super-planned profit) but by truth-telling (reporting the true estimates of the cost price).

However, a problem arises because the Principal announces the value q before the agents report their estimates of cost prices (before he calculates the value k).

To study manipulability, we write formula (5) as

$$P_i = kl_i - qc_i - \left(k - q\right)s_i,\ i = 1,\dots, n. \qquad (7)$$

Recall that $C < R$.

Case $S < R$. Here, formula (2) implies $0 < k < 1$. Due to the expression (7), for $q < k$, the profit of agent i decreases with the growth of his cost price estimate. Therefore, the agent's optimal strategy has the form $s_i^* = c_i$, i = 1,…, n. For $q = k$, the agent's strategy $s_i^* = c_i$, i = 1,…, n, is also optimal. Really—see (7) — if the agent is benevolent to the Principal (i.e., the hypothesis of benevolence holds [31]), he implements an action beneficial for the Principal.

Thus, for receiving reliable information about the projects' cost prices from the agents in the case $S < R$, the Principal should establish $q = \beta k$, where $\beta \le 1$.

Case $S > R$. Here, formula (2) implies $k < 0$. Due to the expression (7), for $k - q > 0$, the profit of agent i becomes negative (a loss occurs). For reducing this loss, the agent benefits by underestimating the cost price. Therefore, the agent's optimal strategy has the form $s_i^* = c_i$, $i = 1,\ldots, n$. Note that $q < k$ corresponds to $q < 0$: the agent is penalized for distorting the true information. For $q = k$, the agent's strategy $s_i^* = c_i$, $i = 1,\ldots, n$, is also optimal. This fact follows from the considerations in the case $S < R$.

Thus, if $q = \beta k$, where $\beta \le 1$, the mechanism will be cost-reducing in both cases.

Case $S = R$. Here, $k = 0$, and the profit of agent i is given by

$$P_i = q(s_i - c_i),\ i = 1,\ldots, n.$$

According to this expression, for $q < 0$, the agent benefits by underestimating the cost price. Therefore, the agent's optimal strategy has the form $s_i^* = c_i$, $i = 1,\ldots, n$. For $q > 0$, the agents should be interested in overestimating the cost prices. However, they cannot realize this scenario: any increase in the cost price estimate will immediately cause transition from $S = R$ to $S > R$, where the constraint $q \le k$ should be satisfied.

Note that the mechanism remains cost-reducing for one agent. Indeed, for one agent,

$$k^{(1)} = \frac{R - s_1}{l_1 - s_1} \tag{8}$$

and $P_1 = R - qc_1 - (1-q)s_1$. Since $q \in (0, 1]$, the optimal strategy has the form $s_1^* = c_1$.

3. STOCHASTIC CASE

Suppose that when planning the project's cost, each agent cannot accurately determine its cost price but knows the cost price distribution function $F(x_i)$, $F(l_i) = 1$, and the density function $f(x_i) = F'(x_i)$. As mentioned above, $s_i \le l_i$. In the sequel, all cost price estimates reported by the agents to the Principal satisfy the conditions $s_i \in [d_i, l_i]$, $i = 1,\ldots, n$, and $c_i \in [d_i, l_i]$, $i = 1,\ldots, n$. Therefore, the Principal and agents know that the project's cost cannot be smaller than d_i, $i = 1,\ldots, n$. Recall that in the deterministic case, the mechanism is cost-reducing if the agent's optimal strategy is reporting the true cost price $s_i^* = c_i$, $i = 1,\ldots, n$. In the stochastic case, in contrast, the mechanism is cost-reducing if the agent's optimal strategy is reporting a planned cost price less than the limit price.

First, consider the problem with one agent ($n = 1$).

Since the value $k^{(1)}$ is given by (8), the agent's profit can be written as

$$P_1(s_1) = R - s_1 + \begin{cases} q(s_1 - x_1) & \text{for } x_1 \le s_1, \\ s_1 - x_1 & \text{for } x_1 > s_1. \end{cases}$$

We calculate the expected profit:

$$M\left[P_1(s_1)\right] = R - s_1 + q\int_{d_1}^{s_1} (s_1 - x_1) f(x_1) dx_1 + \int_{s_1}^{l_1} (s_1 - x_1) f(x_1) dx_1. \tag{9}$$

From the expression (9) it follows that

$$\frac{dM\left[P_1(s_1)\right]}{ds_1} = -(1-q)F(s_1).$$

Hence, $\frac{dM\left[P_1(s_1)\right]}{ds_1} < 0$. In this case, the expected profit achieves maximum at $s_1 = d_1$, which corresponds to the that that the mechanism is cost-reducing:

$$M\left[P_1(d_1)\right] = R - \int_{d_1}^{l_1} x_1 f(x_1) dx_1 .$$

For example, let the random value x_i, $i = 1$, obey the uniform distribution on the interval $[d_i, l_i]$. Then the density function $f(x_i)$ has the form

$$f(x_i) = \frac{1}{l_i - d_i}. \tag{10}$$

The expected profit achieves the maximum value

$$M\left[P_1(d_1)\right] = R - \frac{l_1 + d_1}{2}.$$

Consider the case of n agents. In view of the expressions (4) and (5), the profit of agent i can be written as

$$P_i = k(L_i - s_i) + \begin{cases} q(s_i - x_i) & \text{for } x_i \le s_i, \\ s_i - x_i & \text{for } x_i \ge s_i, \end{cases} \quad i = 1,\ldots, n.$$

We calculate the expected profit:

$$M\left[P_i\right] = k(l_i - s_i) + s_i - (1-q) \times \left(s_i F(s_i) + \int_{s_i}^{l_i} x_i f(x_i) dx_i \right) - q\int_{d_i}^{l_i} x_i f(x_i) dx_i.$$

First, assume that the hypothesis of weak contagion [5] holds: the agent's estimate s_i has negligible effect on the value k, i.e., $\frac{\partial k}{\partial s_i} = 0$. Then

$$\frac{\partial M\left[P_i\right]}{\partial s_i} = (1-k) - (1-q)F(s_i),\ i = 1,\ldots, n. \tag{11}$$

The inequality $\frac{\partial M\left[P_i\right]}{\partial s_i} > 0$ holds if $(1-k) > (1-q)$, or equivalently, $S > \frac{R - qL}{1-q}$. The latter inequality is valid under $D > \frac{R - qL}{1-q}$, yielding

$$q > 1 - \frac{L-R}{L-D}. \quad (12)$$

If the value q satisfies (12), the expected profit will tend to the maximum value as $s_i \to l_i$, $i = 1,\ldots, n$. In other words, the mechanism is not cost-reducing.

Now consider the case $q < 1 - \frac{L-R}{L-D}$. To find the agents' estimates s_i, $i = 1,\ldots, n$, maximizing the expected profit, we solve the system of equations

$$(1-k) - (1-q)F(s_i) = 0,\ i = 1,\ldots, n.$$

If the random value x_i, $i = 1,\ldots, n$, obeys the uniform distribution on the interval $[d_i, l_i]$, then this system (see formula (10)) can be written as

$$\frac{s_i - d_i}{l_i - d_i} = \frac{1-k}{1-q},\ i = 1,\ldots, n. \quad (13)$$

The solution of (13) is given by

$$s_i^{(1)} = \frac{l_i + d_i}{2} + \frac{l_i - d_i}{2}\sqrt{1-V},\ i = 1,\ldots, n,$$

and

$$s_i^{(2)} = \frac{l_i + d_i}{2} - \frac{l_i - d_i}{2}\sqrt{1-V},\ i = 1,\ldots, n,$$

where $V = \frac{4}{1-q}\frac{L-R}{L-D}$ and $D = \sum_{i=1}^{n} d_i$.

Hence, the system (13) is solvable if

$$q \le 1 - 4\frac{L-R}{L-D}. \quad (14)$$

As noted, for $q \in (0, 1]$, the Principal allocates part of the super-planned profit to the agent. From inequality (14) it follows that part of the super-planned profit is at the agent's disposal if $q > 0$, or $R > \frac{1}{4}(3L + D)$.

Since

$$\frac{\partial^2 M[P_i]}{\partial s_i^2} = -(1-q) f(s_i),\ i = 1,\ldots, n,$$

the expected profit has two local maxima at the points $\{s_i^{(1)}\}$ and $\{s_i^{(2)}\}$, $i = 1,\ldots, n$.

The expected profit takes the following values:

– at the point $\{s_i^{(1)}\}$, the value

$$M\left[P_i\left(s_i^{(1)}\right)\right] = \frac{l_i - d_i}{2} \times \left[\frac{1+q}{2} - \frac{1-q}{2}\left(\sqrt{1-V} + \frac{1}{2}V\right)\right],\ i = 1,\ldots, n; \quad (15)$$

– at the point $\{s_i^{(2)}\}$, the value

$$M\left[P_i\left(s_i^{(2)}\right)\right] = \frac{l_i - d_i}{2} \times \left[\frac{1+q}{2} + \frac{1-q}{2}\left(\sqrt{1-V} - \frac{1}{2}V\right)\right],\ i = 1,\ldots, n. \quad (16)$$

Comparing the expressions (15) and (16), we observe the following: the agents gain the maximum expected profit at the point $\{s_i^{(2)}\}$. In other words, the mechanism is cost-reducing.

The paper [5] introduced the concept of a reliable estimate of the element's plan when studying the interaction between the Principal and one stochastic element. By analogy with reliability, let us define the probability that the project's random cost price estimate $\{s_i\}$ will take a value not exceeding $\{s_i^{(2)}\}$. Due to the distribution function formula, the probability $p\left(s_i \le s_i^{(2)}\right)$ is given by

$$p\left(s_i \le s_i^{(2)}\right) = \frac{1-\sqrt{1-V}}{2}.$$

This probability (the reliability of the estimate $\{s_i\}$) will be not smaller than u under the condition

$$\frac{1-\sqrt{1-V}}{2} \ge u.$$

Hence, the maximum value of u never exceeds 0.5. This reliability can be ensured by an appropriate choice of q. Indeed, it suffices to choose q so that inequality (14) turns into equality. Note that the choice of q determines the value $\{s_i^{(2)}\}$. For $q = 1 - 4\frac{L-R}{L-D}$, we obtain $s_i^{(2)} = \frac{l_i + d_i}{2}$, and moreover, $p\left(s_i \le \frac{l_i + d_i}{2}\right) = \frac{1}{2}$.

Next, consider the case $\frac{\partial k}{\partial s_i} \ne 0$. Here, formula (11) can be written as

$$\frac{\partial M[P_i]}{\partial s_i} = (1-k)\left(1 - \frac{l_i - s_i}{L-S}\right) - (1-q)F(s_i),\quad i = 1,\ldots, n. \quad (17)$$

Assuming that

$$\frac{l_i - s_i}{L-S} = \frac{l_i - s_i}{\sum_{j=1}^{n}(l_j - s_j)} \approx \frac{1}{n}$$

for sufficiently great n, we arrive at

$$\frac{\partial M[P_i]}{\partial s_i} = \frac{(n-1)(1-k)}{n} - (1-q)F(s_i),\ i = 1,\dots, n.$$

Clearly, the inequality $\frac{\partial M[P_i]}{\partial s_i} > 0$ will hold if $\frac{(n-1)(1-k)}{n} > (1-q)$, or equivalently, $S > \frac{(n-1)R-(nq-1)L}{n(1-q)}$. The latter inequality is the case under $D > \frac{(n-1)R-(nq-1)L}{n(1-q)}$, yielding

$$q > 1 - \frac{n-1}{n}\frac{L-R}{L-D}. \tag{18}$$

Well, if the value q satisfies (18), the expected profit will tend to the maximum as $s_i \to l_i$, $i = 1,\dots, n$. In other words, the mechanism is not cost-reducing.

Now consider the case $q < 1 - \frac{n-1}{n}\frac{L-R}{L-D}$. To find the agents' estimates s_i, $i = 1,\dots, n$, maximizing the expected profit, we solve the system of equations

$$\frac{(n-1)(1-k)}{n} - (1-q)F(s_i) = 0,\ i = 1,\dots, n.$$

If the random value x_i, $i = 1,\dots, n$, obeys the uniform distribution on the interval $[d_i, l_i]$, then its density function has the form (10). Therefore, this equation can be written as

$$\frac{s_i - d_i}{l_i - d_i} = \frac{n-1}{n(1-q)}\frac{L-R}{L-S},\ i = 1,\dots, n. \tag{19}$$

The solution of (19) is given by

$$\hat{s}_i^{(1)} = \frac{l_i + d_i}{2} - \frac{l_i - d_i}{2}\sqrt{1 - \frac{n-1}{n}V},\ i = 1,\dots, n,$$

and

$$\hat{s}_i^{(2)} = \frac{l_i + d_i}{2} + \frac{l_i - d_i}{2}\sqrt{1 - \frac{n-1}{n}V},\ i = 1,\dots, n.$$

Hence, the system (19) is solvable if

$$q \le 1 - 4\frac{n-1}{n}\frac{L-R}{L-D}. \tag{20}$$

Recall that the Principal allocates part of the super-planned profit when $q > 0$. From inequality (20) it follows that part of the super-planned profit is at the agent's disposal if $R > \frac{(3n-4)L + nD}{4(n-1)}$.

Since

$$\frac{\partial^2 M[P_i]}{\partial s_i^2} = -\frac{n-1}{n}\frac{\partial k}{\partial s_i} - (1-q)f(s_i),$$

we have

$$\left.\frac{\partial^2 M[P_i]}{\partial s_i^2}\right|_{s_i = \hat{s}_i^{(1)}} = \frac{1-q}{l_i - d_i} \times$$

$$\times\left(\frac{n}{n-1}\frac{l_i - d_i}{(L-D)V}\left(1 - \sqrt{1 - \frac{n-1}{n}V}\right)^2 - 1\right) < 0.$$

Consequently, the expected profit achieves maximum at the point $\{\hat{s}_i^{(1)}\}$, $i = 1,\dots, n$. This maximum is equal to

$$M[P_i] = \frac{l_i - d_i}{2} \times$$

$$\times\left[\frac{1+q}{2} + \frac{1-q}{2}\left(\sqrt{1 - \frac{n-1}{n}V} - \frac{n+1}{2n}V\right)\right], \tag{21}$$

$$i = 1, \dots, n.$$

In other words, the mechanism is cost-reducing.

For the case under consideration, we also define the probability that the project's random cost price estimate $\{s_i\}$ will take a value not exceeding $\{s_i^{(1)}\}$. As is easily shown, in this case, the maximum value of u will not exceed 0.5. To achieve this value, it suffices to choose q so that inequality (20) turns into equality.

Due to (15), (16), and (21), the maximum expected profits of agents diverge from each other only under unequal differences between the limit price l_i and the minimum cost estimate d_i.

Let the cost prices of agents' projects vary insignificantly; in this case, assume that the difference between the limit price l_i and the minimum cost estimate d_i is the same for all agents: $l_i - d_i = w$, $i = 1,\dots, n$. Under the uniform distribution of the random value x_i, $i = 1,\dots, n$, on the interval $[d_i, l_i]$, the derivative (17) can be written as

$$\frac{\partial M[P_i]}{\partial s_i} = (1-k)\left(1 - \frac{l_i - s_i}{L-S}\right) - (1-q)\frac{s_i - d_i}{w},$$

$$i = 1,\dots, n.$$

To find the agents' estimates s_i, $i = 1,\dots, n$, maximizing the expected profit, we solve the system of equations

$$(1-k)\left(1-\frac{l_i-s_i}{L-S}\right)-(1-q)\frac{s_i-d_i}{w}=0,\ i=1,\ldots,n. \quad (22)$$

The solution of (22) is given by

$$\tilde{s}_i^{(1)}=l_i-\frac{w}{2}\left(1-\sqrt{1-\frac{n-1}{n}V}\right),\ i=1,\ldots,n,$$

and

$$\tilde{s}_i^{(2)}=l_i-\frac{w}{2}\left(1+\sqrt{1-\frac{n-1}{n}V}\right),\ i=1,\ldots,n. \quad (23)$$

Clearly, the system (22) is solvable under inequality (20). For the case $l_i-d_i=w$, $i=1,\ldots,n$, inequality (20) reduces to $q\le 1-\frac{4}{n-1}\frac{L-R}{w}$.

Due to

$$\frac{\partial^2 M[P_i]}{\partial s_i^2}=2\left(1-\frac{l_i-s_i}{L-S}\right)\frac{(1-k)}{L-S}-(1-q)f(s_i), \quad (24)$$

$$i=1,\ldots,n,$$

and formulas (2), (10), and (23), the expression (24) can be written as

$$\left.\frac{\partial^2 M[P_i]}{\partial s_i^2}\right|_{s_i=\tilde{s}_i^{(2)}}=2\frac{1-q}{w}\left[\frac{1}{(n-1)V}\left(1-\sqrt{1-\frac{n-1}{n}V}\right)^2-\frac{1}{2}\right],$$

$$i=1,\ldots,n.$$

Obviously, $\left.\frac{\partial^2 M[P_i]}{\partial s_i^2}\right|_{s_i=\tilde{s}_i^{(2)}}<0$, and the expected profit achieves maximum at the point $\tilde{s}_i^{(2)}$. Therefore, the mechanism is cost-reducing.

The expected profit takes the value

$$M\left[P_i\left(\hat{s}_i^{(2)}\right)\right]=$$

$$=\frac{w}{4}\left[1+q-(1-q)\left(\frac{n+1}{2n}V-\sqrt{1-V\frac{n-1}{n}}\right)\right], \quad (25)$$

$$i=1,\ldots,n.$$

Hence, if the difference between the limit price l_i and the minimum cost estimate d_i is the same for all agents, the maximum expected profits of agents do not diverge from each other.

The expression (16) with $l_i-d_i=w$, $i=1,\ldots,n$, can be written as

$$M\left[P_i\left(s_i^{(2)}\right)\right]=\frac{w}{4}\left[1+q-(1-q)\left(\frac{V}{2}-\sqrt{1-V}\right)\right]. \quad (26)$$

Comparing formulas (25) and (26), we establish that

$$M\left[P_i\left(\tilde{s}_i^{(2)}\right)\right]>M\left[P_i\left(s_i^{(2)}\right)\right].$$

Thus, under the hypothesis of weak contagion, the expected profit is smaller compared to the case when the agents disregard it.

CONCLUSIONS

The problem of determining the prices of individual projects within a single program has been considered in the deterministic and stochastic statements. In the deterministic case, the cost reduction property of the pricing mechanism is ensured by choosing the super-planned profit q allocated to the agent ($q \le k$). As for the stochastic case, the cost reduction conditions, for known reasons, can no longer ensure the coincidence of the planned cost of the project with the actual cost but encourage agents to report the planned cost prices below the limit prices. An appropriate choice of q yields the cost price estimates below the limit prices and, moreover, the conditions to calculate the expected profits. In addition, note that the parameters of the pricing mechanisms for projects with budget constraints can be expressed analytically under the hypothesis of weak contagion (when the actions of one agent negligibly affect the performance of the entire system). Weak contagion holds under very many agents. At the same time, with an increase in the number of projects (agents), the agents can obtain part of the super-planned profit under more stringent constraints on the value q.

REFERENCES

1. Burkov, V.N., Danev, B., Enaleev, A.K., et al., *Bol'shie sistemy: modelirovanie organizatsionnykh mekhanizmov* (Large-Scale Systems: Modeling of Organizational Mechanisms), Moscow: Nauka, 1989. (In Russian.)
2. Burkov, V.N. and Kashenkov, A.R., Cost-reducing Mechanisms for R&D Management, in *Sovershenstvovanie organizatsionno-ekonomicheskogo mekhanizma upravleniya deyatel'nost'yu nauchno-issledovatel'skikh i proektno-konstruktorskikh uchrezhdenii* (Improving the Organizational and Economic Mechanism for Managing the Activities of R&D Institutions), Moscow: MDNTP, 1988, p. 45 (In Russian.)
3. Shchepkin, A.V., Cost-Reducing Funding Mechanisms, *Control Sciences*, 2018, no. 3, pp. 17–25. (In Russian.)
4. Burkov, V.N., Dorri, M.Kh., Shchepkin, A.V., and Kashenkov, A.R., Manipulability of the Cost-Reducing Funding Mechanism in Determining the Cost of a Development Consisting of Several Projects, *Informatsionnye Tekhnologii Modelirovaniya i Upravleniya*, 2018, vol. 112, no. 4, pp. 244–252. (In Russian.)
5. Burkov, V.N., *Osnovy matematicheskoi teorii aktivnykh sistem* (Fundamentals of the Mathematical Theory of Active Systems), Moscow: Nauka, 1977. (In Russian.)
6. Hayek, F.A., *Law, Legislation and Liberty. A New Statement of the Liberal Principles of Justice and Political Economy*, Routledge, 1982.
7. Iskoskov, M.O., Development of an Algorithm for Managing Investment Costs Based on the Goal Deployment Mechanism, *Economic and Law Issues*, 2011, no. 37, pp. 102–107. (In Russian.)

8. Kuzov, M. Cost Management: Practice, Ideas, Approaches, *Upravlenie Kompaniei*, 2007, no. 1, pp. 24–26. (In Russian.)
9. Lapygin, Yu.N. and Prokhorova, N.G., *Upravlenie zatratami na predpriyatii: planirovanie i prognozirovanie, analiz i minimizatsiya zatrat: prakticheskoe rukovodstvo* (Cost Management at the Enterprise: Planning and Forecasting, Cost Analysis and Minimization: A Practical Guide), Moscow: Expo, 2008. (In Russian.)
10. Mitskevich, A.A., *Upravlenie zatratami i pribyl'yu* (Cost and Profit Management), Moscow: Olma-Press Invest, Institute of Economic Strategies, 2003. (In Russian.)
11. Saraev, A.L. and Tatarskikh, B.Ya., Expenditures Rate Setting as a Tool of Management of Industrial Enterprises Output of Products, *Vestnik of Samara State University of Economics*, 2012, no. 4, pp. 79–85. (In Russian.)
12. Khamidullina, G.R., *Upravlenie zatratami: planirovanie, uchet, kontrol' i analiz izderzhek obrashcheniya* (Cost Management: Planning, Accounting, Control and Analysis of Circulation Costs), Moscow: Ekzamen, 2004. (In Russian.)
13. Hylton, K.N, *Antitrust Law: Economic Theory and Common Law Evolution*, Cambridge University Press, 2003.
14. Rockefeller, E.S., *Antitrust Religion,* Cato Institute, 2007.
15. Connor, J.M., Effectiveness of Antitrust Sanctions of Modern International Cartels, *Journal of Industry, Competition and Trade*, 2006, vol. 6, no. 3, pp. 195–223.
16. Manne, G.A. and Wright, J.D., Innovation and the Limits of Antitrust, *Journal of Competition Law and Economics*, 2010, vol. 6, no 1, pp. 153–202.
17. Spedding, T.A. and Sun, G.Q., Application of Discrete Event Simulation to the Activity Based Costing of Manufacturing Systems, *International Journal of Production Economics*, 1999, vol. 58, no. 3, pp. 289–301.
18. Mevellec, P., *Cost Systems Design*, Springer International Publishing, 2009.
19. Berk, J., *Cost Reduction and Optimization for Manufacturing and Industrial Companies*, Hoboken: John Wiley & Sons, and Massachusetts: New Jersey and Scrivener Publishing LLC, 2010.
20. Burkov, V.N., Enaleev, A.K., and Novikov, D.A., Probabilistic Incentive Problem, *Autom. Remote Control*, 1993, vol. 54, no. 12, pp. 1846–1851.
21. Enaleev, A.K. and Novikov, D.A., Optimal Incentive Mechanisms in an Active System with Probabilistic Uncertainty. I, *Autom. Remote Control*, 1995, vol. 56, no. 9, pp. 1297–1304.
22. Burkov, V.N. and Novikov, D.A., Optimal Incentive Mechanisms in an Active System with Probabilistic Uncertainty. II, *Autom. Remote Control*, 1995, vol. 56, no. 10, pp. 1455–1458.
23. Hart, O. and Holmstrom, B., Theory of Contracts, *5th World Congress on Advances in Economic Theory*, Cambridge, 1987, pp. 71–155.
24. Hart, O., *Firms, Contracts and Financial Structure*, Oxford: Oxford University Press, 1995.
25. Hart, O. and Holmström, B., Theory of Firm Scope, *Quarterly Journal of Economics,* 2010, vol. 125, no. 2, pp. 483–513.
26. Bolton, P., Corporate Finance, Incomplete Contracts, and Corporate Control, *Journal of Law, Economics, and Organization*, 2014, vol. 30, no. 1, pp. 64–81.
27. Kreitner, R., Voicing the Market: Extending the Ambition of Contract Theory, *University of Toronto Law Journal*, 2019, vol. 69, no. 3, pp. 295–336.
28. Dagan, H. and Heller, M., *The Choice Theory of Contracts*, Cambridge: Cambridge University Press, 2017.
29. Osadská, V., Stochastic Methods in Risk Analysis, *Transactions of the VSB – Technical University of Ostrava Safety Engineering Series*, 2017, vol. 12, no. 1, pp. 61–67.
30. Moncef, E., Stochastic Calculus in a Risk Model with Stochastic Return on Investments, *An International Journal of Probability and Stochastic Processes*, 2021, vol. 93, pp. 110–129.
31. Burkov, V.N., Goubko, M.V., Korgin, N.A., and Novikov, D.A., *Introduction to Theory of Control in Organizations*, CRC Press, 2019.

This paper was recommended for publication by G.A. Ougolnitsky, a member of the Editorial Board.

Received February 8, 2021, and revised March 9, 2021.
Accepted March 10, 2021.

Author information

Burkov, Vladimir Nikolaevich. Dr. Sci. (Eng.), Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia ✉ vlab17@bk.ru

Shchepkin, Aleksandr Vasil'evich. Dr. Sci. (Eng.), Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia, ✉ av_shch@mail.ru

Cite this article

Burkov V.N., Shchepkin A.V. Pricing Mechanisms for Cost Reduction under Budget Constraints. *Control Sciences* **3**, 37–43 (2021). http://doi.org/10.25728/cs.2021.3.5

DOI: http://doi.org/10.25728/cs.2021.3.6

STUDYING THE INDICATORS OF REGIONAL SPORTS DEVELOPMENT IN RUSSIAN FEDERATION

V.N. Prokofiev[1], K.V. Akimova[1], and A.L. Myachin[1,2]

[1]National Research University Higher School of Economics, Moscow, Russia
[2]Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia

[1]✉ vprokofiev@hse.ru, [2]✉ Aksenia1998@yandex.ru, [3]✉ amyachin@hse.ru

Abstract. The indicators of regional sports development in the Russian Federation are analyzed to find regions with a similar sports development strategy (according to the chosen methodology and measures of closeness) and identify dynamic groups in a four-year period. Some clustering and pattern analysis methods are described, and their use in the study is validated. The results obtained by classical clustering and ordinal-invariant pattern clustering methods are compared. The main state programs in the field of sports in the Russian Federation are highlighted and analyzed. The key aspects and problems of the state regulation of sports activities in the Russian Federation are indicated. Some ways for improving the existing regulatory and legal acts based on the dynamic analysis of regional patterns are proposed.

Keywords*:* sports, sport life, physical education, state regulation, cluster analysis, pattern analysis, ordinal-invariant pattern clustering.

INTRODUCTION

Over the past 15 years, reforms have been carried out in the Russian Federation to improve the sports infrastructure and increase the generalized indicators characterizing the health level of the citizens. As has long been known, sports improve well-being and increase the quality and duration of life.

When studying this topic, a problem consists in the operationalization of the very concept of "sports": "sports" and "physical culture" are often mixed. Several factors contribute to this ambiguity. As noted in the monograph [1], misunderstanding arises due to the similarity of these terms. In 2007, Federal Law no. 329 "On Physical Culture and Sports in the Russian Federation" was adopted in Russia to eliminate confusion and give precise definitions. This law contains sports concepts and their explanations and the rights and obligations of sports entities, forming a system of sports federations at all public authority levels. For example, the concept of "sports" is defined as "the sphere of social and cultural activities as a set of particular sports in the form of competitions and special practice of training a person for them" [2]. Hence, we conclude the following: although sports are part of physical culture, it differs significantly by content. This paper uses a basic system of indicators for a comprehensive study of sports development efficiency. To the best of authors' knowledge, only single indicators were estimated previously: infrastructure development [3], personnel policy [4], and the efficiency of state programs and regulatory legal acts [5].

Nowadays, sports, physical culture, and health care are socially popular and "in a growth trend" [6]: compared to 19% in 2011, the share of Russians involved in sports increased to 39.4% in 2018. This result is due to the state socio-economic policy considering the population's needs in sports and physical education in Russia. The government approves state programs encouraging people to participate in sports and improving the sports infrastructure in the country. An example is the Federal Target Program "Development of Physical Culture and Sports in the Russian Federation for 2006–2015," aimed at involving all segments of the population in sports by creating necessary conditions (developing sports infrastructure).

Table 1

Predicted and factual results of FTP "Development of Physical Culture and Sports in the Russian Federation in 2006-2015"

Indicator	Predicted results [7]	Factual results[1]
Share of citizens of the Russian Federation systematically involved in sports (%)	Increase to 30% of the total population of the Russian Federation	Before: 15.9 After: 31.7
Provision of population of the Russian Federation with sports objects (%)	Increase to 30 objects per 100 000 people (one object per 3 000 people)	Before: 22.7 After: 30.1
Number of qualified trainers occupied in sports and health work (in 1000 people)	Increase to 300 000 people (approximately 1 active trainer per 420 people)	Before: 295.6 After: 361.3
Share of sports events (%)	Increase from 3% in 2012 to 6% in 2020	Within the Unified Schedule of interregional, all-Russian, and international sports events, the number of such events increased by almost 6% compared to 2014. (More than 11 600 in total.)

Despite relatively high values achieved for all indicators (see Table 1), the following problems were identified in the report[2] of the Ministry of Sports following the results of this state program:

– no effective system for the development of children's and youth sports,

– a weak level of competition in elite sports,

– average statistical opportunities for improving physical development and health among citizens,

– lagging in innovative sports technologies.

As mentioned earlier, the values of the sports development indicators were nevertheless successfully increased. Thus, some problems have already been partially solved. For example, a technologically innovative sports base and infrastructure have been created (thanks to hosting the 2018 FIFA World Cup, which improved the investment climate and enhanced Russia's image in the world sports arena). However, there are some problems [8]:

– high market entry barriers for new manufacturers of sports goods,

– the lack of research and development projects in the field of sports,

– an insufficient demand for domestic sports goods and an insufficient share of Russian goods at the world level.

For solving them in 2019, the Government of the Russian Federation approved the Strategy for the Development of the Sports Industry until 2035 [9], with a focus on creating a new system of sports education for the population and elaborating and implementing active promotion measures for a healthy lifestyle, and providing efficient conditions for the development of physical education in educational institutions. The strategy is expected to increase to 55% the share of the population systematically involved in sports and to 67 years the average healthy life of the population. In addition, it is planned to renovate about 20 000 sports grounds, providing them with modern equipment, and increase the exports of Russian sports goods by 30% by 2024 (up to approximately $113 million).

1. METHODOLOGY FOR DETECTING HOMOGENEOUS GROUPS OF REGIONS BY SPORTS DEVELOPMENT STRUCTURE

Even under a fixed system of indicators, detecting structurally similar regions by the level of sports development is difficult. Various detection methods have certain features determining the difference in the final results. According to the problem statement, the following approaches can be used: linear convolution and threshold aggregation; formation of aggregate ratings (when applying the theory of individual and social choice); data classification, data clustering, and pattern analysis methods. Let us discuss each approach in detail.

A unified aggregate rating based on linear convolution and threshold aggregation allows identifying single groups of regions and has several positive aspects. First of all, it is simple and transparent calculations. Assigning each region a particular numerical

[1]https://www.minsport.gov.ru/activities/reports/9/28555/ (Accessed January 17, 2021).
[2]https://minsport.gov.ru/activities/reports/fiz-ra-i-sport-skryt/26361/ (Accessed January 17, 2021).

characteristic (linear convolution) or a particular rank (threshold aggregation) shows how much regions lag behind others. In addition, the time costs are minimized since these methods do not require intensive computing (given a basic system of indicators). However, the complexity of such an approach lies in determining the weight of each indicator (justifying their equal significance), which is often impossible. When compiling a unified aggregate characteristic, linear convolution allows compensating the small values of some indicators with the high values of others. Note that building a unified rating of regions by the level of sports development goes beyond the scope of this study.

Another approach involves aggregate ratings (if possible, within the theory of individual and social choice). Different quantitative indicators allow using different procedures (e.g., Borda's rule, Hare's rule, Nanson's procedure) to compile the unified aggregate rating. Thus, the closeness of the regions can be determined by their rank in the rating.

The practical application of data classification methods in this study is complicated by the uncertainty about the finite number of classes and their typical representatives. It is difficult to compile a learning sample, and using such methods is therefore complicated.

The main difference between the two remaining approaches (cluster analysis and pattern analysis) is that the latter methods are independent of the absolute values of indicators. Data clustering methods were considered in many surveys; for example, see [10, 11]. This study aims at dividing all regions into groups, each containing elements with a similar set of features. Hence, clustering methods seem to be one of the possible solutions. We adopt the concept of a "cluster" defined in the paper [11, p. 4]: "a piece of data…standing out from the rest of the data by some homogeneity of elements." Among the numerous methodologies, we employ two approaches with good interpretations of the final results: hierarchical clustering and *k*-means. All calculations were performed in Orange (Fig. 1) with the Manhattan distance L_1 and the Euclidean distance L_2 as the measures of closeness. For comparison, in the applied pattern analysis methods, the Hamming distance is used to estimate the proximity of objects.

Pattern analysis is less known in the Russian-language literature than cluster analysis. (On February 25, 2021, the Google Academy resource provided 76 000 search results for the user query "кластерный анализ" vs. 42 000 search results for the user query "анализ паттернов" and more than 4.5 million search

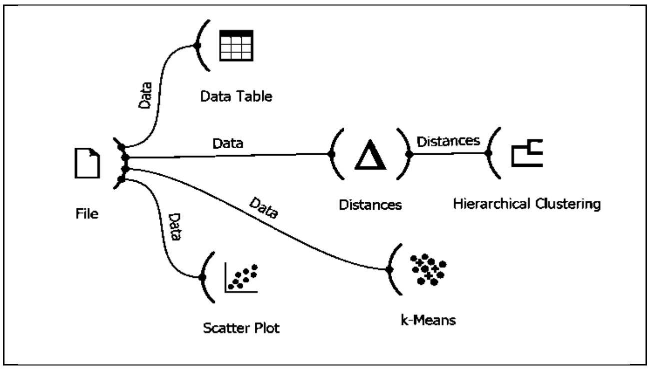


Fig. 1. Cluster methods realization in Orange.

is understood as "a combination of certain qualitatively similar features" [12, p. 139]. The methods should group objects with a similar structure of indicators under the endogenous composition and number of the groups. We select pattern analysis methods with the paired comparison of indicators (ordinal-fixed and ordinal-invariant pattern clustering). Their algorithmic implementation was presented in the papers [12–14]. Ordinal-fixed pattern clustering involves a predetermined sequence of the indicators and is often used to obtain preliminary ("rough") results. Ordinal-invariant pattern clustering assumes that the results are independent of the chosen initial sequence of indicators. A complete description of this methodology, including main properties, was given in the paper [12].

At present, pattern analysis is used in various fields: science, education, and innovations [15], macroeconomic analysis, political science, and regional innovative development assessment [16], and electoral behavior analysis [17]. In the paper [18], the behavior of commercial banks was analyzed, and the state capacity was estimated in [19].

This method prevails due to the possibility of analyzing both the current position of the objects (statics) and the dynamic trajectories of development (when studying regional pattern changes over time).

2. INDICATORS OF REGIONAL SPORTS DEVELOPMENT IN RUSSIAN FEDERATION

The efficiency of sports and physical culture development in the country depends on many factors. Among them, we mention, e.g., economic (wages, construction of the necessary infrastructure), demographic (the standard of living of the population), and political factors (state programs and funding, state support for sports life). The Ministry of Sports of the Russian Federation formed a system of indicators characterizing all sport life areas. Regular reports on these indicators are published on the official website of the Ministry of Sports of the Russian Federation.

The initial variables are the indicators for the period 2014–2017 (see the annual reports of the Ministry of Sports of the Russian Federation). For clarity, each region r_i in year t is described by the vector $r_i^t = (r_{i1}^t, r_{i2}^t, r_{i3}^t, r_{i4}^t, r_{i5}^t, r_{i6}^t, r_{i7}^t)$ with the following notations:

r_{i1}^t is the factual share of citizens systematically involved in sports in year t;

r_{i2}^t is the factual share of schoolchildren and students systematically involved in sports in year t;

r_{i3}^t is the number of sports objects per 100 000 people in year t;

r_{i4}^t is the factual provision of population with sports objects, based on the capacity of sports objects in year t;

r_{i5}^t is the number of rated sportsmen trained in year t;

r_{i6}^t is the amount of sports funding per one resident in year t (RUB);

r_{i7}^t is the number of sports staff in year t.

The data for the period 2018–2020 were not considered due to some differences in their acquisition and the methodology for calculating some indicators. These data will be studied in future as one line of further research.

Before analysis, the data were preprocessed. Correlation analysis with the Pearson correlation coefficient was carried out; see Table 2 presenting the results for the year 2017. Each indicator was normalized using linear scaling by the formula

$$r_{ij}^t = \frac{r_{ij}^t - r_{j\_\min}^t}{r_{j\_\max}^t - r_{j\_\min}^t},$$

where $r_{j\_\max}^t$ and $r_{j\_\min}^t$ denote the maximum and minimum values of indicator j in year t, respectively.

Correlation analysis for the four years showed a strong correlation (at the level of 0.83) between the factual share of citizens and the factual share of schoolchildren and students systematically involved in sports. Therefore, one of these indicators should be excluded from further analysis: a strong correlation may negatively affect the grouping of patterns. Also, note a positive correlation between the number of sports objects per 100 000 people and the factual provision of population with sports objects (0.73) and a weak correlation between the factual share of citizens systematically involved in sports and the factual provision of population with sports objects (0.28). This fact is partially explained by the relative efficiency of the government's policy to promote sports development and involve the population in sports. However, despite the increasing demand from the population for sports, the number of sports objects is not enough.

Table 2

Correlation analysis of indicators for year 2017

	r_{i1}^{2017}	r_{i2}^{2017}	r_{i3}^{2017}	r_{i4}^{2017}	r_{i5}^{2017}	r_{i6}^{2017}	r_{i7}^{2017}
r_{i1}^{2017}	1						
r_{i2}^{2017}	0.83	1					
r_{i3}^{2017}	0.31	0.37	1				
r_{i4}^{2017}	0.34	0.36	0.76	1			
r_{i5}^{2017}	0.26	0.21	–0.17	–0.23	1		
r_{i6}^{2017}	0.44	0.37	0.14	0.12	0.46	1	
r_{i7}^{2017}	–0.16	–0.20	–0.14	–0.17	0.02	–0.07	1

3. ANALYSIS RESULTS

We studied 340 objects: 85 Russian regions over four years. The results yielded by pattern analysis methods were compared with those of cluster analysis. Then groups of regions were identified and combined using different methodologies. Note that 1/3 of the studied objects were classified as "unique" (one pattern for one region). Let us describe some of the groups obtained.

Group 1. A distinctive feature of this regional group is a relatively great number of people involved in sports, although sports funding and staffing are relatively low. Possible explanations are as follows: similar to the availability of sports objects, the share of citizens involved in sports is high since the regional policy aims at increasing this share by actively promoting a healthy lifestyle and increasing the number of free and outdoor sports grounds. (For example, more than 800 sports events were organized in Omsk oblast.) The regional data are visualized below by piecewise linear functions in a parallel coordinate system for ease of comparison.

Figure 3 visualizes **Group 2**. In Vologda oblast, the share of citizens involved in sports is sufficient, but the number of sports staff has decreased compared to the year 2014. There was a massive downsizing due to small funding in the subsequent years. The regional authorities should support sportsmen who graduated from professional sports institutions and sports clubs and include the corresponding articles in the budget). In Kursk oblast, the provision of population with sports objects lags the normative values. Note that the municipal administration of this region and PJSC Gazprom concluded a cooperation agreement, increasing the number of sports objects in the region and improving the availability of sports for the population.

Group 3 includes only one region for two years: Krasnodarskii krai (Fig. 4). This region has medium and relatively high values of most indicators. Note that the 2014 Winter Olympic Games were held there, and the level of sports development sharply increased.

The second task of this study is to construct dynamic trajectories for the development of different regions of the Russian Federation in the period 2014–2017 based on the selected structure of sports indicators. The following question arises: Are there any changes in the sports development trajectory of these regions and their belonging to patterns? It can be answered using dynamic pattern analysis, the methods described in detail in [14, 15, 17]. In short, after partitioning the regions into groups, each region is assigned a vector $y_{ri} = \left(y_{2014}, y_{2015}, y_{2016}, y_{2017}\right)$, where y_t is the number of the group to which region *i* belongs in year *t*. Thus, dynamic groups are identified according to the change in strategies:

– absolutely stable (the region stays in the same group during the period),

– stable (the region changes its group only once);

– unstable (the region changes its strategy annually).

The dynamic groups were formed according to this approach. For example, consider three regions—Vladimir, Kostroma, and Orenburg oblasts—that changed their groups once (the stable type of dynamic groups, see Fig. 5).

The unstable group contains 32 regions. However, some regions returned to the original sports development strategy after changing it. Note that changes from "best" to "worst" strategies (and conversely) were not considered: ranking of regions goes beyond the scope of this study.

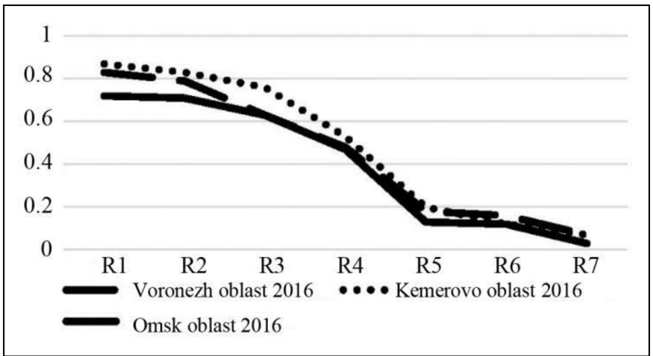


Fig. 2. Visualization of Group 1 in parallel coordinates.

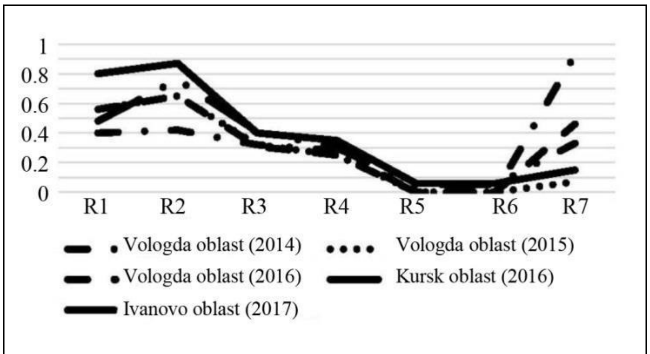


Fig. 3. Visualization of Group 2 in parallel coordinates.

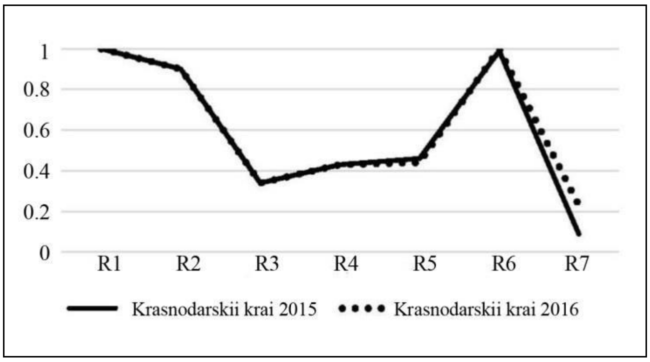


Fig. 4. Visualization of Group 3 in parallel coordinates.

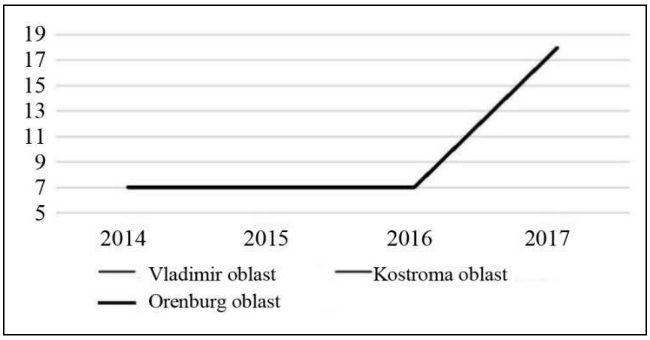


Fig. 5. Example of regions with strategy change by the system of indicators.

4. PROPOSALS FOR IMPROVING THE EXISTING LEGISLATION

According to the study results, it seems reasonable to revise the sports law completely and develop a single codified act—the Sports Code of the Russian Federation (SC RF). The code will fill the existing gaps in law and regulate public relations in sports as widely as possible: in the absence of relevant legal norms, these relations cannot develop properly. (In particular, the matter concerns the mechanisms for replicating and deploying best regional practices among the "lagging" regions of the Russian Federation.) Positive expectations from this code are based on foreign experience: for example, a similar document was adopted in France. The French Sporting Code compiles all existing normative and legal acts regulating sports activities, and new laws are not adopted.

Social relations in sports and associated areas have substantial and unique specifics. In interaction and interconnection, they differ by completeness, and most importantly, by systemic integrity: they can be treated as an independent area of state-legal management and regulation. The SC RF is the most efficient way of regulating and systematizing sports norms into a single whole [20].

Today, sports undergo various changes: new kinds of sports appear in the "snowball" (e.g., extreme sports and "active" cybersport), which are not included in the state classification of physical culture. It is very difficult to register a sport or a sports federation at the state level. This procedure should be transparently regulated and simplified as much as possible.

CONCLUSIONS

The previous studies raised the issue of analyzing sports clusters only. The definition of this concept included the construction of a sports infrastructure [21]. Increasing the number of sports objects that meet all technical requirements is a priority task of the state program, but some adjustments are needed. In this paper, "sports clusters" have been therefore defined as groups of regions with a similar sports development strategy. They have been combined using cluster and pattern analysis. According to the obtained results, each group contains one region with the indicator values differing from those of all other regions in the group. The sports development strategy in such a region is most efficient; therefore, the authorities of all other regions in the group should pay attention to the lagging indicators, undertaking some measures to improve them based on the best regional practice.

To the best of authors' knowledge, this paper applied pattern analysis methods to study sports activities in the Russian Federation for the first time in the literature. Some features of this method have been highlighted. Although pattern analysis is a relatively young data analysis method, it has already demonstrated high effectiveness in many areas. Identifying regions with a similar structure of indicators allows modifying the state and regional policies according to the regional sport life characteristics. The recommendations have been supplemented with some proposals for improving the current legislation.

REFERENCES

1. Matveev, L.P., *Obshchaya teoriya sporta i ee prikladnye aspekty* (General Theory of Sports and Its Applied Aspects), St. Petersburg: Lan', 2005. (In Russian.)
2. Federal Law no. 329-FZ dated December 4, 2007, "On Physical Culture and Sports in the Russian Federation" (as amended on August 2, 2019). (In Russian.)
3. Bobrovskii, E.A., Improvement of Sports Infrastructure as a Tool of Sports Popularization in a Region, *Mezhdunarodnyi Zhurnal Prikladnykh i Fundamental'nykh Issledovanii*, 2017, no. 3-2, pp. 297–301. (In Russian.)
4. Bobrovskii, E.A., Staff Policy in the Sphere of Physical Culture and Sports in the Russian Federation, *Karelian Scientific Journal*, 2017, vol. 6, no. 4 (21), pp. 315–318. (In Russian.)
5. Kosova, Yu.A., Perspective Guidelines of Governmental Policy in the Field of Physical Training and Sports in Russia, *Transport Business in Russia*, 2011, no. 8, pp. 115–116. (In Russian.)
6. Russian Public Opinion Research Center (VCIOM) Named the Main Life Values of Russians. https://www.dp.ru/a/2017/06/07/VCIOM_nazval_glavnie_zhizn (In Russian.)
7. Resolution of the Government of the Russian Federation no. 7 dated January 11, 2006, "On the Federal Target Program "Development of Physical Culture and Sports in the Russian Federation for 2006–2015." (In Russian.)
8. Decree of the Government of the Russian Federation no. 1101-r dated August 7, 2009, "On Approval of the Strategy for the Development of Physical Culture and Sports in the Russian Federation for the period up to 2020." (In Russian.)
9. Decree of the Government of the Russian Federation no. 1188-r dated June 3, 2019, "On Approval of the Strategy for the Development of the Sports Industry until 2035 and a Set of Measures for the Formation of a Modern Branch of the Sports Industry for 2019–2020." (In Russian.)
10. Kim, D.O., *Faktornyi, diskriminantnyi i klasternyi analiz* (Factor, Discriminant and Cluster Analysis), Ripol Klassik, 1989. (In Russian.)
11. Mirkin, B.G., Review of Cluster Analysis Methods for Decision-making Support, Working paper no. WP7/2011/03, National Research University Higher School of Economics, Moscow, 2011. (In Russian.)
12. Myachin, A.L., Pattern Analysis in Parallel Coordinates Based on Pairwise Comparison of Parameters, *Autom. Remote Control*, 2019, vol. 80, pp. 112–123.

13. Myachin, A.L., Pattern Analysis: Diffusion-Invariant Pattern Clustering, *Control Sciences*, 2016, no. 4, pp. 2–9. (In Russian.)
14. Myachin, A.L., Pattern Analysis: Ordinal-Invariant Pattern Clustering, *Large-Scale Systems Control*, 2016, vol. 61, pp. 41–59. (In Russian.)
15. Aleskerov, F.T., Gokhberg, L.M., Myachin, A.L., et. al., Data Analysis of Science, Education and Innovative Activities Using Pattern Analysis Methods, Working paper no. WP7/2012/07, National Research University Higher School of Economics, Moscow, 2012. (In Russian.)
16. Aleskerov, F.T., Belousova, V.Yu., Mirkin, B.G., et al., Methods of Pattern Analysis in Statics and Dynamics, Part 1 : Literature Review and Clarification of the Term, *Business Informatics*, 2013, no. 3 (25), pp. 3–18. (In Russian.)
17. Aleskerov, F.T., Borodin, A.D., and Kaspe, S.I., Analysis of Electoral Preferences in Russia in 1993–2003: Polarization Index Dynamics, *HSE Economic Journal*, 2005, vol. 9, no. 2, pp. 173–184. (In Russian.)
18. Aleskerov, F.T., Solodkov, V.M., and Chelnokova, D.S., A Dynamic Analysis of Behavioral Patterns of Russian Commercial Banks, *HSE Economic Journal*, 2006, vol. 10, no. 1, pp. 48–62. (In Russian.)
19. Akhremenko, A.S. and Myachin, A.L., Pattern Analysis and Clustering in the Study of State Capacity: "Adaptive Optics" for Political Science, *Political Science* (*RU*), 2019, no. 3, pp. 112–139. (In Russian.)
20. Solov'ev, A.A., Concept of Sports Code of the Russian Federation, Sports Right Commission of Russian Lawyers Association, Moscow, 2009. (In Russian.)
21. Ananishnev, V.V. and Korepova, V.V., Problems of Formation of Sports Clusters in Russia, *Clusters. Research and Development*, Corporate Scientific Journal of the Moscow Cluster of Business Initiatives (Moscluster), 2016, vol. 2 (3), pp. 35–38. (In Russian.)

This paper was recommended for publication by P.Yu.Chebotarev, a member of the Editorial Board.

Received July 30, 2020, and revised March 16, 2021.
Accepted April 6, 2021.

Author information

Prokofiev, Vadim Nikolaevich. Dr. Sci. (Eng.), National Research University Higher School of Economics, Moscow, Russia, ✉ vprokofiev@hse.ru

Myachin, Aleksei Leonidovich. Dr. Sci. (Eng.), Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia; National Research University Higher School of Economics, Moscow, Russia, ✉ amyachin@hse.ru,

Akimova, Kseniya Vital'evna. Student, National Research University Higher School of Economics, Moscow, Russia, ✉ Aksenia1998@yandex.ru.

Cite this article

Prokofiev V.N., Myachin A.L., Akimova K.V. Studying the Indicators of Regional Sports Development in Russian Federation. *Control Sciences* **3**, 44–50 (2021). http://doi.org/10.25728/cs.2021.3.6

DOI: http://doi.org/10.25728/cs.2021.3.7

AN UPDATING METHOD FOR THE DYNAMIC MIMO MODEL OF A CONTROLLED TECHNOLOGICAL OBJECT

M.A. Rabotnikov

OOO ZapSibNeftekhim, Tobolsk, Russia

✉ rabotnikovma@tobolsk.sibur.ru

Abstract. This paper considers the degradation of MIMO models of controlled industrial processes. We propose a method for solving a nonlinear programming problem with an objective function formed by the least squares method according to technological object data. The method involves dynamic process modeling algorithms based on the imposition of the step response of the process. The advantage of this method is the possibility of using passive experiment data to construct an appropriate multi-loop model of a technological object. The method is applied to update the multiparameter controller model for the Advanced Process Control system of butyraldehydes oxo synthesis. For real controlled technological objects, this method allows improving the accuracy of process modeling and the performance of automatic control, reducing the human factor, and increasing the overall economic efficiency of the production process.

Keywords: MIMO system, transfer function, Model Predictive Control, Advanced Process Control.

INTRODUCTION

For reaching maximum production efficiency, many high-tech solutions are being actively implemented at modern enterprises to optimize processes at all production stages and provide the high adaptability of production units to the current economic and technological conditions [1–4]. One of such solutions is Advanced Process Control (APC) systems [5].

In the overwhelming majority of cases, the concept of advanced control of continuous industrial processes involves dynamic models predicting the behavior of controlled process parameters. Predictive control algorithms are implemented based on the calculations to stabilize an industrial process, compensating for the effect of disturbances on the controlled parameters, and bring the technological object to an optimal mode by an economic criterion under given technological constraints [6–8].

1. DEGRADATION OF DYNAMIC MODELS: PROBLEM STATEMENT

A key requirement for MIMO control systems is the accuracy of the structural and parametric identification of control loops [9, 10]. According to the classical approach, the dynamics are described by a system of linear differential equations written as a matrix transfer function of dimensions $p \times q$, where p is the number of controlled variables and q is the total number of manipulated variables and observed disturbance variables. The model inputs are the manipulated variables of the process and the detected disturbance variables; at its output, the model predicts the behavior of the controlled variables of the process. In general, such a model has the form:

$$W(s) = \frac{A(s)}{B(s)} e^{-\tau s} = = \frac{a_n s^n + a_{n-1} s^{n-1} + \ldots + a_1 s + a_0}{b_m s^m + b_{m-1} s^{m-1} + \ldots + b_1 s + b_0} e^{-\tau s}, \quad (1)$$

where s denotes the Laplace operator; $A(s)$ and $B(s)$ are the polynomial representations of the numerator and denominator, respectively; n and m are the degrees of the polynomials $A(s)$ and $B(s)$, respectively. The physical implementability condition is given by $n \le m$.

This approach to describing the controlled object is accompanied by the degradation of the MIMO model (1) due to nonlinear physical and chemical processes

[11]. Therefore, control models for continuous industrial processes need to be periodically adjusted to the current technological mode [12]. A typical solution is to conduct an active experiment (step-by-step testing) to read the response characteristics of the process and update the MIMO model of the controlled object. A drawback of this approach is applying strong controls to a continuously operating object, which may destabilize the technological mode and bring the controlled variables of the process and key quality indicators of the products beyond the permissible limits.

2. DESCRIPTION OF THE UPDATING METHOD

Considering the range of problems mentioned, we develop an updating method for the dynamic MIMO model based on passive experiment data analysis. This method involves a nonlinear programming problem with an objective function formed by the least squares method according to technological object data:

$$\min_{k,a,b,\tau} \sum_{i=1}^{N} \left(y_i^e - y_i^{ap}(k, a, b, \tau) \right)^2 \to \to k^0, a^0, b^0, \tau^0, \qquad (2)$$

where N denotes the size of the initial data sample; y_i^e are the real values of controlled variables; $y_i^{ap}(k, a, b, \tau)$ are the values predicted by the model; k is the gain matrix of dimensions $1 \times q$; a is the coefficient matrix of the polynomial $A(s)$ of dimensions $n \times q$; b is the gain matrix of the polynomial $B(s)$ of dimensions $m \times q$; τ is the time delay matrix of dimensions $1 \times q$; finally, k^0, a^0, b^0, and τ^0 are the estimates of the transfer function parameters.

The predicted values are calculated using the imposition of the step response of the process:

$$y_i^{ap} = \sum_{j=1}^{i-1} \Delta x_j^{\mathrm{T}} h\left(\Delta T\left[i-j\right]\right) + y_1^e\,,\; i = \overline{2, N}\,,$$

where $\Delta x = x_j - x_{j-1}$ is the step excitation matrix of the input signals of dimensions $q \times 1$; $h(t)$ is the step responses of the control loops; ΔT is the sampling interval.

According to the residue theorem, the step response of the loop has the form

$$h(t) = \sum_{r=1}^{n} \mathrm{Res}_{s=s_r} \left[W(s_r) e^{s_r t} \right],$$

where s_r is a pole of the function $W(s)$.

In the general case of n simple nonzero poles (the transfer function has one zero pole and n simple poles), the step response can be calculated as

$$h(t) = \frac{B(0)}{A(0)} + \sum_{r=1}^{n} \frac{B(s_r) e^{s_r t}}{s_r A'(s_r)}\,, \qquad (3)$$

where $A'(s_r)$ denotes the derivative of the characteristic polynomial $A(s)$ calculated at the pole r [13].

For advanced process control, the dynamics of the control loops can be described with sufficient precision by a second-order differential equation with the transfer function [14]

$$W(s) = k \frac{a_1 s + 1}{b_2 s^2 + b_1 s + 1} e^{-\tau s}\,. \qquad (4)$$

For the dynamics (4), the step response formula (3) reduces to

$$h(t) = k \left(1 + \frac{(c_1 - a_1) e^{-\frac{t-\tau}{c1}} - (c_2 - a_1) e^{-\frac{t-\tau}{c_2}}}{c_2 - c_1} \right),$$

where $c_1 = -s_1^{-1}$, $c_2 = -s_2^{-1}$, and $s_{1,2}$ are the poles of the transfer function (4).

After the transformations of the objective function, the nonlinear programming problem (2) is written as

$$\min_{k,a_1,c_1,c_2,\tau} \sum_{i=2}^{N} \left(y_i^e - y_1^e - \sum_{j=1}^{i-1} \Delta x_j^{\mathrm{T}} h\left(\Delta T\left[i-j\right]\right) \right)^2 \to \to k^0, a_1^0, c_1^0, c_2^0, \tau^0 \qquad (5)$$

where

$$h_r\left(\Delta T\left[i-j\right]\right) = k_r \left(1 + \frac{(c_{1r} - a_{1r}) e^{-\frac{\Delta T(i-j)-\tau_r}{c_{1r}}} - (c_{2r} - a_{1r}) e^{-\frac{\Delta T(i-j)-\tau_r}{c_{2r}}}}{c_{2r} - c_{1r}} \right),$$

$$r = \overline{1, q}.$$

When necessary, we can easily pass from the poles c_1 and c_2 to the desired coefficients of the polynomial $B(s)$:

$$b_{1r} = c_{1r} + c_{2r}\,,\; b_{2r} = c_{1r} c_{2r}\,,\; r = \overline{1, q}\,.$$

The starting point for object identification can be the current parametric configuration of the MIMO model and a given set of parameters based on the appropriation of the passive experiment data.

Due to the amount of required calculations and the parametric dimension of the objective function, problem (5) should be solved using methods with a high convergence rate. In particular, the class of quasi-Newtonian methods with the Hessian calculated by *the*

Broyden–Fletcher–Goldfarb–Shanno algorithm [15] can be employed.

In practice, the method is implemented by adjusting the object model in automatic mode according to a given quantitative criterion or in manual mode following the request of the operating personnel.

3. APPLICATION OF THE UPDATING METHOD

The updating method for the MIMO model of controlled objects was tested on a real technological object: an industrial installation for butyraldehydes oxo synthesis to produce butyl alcohols. The installation has a continuous scheme: its key part is a hydroformylation unit to produce normalbutyraldehydes and isobutyraldehydes from propylene and synthesis gas in the presence of cobalt hydrocarbonyl in series reactors R-3 and R-4 with a catalyst from Cobalt 2- ethylhexanate and synthesis gas prepared in reactor R-1. The catalytic mixture is purified from fuel gases in high- and low-pressure separators (S-1 and S-2, respectively) and enters the oxidative decobaltization unit for the further processing of the key catalyzate components.

Butyl alcohol production has an APC system. A fragment of this system—the hydroformylation unit with the process parameters involved—is shown in the general process flow diagram of the installation (Fig. 1 and Table 1).

Table 1

Variables of APC system controller

Tag	Description
1. Controlled variables	
cv1	Propylene conversion
cv2	Propylene losses in blow-off
2. Manipulated variables	
mv1	Consumption of propane-propylene fraction
mv2	Temperature in reactor R-3
mv3	Temperature in reactor R-4
mv4	Gas pressure in high-pressure separator S-1
mv5	Consumption of catalytic mixture
mv6	Gas pressure in low-pressure separator S-2
3. Disturbance variables	
dv1	Carbon monoxide (CO) content in synthesis gas
dv2	Hydrogen concentration in synthesis gas

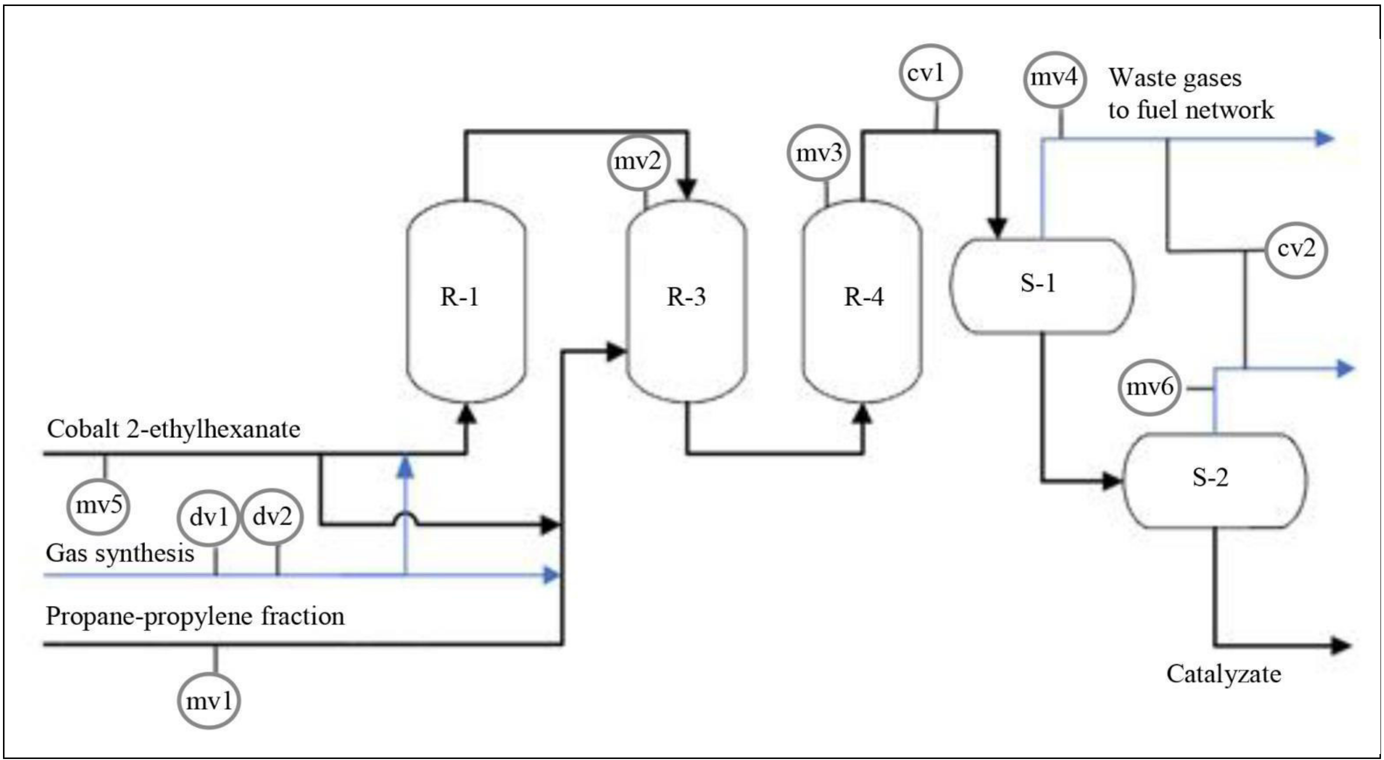


Fig. 1. Hydroformylation unit.

For this part of the APC controller of the hydroformylation unit, the developed method was applied to identify the MIMO model and update its parametric and structural configuration according to the current technological mode. Table 2 shows the results of updating the dynamic controller model.

The identification procedure was assessed by testing the model on a sample of historical industrial process data. Based on the results of model tests, the graphs of propylene conversion (Fig. 2) and propylene losses in blow-off (Fig. 3) predicted by the original and updated models on the same set of initial data were obtained.

Table 2

Updating the dynamic model of the hydroformylation unit

		Original model		Updated model	
		Output signal			
		cv1	cv2	cv1	cv2
Input signal	mv1	$0.0353\frac{1}{20s+1}$	$8.82\frac{1}{20s+1}$	$0.0353\frac{1}{20s+1}$	$12.36\frac{1}{18.93s+1}$
	mv2	$0.0678\frac{1}{15s+1}$	$-13\frac{0.51s+1}{3.89s^2+2.36s+1}$	$0.0348\frac{1}{18.82s+1}$	$-4.5\frac{1.41s+1}{3.76s^2+1.54s+1}$
	mv3	$0.0025\frac{1}{15s+1}e^{-4s}$	$-15.3\frac{1}{130s^2+22.8s+1}$	$0.0016\frac{1}{11.34s+1}e^{-6s}$	$-7.01\frac{0.21s+1}{130s^2+23s+1}$
	mv4	$0.007\frac{1}{5s+1}e^{-3s}$	–	$0.009\frac{1}{5s+1}e^{-3s}$	–
	mv5	$0.021\frac{1}{6.25s^2+5s+1}$	$-68.2\frac{1}{10.1s+1}e^{-6s}$	$0.021\frac{1}{6.25s^2+5s+1}$	$-68.2\frac{1}{10.1s+1}e^{-6s}$
	mv6	–	$-86.1\frac{1}{15s+1}$	–	$-83\frac{1}{15s+1}$
	dv1	$0.0021\frac{1}{12s+1}e^{-2s}$	$19.4\frac{1}{22.4s+1}e^{-24s}$	$0.0021\frac{1}{12s+1}e^{-2s}$	$19.19\frac{1}{22.4s+1}e^{-23s}$
	dv2	$0.01\frac{2.57s+1}{9.22s^2+6.72s+1}$	$0.659\frac{1}{12s+1}e^{-2s}$	$0.01\frac{2.57s+1}{9.22s^2+6.72s+1}$	$1.132\frac{1}{12s+1}e^{-2s}$

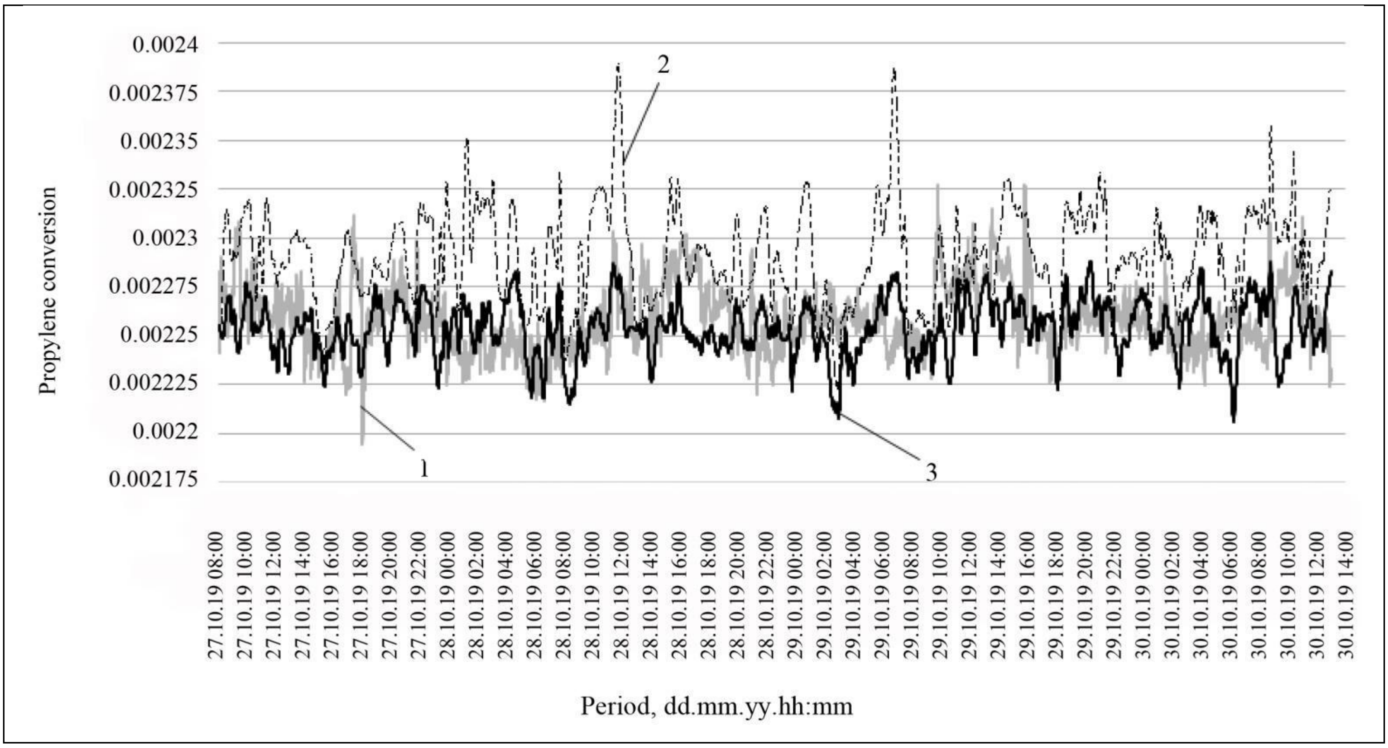


Fig. 2. Propylene conversion: (*1*) real values, (*2*) original model, and (*3*) updated model.

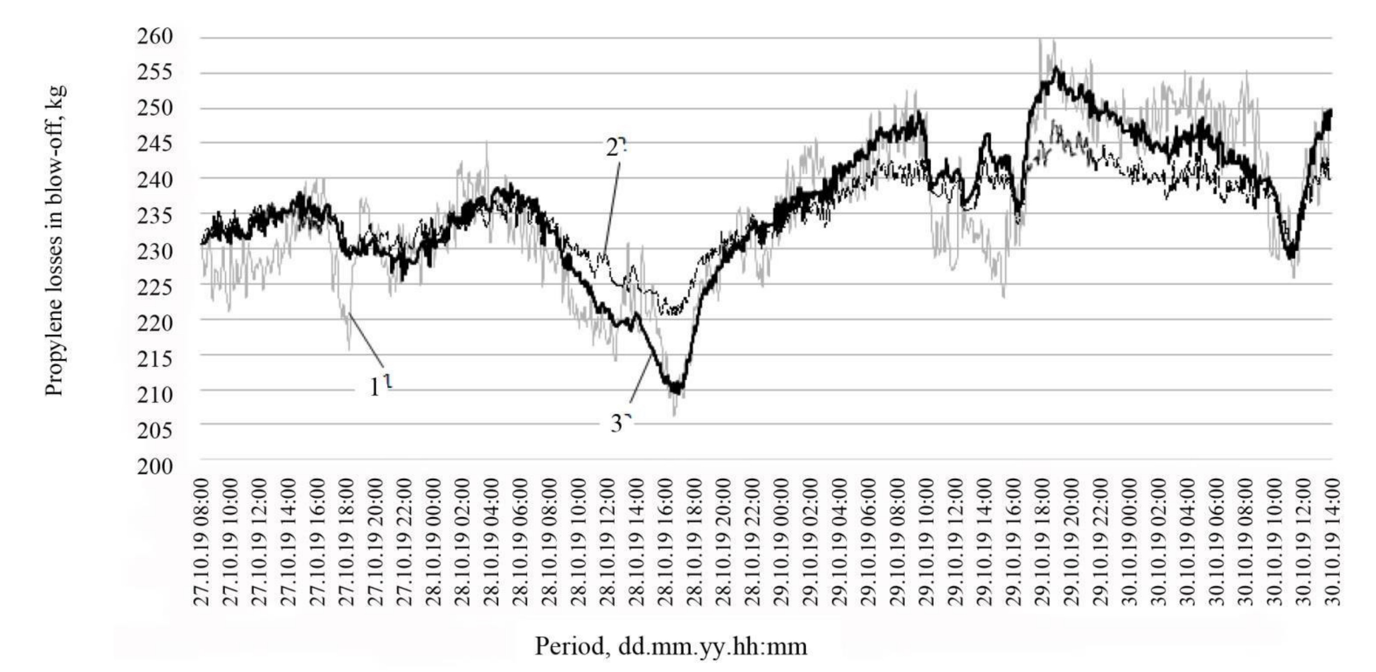


Fig. 3. Propylene losses in blow-off: (*1*) real values, (*2*) original model, and (*3*) updated model.

As the result of updating, the standard deviation of the predicted values from the real ones decreased from 4.159×10^{-5} to 2.045×10^{-5} (propylene conversion) and from 6.483 kg to 5.112 kg (propylene losses in blow-off).

CONCLUSIONS

This paper has presented a new method for updating the dynamic MIMO model of a controlled technological object. The method involves a nonlinear programming problem with an objective function formed using the least squares method. A distinctive feature of the developed method is the ability to model the transient processes of the control loops based on the previous structural and parametric identification. An advantage of the method is that there is no need to conduct an active experiment at the industrial installation: historical data of the technological process are used to update the model.

The developed method has been applied to the MIMO model of the APC system of the hydroformylation unit within an industrial installation for butyraldehydes oxo synthesis to produce butyl alcohols. As the result of updating, the standard deviation of the predicted values from the real ones has been reduced by factors of 2.033 (propylene conversion) and 1.268 (propylene losses in blow-off).

The new method improves the modeling accuracy of ongoing physical and chemical processes. In advanced process control, it enhances the performance of automatic control of an industrial process, stabilizing its key controlled variables and decreasing the dispersion of laboratory readings. Consequently, the overall economic efficiency of the production process is increased, and the human factor in an industrial process is reduced.

REFERENCES

1. Vladov, R.A., Dozortsev, V.M., Shaidullin R.A., et al., Practical Aspects of the Fourth Industrial Revolution, *Automation in Industry*, 2017, no. 7, pp. 7–13. (In Russian.)
2. Kueppers, S. and Haider, M., Process Analytical Chemistry – Future Trends in Industry, *Analytical and Bioanalytical Chemistry*, 2003, vol. 376, pp. 313–315.
3. Bakhtadze, N.N. and Lototskii, V.A., Contemporary Methods of Production Process Control, *Control Sciences*, 2009, no. 3, pp. 56–63. (In Russian.)
4. Seleznev, A.Yu., Possible Applications to Improve the Operational Efficiency of Oil Refineries, *Neft'. Gaz. Novatsii*, 2020, no. 3, pp. 35–42. (In Russian.)
5. Lehman, K.A., Implement Advanced Process Control, *Chemical Engineering Progress*, 2018, vol. 114, no. 1, pp. 60–66.
6. Dombrovskii, V.V. and Pashinskaya, T.Yu., Model Predictive Control for Nonlinear Stochastic Systems with Serially Correlated Parameters under Constraints, *Tomsk State University Journal of Control and Computer Science*, 2018, no. 42, p. 4–11. (In Russian.)
7. Torgashov, A.Yu., Goncharov, A.A., and Samotylova, S.A., Modern Methods of Evaluation of Advanced Process Control Systems, *Vestnik of the Far East Branch of the Russian Academy of Sciences*, 2016, no. 4 (133), pp. 102–107. (In Russian.)

8. Rabotnikov, M.A., Vialykh, I.A., and Nemtin, A.M., Technical Design Advanced Control System of the Catalytic Reforming Unit, *Bulletin of the Perm National Research Polytechnic University. Electrotechnics, Information Technologies, Control Systems*, 2019, no. 2, pp. 119–132. (In Russian.)
9. Berlik, S. and Nasiri, M., An Adaptive Fuzzy Model Predictive Control System for the Textile Fiber Industry, in *Communications in Computer and Information Science*, 2010, vol. 81, pp. 729–736.
10. Pan, J., Jiang, X., Wan, X.K., et al., A Filtering Based Multi-Innovation Extended Stochastic Gradient Algorithm for Multivariable Control System, *International Journal of Control Automation and Systems*, 2017, vol. 15, no. 3, pp. 1189–1197.
11. Cannon, M. and Kouvaritakis, B., Continuous-Time Predictive Control of Constrained Nonlinear Systems, in *Progress in Systems and Control Theory*, 2000, vol. 26, pp. 205–215.
12. Le Lann, M.V., Cabassud, M., and Casamatta, G., Adaptive Model Predictive Control, in *NATO ASI Series (Series E: Applied Sciences)*, 1995, vol. 293, pp. 427–457.
13. Konovalov, B.I. and Lebedev, M.Yu., *Teoriya avtomaticheskogo upravleniya* (Theory of Automatic Control), Tomsk: TUSUR, 2011. (In Russian.)
14. Tyatyushkin, A.I., Zarodnyuk, T.S., and Gornov, A.Yu., Algorithms for Nonlinear Optimal Control Problems Based on the First and Second Order Necessary Conditions, *Journal of Mathematical Sciences*, 2019, vol. 239, pp. 185–196.
15. Zakharova, E.M. and Minashina, I.K., Review of Multidimensional Optimization Techniques, *Information Processes*, 2019, vol. 14, no. 3, pp. 256–274. (In Russian.)

This paper was recommended for publication by N.N. Bakhtadze, a member of the Editorial Board.

Received January 11, 2021, and revised February 9, 2021.
Accepted March 10, 2021.

Author information

Rabotnikov, Mikhail Alekseevich. OOO ZapSibNeftekhim, Tobolsk, ✉ rabotnikovma@tobolsk.sibur.ru

Cite this article

Rabotnikov, M.A. An Updating Method for the Dynamic MIMO Model of a Controlled Technological Object. *Control Sciences* **3,** 51–56 (2021). http://doi.org/10.25728/cs.2021.3.7

DOI: http://doi.org/10.25728/cs.2021.3.8

A COMPUTER SIMULATION COMPLEX FOR ANALYSIS OF MAGNETIC GRADIOMETRY SYSTEMS

A.K. Volkovitskiy[1], A.I. Gladyshev[2], D.A. Goldin[3], E.V. Karshakov[4], B.V. Pavlov[5], and M.Yu. Tkhorenko[6]

[1, 3–5]Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia
[2]Chairman, the Defense Problems Section of the Ministry of Defense under the RAS Presidium, Moscow

[1]✉avolkovitsky@yandex.ru, [2]✉ tolyagladyshev@yandex.ru, [3]✉ goldind@ipu.ru, [4]✉ karshak@mail.ru,
[5]✉ pavlov@ipu.ru, [6]✉ tkhorenkom@mail.ru

Abstract. A computer simulation complex for magnetic gradiometry systems is described. This complex simulates the estimation procedure for the magnetic dipole moment of a moving object according to the magnetic gradiometry data. The paper considers the software architecture and intended purpose of the complex and the algorithms of its modules, including the magnetic field module, the ambient magnetic field module (the object's magnetic field, the main magnetic field, the magnetic anomaly, and industrial magnetic noise), and the magnetic dipole moment module. Some numerical experiments with the simulation complex are briefly described. This complex can be used to design degaussing systems for the magnetic field of moving objects.

Keywords: numerical modeling, magnetic gradiometry, degaussing.

INTRODUCTION

Several application-relevant problems require active control of the magnetic field of moving objects. Among them, we mention the following problems: (1) suppressing the effect of the carrier vehicle's field on the readings of the onboard magnetic measuring equipment during aeromagnetic survey [1] and (2) reducing the underwater vehicle's magnetic field to secure against magnetic detection means [2]. The first problem can be solved analytically when processing the received measurement information, whereas the second problem requires a special onboard degaussing system for the moving object. This system represents a set of coils with an applied current to compensate, to a certain extent, the moving object's magnetic field. Obviously, the performance of such a degaussing system strongly depends on the choice of control laws for coils currents. The paper [3] proposed an original control scheme for compensating currents based on estimating the parameters of the magnetic dipole moment (MDM) of a moving object using real-time magnetic gradiometry data from a special measuring system; the resulting MDM estimates were considered when controlling the degaussing device.

The magnetic gradiometry system can be designed in a compact form, ranging from 20 cm to 2 m, depending on the selected magnetically sensitive elements. Hence, it can be implemented on a stationary bench (like all standard means for determining the MDM currently used) and in a mobile or towed configuration. Adopting such systems, which estimate the MDM in real time during motion and can be used to control coils currents, will significantly improve the quality and reliability of degaussing of a moving object (MO). The advantages seem obvious: a closed loop control system with an MDM value feedback control is used instead of compensating the MO's magnetic field based on the static bench data. This approach considers a slow change in MO's magnetization and its magnetization reversal due to the movements of ferromagnetic masses in the Earth's magnetic field.

To assess the effectiveness of this compensation scheme when controlling the MO's magnetic field, we developed a computer simulation complex for the magnetic gradiometry system, considering both the motion features of the MO and the measuring unit and various natural and artificial noises arising when measuring the magnetic field gradient. Thus, the complex is applied to analyze the operation of the MO's magnetic field control system. This paper describes the complex in detail.

1. STRUCTURE OF COMPUTER SIMULATION COMPLEX

The computer simulation complex for the magnetic gradiometry system is a set of software modules united in an information network that simulates the interaction of ferromagnetic objects with the Earth's magnetic field (EMF), the operation of various-type magnetic gradiometers under the motion of the object and the measuring unit, and remote determination algorithms for the MDM parameters. Simulations can be performed in real time.

Information interaction in the complex is carried out in the client-server mode. In this scheme, the central element (server) is a special software module that controls and coordinates the entire complex and writes model data to files for subsequent analysis.

In addition to the server, the complex has service modules (clients) of two fundamentally different types:

– modules implementing mathematical modeling algorithms for virtual objects, systems, and processes (virtual devices);

– user interface modules visualizing the necessary parameters and allowing the operator to manage the virtual experiment parameters in real time (indicators).

The virtual device modules include a module for determining the navigational parameters of the moving object and measuring unit, a module for calculating the magnetic field and its gradient, and a module for estimating the MDM parameters.

The composition of the indicator modules depends on the goals of the virtual experiment and can be selected individually by the operator.

The simulation mode is controlled by the system timer. Each virtual device operates independently, notifying the server about the arrival of a new data portion. In turn, the server notifies the corresponding indicators about the possibility of visualizing new data and notifies the virtual devices about the corresponding control commands received from the operator. The information generated by each virtual device is saved and processed in its data storage area; the server transfers this information to the indicator modules. During visualization, indicators and virtual devices can have a unique relationship: an indicator reflects the virtual behavior of a specific virtual device. In the general case, the data of all virtual devices are available for single visualization control. Virtual devices are always managed individually.

2. MODULE FOR DETERMINING NAVIGATIONAL PARAMETERS

This module is intended for generating data on the position, speed, and orientation of the MO and measuring unit at different time instants. Two sets of motion parameters have to be determined (for the MO and measuring unit). Therefore, the module for determining the navigational parameters consists of two parts, each generating parameters for one object (the MO or measuring unit). The operator can choose between independent and dependent laws of motion of the MO and measuring unit. The latter case is implemented if the MDM parameters are determined using a magnetic gradiometer towed behind the MO, as described in [3]. On the one hand, the simulation results can be reliably analyzed only if the navigation plan (scenario) for each computational experiment is exactly specified and strictly followed. On the other hand, the maximum degree of correspondence of the model to real physical processes is needed; particularly, the virtual MO should be controlled so that its motion matches that of the real one.

To meet these requirements, we developed a module for determining the navigational parameters in which the virtual motion is controlled according to the navigational plan (task). Moreover, motion control is largely similar to that of the ship's crew (the navigator and helmsman).

The navigational plan is a sequence of given path lines, each specified by a set of points with known geographic coordinates (latitude, longitude, and altitude), traversed by the moving object with a given linear velocity. The navigational plan is described by a structured sequence of points and saved as a text file. The navigational plan is prepared for each computational experiment separately and loaded before its start. At the same time, the operator can interactively modify the navigational plan after loading using special indicators.

In the simulation complex, the virtual object is moving along the trajectories specified by the navigational plan in two modes: manual (the operator controls the object's motion from the keyboard) and automatic (the program imitates the operator's actions).

For the convenience of virtual motion control, auxiliary control information is formed as a hint for the operator (manual mode) or a guideline for the virtual autopilot (automatic mode). This control information is a scalar course deviation (correction) signal for both the operator and the autopilot. For the successful execution of the navigational task, the operator (autopilot) should maintain a motion mode in which the correction parameter is close to 0. If the object needs to deviate to the right (left) to follow the specified trajectory, then the correction value will take a positive value (a negative value, respectively). The orientation is calculated based on the parameters of the MO' s trajectory considering the sinusoidal disturbance caused by the natural waves of the marine environment. The algorithm for calculating the navigational parameters was described in detail in [4, 5].

3. MAGNETIC FIELD MODULES

Assume that the magnetic field near the moving object is induced by the following sources:

– the constant component of the MO's magnetic field caused by the "hard" (constant) magnetization of the casing materials and the operation of the onboard electrical equipment;

– the variable component of the MO' magnetic field caused by the "soft" (variable) magnetization of the casing;

– the main magnetic field;

– the short-term magnetic variation;

– the magnetic anomaly;

– industrial magnetic noise (when the magnetic measuring system is located on the coast).

The resulting magnetic field is the sum of the fields from these sources. Consider the process of calculating each of these components in detail.

When calculating the constant component of the MO's magnetic field, we assume that the MI's casing contains virtual sources of constant (hard) magnetization described by a system of local dipole transmitters in a coordinate system rigidly connected with the MO's casing. The model parameters are their magnetic dipole moments M_i and the centers' coordinates r_i. The total field has the following formula [6]:

$$B_p = \sum_i \frac{\mu_0}{4\pi} \frac{3\left(M_i^{\mathrm{T}} R_i\right) R_i - M_i R_i^2}{R_i^5}, R_i = r_i - r, \quad (1)$$

where r denotes the radius vector of the observation point. (All vectors in the paper are three-dimensional columns.)

To determine the magnetization of the ellipsoidal casing, we emphasize that the magnetic field induced by the shell can be written as a gradient of some scalar potential [7]:

$$B_s = -\nabla u. \quad (3)$$

It can be shown that the scalar function u introduced in this way satisfies the Laplace equation both inside the casing and outside it. Moreover, the boundary conditions for Maxwell's equations hold on the casing surfaces. Hence, the solution functions u_1 (outside the casing), u_2 (in the casing walls) and u_3 (inside the casing) are given by [8]:

$$\begin{aligned} u_1 &= \sum_{n=0}^{\infty} \sum_{m=-n}^{n} D_{nm} Q_n^m(\eta) Y_n^m(\theta,\phi), \\ u_2 &= \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \left(B_{nm} Q_n^m(\eta) + C_{nm} P_n^m(\eta)\right) \times \\ &\times Y_n^m(\theta,\phi), \\ u_3 &= \sum_{n=0}^{\infty} \sum_{m=-n}^{n} A_{nm} P_n^m(\eta) Y_n^m(\theta,\phi), \end{aligned} \quad (4)$$

where n is the maximum order of the expansion; η, θ, and ϕ are the prolate spheroidal coordinates; $P_n^m(\eta)$ and $Q_n^m(\eta)$ are the associated Legendre functions of the first and second kinds, respectively; finally, $Y_n^m(\theta,\phi)$ are the spherical functions [9]. The constants A_{nm}, B_{nm}, C_{nm}, and D_{nm} are obtained from the system of linear equations

$$\begin{aligned} &B_{nm} Q_n^m(\eta) + \left(C_{nm} - A_{nm}\right) P_n^m(\eta) = 0, \eta = \eta_2, \\ &\left(B_{nm} - D_{nm}\right) Q_n^m(\eta) + C_{nm} P_n^m(\eta) = u_{nm}, \eta = \eta_1, \\ &\mu B_{nm} Q'^m_n(\eta) + \left(\mu C_{nm} - \mu_0 A_{nm}\right) P'^m_n(\eta) = 0, \eta = \eta_2, \\ &\left(\mu B_{nm} - \mu_0 D_{nm}\right) Q'^m_n(\eta) + \mu C_{nm} P'^m_n(\eta) = \\ &= \mu_0 w_{nm}, \eta = \eta_1, \end{aligned} \quad (5)$$

where η_1 and η_2 are the geometrical parameters determining the outer and inner ellipsoidal surfaces, and the coefficients u_{nm} and w_{nm} are given by

$$\begin{aligned} u_{nm} &= \frac{1}{\| Y_n^m \|^2} \int_0^{2\pi} d\phi \int_0^{\pi} d\theta sin\theta u_e(\eta,\theta,\phi) Y_n^m(\theta,\phi), \\ w_{nm} &= \frac{1}{\| Y_n^m \|^2} \int_0^{2\pi} d\phi \int_0^{\pi} d\theta sin\theta \frac{\partial u_e(\eta,\theta,\phi)}{\partial\eta} Y_n^m(\theta,\phi), \end{aligned} \quad (6)$$

where u_e is the scalar potential of external sources (the EMF, the magnetic field of the system of constant dipoles). In the simulation complex, the integrals (6) can be calculated in two ways: using the Gauss–Legendre quadrature method [10] and a generalization of the sampling theorem [11]. With the coefficients A_{nm}, B_{nm}, C_{nm}, and D_{nm} obtained from (5), the scalar potential and the magnetic field can be calculated by formulas (4) and (3). Thus, the magnetic field of the casing can be expanded into a series with respect to the spherical functions: the input parameters are the geometric dimensions of the casing and the magnetic permeability of its material, and the expansion coefficients depend on the EMF. They are recalculated at each step of the algorithm.

In this complex, the main magnetic field is calculated using the IGRF 13 model [12]. Similar to the casing field, the main magnetic field is written as the gradient of the scalar potential given by

$$u = a\sum_{n=1}^{N}\sum_{m=0}^{n}\left(\frac{a}{r}\right)^{n+1}(g_{mn}\cos m\phi + h_{mn}\sin m\phi)\times$$
$$\times P_n^m(\cos\theta), \qquad (7)$$
$$g_{mn} = g_{mn}^0 + \dot{g}_{mn}(t - t_0),$$
$$h_{mn} = h_{mn}^0 + \dot{h}_{mn}(t - t_0),$$

where $r, \theta,$ and ϕ are the spherical coordinates of the observation point, and t denotes time. The components of the magnetic induction vector are calculated by formula (3). The constants $a, t_0, g_{mn}^0, \dot{g}_{mn}, h_{mn},$ and $\dot{h}_{mn}$ are known and loaded from the file before the operation starts.

The short-term magnetic variation is written as the sum of sinusoidal components

$$B_v = \sum_i \left(C_i \cos\omega_i t + S_i \sin\omega_i t\right). \qquad (8)$$

The values $C_i, S_i,$ and ω_i are input parameters specified in the settings file.

The magnetic anomaly is described by a synthetic model built as the potential of dipoles randomly distributed in a square of 10×10 km, with a periodically repeating palette. During the simulation, the depth of the dipoles' layer is set to 1 km; the range of the maximum dipole moment amplitude, to ± 25 MA · m^2. The anomaly field does not reflect the true structure of the EMF of the indicated area, just simulating the characteristic magnetic disturbances and inhomogeneities in the magnetization of geological structures. Nevertheless, the model has an adequate behavior: as the height increases (deep water), the amplitude of anomalies decreases, and the characteristic period increases. The total field of anomalies is calculated using a formula similar to (1). The correspondence of the magnetic anomaly model to the true magnetic field distribution is not significant for the algorithm for determining the MDM: the very presence or absence of an anomalous component affecting the operation of this algorithm is essential. The synthetic magnetic anomaly model allows estimating this effect in the simulation mode and establishing the conditions under which the remote determination of the MDM during sailing can be considered correct. The model does not evolve, being stable over time. The model parameters are "hardly" embedded into the program code of the only available function. Calling this function generates the parameters of the model field.

Finally, the industrial noise is simulated using the field of a system of infinitely thin wires with a sinusoidal current of industrial frequency (50 Hz or 60 Hz) applied. The magnetic field of a wire with an applied current has a simple analytical expression [6]: the total wire field is simply the vector sum of the individual wire fields. The input data for simulation are the geometric parameters of the system and the frequency and amplitude of the wires currents.

4. MAGNETIC DIPOLE MOMENT MODULE

This module simulates the process of magnetic measurements and determines the MDM by these measurements. First, we describe the simulation of measurements.

The complex simulates the operation of the measuring unit with a spaced layout of the sensors (Fig. 1). The virtual unit represents a regular tetrahedral pyramid with magnetic sensors *S*1–*S*4. In addition, the unit includes blocks of an orientation system (indicated by GIS) to determine the spatial and angular position.

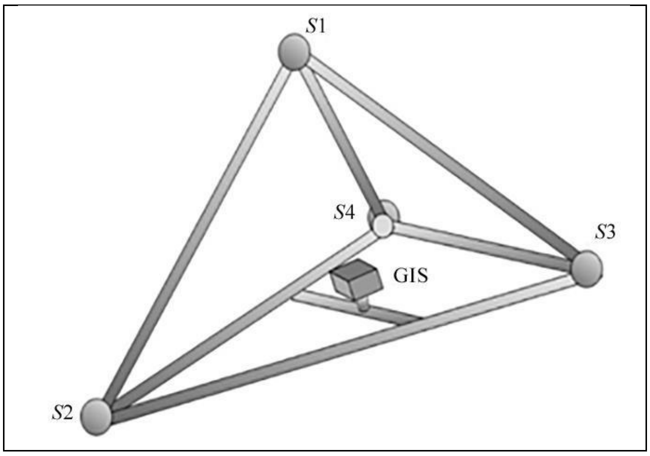


Fig. 1. Magnetosensitive measuring unit on spaced sensors.

The complex simulates the operation of the measuring unit using magnetic sensors of three different types:

– scalar sensors (quantum sensors) [13],

– vector three-component sensors of high sensitivity and accuracy (superconducting quantum interference devices, SQUIDs) [14],

– vector three-component sensors of low sensitivity and accuracy (fluxgate magnetometers) [13].

The first type corresponds to a vector magnetic gradiometer; the second and third ones, to a tensor magnetic gradiometer.

The operation of magnetosensitive elements of all three types is simulated by the same scheme. At each cycle of the system time, the following actions are performed:

- The model (“true”) parameters of the position—coordinates and orientation—are calculated for each sensor of the virtual magnetic measuring unit using the procedures of the navigation module of the measuring unit.
- The new (“distorted”) values of the spatial coordinates of the measuring units are calculated. (The deformation of the magnetic gradiometer unit’s platform is simulated.)
- The components of the model (“true”) magnetic field vector are calculated for the “distorted” position parameters in the coordinate system associated with each sensor.
- The parameters of the “true” field vector are virtually measured in accordance with the model parameters of the sensor’s measuring properties: the parameters of the new “measured” field vector, distorted relative to the “true” one, are calculated. The distortions result from the systematic zero-drift errors, deviations from the scale factor from 1 (for each component of the vector), the non-orthogonality of the sensitive axes, and measurement noises.
- If the scenario assumes simulating the measuring unit using scalar magnetosensitive elements, then for each sensor, the “measured” vector parameters are transformed to scalars by taking the absolute value.
- Depending on the type of sensors, the type of the magnetic gradiometer unit is automatically determined: the unit will “measure” the gradient vector of the field’s absolute value in the case of scalar sensors and the tensor of the second derivative of the field potential in the case of vector sensors.
- The orientation angles of the “true” position parameters block are also measured virtually: the “measured” parameters of the angular orientation vector are calculated. They differ from the “true” ones due to the systematic distortions under the specified accuracy parameters of the virtual orientation system of the measuring unit.

The calculated “measured” parameters are input data for the algorithm for determining the MDM parameters.

Note that this module may operate in the calibration and measurement modes. The former mode activates the algorithm for calculating the corrections for the virtual measurement data. The corrections are introduced by “measuring” the alternating field induced by the virtual reference source. The difference between the measurement and calibration modes is the absence of the reference field: during virtual measurements, the reference field source is switched off.

Let us briefly describe the algorithm for estimating the MDM parameters by magnetic field measurements. For the sake of definiteness, consider a tensor measuring unit. Assume that a single dipole is located at the origin of the coordinate system. Then formula (1) yields

$$B = \frac{\mu_0}{4\pi r^5}\left(3rr^{\mathrm{T}} - r^2 I\right)M. \tag{9}$$

Here M is the magnetic dipole moment; I denotes an identity matrix of dimensions 3 × 3; rr^{T} is the outer product of the vectors. We will determine r and M by the known values of the left-hand side of (9) at several measurement points and the distance between them. For each of 4 vector field sensors, we calculate

$$B_i = \frac{\mu\mu_0}{4\pi\left|\tilde{r} + \delta r_i\right|^3}\left(3\frac{\left(\tilde{r} + \delta r_i\right)\left(\tilde{r} + \delta r_i\right)^{\mathrm{T}}}{\left(\tilde{r} + \delta r_i\right)^T\left(\tilde{r} + \delta r_i\right)} - I\right)\tilde{M}. \tag{10}$$

Here $\tilde{M}$ and $\tilde{r}$ are the estimates of M and r; δr_i are the known displacement vectors of the vector magnetic sensors relative to the system’s measuring center; B_i is the calculated value of the field vector under the current hypothesis for sensor i.

Since the field vector measurements are subject to variations, we should consider the differences in the field components. The magnetic field variations are spatially homogeneous enough to neglect their spatial variations at distances of up to several kilometers.

We introduce the vector of estimated parameters

$$x = \left(r_1, r_2, r_3, m_1, m_2, m_3\right),$$

where r_i are the components of the corrections vector for some prior estimate of the radius vector r of the dipole location point; m_i are the components of the corrections vector for some prior estimate of the dipole moment vector M:

$$X = X_0 + x, X_0 = \left(\tilde{r}_1, \tilde{r}_2, \tilde{r}_3, \tilde{M}_1, \tilde{M}_2, \tilde{M}_3\right).$$

Next, we introduce the vector of the measured componentwise differences between the vector sensors readings:

$$z = \left(z_{121}, z_{122}, z_{123}, z_{231}, z_{232}, z_{233}, z_{241}, z_{242}, z_{243}\right).$$

Here the first subscript corresponds to the number of the “minuend” sensor; the second subscript, to the number of the “subtrahend” sensor; the third subscript, to the number of the sensitivity axis along which the difference is measured. Despite that 18 such differences are measured for 4 sensors in total, only 9 of them can be considered independent: any other component of the differences can be expressed through the parameters of the vector z. Now we introduce a vector containing the parameters of the gradient (the differences of the field components):

$$G = \left(G_{121}, G_{122}, G_{123}, G_{231}, G_{232}, G_{233}, G_{241}, G_{242}, G_{243}\right),$$
$$G_{ijk} = B_{ik} - B_{jk}.$$

Note that G_{ijk} are expressed through r and M using formula (10). Thus, the problem is finding r and M satisfying the relations

$$G_{ijk} = f_{ijk}\left(r, M\right),$$

where f_{ijk} are nonlinear functions of r and M, expressed through formula (10). This class of problems can be solved using a nonlinear generalization of the Kalman filter—the Iterated Extended Kalman Filter (IEKF) [15–18].

A similar approach applies to determining the MDM parameters by the gradient vector measurements. The only difference is that the measurements equation will include not the differences between the field vectors components but the differences between their absolute values. However, in the configuration of the measuring system with 4 sensors, a vector gradiometer yields 3 independent measurements: we can determine only the MDM parameters but not the MDM radius vector.

Numerical experiments show that under moderate errors in the prior estimates of r and M and measurement noises, this algorithm always converges to the true values of the MDM parameters. Its domain of convergence is presented in Fig. 2. The corresponding initial conditions are from a cube of dimensions 40 × 40 × 40 m (from –20 to +20 for each coordinate), and the dipole position is (5, 0, 0.306). The dipole moment is equal to 100 Am^2, and the measurement noise of the field components is nT. The shade of gray indicates the number of iterations for reaching the solution point. Due to the gradiometer's design, the plane passing through its center (with the normal directed to the dipole) is the boundary of this domain; see Fig. 2.

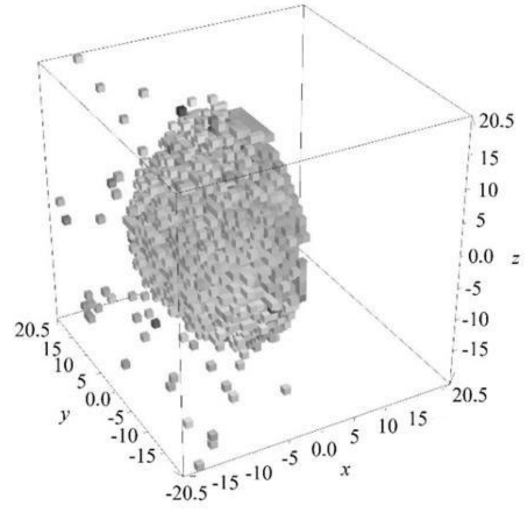


Fig. 2. The domain of convergence for iterative algorithm.

5. EXAMPLES OF NUMERICAL EXPERIMENTS WITH THE SIMULATION COMPLEX

To demonstrate the effectiveness of the algorithms for determining the MDM by the magnetic field gradient parameters at the detector's location, we carried out several numerical experiments with the simulation complex. The results of some experiments are presented below. To control the accuracy of determining the MDM by the detector, we applied the following approach: given the parameters of the object's model, we calculated the asymptotic behavior of the magnetic field at infinity, corresponding to the field of a point magnetic dipole (decreasing no faster than $1/r^3$). Note that the main magnetic field, the magnetic anomaly, and industrial noise were not included in the model field expression.

The first experiment simulated the simplest possible situation: the object's motion with a detector towed behind it on a flexible cable along a rectilinear route. The field of a point magnetic dipole was the object's field, and the fields of magnetic anomaly and industrial noise were neglected. The results of this experiment are shown in Figs. 3–5. (All values on the right scale of Fig. 5 are multiplied by 10^6.)

Clearly, the algorithm for calculating the MDM converges rather quickly to the true dipole moment value (the coefficient at the expansion term decreasing with a rate of $1/r^3$).

The second experiment simulated the estimation procedure of the MDM considering the sensors' noises and the inaccurately known geometry of the measuring unit. The results are shown in Figs. 6–8. (All values on the right scale of Fig. 8 are multiplied by 10^6.) This experiment was carried out with the following mean-squared errors of the geometric parameters: 10 cm (the distance between the object and detector), 2 mm (the distance between the individual sensors in the detector, with a sensor spacing of 2 m), and 10^{-4} rad (the orientation of the sensitivity axes of the magnetic measuring system). Quite expectedly, the estimation quality decreased compared to the first experiment.

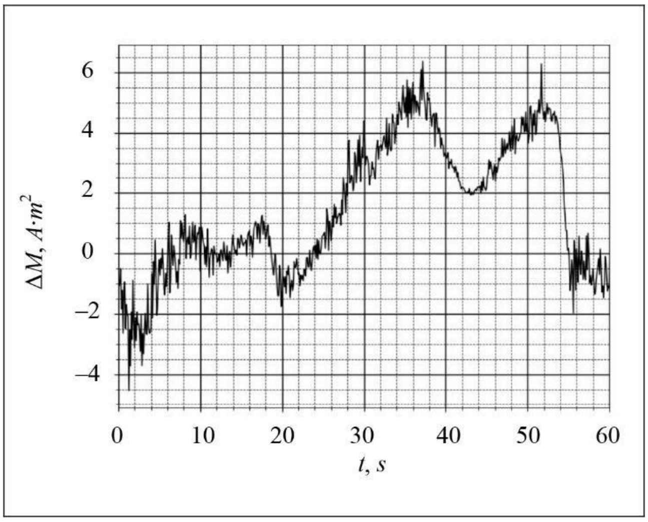


Fig. 3. Estimation error for the absolute value of MDM.

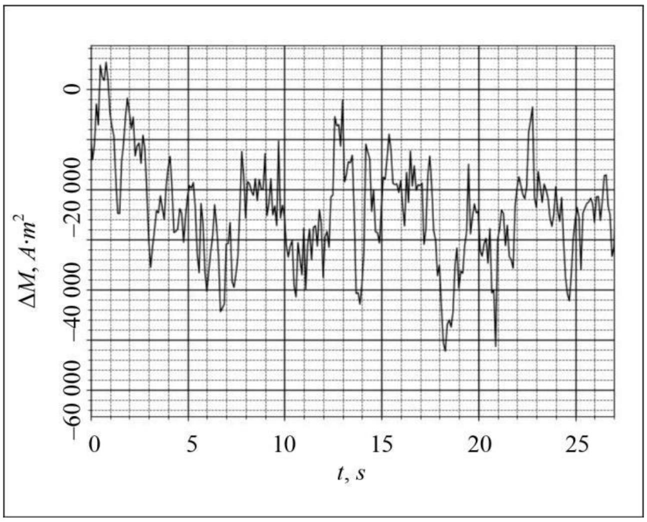


Fig. 6. Estimation error for the absolute value of MDM.

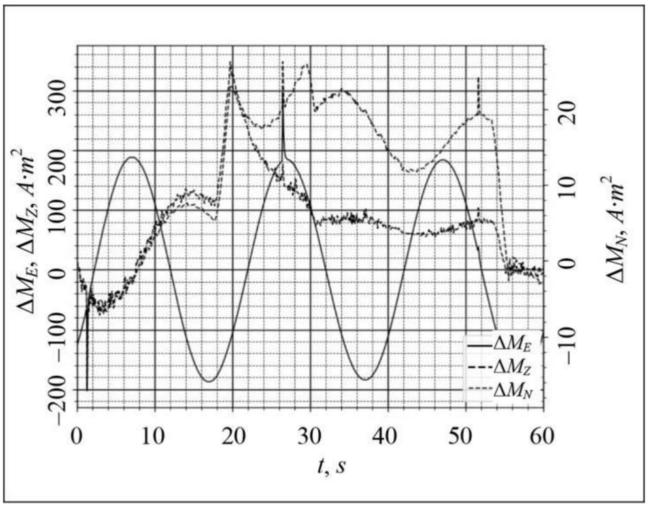


Fig. 4. Estimation errors for MDM components.

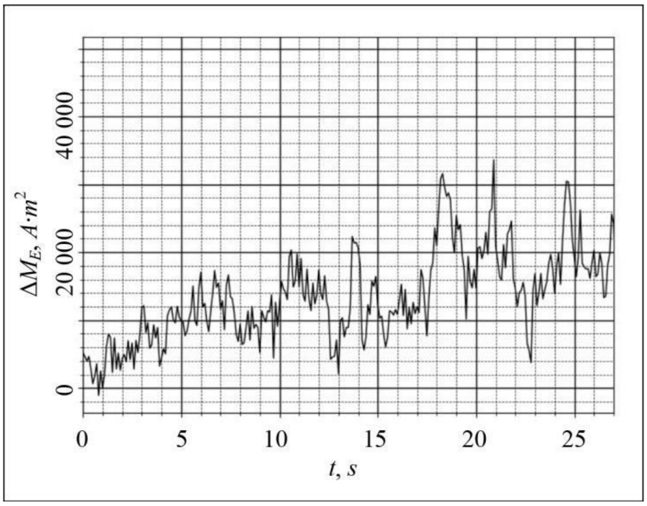


Fig. 7. Estimation error for MDM component.

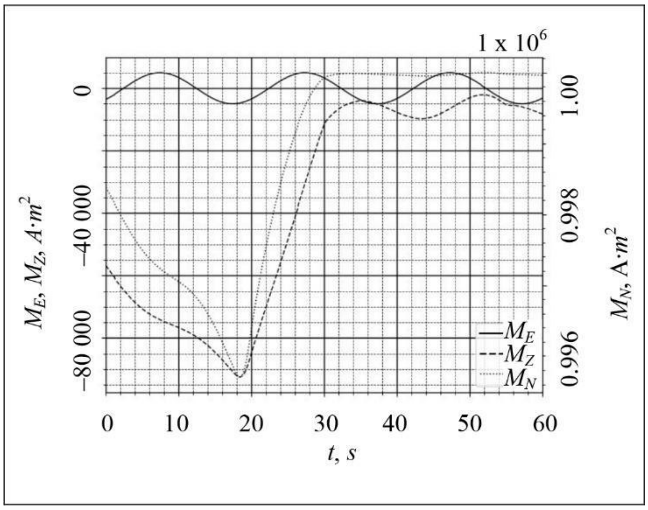


Fig. 5. MDM components.

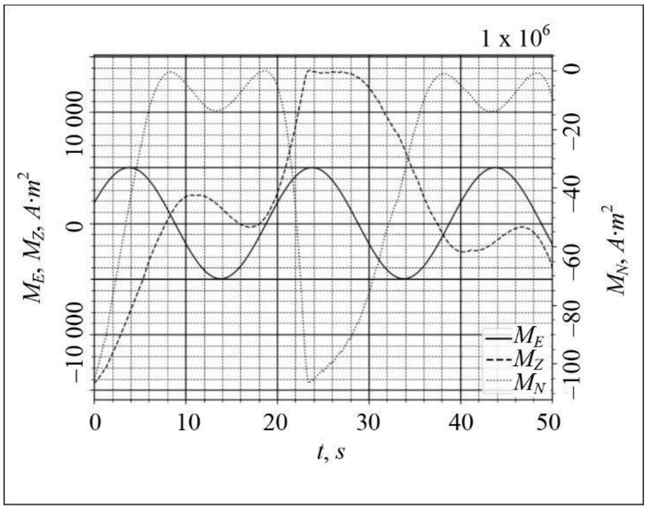


Fig. 8. MDM components.

In the third experiment, we calculated the object's magnetic field as the sum of the fields induced by a system of point dipoles and a homogeneous casing. The results of this experiment are shown in Figs. 9–11. (All values on the right and left scales of Fig. 11 are multiplied by 10^6.) The quality of estimating the MDM is worse than in the first experiment. This fact is explained both by the algorithm's peculiarities (it is based on the point dipole field formula and determines the MDM of more complex systems only approximately) and by the presence of the simulated measurement noises. Moreover, it was assumed that the object was moving along a curved route with maneuvering. Note that the estimation errors in the third experiment are not much worse than in the second experiment. Therefore, the algorithms under consideration are applicable for estimating the MDM of the objects with a rather complex structure of their magnetic field.

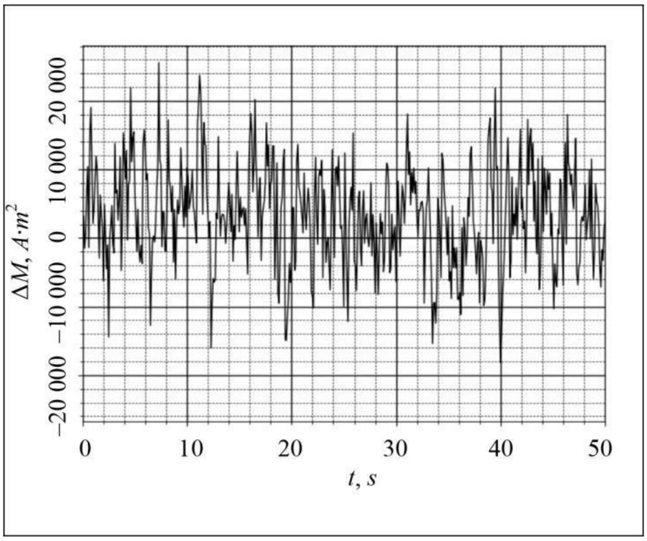


Fig. 9. Estimation error for the absolute value of MDM.

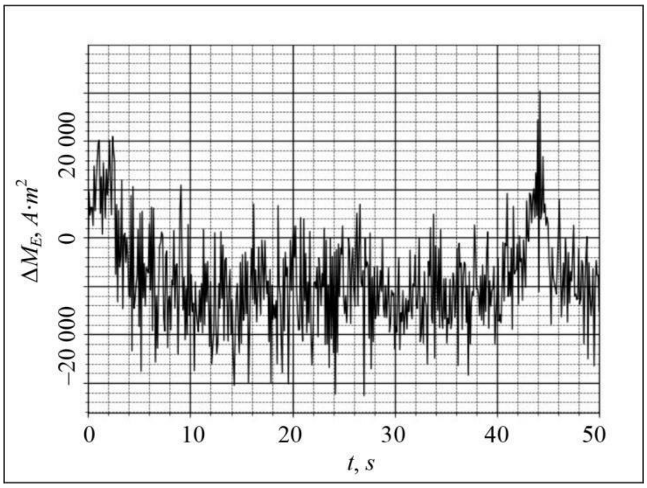


Fig. 10. Estimation error for MDM component.

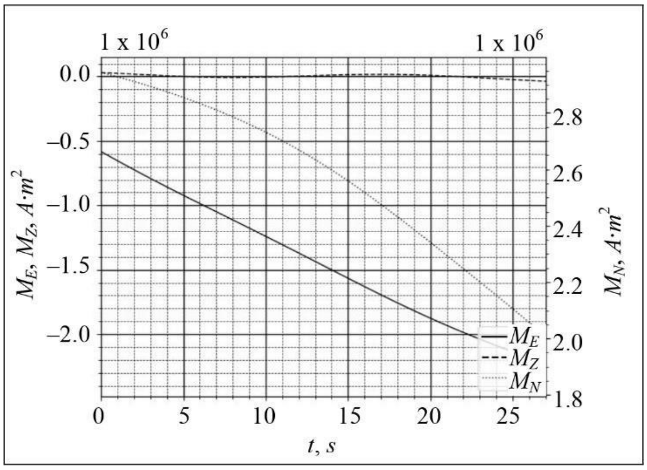


Fig. 11. MDM components.

CONCLUSIONS

This paper has considered a computer complex developed by the authors to simulate the operation of a magnetic gradiometry system in the process of estimating the MO's MDM. The general structure and principles of the complex and the purpose of individual service modules have been described in detail. In addition, three algorithms have been presented: an algorithm for simulating the MO's magnetic field that includes two components (the constant field and the field of an ellipsoidal casing), an algorithm for simulating the main magnetic field and various noises (the industrial noises, the magnetic anomaly, and the short-term magnetic variation), and a nonlinear algorithm for estimating the MDM parameters by measurements of the magnetic field vector components at several points with fixed spacing. Finally, several numerical experiments with the simulation complex have been provided.

The complex estimates the MO's MDM in the stationary, mobile, and towed configurations. The complex can be used to assess the effectiveness of implementing the described MDM estimation systems in real time during motion and test coil current control during the degaussing of the MO, which may significantly improve its quality and reliability.

REFERENCES

1. Noriega, G., Aeromagnetic Compensation in Gradiometry – Performance, Model Stability, and Robustness, *IEEE Geoscience and Remote Sensing Letters*, 2015, vol. 12, iss. 1, pp. 117–121.
2. Bayens, T.M., Analysis of the Demagnetisation Process and Possible Alternative Magnetic Treatments for Naval Vessels, *PhD Dissertation*, Sydney: University of New South Wales, 2002.
3. Volkovitskiy, A.K., Karshakov, E.V., Tkhorenko, M.Yu., and Pavlov, B.V. Application of Magnetic Gradiometers to Control Magnetic Field of a Moving Object, *Automation and Remote Control*, 2020, no. 2, vol. 81, pp. 333-339.
4. Volkovitsky, A.K., Karshakov, E.V., and Pavlov, B.V., The Structure of the Control Algorithms of Aircraft Navigating, *Izv. SFedU. Engineering Sciences*, 2013, no. 3 (140), pp. 217–225. (In Russian.)
5. Karshakov, E.V., Shevchenko, A.M., and Garakoev, A.M., Formation of a Director Index to Assist the Pilot in Conducting Airborne Geophysical Survey, *IOP Conference Series: Materials Science and Engineering - International Workshop on Navigation and Motion Control* (*NMC 2020*), 2020, vol. 984, p. 012015 (1-10).
6. Purcell, E., *Electricity and Magnetism*, 3rd. ed., Cambridge: Cambridge Univ. Press, 2013.
7. 7. Landau, L.D., Lifshitz, E.M., and Pitaevskii, L.P., *Course of Theoretical Physics. Vol. 8: Electrodynamics of Continuous Media*, 2nd ed., Butterworth-Heinemann, 1984.
8. Moon, P. and Spencer, D.E., *Field Theory Handbook*, Springer-Verlag, 1988.
9. Abramowitz, M. and Stegun, I.A., *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*, National Bureau of Standards, 1964.
10. Press, W.H., Teukolsky, S.A., Vetterling, W.T., and Flannery, B.P., *Numerical Recipes: The Art of Scientific Computing*, Cambridge: Cambridge University Press, 2007.
11. Dricsoll, J.R. and Healy, D.M., Computing Fourier Transforms and Convolutions on 2-Sphere, *Advances in Applied Mathematics*, 1994, vol. 15, pp. 202–250.
12. Thebault, E., Finlay, C.C., Beggan, C.D., et al., International Geomagnetic Reference Field: the 12th Generation, *Earth, Planets and Space*, 2015, vol. 67, article no. 79, pp. 1–19.
13. Telford, W.M., Geldart, L.R. and Sherif, R.E., *Applied Geophysics*, Cambridge: Cambridge University Press, 2004.
14. Foley, C.P., Tilbrook, D.L., Leslie, K.E., et al., Geophysical Exploration Using Magnetic Gradiometry Based on HTS SQUIDs, *IEEE Trans. Appl. Superconductivity*, 2001, vol. 11, no. 1, pp. 1375–1378.
15. *Handbook of Applicable Mathematics. Vol. 6: Statistics*, Ledermann, W. and Lloyd, E., Eds., Wiley, 1984.
16. Stepanov, O.A., *Primenenie teorii nelineinoi fil'tratsii v zadachakh obrabotki navigatsionnoi informatsii* (Application of Nonlinear Filtering in Navigational Data Processing), St. Petersburg: Concern CSRI Elektropribor, 2003. (In Russian.)
17. Simon, D., *Optimal State Estimation. Kalman, H_∞ and Nonlinear Approaches*, Wiley, 2006.
18. Xu, Y., Chen, X., and Li, Q., Adaptive Iterated Extended Kalman Filter and Its Application to Autonomous Integrated Navigation for Indoor Robot, *The Scientific World Journal*, 2014, vol. 2014, p. 7.

This paper was recommended for publication by L.B. Rapoport, a member of the Editorial Board.

Received April 7, 2021, and revised April 22, 2021.
Accepted May 13, 2021.

Author information

Volkovitskiy, Andrei Kirillovich. Senior researcher, Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia, ✉ avolkovitsky@yandex.ru

Gladyshev, Anatoly Ivanovich. Chairman, the Defense Problems Section of the Ministry of Defense under the RAS Presidium, Moscow, Russia, ✉ tolyagladyshev@yandex.ru

Goldin, Dmitry Alekseevich. Senior researcher, Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia, ✉ goldind@ipu.ru

Karshakov, Evgeny Vladimirovich. Leading researcher, Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia, ✉ karshakov@ipu.ru

Pavlov, Boris Viktorovich. Chief researcher, Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia, ✉ pavlov@ipu.ru

Tkhorenko, Maksim Yur'evich. Senior researcher, Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia, ✉ tkhorenkom@mail.ru

Cite this article

Volkovitskiy, A.K., Gladyshev, A.I., Goldin, D.A., Karshakov, E.V., Pavlov, B.V., Tkhorenko, M.Yu. A Computer Simulation Complex for Analysis of Magnetic Gradiometry Systems. *Control Sciences* **3,** 57–65 (2021). http://doi.org/10.25728/cs.2021.3.8

DOI: http://doi.org/10.25728/cs.2021.3.9

13TH INTERNATIONAL CONFERENCE ON MANAGEMENT OF LARGE-SCALE SYSTEM DEVELOPMENT (MLSD'2020)

The 13th International Conference on Management of Large-Scale System Development took place on September 28–30, 2020. The event's organizer was Trapeznikov Institute of Control Sciences, Russian Academy of Sciences (ICS RAS), Moscow, supported by the IEEE Russia Section. Due to the special COVID-2019 pandemic conditions, the conference sessions were held online. With the virtual format, there was no interruption in regular publications and the presentation of new scientific results at the conference organized annually.

MLSD'2020 was attended by 400 participants from different institutions of the Russian Academy of Sciences, Russian universities, management and commercial organizations, and foreign research institutions (France, Norway, Burma, Vietnam, Belarus, and Kazakhstan). Note that t144 papers were included in the conference proceedings.[1]

Traditionally, on the first day of the conference, a plenary session was dedicated to fundamental research into the theoretical and methodological platform for strategic management of large-scale system development, with the presence of all conference participants. The following papers of leading scientists were presented at the plenary session of MLSD'2020:

- "Research principles in the theory of control of organizational and technical systems" by *D.A. Novikov*. The paper examined the epistemological principles of the theory. The author identified the principles of rationality, coordination, and decomposition, which reflect the specifics of organizational and technical systems as controlled objects with a complex structure (logical, causal, process, etc.). Such systems include active elements (subjects with active behavior), and their lifecycles should be coordinated.
- "Management of large-scale system development in new conditions" by *A.D. Tsvirkun*. The paper considered a comprehensive analysis methodology and tools for managing the development of large-scale systems intended for solving strategic tasks and management problems in Russia, particularly in emergency conditions.
- "A multilevel system for modeling regional budget revenues in the coronavirus crisis conditions" by *O.I. Dranko*. The paper proposed express modeling of regional budget revenues depending on the crisis depth. A set of support measures for organizations was considered, and their effect on budget revenues was studied.
- "New forms of public-private partnership in Russia's scientific and technological development" by *V.G. Varnavskii*. Particular attention was paid to large-scale projects. The foreign experience of implementing the concept of public-private partnership was presented, and Russia's regulatory and legal base that appeared in recent years was analyzed. According to the author's conclusion, public-private partnership has significant potential for Russia's scientific and technological development.
- "Towards a platform for managing the infrastructural development of a large-scale region in extreme climatic and geographical conditions" by *V.V. Tsyganov*. The paper proposed a theoretical and methodological platform for strategic management of the infrastructural development of a large-scale region in extreme climatic and geographical conditions. Several platform elements were designed: organizational systems and mechanisms for managing infrastructural development, a cost-effect decision methodology for resource exploitation and the spatial development of territories, and a scenario-based simulation model for the development of transport infrastructure during the evolution of the regional socio-economic system.
- "Scenario analysis of the problems of ensuring public security in digitalization conditions" by *V.V. Kulba, A.B. Shelkov, I.V. Chernov, and L.V. Bogatyreva*. The paper considered the problems of increasing the efficiency of public security management processes and the transformation of legislative regula-

[1] Materialy 13-oi Mezhdunarodnoi konferentsii "Upravlenie razvitiem krupnomasshtabnykh sistem" (Proceedings of 13th International Conference on Management of Large-Scale System Development (MLSD'2020), September 28–30, 2020, Moscow, Tsvirkun, A.D., Ed., Moscow: Trapeznikov Institute of Control Sciences, 2020. (In Russian.)

tion and law enforcement systems in a developed information society. A multigraph model was constructed, and the feasibility of creating a single investigative authority was studied using scenario analysis based on the model.

- "New analysis assessment and control methods for complex power systems based on spectral and structural analysis" by *I.B. Yadykin and A.B. Iskakov*. The paper discussed new stability assessment and control methods for large-scale power systems within project no. 19-19-00673 supported by the Russian Science Foundation. The main theoretical result of the project is the development of a new concept of Lyapunov modal analysis (LMA), which combines two approaches to assess the stability of dynamical systems: selective modal analysis and spectral factorizations of Lyapunov functions.
- "The use of existing managerial innovations is enough for breakthrough development and recession reversal" by *V.A. Irikov*. The practical proposals outlined in the paper are a response to the initiative of the President of the Russian Federation for breakthrough development and the goals set by him in the Presidential Address to the Federal Assembly on March 1, 2018. As a result, the implementation of the described breakthrough technologies of the third generation will accelerate by 1.5-2 times the achievement of the socio-economic targets indicated in the Budget Law for 2018–2020 and will even double the increase in the rates of social and economic growth in the next 2–3 years.
- "Prospects for coordinating technological development with national projects" by *N.I. Komkov*. The paper focused on increasing the requirements to the competitiveness of domestic innovative solutions and technologies: diversifying technological potential throughout the entire technological cycle, reducing costs and economic losses at all stages and technological redistributions, coordinating the potential, quality, and interests of all links of technological chains, etc.
- "New functions for designing supervisory control systems of nuclear power plants" by *A.G. Poletykin*. The paper proposed adding modern computer technology-based functions into the upper level of the industrial process control system of new-generation nuclear power plants (NPPs). The list of functions, their placement at NPPs, required personnel, and implementation methods were considered. In addition, the issues of ergonomics and cybersecurity were discussed.
- "Structural dynamics and macroeconomic policy in Russia" by *O.S. Sukharev*. As shown in the paper, the structural policy is an indispensable way to create a new growth model for the Russian economy; it reduces to institutional changes balancing the profitability of the transactional, raw material, and manufacturing sectors. The research methodology was structural analysis and the pairwise correlation method to determine the contribution of macroeconomic tools to the growth rates of Russia's GDP components and their effect on GDP on the time interval under consideration.

The program of MLSD'2020 covered 234 papers distributed in the following sections:

Section 1. Problems of managing large-scale system development, including multinational corporations, state holdings, and state corporations.

Section 2. Methods and tools for managing investment projects and programs.

Section 3. Managing the development of a digital economy. Design offices, situational and prediction and analytical centers, institutes of large-scale system development.

Section 4. Simulation and optimization in problems of managing large-scale system development.

Section 5. Nonlinear processes and computing methods in problems of managing large-scale systems.

Section 6. Managing the development of banking and financial systems.

Section 7. Management of fuel, power, infrastructure, and other systems.

Section 8. Management of transport systems.

Section 9. Managing the development of aerospace and other large-scale organizational-technical complexes.

Section 10. Managing the development of regional, urban, and municipal systems.

Section 11. Management of nuclear power objects and other objects of increased danger.

Section 12. Infoware and software for management systems of large-scale production.

Section 13. Methodology, methods, software, and knoware for big data processing and intelligent analysis.

Section 14. Monitoring in managing the development of large-scale systems.

Section 15. Managing the development of large-scale health systems, biomedical systems, and technologies.

Section 16.1 Managing the development of socio-economic networks.

Section 16.2 Mechanisms for managing the development of socio-economic systems.

The sections were conducted on the second and third days of the conference. Of major interest is to group the sections papers by the relevant problems of managing large-scale system development. Among them, we will identify seven areas described below.

Industrial and commercial use of resources (energy, transport, production, etc.) requires optimal **investment** in their development. The following papers presented interesting solutions for the problems of this area:

- "Methods to model and optimize the choice of investment decisions" by *V.K. Akinfiev and A.D. Tsvirkun*. The authors considered models for selecting a portfolio of projects for industrial companies with a complex asset structure based on a simulation and optimization approach. In addition, portfolio selection by maximizing the market value of a holding company was described.
- "Assessing the effectiveness of large-scale investment projects in various conditions" by *V.N. Livshits, I.A. Mironova, T.I. Tishchenko, and M.P. Frolova*. The paper proposed an unconventional model for assessing the social effectiveness of a large-scale infrastructural project to maximize the investor's capital at the end of the calculation period, considering the scenarios of an efficient use (reinvestment) of the capital formed from the project profit and loss of profits.
- "The role of investment projects for large-scale systems in crisis conditions and a minimax approach to implement them" by *E.V. Popova*. As shown by the author, minimax strategy, property strategy, reform methodology, target models, and value drivers are effective tools to revise large projects and programs.
- "Comparative analysis of efficiency criteria for share portfolios under different approaches to form them" by *D.S. Sizykh*. The paper presented a modified version of the Markowitz portfolio optimization model using the stable growth indicator of stock prices. In addition, the efficiency of various portfolio formation approaches was compared on an example of portfolios consisting of the shares of leading IT and telecommunications companies for the period 2015–2020.
- "A mixed investment portfolio with a limited choice of assets" by *A.S. Syrovatkin*. As noted in the paper, modern economic reality is increasingly moving away from classical market regulation. As a result, the concept of a free market becomes less and less applicable, and an opportunity to use it may arise only in narrow market niches. In other cases, in one form or another, there is an external impact breaking the natural balance. The author proposed a method to form and optimize a mixed investment portfolio under restrictions (the obligation to launch a given share of real projects).

An important tool for managing the development of large-scale systems is **intelligent computer modeling** associated with knowledge acquisition. The following papers were devoted to various aspects of this area:

- "The concept of forming a single information Internet-based space for the scientific and educational resources of a country" by *F.I. Ereshko, V.I. Medennikov, and Yu.A. Flerov*. The authors proposed a methodology for assessing the effectiveness of using scientific and academic information resources on the Internet under the transition to the digital economy and innovative development.
- "Mathematical models and algorithms for the predictive assessment of national security in training situational centers" by *N.V. Yandybaeva, A.F. Rezchikov, E.A. Gorshkov, A.S. Bogomolov, V.A. Kushnikov, and A.D. Tsvirkun*. A mathematical system dynamics-based model was developed to carry out an express analysis of the current state and predict the national security of countries. In addition, an algorithm for using the developed knoware for assessing the national security of countries was presented. Finally, a methodology to train national security specialists in situational centers was demonstrated.
- "Features of managing large-scale cyber-physical municipal water supply systems in different countries" by *N.A. Fomin and R.V. Meshcheryakov*. The paper considered the models of water supply management in cities of different countries: Asia (China and Singapore), the USA, Europe (particularly England), and Russia. Some drawbacks in the management models were identified due to the existing potential threats of water shortage, degradation of water sources, and chemical and biological pollution risks.
- "Assessing the level of digital transformation in Russia's agricultural sector" by *V.V. Kul'ba and V.I. Medennikov*. The paper introduced a mathematical model of the readiness of Russia's agricultural sector for digital transformation. In addition, several scenarios for digitalization were calculated, and the corresponding results were presented.

The following papers were devoted to the problem of **sustainable development**:

- "Decision-making in development management systems" by *V.V. Baranov*. The author formulated the fundamentals of feasible and sustainable development management processes in a degrading material and reflexive social environment. Sustainability was related to conflict resolution through sustainable compromise. The structural configurations of the systems were determined, taking into account the environment values. Finally, a structure, methodology, and constructive tools for development management through the planning cycles of strategic renewal and real management were described.

- "An entropy approach to problems of sustainable regional development" by *A.N. Solomatin*. The paper proposed various ways to fight the growth of entropy. Necessary conditions for sustainable development were formally defined. In addition, the technological, economic, environmental, and management aspects of counteracting the growth of entropy were analyzed in detail.
- "Stable economic dynamics of large-scale systems" by *V.V. Glazunova*. The author introduced the concept of the stable operation of an economy and methods for assessing it. By stability, the author understood the system's ability to return to the initial state or maintain the trajectory of development under disturbances. If the system cannot return to the previous state or deviates from the given trajectory, it is unstable. For economic systems, the most important ability is maintaining the trajectory of development under exogenous disturbances (the so-called dynamic equilibrium).

Of great interest are original papers expanding the existing knowledge about the control of **nonlinear systems** for industry and cross-industry applications. Among them, we mention the following:

- "Optimal thermodynamic processes for ideal gases" by *A.G. Kushner, V.V. Lychagin, and M.D. Roop*. The problem of optimal control in equilibrium thermodynamics of ideal gases was solved. A thermodynamic process curve maximizing the work functional was found on the Legendre manifold of an ideal gas. Furthermore, constraints on the control parameters were obtained. As shown in the paper, in the case of an ideal gas, the corresponding Hamiltonian system is Louisville integrable, and the controllability of such a system was proved.
- "Singular control for enhanced oil recovery in natural oil reservoirs" by *A.V. Samokhin, A.V. Akhmetzyanov, and E.I. Krupina*. The paper considered sawtooth waves with periodic shock fronts generated in wellbores. The approach proposed by the authors allows predicting the dynamics of phase transitions adequately and optimizing oil production control.
- "Identification of integrated ranking mechanisms as an optimization problem" by *V.N. Burkov, N.A. Korgin, and V.A. Sergeev*. The authors demonstrated that with the single-coded transformation of mechanisms, the identification of any integrated ranking mechanism reduces to an optimization problem, and a wide range of methods apply to it. The proposed approach was illustrated by several examples of identification or approximation of learning sets generated by several Boolean functions.
- "Stabilization of a two-rotor electromechanical system based on the principle of decomposition" by *A.S. Antipov, S.A. Krasnova, and S.V. Pivneva*. The paper considered a two-rotor electromechanical system as a plant under parametric uncertainties and cross-links between the rotors. A block design procedure was developed for nonlinear local links and discontinuous controls stabilizing the angular positions under the design constraints. Finally, the simulation results were presented.

Robotics and related computer simulation technologies are of constant interest. In this area, note the following papers:

- "Cybernetics systems: algorithmization in the problems of the primary assessment of objects in a complex unmanned aerial vehicle" by *S.S. Semenov, A.V. Poltavskii, and E.Yu. Rusyaeva*. The paper suggested an approach to substantiating the utility function for obtaining initial estimates of the performance criteria of complex technical systems in the form of an information model considering the consistency of expert judgments.
- "Some methods to classify and recognize targets in modeling the target environment of unmanned aerial vehicles" by *V.Tr. Nguyen, Ch. A. Bui, F.F. Pashchenko, A.F. Pashchenko, and Yu.I. Kudinov*. The paper addressed the issues of modeling unmanned aerial vehicles under multifactor uncertainties. Approaches to the selection of target tasks were proposed. An algorithm for classifying air targets based on dimensional features and an algorithm for classifying air targets based on moment invariants of images were developed. These algorithms can be used for object image recognition and statistical decision-making on target tasks.
- "Estimating the derivatives of reference signals in a control system for UAVs" by *Yu.G. Kokun'ko and S.A. Krasnova*. The authors developed a procedure for designing a dynamic feedback control under which the center of mass of an unmanned aerial vehicle will track a given trajectory invariantly with respect to exogenous disturbances with a desired accuracy under incomplete information about the state variables and the derivatives of the reference signals. In addition, the simulation results were provided.
- "Design and prototype of a six-legged walking machine" by *V.A. Danilov and V.I. Goncharenko*.

The paper presented a mathematical model of a walking hexapod robot. The model provides visual information about its spatial movements and can be used to develop control algorithms.

- "Development of a flight simulator in the conditions of UAV group control" by *P.M. Trefilov, M.V. Mamchenko, and K.A. Kulagin*. The authors described a generalized process of creating a virtual simulator with an integrated geographic information system for simulating flight tasks of unmanned aerial vehicles. The simulator can be used to create flight tasks for single UAVs and their groups and export data on completed missions for performing real flights.
- "Logical control of a gantry robot based on regular grammars in the presence of non-stationary obstacles in the working area" by *O.S. Tkacheva, A.V. Utkin, and M.S. Vinogradova*. The paper was devoted to designing logical control of a gantry robot facing non-stationary obstacles in the working area. For a two-link gantry robot, a gripper motion control model was constructed based on a hybrid automaton with a finite set of states. The logic control algorithm was tested using numerical simulation methods.

Many studies on cybersecurity analysis and software outlined the key tenets of **risk management**. Among such papers, we mention the following:

- "Analyzing the cybersecurity of a significant object of critical information infrastructure" by *E.A. Sakrutina and A.O. Kalashnikov*. As noted in the paper, for effective risk analysis, it is crucial to identify objects, threats, and vulnerabilities and understand the nature of cyberattacks. Moreover, it is crucial to determine the risk as accurately as possible, identifying its causes, scale, and limitations, and the type of potential threats affecting the object's goals. In the proposed approach, identifying and managing potential risks was treated as a continuous process of an ordered sequence of events, actions, and decisions ("threats–vulnerabilities–consequences").
- "Application of cloud service technology to ensure the cybersecurity of an industrial control system" by *A.I. Samoshina, V.G. Promyslov, S.B. Kamesheva, and R.R. Galin*. The paper introduced a mathematical graph theory-based model to describe access relations between the objects and subjects of security policy. Algorithms for traversing the graph vertices were compared to select a suitable method for identifying security zones. Finally, the algorithm for calculating security zones was implemented and added into to the omole.ws cloud service.
- "An example of initial data verification for designing the Information base of a nuclear power plant" by *E.R. Budynkova and A.A. Baibulatov*. A complete code parsing algorithm for the Kraftwerk Kennzeichen system was presented, and all its sectors were studied. In addition, a list of the most common inconsistencies was provided.
- "An arbitration model of information risk management for significant objects of information infrastructure" by *A.O. Kalashnikov and E.V. Anikina*. The paper considered an effective limited resource allocation method for managing information risks of the significant objects of critical information infrastructure based on game-theoretic models (arbitration schemes).

In the theory of managing the development of large-scale systems, of considerable interest is research into methods, models, and tools for **intelligent analysis of big data**. In this area, we highlight the following papers:

- "Digraph clustering methods based on the Laplace matrix and its eigenprojector" by R.P. Agaev. The paper was devoted to the topical problem of "meaningful" clustering of oriented networks. As noted by the author, adequate clustering of digraph vertices is impossible when ignoring the direction of graph edges (considering the graph to be undirected).
- "An algorithm for constructing a regression decision tree using additional functions" by *S.A. Saltykov*. The paper examined the problem of constructing accurate and intuitively plausible analytical models clear for the analyst. An algorithm for constructing a regression decision tree with additional functions was presented. Also, a reliability condition for a two-level decision tree was described.
- "Methods for assessing the effectiveness of integrating software and technology solutions into digital platforms" by *D.Yu. Il'in and E.V. Nikul'chev*. In the paper, a digital platform was understood as a technology for acquiring and exchanging information between very many users. The authors developed methods and virtual simulation information infrastructures for assessing the effectiveness of integrating software and technological solutions into digital platforms (on an example of a digital platform for mass research in the education system). The "infrastructure as code" concept reduces the cost of computational experiments on a given technology stack and allows an adequate assessment of a technological solution when technologies are integrated into the stack. Moreover, the infrastructure and operating conditions of the developed digital platform are taken into account. The proposed approach allows assessing the effectiveness of technologies and reducing the cost of computational experiments at the stage of selecting

technologies. In addition, the implementation results were presented to demonstrate the effectiveness of the proposed methods.

The Deputy Chair of the Program Committee, A.D. Tsvirkun, delivered welcoming and concluding remarks for the conference participants. He emphasized the steadily growing scientific potential of the annual conference for 15 years of its history and expressed confidence in future conferences. Moreover, he noted that large-scale systems (corporations, financial and industrial groups) are the locomotive ensuring the competitiveness of the national and transnational economy and supporting them in the global market space. The Government and Science need to undertake the elaboration, renewal, and mutual coordination of general schemes for the development and placement of industries considering the lines of international and regional development. Investment projects should be carried out within a comprehensive long-term cross-sectoral program of the country's socio-economic development and the territorial placement of production. Government programs and large business projects should be implemented after careful consideration.

Chair of the Organizing Committee
A.D. Tsvirkun

Secretary of the Organizing Committee
I.A. Stepanovskaya

Author information

Tsvirkun, Anatolii Danilovich. Dr. Sci. (Eng.), Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia, ✉ tsvirkun@ipu.ru

Stepanovskaya, Iraida Aleksandrovna. Cand. Sci. (Eng.), Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia, ✉ irstepan@ipu.ru

Cite this article

Tsvirkun, A.B., Stepanovskaya, I.A. 13th International Conference on Management of Large-Scale System Development. *Control Sciences* **3**, 66–71 (2021). http://doi.org/10.25728/cs.2021.3.9