ISSN 2782-2427

CONTROL SCIENCES

6/2022

[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: December 30, 2022

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

6.2022

CONTENTS

DOI: http://doi.org/10.25728/cs.2022.6.1

DESIGN OF INTEGRATED RATING MECHANISMS BASED ON SEPARATING DECOMPOSITION

V.A. Sergeev

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

✉ sergeev.bureau@gmail.com

Abstract. This paper proposes an approach to reducing significantly the computational complexity of optimization problems in the design of integrated rating mechanisms (IRMs). The background concepts are introduced. The representability of a given discrete function as some IRM is proved. The decomposition procedure for a particular training example on some partition of input parameters is considered, and the following results are established under some restrictive conditions. First, an IRM matrix for a particular example of an input data set can be designed by maximizing a certain polynomial. Second, a set of given examples can be implemented by some IRM matrix. Third, an IRM can be implemented on a training data set in a certain complete binary tree based on the decomposition method. Fourth, some discrete function is implemented through a given complete binary tree if the discrete functions represented by convolution matrices are implemented in each node of this tree. All these results are rigorously formulated and proved. An illustrative example of the decomposition procedure based on a complete binary tree on three leaves is given. We propose a method for finding IRMs that implement a given training set in the space of all possible complete binary trees based on the branch table. In addition, we describe the decomposition procedure according to the branch table for each partition of input parameters. Finally, the advantages of the proposed method are outlined.

Keywords: integrated rating mechanism, discrete function, assessment, decomposition.

INTRODUCTION

Training models based on precedents is a well-known practice [1–3] going beyond the field of machine learning. In recent years, the development of training procedures for integrated rating mechanisms (IRMs) has attracted the attention of researchers.

IRMs are widely used as multidimensional assessment and ranking systems for management and control in organizational and production systems [4–9]. When used for complex systems (e.g., organizational and production systems), the integrated rating procedure allows dealing with the typical difficulties of complex object assessment [10, 11]. The basic application of IRMs is ordinal ranking or classification with a predetermined number of classes for a finite set of multicriteria alternatives [12–14]. The main components of IRMs are a binary tree and convolution matrices, which yield a complex assessment based on the values of several input indicators. Recently, several approaches have been proposed to design (in other words, identify) convolution matrices by a particular binary tree [15, 16]. This paper introduces a design approach that further develops the method outlined in [15]. The approach under consideration is intended to settle the difficulties associated with the complexity of solving the optimization problem during IRM matrix design. For this purpose, we adopt the decomposition method of discrete functions. Also, a topical problem is finding a set of IRMs implementing a given data set.

Many researchers showed interest in the possibility of functional decomposition. For example, A.N. Kolmogorov [17] and V.I. Arnold [18] studied the decomposability of continuous functions. For the class of discrete functions, V.S. Vykhovanets [19] constructed the decomposition procedure of algebraic functions and analyzed the identification problem of a discrete system using a spectral decomposition; for example,

see [20]. The complexity of the representation of Boolean functions was investigated by S.V. Yablonsky [21]. In the paper [22], A.V. Kuznetsov considered repetition-free Boolean functions. Also, we emphasize the work [23] on multicriteria assessment by V.A. Glotov and V.V. Pavel'ev; the authors described the application of decomposition to construct a criterion-target structure. V.N. Burkov and I.V. Burkova with colleagues studied the dichotomous function representation [24, 25] in terms of solving discrete optimization problems, including application to complex assessment.

1. BASIC NOTIONS AND DEFINITIONS

Consider a finite set of indicators $L \subset \mathbb{N}$, $|L| = l$, to rate some object on a discrete scale or rank several objects. For the IRM identification problem, assume that there is a finite set $K_i \subset \mathbb{N}$ of possible values of each indicator $i \in L$, where $k_i \in K_i$ is an assessment of an individual parameter. The vector $k = (k_1,..., k_l)^{\mathrm{T}}$, the set of all assessments, describes any possible state of the assessed objects. Also, there is a finite set $K_L \subset \mathbb{N}$ of possible integral values (ranks or classes) $k_L \in K_L$ for any k. Thus, we have some discrete function $f(\cdot): K_D \to K_L$. Here, $K_D = K_1 \times K_2 \times ... \times K_l$ is the (definitional) domain, with × denoting the Cartesian product of sets, and K_L is the codomain (range) of the function. This paper focuses on the discrete scales of indicators and values obtained in the nodes of a convolution tree [12]. A function f defined on a set K_D and taking values in a set K_L is a mapping of K_D into K_L such that each element x of the domain K_D is related to at most one element of the codomain K_L.

Definition 1. An IRM with a binary tree and matrix convolutions is a function $f(\cdot): K_D \to K_L$ for which indicators L are leaves of a complete binary tree, i.e., a digraph $G = (V, E)$:

- $V = L \cup \hat{L}$, where $\hat{L} = \{l+1,...,2l-1\}$.
- $E = \{e_{ij}\} \subseteq V \times V$:
  - $\forall i \in V \setminus \{2l-1\}\ \exists! j \in \hat{L} \setminus \{i\}\ \ e_{ij} = 1,\ \forall t \in V \setminus j\ e_{it} = 0$;
  - $\forall j \in L\ \forall i \in V\ e_{ij} = 0$;
  - $\forall j \in \hat{L}\ \exists! \{r,c\} \in V \setminus \{j\} \times V \setminus \{j\} : e_{rj} = 1, e_{cj} = 1$;

and $\forall j \in \hat{L}$ (an inner node of the tree, including its root):

- a finite set $K_j \subset \mathbb{N}$ with possible values $k_j \in K_j$, $K_{2l\text{-}1} = K_L$, and
- a convolution matrix $M_j = [m_{jrc} \in K_j]_{r \in \{0,...,|K_l-1|\}, c \in \{0,...,|K_r-1|\}}$,

$\{r, c\} \in V \setminus \{j\} \times V \setminus \{j\}\ : e_{lj} = 1,\ e_{rj} = 1$, are given. ♦

Potentially, this definition can be extended to fuzzy [26] or continuous scales. For some IRM, by analogy with [15], let $M_f = \{M_j\}_{j \in L}$ denote the set of all its convolution matrices. This paper is devoted to IRMs with a single scale such that $\forall j \in V\ K_j = K_L$. For $L \subset \mathbb{N}$, we introduce the following notations: $\Gamma_2(L)$ is the set of all complete binary trees on named leaves from the indicator set L; $IRM_{L,2}$ is the set of all IRMs for any particular binary tree $G \in \Gamma_2(L)$; $IRM_{L,G} \subseteq IRM_{L,2}$ is the set of all IRMs with such a tree. According to Definition 1, a complete binary tree in this paper is understood as a tree in which each node has either none or two child nodes.

Based on the definitions given in [15], we denote by $q = (k, k_L)$ an individual training example consisting of the assessments for each indicator and the integrated rating for a given set of indicator values and by $Q \subset K_D \otimes K_L$ a training set of the provided examples. A training set is compatible if $\forall \{q, \tilde{q}\} \subseteq Q\ \ k \neq \tilde{k}$. A training set is complete if $\forall k \in K_D\ \ \exists q \in Q: q = (k, k_L)$. A training set is given in a single scale if $\forall i \in L\ K_i = K_L$. For arbitrary elements $\{k, \tilde{k}\} \subseteq K_D$, the relation $k \succ \tilde{k}$ means $\forall i \in L\ \ \tilde{k}_i \le k_i$. For an arbitrary set $Q \subset K_D \otimes K_L$, we present key notions concerning the identification problem. First of all, we formalize the implementability problem of a training set.

Definition 2. A function $f(\cdot) \in IRM_{L,2}$ implements a set Q if and only if $\forall q \in Q\ f(k) = k_L$. ♦

We introduce the following notations: $IRM_{L,2}(Q)$ is the set of all IRMs implementing a set Q; $IRM_{L,G}(Q)$ is the set of all IRMs that implement Q and are based on a binary tree $G \in \Gamma_2(L)$. If $IRM_{L,2}(Q) \neq \emptyset$, then the set is implementable based on an IRM; if $IRM_{L,G}(Q) \neq \emptyset$, then the set Q is implementable based on an IRM with a structure G. Definition 2 can be narrowed to one particular training example: a function $f(\cdot) \in IRM_{L,2}$ implements some example $q \in Q$ if and only if $f(k) = k_L$; the sets $IRM_{L,2}(q)$ and $IRM_{L,G}(q)$ are defined by analogy.

For some finite set $K \subset \mathbb{N}$, we write its normalized representation: $\bar{K} = \{0,...,s\}$, $s = |K| - 1$. Then $\forall x \in \bar{K}$ the unitary representation is given by $\tilde{x} = (\underset{x}{0},...,0,1,\underset{s-x}{0},...,0)^{\mathrm{T}}$. This paper deals with IRMs with a single scale. Then for any pair $\{x, y\} \in \{0,...,\bar{K}\}^2$, where x and y are the chosen matrix column and row, respectively, the convolution result with a matrix $M = [m_{rc} \in K]_{\{r,c\} \in \bar{K}^2}$ is described by the matrix equation

$\tilde{y}^{\mathrm{T}} M\tilde{x}$. Below, we also adopt the so-called quadratic representation $(M\tilde{x}, \tilde{y})$, simplifying it to $M\tilde{x}\tilde{y}$.

2. DISCRETE FUNCTION DECOMPOSITION

The paper [15] proposed an approach to identifying IRMs with a training mechanism on discrete data. It involves a mechanism for constructing an optimization functional based on input data and a complete binary tree. This approach to the identification problem causes difficulties when solving the optimization problem with large-dimension input data. The degree of the optimization polynomial grows linearly with the number of input parameters, and the number of general constraints of the optimization problem grows exponentially: $Cn = \kappa^{l-2} ex\_num$, where Cn is the number of general constraints; $\kappa = |K|$ is the indicator value scale for the parameters in a single scale k_L; *ex_num* is the number of examples in the set Q; finally, l is the number of parameters.

As an alternative, the identification problem can be solved in steps using the separating decomposition of the function $f(X) = \Sigma(X_1, a(X_2))$, where Σ and a are some functions and $X_1 \cap X_2 = \emptyset$ are the subsets resulting from the separation of the set X; for details, see [19]. The decomposability of any continuous function of n variables into a superposition of continuous functions of fewer variables was studied by A.N. Kolmogorov and V.I. Arnold. In particular, for two variables, it was proved in [17, 18]. The paper [17] obtained the following theoretical result: any continuous function of $n \geq 3$ variables can be represented as a superposition of some continuous functions. V.A. Glotov and V.V. Pavel'ev [23] established the representability of a discrete function of n variables in the binary (separating) form. In the case of IRMs, it is easy to show (see the Appendix) the decomposability of a given function in a complete binary tree under unfixed value scales $k_i \in K_i \; \forall i \in \{1,..., l-1\}$: due to the finiteness of k_L, the value scale of functions decomposing a given discrete function is also finite. Thus, identification problems are sequentially formulated and solved for the decomposition procedure in each node of the tree from the set $\Gamma_2(L)$.

Following [27], let the indicator decomposition structure of an arbitrary tree $G \in \Gamma_2(L)$ be denoted by $\Lambda(G) = \{L_i\}_{i \in \{1,..,l-1\}}$, where $\forall i \in \{1,...,l-1\}$ $L_i \subseteq L$ is the set of leaves (indicators) of a subtree with root node i. Then $L_1 = L$ and $\forall i \in \{1,...,l-1\}$ $|L_i| \geq 2$. Consider a given complete set Q and a given tree $G \in \Gamma_2(L)$. For any set $L_i \in \Lambda(G)$ such that $L_i \geq 2$ and its subgroup $\{L_{ir}; L_{ic}\} \subset \Lambda(G)$: $L_{ir} \bigcup L_{ic} = L_i$. Consider tuples $k_{(L_{ir})}, \tilde{k}_{(L_{ir})}$ and $k_{(L_{ic})}, \tilde{k}_{(L_{ic})}$ of any admissible indicator values from the sets L_{ir} and L_{ic}, respectively. For some subset of indicators $\tilde{L} \subseteq L$, we denote by $\lambda = \left(k_{(\tilde{L})}, k_{(L \setminus \tilde{L})}\right)$ the partition of the tuple of indicators k of some training example $q = (k, k_L)$ into two tuples. Due to its complete binary structure, each tree from $G \in \Gamma_2(L)$ can be assigned the set of indicators L_i; then $\lambda_i = \left(k_{(L_{ir})}, k_{(L_{ic})}\right)$ is the partition of the tuple of indicators in node i of the tree G. In each tree node, the component functions decomposing the discrete function $\varphi_i(k_{(L_i)})$ will be named in accordance with the partition and the numbering of matrices located in the tree nodes implementing the component functions of $\varphi_i(k_{(L_i)})$. For example, consider the partition $\lambda = \left(k_{(1,2)}, k_{(3,4)}\right)$; the components of the two subfunctions are named as follows: $\varphi_{i+1\_}k_1k_2$ and $\varphi_{i+2\_}k_3k_4$. For convenience, we may also name the components of the discrete function $\varphi_i(k_{(L_i)})$ in accordance with the partition λ_i as $\varphi_r(k_{(L_{ir})})$ and $\varphi_c(k_{(L_{ic})})$. If the partition consists of an individual leaf and a group, i.e., $\lambda = \left(k_{(1)}, k_{(3,4)}\right)$, we encode only the components $\varphi_{i+1\_}k_3k_4$, where $i \in \{1,..., l-1\}$. Consider an illustrative example for the proposed approach.

Example 1. Consider the case $|L| = 3$, $|K_L| = 2$ and the training example $q = ((0, 0, 0), 0)$. The unitary representation is $\tilde{q} = \left(\left(\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}\right), \begin{pmatrix}1\\0\end{pmatrix}\right)$.

First, we analyze the implementability of some discrete function $\varphi_1(\tilde{k}_1, \tilde{k}_2, \tilde{k}_3)$. Let the decomposition functions be named using the partition $\lambda_1 = \left(k_{(1)}, k_{(2,3)}\right)$

$$\left(\begin{pmatrix}\varphi^0_{2\_00}\\ \varphi^1_{2\_00}\end{pmatrix}^{\mathrm{T}} \begin{bmatrix}\begin{pmatrix}ml^0_{00}\\ml^1_{00}\end{pmatrix} & \begin{pmatrix}ml^0_{01}\\ml^1_{01}\end{pmatrix}\\ \begin{pmatrix}ml^0_{10}\\ml^1_{10}\end{pmatrix} & \begin{pmatrix}ml^0_{11}\\ml^1_{11}\end{pmatrix}\end{bmatrix} \begin{pmatrix}1\\0\end{pmatrix}\right) = \begin{pmatrix}1\\0\end{pmatrix}$$

with the following unitary conditions: $\forall \{i, j\} \in \{0, 1\}^2$ $ml^0_{ij} + ml^1_{ij} = 1, ml^t_{ij} \in \{0,1\}$, $\varphi^0_{2\_00} + \varphi^1_{2\_00} = 1$, $\varphi^t_{2\_00} \in \{0, 1\}$, $\forall t \in \{0,1\}$. Simple transformations lead to the system

$$\varphi^0_{2\_00} ml^0_{00} + \varphi^1_{2\_00} ml^0_{10} = 1, \quad (1)$$

$$\varphi^0_{2\_00} ml^1_{00} + \varphi^1_{2\_00} ml^1_{10} = 0. \quad (2)$$

System (1), (2) is easily solved using binarity constraints, e.g., $\varphi^0_{2\_00}=1$ and $ml^0_{00}=1$. We have a part of the optimization problem corresponding to the example q = ((0, 0, 0), 0). ♦

Let $P(\lambda_i, q)$ denote the function on the left-hand side of equation (1). Due to the unitary approach, there exists a unique such function for any example q. In addition, Q_i is the data set obtained from the original one Q by selecting only the columns where the function $\varphi_i(k_{(L_i)})$ is defined.

Proposition 1. *For any sets $L \subset \mathbb{N}$ and $K \subset \mathbb{N}$ and any possible example q in a single scale, there exists a homogeneous polynomial $P(\lambda_i, q)$ of a degree not exceeding* 3 *that can be represented as the sum of k^{φ_num} unique components*:

$$P(\lambda_i, q)=\sum_{j=1}^{\kappa^{\varphi\_num}} p_j,\ \forall j\in\{1,...,\kappa^{\varphi\_num}\},$$

$$p_j = m_j \prod_{d=1}^{\varphi\_num}\varphi_d\ ,\ \forall d\in\{1,...,\varphi\_num\}.$$

Here, the notations are the following: φ_d is the function components decomposing the function $\varphi_i(k_{(L_i)})$; m_j is one tuple component in some cell of the unitary encoded matrix $\tilde{M}_i$; q is an example from Q_i; φ_num = 1 when a branch and a leaf are connected to the matrix; φ_num = 2 when a branch pair is connected to the matrix;

$P(\lambda_i, q)\in\{0, 1\}$;

$\varphi_i(k_{(L_i)})$ *implements* $q \Leftrightarrow P(\lambda_i, q) = 1$.

The proof of Proposition 1 is given in the Appendix.

For the function $\varphi_i(k_{(L_i)})$, at each step of the decomposition procedure, we form an appropriate set Q_i from the set Q based on the leaves corresponding to the partition λ_i.

Example 2. Within the conditions of Example 1, we add another training example q_2 = ((0, 1, 0), 1) based on the same partition $\lambda_1=\left(k_{(1)},k_{(2,3)}\right)$. Its unitary representation is $q_2=\left(\left(\begin{pmatrix}1\\0\end{pmatrix},\begin{pmatrix}1\\0\end{pmatrix},\begin{pmatrix}1\\0\end{pmatrix},\begin{pmatrix}1\\0\end{pmatrix}\right)\right)$. Then we have the set of operations

$$\left(\begin{pmatrix}\varphi^0_{2\_00}\\ \varphi^1_{2\_00}\end{pmatrix}^{\mathrm{T}}\begin{bmatrix}\begin{pmatrix}ml^0_{00}\\ml^1_{00}\end{pmatrix} & \begin{pmatrix}ml^0_{01}\\ml^1_{01}\end{pmatrix}\\ \begin{pmatrix}ml^0_{10}\\ml^1_{10}\end{pmatrix} & \begin{pmatrix}ml^0_{11}\\ml^1_{11}\end{pmatrix}\end{bmatrix}\begin{pmatrix}1\\0\end{pmatrix}\right)=\begin{pmatrix}1\\0\end{pmatrix},$$

$$\left(\begin{pmatrix}\varphi^0_{2\_00}\\ \varphi^1_{2\_00}\end{pmatrix}^{\mathrm{T}}\begin{bmatrix}\begin{pmatrix}ml^0_{00}\\ml^1_{00}\end{pmatrix} & \begin{pmatrix}ml^0_{01}\\ml^1_{01}\end{pmatrix}\\ \begin{pmatrix}ml^0_{10}\\ml^1_{10}\end{pmatrix} & \begin{pmatrix}ml^0_{11}\\ml^1_{11}\end{pmatrix}\end{bmatrix}\begin{pmatrix}1\\0\end{pmatrix}\right)=\begin{pmatrix}0\\1\end{pmatrix}$$

with the following unitary conditions: $\forall\{i, j\}\in\{0,1\}^2$ $ml^0_{ij}+ml^1_{ij}=1$, $ml^t_{ij}\in\{0,1\}$, $\varphi^0_{2\_00}+\varphi^1_{2\_00}=1$. A comparison of the equations written for the first and second training examples shows that the values φ_{2_00} and φ_{2_01} must be compatible. Therefore, when constructing the function $\varphi_i(k_{(L_i)})$ implementing the examples q_1 and q_2 simultaneously, we require $\varphi_{2\_00}{}^{\mathrm{T}}\varphi_{2\_01}=0$. ♦

Corollary 1. *For any sets $L \subset \mathbb{N}$ and $K \subset \mathbb{N}$ and any possible $Q_i \subset K^{l+1}$ with a single scale, the discrete function $\varphi_i(k_{(L_i)})$ implements the set Q_i if $\sum_{q\in Q_i} P(\lambda_i, q)=|Q_i|$ for any possible example q in the single scale considering the compatibility of the functions $\varphi_r(k_{(L_{ir})})$ and $\varphi_c(k_{(L_{ic})})$ for all examples in the set Q_i.*

The proof of Corollary 1 is given in the Appendix.

If two branches join the matrix under consideration, then φ_num = 2 and the compatibility of the functions $\varphi_r(k_{(L_{ir})})$ and $\varphi_c(k_{(L_{ic})})$ should be verified for each branch.

Proposition 2. *For any sets $L \subset \mathbb{N}$ and $K \subset \mathbb{N}$, the mechanism $IRM_{G,2}(q)$ is represented as a decomposition of the function $f(\tilde{k}_1,...,\tilde{k}_l)$ and implements a set $Q \subset K^{l+1}$ if the function $\varphi_i(k_{(L_i)})$ for some partition sequence $\Lambda(G)$ corresponding to the tree G implements the set Q_i, for $i \in \{1,..., l-1\}$.*

Corollary 2. *If $\sum_{q\in Q_i} P(\lambda_i, q)<|Q_i|$, $\forall i \in \{1,..., l-1\}$, then $IRM_{G,2}(Q) = \varnothing$.*

The proofs of Proposition 2 and Corollary 2 are given in the Appendix.

Thus, if the optimization problem $\underset{m_{rc},\varphi_r,\varphi_c}{\mathrm{Arg\,max}}\sum_{q\in Q_1} P(\lambda_1, q)$ has a solution such that $\sum_{q\in Q_1} P(\lambda_1, q)=|Q_1|$, we can continue the decomposition procedure to the next tree node with the found components of the discrete function $\varphi_1(k_{(L_1)})$. On this way, $\varphi_1(k_{(L_1)})$ and the found values of the matrix M_1 can be used for the decomposition procedure on the subtrees $\{L_{1r}; L_{1c}\}\subset\Lambda(G)$ on the indicator values $k_{(L_{1r})}$ and $k_{(L_{1c})}$, respectively. This approach sequentially yields

the vectors $\varphi_i(k_{(L_i)})$ and the matrix M_i, $\forall i \in \{1,..., l-1\}$ for the structure $\Lambda(G)$.

Example 3. Consider the initial data below.

Table 1

Initial data for the decomposition procedure

q	k_1	k_2	k_3	k_4	k_L
1	0	0	0	0	0
2	0	1	0	0	1
3	1	1	0	0	0
4	0	0	1	0	1
5	1	1	1	0	1
6	0	1	0	1	1
7	1	0	1	1	0
8	1	1	1	1	0

First, we analyze the implementability of the function $f = \varphi_1(\tilde{k}_1, \varphi_2(\tilde{k}_2, \tilde{k}_3, \tilde{k}_4))$; if $\varphi_1(k_{(L_1)})$ is available in the given scale k_L, we proceed to the implementation of the discrete function $\varphi_2(k_{(L_2)})$. Let the first-step partition be $\lambda_1 = \left(k_{(1)},\ k_{(2,3,4)}\right)$. The equations for the first example of the first step have the form

$$\left(\begin{pmatrix}\varphi^0_{2\_000}\\ \varphi^1_{2\_000}\end{pmatrix}^{\mathrm{T}}\begin{bmatrix}\begin{pmatrix}ml^0_{00}\\ ml^1_{00}\end{pmatrix} & \begin{pmatrix}ml^0_{01}\\ ml^1_{01}\end{pmatrix}\\ \begin{pmatrix}ml^0_{10}\\ ml^1_{10}\end{pmatrix} & \begin{pmatrix}ml^0_{11}\\ ml^1_{11}\end{pmatrix}\end{bmatrix}\begin{pmatrix}1\\0\end{pmatrix}\right) = \begin{pmatrix}1\\0\end{pmatrix}$$

with the following unitary conditions: $\forall\{i, j\} \in \{0,1\}^2$ $ml^0_{ij} + ml^1_{ij} = 1,\ ml^t_{ij} \in \{0, 1\}$, $\varphi^0_{2\_000} + \varphi^1_{2\_000} = 1$, $\varphi^t_{2\_000} \in \{0,1\}$, $\forall t \in \{0,1\}$:

$$\varphi^0_{2\_000} ml^0_{00} + \varphi^1_{2\_000} ml^0_{10} = 1,$$
$$\varphi^0_{2\_000} ml^1_{00} + \varphi^1_{2\_000} ml^1_{10} = 0.$$

Next, using the scheme described in Example 1, we derive equations for all examples:

$\varphi^0_{2\_000} ml^0_{00} + \varphi^1_{2\_000} ml^0_{10} = 1$; $\varphi^0_{2\_100} ml^1_{00} + \varphi^1_{2\_100} ml^1_{10} = 1$;
$\varphi^0_{2\_100} ml^0_{01} + \varphi^1_{2\_100} ml^0_{11} = 1$; $\varphi^0_{2\_010} ml^1_{00} + \varphi^1_{2\_010} ml^1_{10} = 1$;
$\varphi^0_{2\_110} ml^1_{10} + \varphi^1_{2\_110} ml^1_{11} = 1$; $\varphi^0_{2\_101} ml^1_{00} + \varphi^1_{2\_101} ml^1_{10} = 1$;
$\varphi^0_{2\_011} ml^0_{01} + \varphi^1_{2\_011} ml^0_{11} = 1$; $\varphi^0_{2\_111} ml^0_{01} + \varphi^1_{2\_111} ml^0_{11} = 1$.

The corresponding optimization problem is

$$\begin{aligned}&\varphi^0_{2\_000} ml^0_{00} + \varphi^1_{2\_000} ml^0_{10} + \varphi^0_{2\_100} ml^1_{00}\\ &+\varphi^1_{2\_100} ml^1_{10} + \varphi^0_{2\_100} ml^0_{01}\\ &+\varphi^1_{2\_100} ml^0_{11} + \varphi^0_{2\_010} ml^1_{00} + \varphi^1_{2\_010} ml^1_{10} + \varphi^0_{2\_110} ml^1_{10}\\ &+\varphi^1_{2\_110} ml^1_{11} + \varphi^0_{2\_101} ml^1_{00} + \varphi^1_{2\_101} ml^1_{10} + \varphi^0_{2\_011} ml^0_{01}\\ &+\varphi^1_{2\_011} ml^0_{11} + \varphi^0_{2\_111} ml^0_{01} + \varphi^1_{2\_111} ml^0_{11} \to \max.\end{aligned} \quad (3)$$

In addition, there are constraints due to conflicts between equations for different steps. For example, from the expression

$$\left(\begin{pmatrix}\varphi^0_{2\_000}\\ \varphi^1_{2\_000}\end{pmatrix}^{\mathrm{T}}\begin{bmatrix}\begin{pmatrix}ml^0_{00}\\ ml^1_{00}\end{pmatrix} & \begin{pmatrix}ml^0_{01}\\ ml^1_{01}\end{pmatrix}\\ \begin{pmatrix}ml^0_{10}\\ ml^1_{10}\end{pmatrix} & \begin{pmatrix}ml^0_{11}\\ ml^1_{11}\end{pmatrix}\end{bmatrix}\begin{pmatrix}1\\0\end{pmatrix}\right) = \begin{pmatrix}1\\0\end{pmatrix}$$

for the first example and

$$\left(\begin{pmatrix}\varphi^0_{2\_100}\\ \varphi^1_{2\_100}\end{pmatrix}^{\mathrm{T}}\begin{bmatrix}\begin{pmatrix}ml^0_{00}\\ ml^1_{00}\end{pmatrix} & \begin{pmatrix}ml^0_{01}\\ ml^1_{01}\end{pmatrix}\\ \begin{pmatrix}ml^0_{10}\\ ml^1_{10}\end{pmatrix} & \begin{pmatrix}ml^0_{11}\\ ml^1_{11}\end{pmatrix}\end{bmatrix}\begin{pmatrix}1\\0\end{pmatrix}\right) = \begin{pmatrix}0\\1\end{pmatrix}$$

for the second one, we write $\begin{pmatrix}\varphi^0_{2\_000}\\ \varphi^1_{2\_000}\end{pmatrix}^{\mathrm{T}}\begin{pmatrix}\varphi^0_{2\_100}\\ \varphi^1_{2\_100}\end{pmatrix} = \begin{pmatrix}1\\0\end{pmatrix}$ or

$\varphi_{2\_000}\varphi_{2\_100} = 0$; $\varphi_{2\_000}\varphi_{2\_010} = 0$; $\varphi_{2\_110}\varphi_{2\_100} = 0$;
$\varphi_{2\_000}\varphi_{2\_101} = 0$; $\varphi_{2\_110}\varphi_{2\_011} = 0$; $\varphi_{2\_110}\varphi_{2\_111} = 0$.

The solution of problem (3) is the matrix $\tilde{M}_1 = \begin{bmatrix}\begin{pmatrix}1\\0\end{pmatrix} & \begin{pmatrix}0\\1\end{pmatrix}\\ \begin{pmatrix}0\\1\end{pmatrix} & \begin{pmatrix}1\\0\end{pmatrix}\end{bmatrix}$ and the vector $\tilde{\varphi}_2 = \left(\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right)$. Based on these values of the discrete function $\varphi_2(k_{(L_2)})$, we compile a data table for the second step of the decomposition procedure of the function $f = \varphi_2(\tilde{k}_3, \varphi_3(\tilde{k}_1, \tilde{k}_2))$, placing the values of the function $\varphi_2(k_{(L_2)})$ in column $\tilde{k}_L$; see Table 2.

Table 2

Data on the second step of the decomposition procedure

q	k_2	k_3	k_4	k_L
1	0	0	0	0
2	1	0	0	1
3	0	1	0	1
4	1	1	0	0
5	1	0	1	1
6	0	1	1	1
7	1	1	1	1

Let the second-step partition be $\lambda_2 = \left(k_{(4)}, k_{(2,3)}\right)$. The equation for this partition has the form

$$\left(\begin{pmatrix}\varphi^0_{3\_00}\\ \varphi^1_{3\_00}\end{pmatrix}^{\mathrm{T}}\left[\begin{pmatrix}m2^0_{00}\\ m2^1_{00}\end{pmatrix}\ \begin{pmatrix}m2^0_{01}\\ m2^1_{01}\end{pmatrix}\right.\right.$$
$$\left.\left.\begin{pmatrix}m2^0_{10}\\ m2^1_{10}\end{pmatrix}\ \begin{pmatrix}m2^0_{11}\\ m2^1_{11}\end{pmatrix}\right]\begin{pmatrix}1\\ 0\end{pmatrix}\right)=\begin{pmatrix}1\\ 0\end{pmatrix}.$$

with the following unitary conditions: $\forall\{i,\ j\}\in\{0,\ 1\}^2$ $m2^0_{ij}+m2^1_{ij}=1$, $m2^t_{ij}\in\{0,1\}$, $\varphi^0_{3\_00}+\varphi^1_{3\_00}=1$, $\varphi^t_{3\_00}\in\{0,\ 1\}$, $\forall t\in\{0,\ 1\}$:

$$\varphi^0_{3\_00}m2^0_{00}+\varphi^1_{3\_00}m2^0_{10}=1,$$
$$\varphi^0_{3\_00}b^1_{00}+\varphi^1_{3\_00}b^1_{10}=0.$$

The resulting set of equations is

$$\varphi^0_{3\_00}m2^0_{00}+\varphi^1_{3\_00}m2^0_{10}=1;\ \varphi^0_{3\_10}m2^1_{00}+\varphi^1_{3\_10}m2^1_{10}=1;$$
$$\varphi^0_{3\_01}m2^1_{00}+\varphi^1_{3\_01}m2^1_{10}=1;\ \varphi^0_{3\_11}m2^0_{00}+\varphi^1_{3\_11}m2^0_{10}=1;$$
$$\varphi^0_{3\_10}m2^1_{10}+\varphi^1_{3\_10}m2^1_{11}=1;\ \varphi^0_{3\_01}m2^1_{10}+\varphi^1_{3\_01}m2^1_{11}=1;$$
$$\varphi^0_{3\_11}m2^1_{01}+\varphi^1_{3\_11}m2^1_{11}=1.$$

The corresponding optimization problem has the form

$$\begin{aligned}&\varphi^0_{3\_00}m2^0_{00}+\varphi^1_{3\_00}m2^0_{10}+\varphi^0_{3\_10}m2^1_{00}\\&+\varphi^1_{3\_10}m2^1_{10}+\varphi^0_{3\_01}m2^1_{00}+\varphi^1_{3\_01}m2^1_{10}\\&+\varphi^0_{3\_11}m2^0_{00}+\varphi^1_{3\_11}m2^0_{10}+\varphi^0_{3\_10}m2^1_{10}\\&+\varphi^1_{3\_10}m2^1_{11}+\varphi^0_{3\_01}m2^1_{10}+\varphi^1_{3\_01}m2^1_{11}\\&+\varphi^0_{3\_11}m2^1_{01}+\varphi^1_{3\_11}m2^1_{11}\to\max.\end{aligned}\qquad(4)$$

In addition, the analysis of conflicts between equations for different examples gives the constraints

$$\varphi_{3\_10}\varphi_{3\_00}=0;\ \varphi_{3\_01}\varphi_{3\_00}=0;\ \varphi_{3\_00}=\varphi_{3\_11}.$$

The solution of problem (4) is the matrix $\tilde{M}_2=\left[\begin{pmatrix}1\\0\end{pmatrix}\ \begin{pmatrix}0\\1\end{pmatrix}\right.$ $\left.\begin{pmatrix}0\\1\end{pmatrix}\ \begin{pmatrix}0\\1\end{pmatrix}\right]$ and the vector $\tilde{\varphi}_3=\left(\begin{pmatrix}1\\0\end{pmatrix},\begin{pmatrix}0\\1\end{pmatrix},\begin{pmatrix}0\\1\end{pmatrix},\begin{pmatrix}1\\0\end{pmatrix}\right)$.

Obviously, the discrete function $\varphi_3(\tilde{k}_1,\tilde{k}_2)$ needs no further decomposition. The result is $\tilde{M}_3=\left[\begin{pmatrix}1\\0\end{pmatrix}\ \begin{pmatrix}0\\1\end{pmatrix}\right.$ $\left.\begin{pmatrix}0\\1\end{pmatrix}\ \begin{pmatrix}1\\0\end{pmatrix}\right]$. ♦

3. BRANCH TABLE

When finding all structures on l leaves from the set $\Gamma_2(L)$, we should check equivalence groups; see the corresponding mechanism in [27]. Based on the analysis results, the number of partitions at each step can be reduced by eliminating leaf combinations non-implementable in a given scale k_L. It is convenient to summarize the results in a branch table. This is a compact step-by-step representation of leaf combinations for implementability analysis within the decomposition procedure.

If some leaf combinations are admissible by the analysis results of equivalence groups, we compile the branch table starting from the groups with two leaves, $|L_i|=2$. Taking only the admissible leaf combinations reduces the number of structures considered. Thus, we list the checked groups of three leaves. If there are no two-leaf groups among the admissible ones, the further procedure becomes pointless: it means that no terminal matrix (a matrix taking values of two leaves) can be designed within the given scale.

The admissible groups are placed in the branch table. The admissible groups consisting of two and individual leaves form combinations of three leaves, $|L_i|=3$. The resulting branches are placed in the table column corresponding to their group name. Next, the admissible groups consisting of three, two, and individual leaves form combinations with $|L_i|=4$, yielding branches of four leaves $|L_i|=4$. Note that starting from $|L_i|=4$, we consider the branches included in the equivalence group analysis results and placed in the branch table. In other words, the branches with $|L_i|>2$ are designed from the admissible branches with fewer leaves. The resulting branches are placed in the table column corresponding to their group name. The procedure continues until reaching the tree root, $|L_i|=l$. As a result, any tree $G\in\Gamma_2(L)$ whose structure $\Lambda(G)$ belongs to the list of admissible leaf groups must be considered in the IRM identification problem for the training set Q.

A certain tree $G\in\Gamma_2(L)$ consisting of admissible subbranches is considered using the branch table as follows. The partitions λ_i are sequentially taken from the table column with the largest number. The optimization problem $\underset{m_{rc},\varphi_r,\varphi_c}{\mathrm{Arg\,max}}\sum_{q\in Q_1}P(\lambda_i,q)$ is constructed for the set Q and each partition λ_i. If it has a solution within the admissible scale, the matrix M_i is saved, and the resulting values of the functions $\varphi_r(k_{(L_{ir})})$ and $\varphi_c(k_{(L_{ic})})$ are used to find a solution for the subbranches of the partitions λ_i. After examining all the partitions λ_i in the table column with the largest number, the consideration proceeds to the column with the lower number.

This approach has the following advantage: if the problem $\underset{m_{rc},\varphi_r,\varphi_c}{\mathrm{Arg\,max}}\sum_{q\in Q_1}P(\lambda_i,q)$ is unsolvable, we can exclude from further consideration the entire family of subbranches generated by the partition λ_i. For example, if there is no solution in the scale k_L for $\lambda_1=\left(k_{(1)},\ k_{(2.3.4)}\right)$ (see Table 3), we can exclude the

Table 3

An example of the branch table

#	0	1	2	3	4	5
L_i	1 2	2 3	3 4	1 2 4	2 3 4	1 2 3 4
λ_i	1 2	2 3	3 4	4, 1 2	4, 2 3 2, 3 4	3, 1 2 4 1, 2 3 4 1 2, 3 4

structures *M*1*l*1*M*2*l*4*M*3*l*2*l*3 and *M*1*l*1*M*2*l*2*M*3*l*3*l*4, neglecting decompositions $\lambda_2 = \left(k_{(3)},\ k_{(1,2)}\right)$ and $\lambda_2 = \left(k_{(1)}, k_{(2,3)}\right)$. For details, we refer to the paper [27].

CONCLUSIONS

This paper has considered an approach to designing integrated rating mechanisms based on separating decomposition. The branch table has been proposed as a decomposition scheme. In contrast to the approach described in [15], optimization problems are sequentially constructed and solved for each IRM node; according to Proposition 1, the optimization polynomial does not depend on the number of input parameters. In each node of the complete binary tree, the degree of the polynomial does not exceed 3. Due to these properties, the optimization problems are quickly solved by an optimizer. For example, Gurobi 9.5.0 [28] solves the optimization problem with 8 quadratic constraints and 16 general constraints (the first step of the second example) in 30 ms on a PC with AMD Ryzen 7 4800H processor and 16GB RAM. The proposed approach has another advantage as follows: if some step of the decomposition procedure of a discrete function yields no solutions in a given scale, this procedure becomes pointless to continue (the problem will be unsolvable in this scale). Further research will deal with sorting the most promising solutions from the general solution pool for a given function.

APPENDIX

P r o o f (the separating decomposability of some discrete function of *n* variables under unfixed parameter value scales).

As mentioned above, in some complete binary tree, an IRM is defined through a set of convolution matrices $M_f = \{M_j\}_{j \in L}$. In the IRM structure, the convolution operation $\tilde{M}_i \tilde{x}_i \tilde{y}_j$ is performed using each matrix M_i. Each individual matrix M_i implements some discrete function $\varphi_i(k_{(L_i)})$. For an arbitrary tree $G \in \Gamma_2(L)$, the indicator decomposition structure consists of $\Lambda(G) = \{L_i\}_{i \in \{1,..,l-1\}}$ so that $\forall i \in \{1,...,l-1\}$ $L_i \subseteq L$ is the set of leaves (indicators) of a subtree with the root node i. The set L_i has some subgroups $\{L_{ir}; L_{ic}\} \subset \Lambda(G) : L_{ir} \cup L_{ic} = L_i$. In general, the dimension of the matrix $\tilde{M}_i$ required to implement the discrete function $\varphi_i(k_{(L_i)})$ on the partition $\lambda_i = \left(k_{(L_{ir})}, k_{(L_{ic})}\right)$ is unknown. Based on the description of equivalence groups [27], the number of equivalence groups for the partition λ_i, and hence the corresponding dimension of the matrix $\tilde{M}_i$, cannot exceed the number of combinations encoded by the indicators of the subsets $K_{(L_{ir})} = \prod_{j \in L_{ir}} K_j$ and $K_{(L_{ic})} = \prod_{j \in L_{ic}} K_j$ for the subsets of row and column indicators, respectively. It may be necessary to place examples in the matrix M_i on different cells, each corresponding to a different equivalence group of the current decomposition step. That is, in the worst case, an individual matrix cell should be provided for each training example. In other words, any subset of variables for a discrete function contains a finite number of combinations of their values; this number is an estimate of the maximum number of matrix rows or columns. Consequently, the dimension of some discrete function $\varphi_i(k_{(L_i)})$ will not exceed $K_{(L_i)}$, $L_{ir} \cup L_{ic} = L_i$. Thus, at each decomposition step of the function *f*, the dimension of the matrix $\tilde{M}_i$, $i \in l-1$, is sufficient to implement the function $\varphi_i(k_{(L_i)})$ by construction. This matrix design approach ensures that the IRM constructed from any admissible complete binary tree will implement the given function *f*. ♦

P r o o f of Proposition 1.

Due to the unitary notation, the operation within an IRM with an individual set of indicator values is represented as some stepwise function $f(\tilde{k}_0,...,\tilde{k}_{l-1})$ decomposable in some set of partitions $\Lambda(G)$. At each decomposition step (at each node of the tree *G*), the discrete function $\varphi_i(k_{(L_i)})$ is defined on a subset of leaves L_i in the partition λ_i based on an example set Q_i constructed from the set Q

by selecting leaves of the corresponding subsets L_i. Therefore, the dimension $k_{(L_i)}$ is determined based on the subset L_i where the function $\varphi_i(k_{(L_i)})$ is defined. The function $\varphi_i(k_{(L_i)})$ is a matrix operation on an individual matrix, $\tilde{y}^{\mathrm{T}}\tilde{M}\tilde{x}$, where the vectors $\tilde{x}$ and $\tilde{y}$ have dimension κ and the matrix $\tilde{M}$ have dimension $\kappa\times\kappa$. Obviously, each such operation yields a vector of dimension $\kappa \in \mathbb{N}$ containing the components of an individual matrix cell. An example of this operation on a matrix M_i is as follows:

$$\begin{pmatrix} y^0(q^i) \\ y^1(q^i) \end{pmatrix}^{\mathrm{T}} \begin{bmatrix} \begin{pmatrix} m1^0_{00} \\ m1^1_{00} \end{pmatrix} & \begin{pmatrix} m1^0_{01} \\ m1^1_{01} \end{pmatrix} \\ \begin{pmatrix} m1^0_{10} \\ m1^1_{10} \end{pmatrix} & \begin{pmatrix} m1^0_{11} \\ m1^1_{11} \end{pmatrix} \end{bmatrix} \begin{pmatrix} x^0(q^i) \\ x^1(q^i) \end{pmatrix}.$$

That is, the final result of the operation $\tilde{y}^{\mathrm{T}}\tilde{M}\tilde{x}$ will also be a vector of dimension κ. Each component of this vector will be represented by a homogeneous polynomial of degree 3 since two branches join the matrix at each step. The vectors $\tilde{x}$ and $\tilde{y}$ can be both some functions (the components of $\varphi_i(k_{(L_i)})$) and leaves, and the values of all leaves are given by unitary vectors. Hence, the final degree of the polynomial will not exceed 3, retaining only the terms not multiplied by the zero components of the leaf vector. Each polynomial term will have the form $m_j \prod_{d=1}^{\varphi\_num} \varphi_d$, where m_j is one component of the tuple in some cell of the unitary encoded matrix $\tilde{M}$. The uniqueness of each term also follows from the essence of the described operation.

The cells of all matrices must contain unitary vectors. Therefore, each component of the vector defined by the operation $\tilde{y}^{\mathrm{T}}\tilde{M}\tilde{x}$ can be either 0 or 1. With the scheme $\tilde{y}^{\mathrm{T}}\tilde{M}\tilde{x}$ and the given values of the function *f*, we obtain the equations

$$\begin{pmatrix} y^0(q^i) \\ y^1(q^i) \end{pmatrix}^{\mathrm{T}} \begin{bmatrix} \begin{pmatrix} m1^0_{00} \\ m1^1_{00} \end{pmatrix} & \begin{pmatrix} m1^0_{01} \\ m1^1_{01} \end{pmatrix} \\ \begin{pmatrix} m1^0_{10} \\ m1^1_{10} \end{pmatrix} & \begin{pmatrix} m1^0_{11} \\ m1^1_{11} \end{pmatrix} \end{bmatrix} \begin{pmatrix} x^0(q^i) \\ x^1(q^i) \end{pmatrix} = \begin{pmatrix} K^0(q^i) \\ K^1(q^i) \end{pmatrix}.$$

That the function $\varphi_i(k_{(L_i)})$ implements a single example q from the set Q_i actually means $\tilde{y}^{\mathrm{T}}\tilde{M}_i\tilde{x} = \varphi_i(q)$. Since the vector $\varphi_i(q)$ is unitary, the resulting vector of the operation $\tilde{y}^{\mathrm{T}}\tilde{M}\tilde{x}$ has only one component equal to 1, the same as in the vector $\varphi_i(q)$; all others components must be 0. That is, $\varphi_i(q)^{\mathrm{T}}\,\tilde{y}^{\mathrm{T}}\tilde{M}\tilde{x} = 1$.

Denoting by $P(\lambda_i, q)$ the polynomial corresponding to the vector component determined by the function $\varphi_i(q)$ (must equal 1), we establish all items of Proposition 1. ♦

P r o o f of Corollary 1.

According to the proof of Proposition 1, the function $\varphi_i(k_{(L_i)})$ implements some training example $q \Leftrightarrow P(\lambda_i, q) = 1$. Hence, if all $|Q_i|$ examples are implemented, we have $\sum_{q\in Q_i} P(\lambda_i, q) = |Q_i|$. ♦

P r o o f Proposition 2.

For any $L \subset \mathbb{N}$ and $K \subset \mathbb{N}$, any set $Q \subset K^{l+1}$ according to some sequence of partitions $\Lambda(G)$ on a tree G from the set of complete binary trees $G_2(L)$, and the function $f(\tilde{k}_1,..., \tilde{k}_l)$ can be represented as a superposition of functions of fewer variables $\varphi_i(k_{(L_i)})$, $\forall i \in \{1,.., l-1\}$, defined on the datasets Q_i obtained from the set Q. Indeed, all components of the function $\varphi_i(k_{(L_i)})$, $\varphi_r(k_{(L_{ir})})$ and $\varphi_c(k_{(L_{ic})})$, are obtained from the function $\varphi_i(k_{(L_i)})$, $\forall i \in \{1,..., l-1\}$ by construction, and the values of the function $\varphi_1(k_{(L_1)})$ are defined through the set Q. ♦

P r o o f of Corollary 2.

By Proposition 2, the function $f(\tilde{k}_1,...,\tilde{k}_l)$ as a superposition of functions of fewer variables $\varphi_i(k_{(L_i)})$, $\forall i \in \{1,..., l-1\}$, can be constructed through the sequential optimization $\max \sum_{q\in Q_i} P(\lambda_i, q)$ with finding at each step the function $\varphi_i(k_{(L_i)})$, $\forall i \in \{1,..., l-1\}$. We have the values of the function $\varphi_1(k_{(L_1)})$ from the set Q; we calculate the values of the functions $\varphi_r(k_{(L_{ir})})$ and $\varphi_c(k_{(L_{ic})})$, $\forall i \in \{2,..., l-1\}$, for each decomposable function based on the function $\varphi_i(k_{(L_i)})$. If the problem is unsolvable at any step *i* for some of the $|Q_i|$ examples, it will have no solution for the components of the function $\varphi_i(k_{(L_i)})$ as well. Consequently, $IRM_{G,2}(Q) = \varnothing$. ♦

REFERENCES

1. Raschka, S., *Python Machine Learning*, Packt Publishing, 2015.
2. Knaeble, M., Nadj, M., and Maedche, A., Oracle or Teacher? A Systematic Overview of Research on Interactive Labeling for Machine Learning, in *WI2020 Zentrale Tracks*, 2020, pp. 2–16. DOI:10.30844/wi_2020_a1-knaeble
3. Simard, P., Amershi, S., Chickering, M., et al., Machine Teaching: A New Paradigm for Building Machine Learning Systems, *arXiv:1707.06742*, 2017.
4. Gorelikov, N.I., Designing a Sectoral Model for Management of Development and Manufacture of New Products, *Automation and Remote Control*, 1984, vol. 45, no. 5, pp. 598–604.
5. Burkov, V.N., Novikov, D.A., and Shchepkin, A.V., *Control Mechanisms for Ecological-Economic Systems*, Springer, 2015.
6. Korgin, N.A. and Rozhdestvenskaya, S.M., Concordant Approach for R&D Projects' Evaluation and Ranking for For-

mation of Programs for the Creation of Scientific and Technological Potential, *Proceedings of the 11th IEEE International Conference on Application of Information and Communication Technologies (AICT2017)*, Moscow, IEEE, 2017, vol. 2, pp. 358–362.
7. Shchepkin, A., Application of Integrated Mechanism in Financing Project Works, *Proceedings of the 13th International Conference on Management of Large-Scale System Development*: *(MLSD'2020)*, Moscow, 2020, pp. 1–4.
8. Zheglova, Y. and Titarenko, B., Methodology for the Integrated Assessment of Design Solutions for Foundation Pit Fences Based on the Theory of Active Systems, *IOP Conference Series: Materials Science and Engineering*, 2020, vol. 869, art. no. 052012.
9. Burkov, V., et al., Models and Management Structure for the Development and Implementation of Innovative Technologies in Railway Transportation. I. Mechanisms of Priority Projects Selection and Resource Allocation, *Automation and Remote Control*, 2020, vol 81, no. 7, pp. 1316–1329.
10. Firsova, E.A., Firsov, S.S., and Mayorova, A.N., Evaluation of the Effectiveness of Enterprise Organizational Restructuring, *Azimuth of Scientific Research: Economics and Administration*, 2017, vol. 6, no. 2(19), pp. 283–286. (In Russian.)
11. Andronnikova, N.G., Burkov, V.N., and Leont'ev, S.V., *Kompleksnoe otsenivanie v zadachakh regional'nogo upravleniya* (Complex Assessment in Regional Management), Moscow: Trapeznikov Institute of Control Sciences RAS, 2002. (In Russian.)
12. Burkov, V.N., Gorelikov, N.I., and Cherkashin, A.M., Methodological Foundations of Integrated Assessment of Enterprises Activities Considering Their Progressiveness in VPO Soyuzelektropribor, *Pribory i Sistemy Upravleniya*, 1982, no. 11, p. 21. (In Russian.)
13. Blachev, R.N., Peculiarities of the Binary Aggregation Procedure for Multi-criterial Expert Decisions, *Automation and Remote Control*, 1997, vol. 58, no. 5, pp. 817–821. (In Russian.)
14. Mariel, P., Hoyos, D., Meyerhoff, J., et al., *Environmental Valuation with Discrete Choice Experiments: Guidance on Design, Implementation and Data Analysis*, Berlin: Springer Nature, 2021.
15. Burkov, V., Korgin, N., and Sergeev, V., Identification of Integrated Rating Mechanisms as Optimization Problem, *Proceedings of the 13th International Conference on Management of Large-Scale System Development* (*MLSD'2020*), Moscow: Trapeznikov Institute of Control Sciences RAS, 2020, pp. 1–5.
16. Alekseev, A., Identification of Integrated Rating Mechanisms Based on Training Set, *Proceedings of the 2nd International Conference on Control Systems, Mathematical Modeling, Automation and Energy Efficiency* (*SUMMA*), IEEE, 2020, pp. 398–403.
17. Kolmogorov, A.N., On the Representation of Continuous Functions of Several Variables by Superpositions of Continuous Functions of Fewer Variables, *Doklady Akad. Nauk SSSR*, 1956, vol. 108, no. 2. (In Russian.)
18. Arnold, V.I., On Functions of Three Variables, *Doklady Akad. Nauk SSSR*, 1957, vol. 114, no. 4, pp. 679–681. (In Russian.)
19. Vykhovanets, V.S., Algebraic Decomposition of Discrete Functions, *Automation and Remote Control*, 2006, vol. 67, no. 3, pp. 361–392.
20. Vykhovanets, V.S., Spectral Identification of Discrete Systems, *Proceedings of the 8th International Conference "System Identification and Control Problems" (SICPRO '09)*, Moscow: Trapeznikov Institute of Control Sciences RAS, 2009, pp. 1500–1517. (In Russian.)
21. Yablonskii, S.V., On Algorithmic Difficulties of Synthesizing Minimal Contact Circuits, in *Problemy kibernetiki* (Problems of Cybernetics), Moscow: Fizmatgiz, 1959, no. 2, pp. 75–121. (In Russian.)
22. Kuznetsov, A.V., Repetition-Free Contact Schemes and Repetition-Free Superpositions of Functions of Algebra of Logic, *Trudy Mat. Inst. Steklov.*, 1958, vol. 51, pp. 186–225. (In Russian.)
23. Glotov, V.A. and Pavel'ev, V.V., *Vektornaya stratifikatsiya* (Vector Stratification), Moscow: Nauka, 1984. (In Russian.)
24. Burkova, I.V., Dranko, O.I., Kryukov, S.V., and Strukov, A.Yu., Dichotomizing Representation at the Complex Estimation of the Enterprises, *Vestnik Voronezh. Gos. Tekh. Univ.*, 2010, vol. 6, no. 11, pp. 133–136. (In Russian.)
25. Burkov, V.N., Burkova, I.V., and Popok, M.V., The Method of Dychotomic Programming, *Large-Scale Systems Control*, 2004, no. 9, pp. 57–75. (In Russian.)
26. Alekseev, A.O., Kalent'eva, A.S., Vychegzhanin, A.V., and Klimets, D.V., Algorithmic Foundations of Fuzzy Integrated Assessment Procedure for Different-Nature Objects, *Fundamental Research*, 2014, no. 3-3, pp. 469–474. (In Russian.)
27. Korgin, N.A. and Sergeev, V.A., Identification of Integrated Rating Mechanisms on Complete Data Sets, in *Advances in Production Management Systems. Artificial Intelligence for Sustainable and Resilient Production Systems: Proceedings of IFIP WG 5.7 International Conference*, Berlin: Springer, 2021, vol. 630, pp. 610–616.
28. *Gurobi Optimizer Reference Manual*, Gurobi Optimization, LLC, 2020. URL: https://www.gurobi.com

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

Received November 23, 2022, and revised December 14, 2022.
Accepted December 26, 2022.

Author information

Sergeev, Vladimir Aleksandrovich. Researcher, Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia
✉ sergeev.bureau@gmail.com

Cite this paper

Sergeev, V.A., Design of Integrated Rating Mechanisms Based on Separating Decomposition. *Control Sciences* **6,** 2–10 (2022). http://doi.org/10.25728/cs.2022.6.1

Original Russian Text© Sergeev, V.A., 2022, published in *Problemy Upravleniya*, 2022, no. 6, pp. 3–13.

Translated into English by *Alexander Yu. Mazurov*,
Cand. Sci. (Phys.–Math.),
Trapeznikov Institute of Control Sciences,
Russian Academy of Sciences, Moscow, Russia
✉ alexander.mazurov08@gmail.com

DOI: http://doi.org/10.25728/cs.2022.6.2

A STRATEGIC MANAGEMENT MODEL FOR RESTRUCTURING THE TECHNOLOGICAL CORE OF AN ECONOMY

V. B. Gusev

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

✉ gusvbr@ipu.ru

Abstract. This paper considers the multi-sector model of the technological core of an economy, mathematical methods for its analysis, and procedures for calculating an indicative plan to restructure the core. The productivity of this core is proposed as a formalized criterion (indicator) for the effectiveness of structural innovations. The following optimization problem is stated: find a balanced state maximizing productivity by planned changes in the output and price indices. An equivalent transformation method is developed for the model considering the achieved values of the indicators. Several propositions concerning the properties of equilibrium and balanced states are proved. As a result, a multistage procedure is constructed to calculate the trajectory bringing the economic system closer to a balanced state. The multi-sector model is analyzed to compare the uncontrolled and controlled modes of development. The uncontrolled mode simulates the state of a market economy: no centralized management of the economy, sustainability, and relatively low GDP growth rates. The controlled mode involves the strategic planning methodology. As shown below, due to indicative strategic planning, the productivity of Russia's economy can significantly increase even at the first plan implementation stages. The proposed indicative planning methodology is mathematically justified. Numerical examples of its implementation on real statistical data are given. According to the paper's results, centralized planning institutions should be established for developing the technological infrastructure of Russia's economy. Such institutions are of current importance due to the international economic and political situation.

Keywords: the technological core of an economy, controlled development mode, productivity optimum, equilibrium state, balanced state, equivalent transformation, indicative strategic planning, plan of restructuring.

INTRODUCTION

In the modern world, similar technological processes used by countries lead to ambiguous results: the per capita GDP of one country may differ significantly from that of another country. To some extent, this difference is due to the structural peculiarities of their economies. As noted in [1], there is an increasing awareness that the economy's structure causes the main limitations of economic growth in Russia; the matter is an ineffective structure of production, an unproductive structure of incomes, an outdated structure of exports, and an irrational regional structure of the distribution of productive forces.

Russia's economy faces different crisis phenomena. At the microlevel, they include unfavorable business conditions in the industrial sphere, a bias toward trade and services, a short lifespan of small enterprises, and a significant share of bankruptcies. At the macrolevel, they include a low GDP growth rate, the economy's critical dependence on oil and gas exports, unstable exchange rates, a small share of the manufacturing sector and the high-tech sector, the insufficient growth of its basic assets, the economy's dependence on external sanctions, and ineffective management mechanisms. At the external level, these are the special military operation in Ukraine and international economic sanctions.

Improving economic management mechanisms is the aim of the following official documents: Federal Law No. 172-FZ "On Strategic Planning in the Russian Federation" dated June 28, 2014, Presidential Decree No. 474 "On National Development Goals of the Russian Federation for the period up to 2030" dated July 21, 2020, and Presidential Decree No. 633 "On Approving the Basics of State Policy in the Sphere of Strategic Planning in the Russian Federation" dated November 8, 2021. They envisage "the introduction of modern methods of forecasting, modeling, indicative planning, balance accounting, and information technology."

The documents also note that:

- "The main tools of the strategic planning system are indicative planning, which forms a set of agreed indicators characterizing the state and goals of socio-economic development and national security, and balance accounting with the development of measures to achieve the set goals and their resource supply."
- "The scientific and methodological support of strategic planning is carried out by a specialized scientific center with the participation of scientific organizations and the federal state budgetary institution 'The Russian Academy of Sciences.'"

As a target criterion, the strategic planning methodology uses a performance indicator of the economy's technological core, further called productivity. This methodology includes methods for calculating the dynamics of several output and value parameters of economic development to increase productivity under the existing technological constraints.

1. INDICATORS AND PLANS FOR SUSTAINABLE DEVELOPMENT

In this paper, we analyze the effective use of the existing technological potential of an economic system and determine ways of its development. A possible approach to accelerate economic growth is finding a preferred structure of economic activities and ways to implement this structure. According to numerical calculations, such a possibility is justified [2].

We consider management models and methods aimed at the sustainable self-sufficient development of the technological core of an economy. For this purpose, we employ the closed input-output model of the Leontief type [3]. This model describes the dependence of costs on the output.

Crisis phenomena and external economic sanctions require effective measures to parry them, particularly by transforming the internal mechanisms of economic activity. *The technological core* of an economic system is a set of economic activities available for observation and measurement and sufficient to adequately represent the system state. According to calculations [2], due to the existing potential of the technological core, the economy's productivity can be significantly increased compared to the current level. When realized, this potential can overcompensate for the possible volume of losses from external sanctions.

The technological core model allows considering two methods of organizing reproduction, corresponding to the controlled and uncontrolled modes. In the limit, both methods lead to equilibrium states, when the structure of prices and outputs is stabilized with the same growth shares for all types of products and services. In the uncontrolled mode, each sector or activity disposes only of the funds constituting a share of its own value added. As a result, an equilibrium state with the lowest productivity value is implemented. In the controlled reproduction mode, an equilibrium state with the highest productivity value is implemented. In this mode, a stage-by-stage (indicative [4]) plan of structural changes in the output and prices is formed for each sector or activity. The funds for implementing this plan can be redistributed between the sectors.

The reproduction model [2, 5] of a multiproduct system allows defining an economic productivity indicator (the output reproduction index) as a function of structural proportions of the output and prices of the products and services of different sectors. This index reflects the ratio of the output and costs. Maximizing it, we calculate a balanced structure of the outputs and prices corresponding to an equilibrium reproduction mode and a stage-by-stage plan to achieve this goal.

Different models of an economy have different productivity. Hence, the problem is to choose the model with the highest productivity value. The planning problem has the following peculiarity: the maximum (potential) productivity of the technological core of the economic system can be achieved in different ways: by changing the output structure of the sectors, by changing the price structure, or by changing the output and price structures jointly.

However, the case of joint changes in the output and prices is of practical interest: these parameters are interconnected through market mechanisms and vary jointly: an increase in the output leads to a relative decrease in the price and vice versa.

2. OUTPUT REPRODUCTION CONDITIONS

Assume that the direct costs Z_{ij} of sector j for the products or services of type i and the outputs V_j of products or services of type j are given. These data are used to calculate the specific cost coefficients

$$a_{ij} = Z_{ij} / V_j . \quad (1)$$

The input-output model can be described by the relation

$$V_i(t) = \gamma_i \sum_{j=1}^{n} a_{ij} V_j(t),$$

where γ_i denotes the reproduction index of sector i.

Let the total costs of products or services be equal to their output in the previous stage. Then the system reproduction index γ is given by the minimum reproduction index over all types of products or services: $\gamma = \min_i \gamma_i$.

An optimization problem for the output structure V_i has the following form: find the maximum reproduction index

$$\max_{\gamma, \mathbf{V}} \gamma \qquad (2)$$

subject to the technological output constraint

$$V_i(t) \geq \gamma \sum_{j=1}^{n} a_{ij} V_j(t)$$

with the specific cost coefficients a_{ij}. If the matrix $\mathbf{A} = [a_{ij}]$ is nonsingular, the solution of this problem will coincide with the eigenvector of the matrix **A**. Indeed, the number of inequalities in the constraint equals the dimension of the output vector. Therefore, the solution of the bilinear programming problem is reached on the equality

$$V_i(t) = \gamma \sum_{j=1}^{n} a_{ij} V_j(t), \qquad (3)$$

coinciding with an eigenvector of the matrix **A**.

The outputs satisfying condition (3) in natural indicators are *equilibrium outputs*.

The constraint on the output dynamics can be written as

$$\theta V_i(t-1) \geq V_i(t) \geq \mu V_i(t-1),\ i = 1, \ldots, n, \qquad (4)$$

where $0 < \mu \leq 1 < \theta$.

The outputs satisfying conditions (3) and (4) are called *balanced*.

If the direct cost coefficient a is given by

$$\min_{a, \mathbf{V}} a,$$

$$aV_i(t) = \sum_{j=1}^{n} a_{ij} V_j(t),$$

then in the equilibrium mode, we have

$$a = 1/\gamma.$$

The equilibrium outputs corresponding to the eigenvector **x** of the matrix **A** with the maximum eigenvalue are not optimal for problem (2)–(4).

3. VALUE REPRODUCTION CONDITIONS

By assumption, the production technology in problem (2) remains invariable during one stage. In other words, the natural specific cost coefficients are supposed constant. The solution of this problem does not depend on the prices of products or services but is difficult to implement in practice: for large production systems, measurements are often performed in value terms.

Let the specific cost coefficients be determined based on value indicators rather than the natural ones (formula (1)). In other words,

$$a^c_{ij} = Z^c_{ij} / V^c_j = Z_{ij} P_i / (V_j P_j) = a_{ij} P_i / P_j,$$

where V^c_j denotes the output in value terms and P_j is the price of products of sector j. Then the output structure balancing problem for the production cycle takes the form

$$\max_{\gamma, \mathbf{V}^c} \gamma, \qquad (5)$$

$$V^c_i(t) \geq \gamma \sum_{j=1}^{n} a^c_{ij} V^c_j(t). \qquad (6)$$

The outputs in value terms satisfying conditions (5) and (6) under fixed prices are *equilibrium*. For such outputs, the sectors have the same reproduction indices.

Assume that both natural volumes and prices can be changed. Problems (5), (6) (on the one part) and (2), (3) (on the other) are not equivalent: in the former, the criterion and the output constraint have a new content. Moreover, the values of the specific cost indicators can initially be determined only at the previous stage (t – 1). But they will change with changing the prices. Therefore, to refine these indicators, we should find possible price variations.

Consider the price structure balancing problem for the production cycle. The volumes of products or services are estimated in natural terms. The total value V^c_i of final product or service i is the sum of the values of all intermediate products or services in the production cycle given the price P_j for intermediate product or service j at this stage multiplied by the value increase coefficient for the cycle period (profitability) r_i:

$$V^c_i(t) = r_i \sum_j Z_{ji}(t) P_j. \qquad (7)$$

Using this value relation and formula (1), we obtain

$$V^c_i(t) = V_i(t) P_i = r_i \sum_{j=1}^{n} a_{ji} V_i(t) P_j.$$

Let the reproduction coefficient be determined by the minimum profitability over all products or services. In this case, the price structure optimization problem is written as

$$\max_{r,\mathbf{p}} r, \tag{8}$$

$$P_i(t) \geq r\sum_{j=1}^{n} a_{ji}P_j(t). \tag{9}$$

The prices satisfying conditions (8), (9) under fixed outputs are *equilibrium prices*. For these prices, all sectors have the same profitability. The additional constraint has the form

$$\theta P_i(t-1) \geq P_i(t) \geq \mu P_i(t-1),\ i=1,...,\ n.$$

Its left-hand side restricts possible price reductions and the rate of inflation to the level θ by the types of products or services. Such prices, also satisfying condition (9), are called *balanced*.

4. THE JOINT BALANCING OF OUTPUT AND PRICE REPRODUCTION

Proposition 1. *The maximum values of the criteria γ and r for problems* (2), (3) *and* (8), (9) *coincide. The solution vectors* **V** *and* **P** *are ambiguous: they are determined within a scalar multiplier.*

This proposition is proved in the Appendix.

To calculate the eigenvector and eigenvalue of the positive Schur matrix **A**, we employ the iterative procedure

$$\mathbf{x}^0 = (1,\ 1,...,\ 1),$$

$$\mathbf{x}^{k+1} = \mathbf{A}\mathbf{x}^k / \|\mathbf{x}^k\|,\ k = 0,\ 1,\ 2,\ ... \tag{10}$$

The eigenvalue of the matrix **A**, calculated as the limit of $\|\mathbf{x}^k\|$, is the largest one among all its eigenvalues [6]. The process (10) simulates the uncontrolled reproduction mode for the technological core of the economy when interpreting $\mathbf{x}^k$ as the vector of output or price indices at stage *k*.

Proposition 2. *Let* **A** *be a nonnegative matrix with a*, $|a| < 1$, *as the maximum eigenvalue by magnitude. The eigenvector* **x** *of this matrix satisfies the equation*

$$\mathbf{A}\mathbf{x} = a\mathbf{x}.$$

The eigenvector **x** with the maximum eigenvalue is found using the iterative procedure

$$\mathbf{x}^{k+1} = \mathbf{A}\mathbf{x}^k / \|\mathbf{x}^k\|,$$

where *k* denotes the iteration number. The stopping condition has the form $\|\mathbf{x}^{k+1} - \mathbf{x}^k\| < \varepsilon$, where ε is a given accuracy of calculations (e.g., ε = 0.001). The eigenvalue *a* of the matrix **A** is

$$a = \lim_{k\to\infty} \|\mathbf{x}^k\|.$$

This proposition is proved in the Appendix.

The outputs and prices yielded by this algorithm are equilibrium.

To obtain the output vector maximizing the productivity estimate π for the technological core, we apply the gradient descent method.

We introduce the following notations: **I** and **E** are a unit vector and an identity matrix of appropriate dimensions, respectively; *h* is the step value;

$$a(\mathbf{x}) = \|\mathbf{A}\mathbf{x}\| / \|\mathbf{x}\|.$$

On the corresponding rectangular domain, the gradient descent-based procedure to find

$$\min_{\mu \leq xi \leq \theta,\ i=1,..n} a\ (\mathbf{x})$$

has the form

$$\mathbf{x}^{k+1} = \mathbf{x}^k + h(a(\mathbf{x}^k)\mathbf{E} - \mathbf{A}^{\mathrm{T}})\mathbf{I} / \|\mathbf{x}^k\|,$$

where *k* is the iteration number.

On the other hand, the resulting optimal solution may be unbalanced.

Proposition 3. *Let*

$$a(\mathbf{x}) = \|\mathbf{A}\mathbf{x}\| / \|\mathbf{x}\|.$$

It is possible to choose a step h so that the gradient projection-based procedure

$$x_i^{k+1} = \min(\theta,\ \max(\mu,\ z_i^k)),\ i = 1,\ n,$$

where

$$z^k{}_i = a(\mathbf{y}^k)\sum_j a_{ij}y^k{}_j,\ i = 1,n,$$

$$\mathbf{y}^k = \mathbf{x}^k + h(a(\mathbf{x}^k)\mathbf{E} - \mathbf{A}^{\mathrm{T}})\mathbf{I} / \|\mathbf{x}^k\|,$$

will weakly converge to the conditional optimum

$$\min_{\mu \leq \mathbf{x} \leq \theta} a(\mathbf{x})\ .$$

This proposition is proved in the Appendix.

The process of optimizing the output structure generates a sequence of increasing estimates of the technological core productivity, accompanied by a sequence of positive output increments for some sectors. At the same time, for some sectors, positive output increments do not lead higher productivity estimates. The rate of convergence depends on the step *h*: initially, the rate grows as the step increases; then, it starts decreasing.

The resulting solution is balanced but not equilibrium.

Note that the specific cost coefficients are measured at the previous stage of the production cycle. Hence, the problem of balancing the price proportions in the nonequilibrium mode can be written as

$$\left.\begin{aligned} &\max_{r,\mathbf{P}} r \\ &p_i(t) \geq r\sum_{j=1}^{n} a^c_{ji} p_j(t) \\ &\theta \geq p_i(t) \geq \mu,\ i = 1,..., n \end{aligned}\right\} \qquad (11)$$

where $p_i(t) = P_i(t)/P_i(t-1)$ is the price index of product i.

In practice, this approach requires stage-by-stage implementation. For each stage, an additional constraint should be introduced on the deviation from the existing structure. This constraint should agree with the admissible rate of socio-economic processes.

The initial information on the structural balancing problem is formed by analyzing economic statistics. In several situations (e.g., regional or sectoral planning, scenario forecasting), there are no standard methods for collecting and processing data. In the case of such difficulties, statistical data can be combined with expert assessments of the cost structure per unit of output.

Let the solution of the balancing problem (11) be used to tune the specific cost coefficients. In this case, the repeated solution will give price indices in the range $[\mu,\ \theta]$; moreover, the closer the solution is to the equilibrium state, the smaller increments the growth coefficient r will have. When recalculating the specific cost coefficients, the sectors with small marginal prices may receive the corresponding increments. In addition, the updated direct cost coefficients can be used in the output balancing problem (2), (3).

5. THE INDICATIVE PLAN-FORECAST OF ECONOMIC DEVELOPMENT IN OUTPUT INDICATORS

As a rule, it is impossible to implement a change in the output structure *per saltum*, making all outputs equilibrium. To determine the most rational plan for sectoral development, consider the problem statement

$$\max_{\gamma,\mathbf{V}} \gamma,$$

$$V_i(t) \geq \gamma\sum_{j=1}^{n} a_{ij} V_j(t),$$

$$\theta V_i(t-1) \geq V_i(t) \geq \mu V_i(t-1),\ i = 1,..., n.$$

It includes a technological constraint on the outputs and an output growth condition at the rate θ per one planning stage.

We introduce the following notations: $\mathbf{D}$ is a diagonal matrix with the diagonal elements $V_1, V_2, .. V_n$; $\mathbf{C}$ is a diagonal matrix with the diagonal elements $1/V_1, 1/V_2, ..., 1/V_n$.

Varying outputs lead to changes in the specific cost estimates a_{ij}. To fix the results V_i of the previous stage, we recalculate the direct cost coefficients:

$$\overline{a_{ij}} = a_{ij} \cdot V_i / V_j \text{ or } \overline{\mathbf{A}} = \mathbf{DAC}.$$

The following result was established in the paper [7]; see Proposition 1 therein.

Proposition 4. *If all* $V_i \neq 0,\ i = 1,.., n$, *then the transformed matrix with the coefficients* $\overline{a_{ij}} = a_{ij} \cdot V_i / V_j$ *has the same eigenvalues as the matrix* $\mathbf{A}$, *and its eigenvector equals the original one up to the deformation* $\mathbf{D}$.

The latter transformation does not affect the spectrum of the direct cost (technological) matrix $\mathbf{A}$, and its eigenvectors are preserved within the deformation $\mathbf{D}$. Repeating the described procedures (finding the optimal solution and recalculating the direct cost matrix) stage-by-stage, we obtain an indicative multistage plan-forecast of the joint development of different sectors in the technological core of the economy.

The indicative planning procedure involves absolute and relative outputs. At stage 1, we use the absolute output vector $\mathbf{V}^0$ for passing to the relative outputs $\mathbf{v}^1$ by transforming the technological matrix:

$\mathbf{A}^0 = \mathbf{D}^0\mathbf{A}\mathbf{C}^0$, where $\mathbf{D}^0 = \text{diag}(\mathbf{V}^0)$ and $\mathbf{C}^0 = (\mathbf{D}^0)^{-1}$.

Then we obtain the relative output vector $\mathbf{v}^k$ by solving the problem

$$\left.\begin{aligned} &\max_{\mathbf{v}^i} \gamma, \\ &\mathbf{v}^k \geq \gamma^k \mathbf{A}^k \mathbf{v}^k, \\ &\mu\mathbf{I} \leq \mathbf{v}^k \leq \theta\mathbf{I},\ k = 1, 2... \end{aligned}\right\} \qquad (12)$$

where $\mathbf{A}^k = \mathbf{D}^{k-1}\mathbf{A}^{k-1}\mathbf{C}^{k-1}$, $\mathbf{D}^i = \text{diag}(\mathbf{v}^i)$, and $\mathbf{C}^i = (\mathbf{D}^i)^{-1}$.

The relative output vector $\mathbf{v}^k$ (the solution of problem (12)) at stage k is called a local equilibrium under the technological output constraint and the relation output growth condition with the rate θ, $\theta > 1$, per one planning stage. (Recall that $\mathbf{I}$ denotes a unit vector of an appropriate dimension.) If the problems have a solution at stage $k \geq 1$, the indicative outputs in absolute terms will be

$$\mathbf{V}^k = \prod_{j=k}^{1} \text{diag}(\mathbf{v}^j)\,\mathbf{V}^0.$$

The following result was established in the paper [7]; see Proposition 3 therein.

Proposition 5. *The sequence* $\mathbf{V}^k$ *tends to the eigenvector of the technological matrix* $\mathbf{A}$, *and the estimate* γ^k *tends to the eigenvalue of this matrix.*

Remark. When transforming the technological matrix $\mathbf{A}$ by the deformation $\mathbf{D} = \mathrm{diag}(\mathbf{v}^*)$, after reaching the output vector $\mathbf{v}^*$, further planning becomes trivial: $\mathbf{v} = c\mathbf{I}$, where $\theta \geq c \geq \mu$. In other words, the output structure will no longer change after reaching the technological equilibrium.

6. THE COMBINED PLAN-FORECAST OF ECONOMIC DEVELOPMENT IN OUTPUT AND PRICE INDICATORS

An isolated indicative change in the prices and outputs is of little practical importance: in reality, these parameters vary simultaneously. Equilibrium outputs are unstable under fixed non-equilibrium prices, like equilibrium prices under fixed outputs. A concerted equilibrium pair of the prices and outputs is stable: in this case, the profitability r and the reproduction indices γ have the same values for all sectors, leading to the sustainable development of the economic system in the long run. However, at the initial stages of eliminating the disproportions, the sectoral indicators change very unevenly.

Consider the process of calculating the indicative dynamics of output and price indicators leading to a joint equilibrium of the output and prices.

The transformation procedures

$$\mathbf{A}^k = \mathbf{D}^{k\text{-}1}\mathbf{A}^{k\text{-}1}\mathbf{C}^{k\text{-}1}$$

for the matrices $\mathbf{A}$ and $\mathbf{A}^{\mathrm{T}}$ and the solution procedures for problems (10), (11) are performed sequentially under a given upper bound $\theta > 1$ for the profitability r and the reproduction index γ. Note that:

$$\mathbf{D}^j = \mathrm{diag}(\mathbf{v}^j) \text{ and } \mathbf{C}^i = (\mathbf{D}^i)^{-1}$$

in the output problem;

$$\mathbf{D}^j = \mathrm{diag}(\mathbf{p}^j) \text{ and } \mathbf{C}^i = (\mathbf{D}^i)^{-1}$$

in the price problem.

At stage k, *the intermediate cost deflator* for the price index vector $\mathbf{p}$ is calculated as

$$d = \sum_{i,j} a_{ij}^k p_i^k \Big/ \sum_{i,j} a_{ij}^k.$$

At stage k, *the intermediate cost index* for the output index vector $\mathbf{v}$ is calculated as

$$w = \sum_{i,j} a_{ij}^k v_j^k \Big/ \sum_{i,j} a_{ij}^k.$$

The productivity π of the economic system (the value added divided by the intermediate costs) has the simple relation

$$\pi = \gamma - 1$$

to the output reproduction coefficient γ; for details, see the paper [7].

The same relation holds for the profitability r:

$$\pi = r - 1.$$

Due to Proposition 1, both values coincide in the equilibrium mode.

7. INDICATIVE PLANNING OF OUTPUT PROPORTIONS

In calculations, the values V_i are interpreted as proportions of the volume of transport services; under the assumed nondecrease property, the constraints imposed on them have the form

$$V_i(t) \geq 1,\ i = 1, \ldots, n.$$

The equilibrium output proportions may significantly differ from the existing ones. Therefore, we solve a series of optimization problems with θ close to unity (θ = 1.2). In other words, the proportions can grow at most by 20%. Consider the statistical data on Russia's multisector economy [8]: the supply and use tables for 2019. For these data, the curves of indicative dynamics of 61 output indices were obtained under fixed prices. In several stages, the curves lead to a balanced price structure. We present a few curves below for demonstration purposes. (These are the first dozen of non-constant curves in succession.)

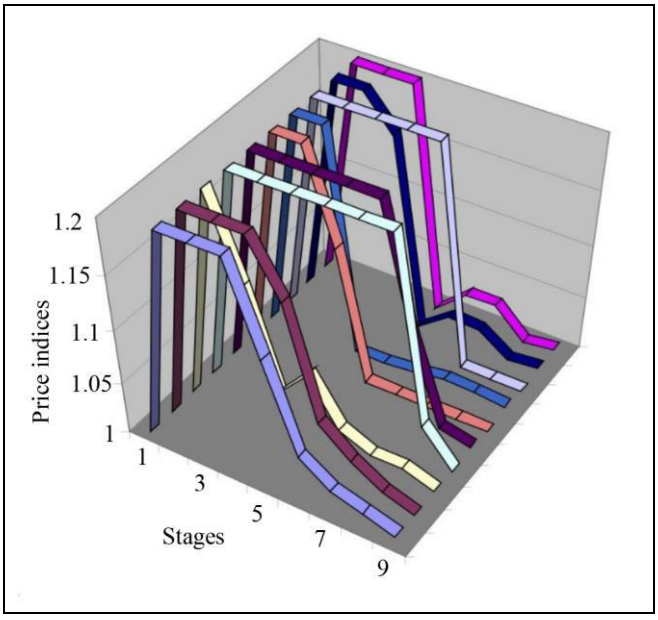

Fig. 1. Indicative dynamics of output indices with an upper bound of 1.2 for different sectors:
- Products of forestry, logging and related services;
- Products of textiles, clothing, and leather;
- Paper and paper products;
- Printing and recording services;
- Basic pharmaceutical products and pharmaceutical preparations;
- Rubber and plastic products;
- Other non-metallic mineral products;
- Electrical equipment;
- Machinery and equipment n.e.c.;
- Repair and installation services of machinery and equipment.

The indicative productivity dynamics are shown below. Figure 2 shows the productivity graph when repeatedly solving problem (7) for Russia's economy data. (The upper bound of the price indices is 1.2, and productivity is measured in percent.)

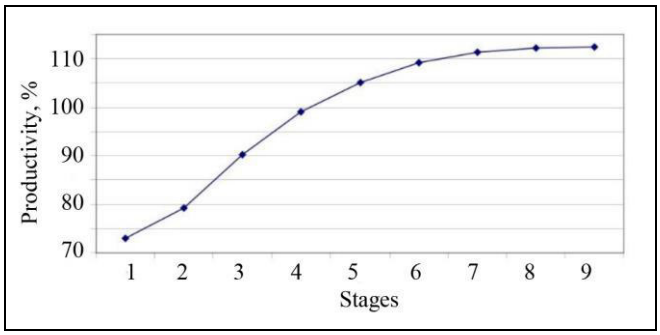

Fig. 2. Productivity under repeated output balancing.

This graph gives an idea of how far the initial state of the economy (the solution at stage 1) is from the equilibrium state (the asymptote for the repeated solution). In addition, the slope of the curve can be used to judge the sustainability of the economy. The closer the economic structure is to the equilibrium state, the higher its tolerance will be to price changes. According to Figs. 1 and 2, the closer the price structure is to the balanced one, the less the prices and productivity index will change in one stage.

Note that the output indices are stabilized at the upper bound of 1.2. This result qualitatively differs from the calculations [7] for 2016, where the output indices were stabilized at a level of 1. The reason lies in the nonunique equilibrium state and changes in the optimization algorithm. In the latter (second) case, the minimum index value was chosen. Due to the nonuniqueness of the optimal solution, both algorithms yield the same optimal value of the objective function. However, the index values of the second algorithm are preferable from the implementation point of view. Moreover, in the first case, the productivity index dynamics reach the target at a lower level (112% vs. 153%). This can be explained as follows: the problem dimension (the number of activities) was 95 (the second case) vs. 61 (the first case), which obviously reduces the choice space of planning.

8. THE JOINT DEVELOPMENT STRATEGY OF THE TECHNOLOGICAL CORE

Several constrained optimization problems with given admissible ranges of the indices are solved to obtain an indicative plan considering the joint change in the outputs and prices. If the stage duration is 1 year, the variation limits of the output indices are chosen depending on the possibility to invest in the stock formation in the current planning year. The variation limits of the price indices are determined by the allowable inflation or deflation requirements. The lower limit of the indices is set equal to 1: according to the plan, the output and prices for the sectors should not decrease.

Consider an example of joint planning of the outputs and prices. Below we present the calculation results for 10 products according to the All-Russian Classification of Products by Economic Activities (OKPD) with index values differing from 1.

For the data [8] on Russia's multisector economy, we applied the method described above to obtain indicative dynamics curves for 60 output indices and 60 price indices. These curves lead to a balanced structure in several stages, which agrees with the highway property of optimization models of economic dynamics [9]. The standard software package for mathematical programming problems [10–13] gives similar results. The plan was calculated under the interval constraint [1, 1.2] for the output indices (Fig. 3) and the price indices (Fig. 4). For demonstration purposes, only the first dozen of non-constant curves in succession is provided for each group of indices.

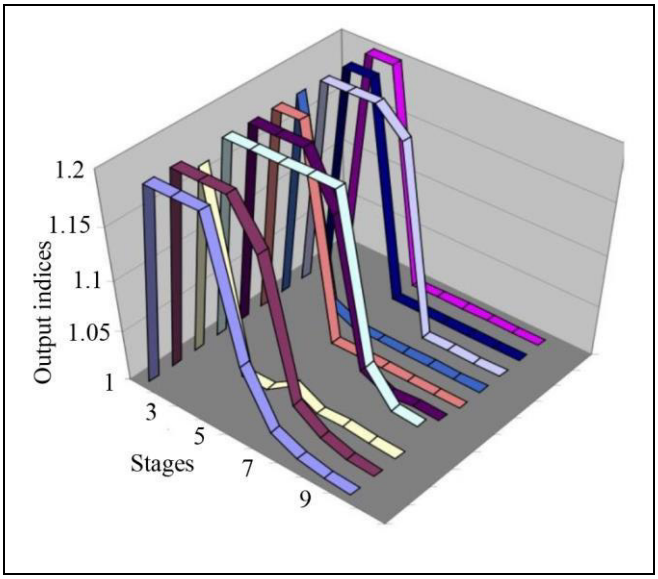

Fig. 3. Dynamics of output indices on range [1, 1. 2] for different sectors:

- Products of forestry, logging and related services;
- Products of textiles, clothing, and leather;
- Paper and paper products;
- Printing and recording services;
- Basic pharmaceutical products and pharmaceutical preparations;
- Rubber and plastic products;
- Other non-metallic mineral products;
- Electrical equipment;
- Machinery and equipment n.e.c.;
- Repair and installation services of machinery and equipment.

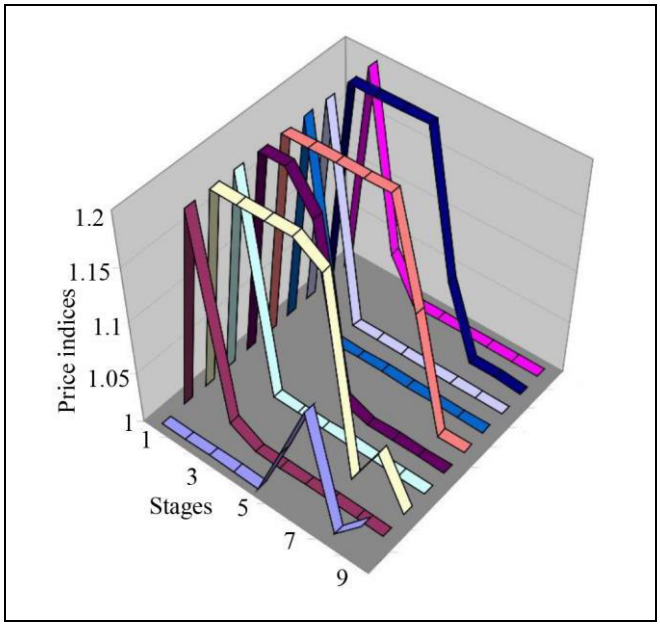

Fig. 4. Dynamics of price indices on range [1, 1.2] for different sectors:
- Products of forestry, logging and related services;
- Products of textiles, clothing, and leather;
- Paper and paper products;
- Printing and recording services;
- Basic pharmaceutical products and pharmaceutical preparations;
- Rubber and plastic products;
- Other non-metallic mineral products;
- Electrical equipment;
- Machinery and equipment n.e.c.;
- Repair and installation services of machinery and equipment.

According to the figures, the prices rose in the sectors with nonincreasing output plans. In other words, market pricing holds considering the allowable inflation.

Figure 5 shows the productivity curve when solving the stage-by-stage planning problem for Russia's economy data. (Productivity was measured in percent.)

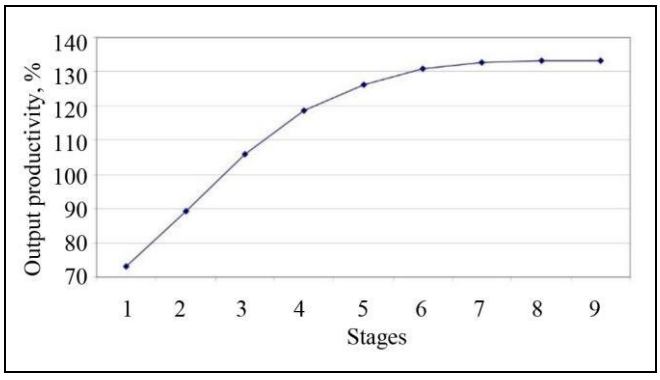

Fig. 5. Productivity under multistage output and price balancing.

The graph gives an idea of how far the initial state of the economy (the solution at the initial step was 90%) from the equilibrium state (the asymptote for the repeated solution is close to 190%). In addition, the slope of the curve can be used to judge the sustainability of the economy.

This graph gives an idea of how far the initial state of the economy (the solution at the initial stage, 90%) is from the equilibrium state (the asymptote for the repeated solution, close to 190%). In addition, the slope of the curve can be used to judge the sustainability of the economy.

According to Fig. 5, the closer the price structure is to the equilibrium state, the higher tolerance the economy will have to price changes. Also, see Fig. 5, the first few stages of the indicative plan are the most effective. Let us compare these results with the productivity dynamics for the isolated output changes (Fig. 2). Obviously, the marginal level of productivity is higher in the case of joint changes in the prices and output. This can be explained as follows: the marginal level of productivity under output variations only corresponds to the optimum on the smaller-dimension admissible set of the parameter space; a larger value of the optimum is obtained when passing to the additional variation of prices.

Price changes cause inflation. Figure 6 shows the deflator dynamics for the entire list of products and services.

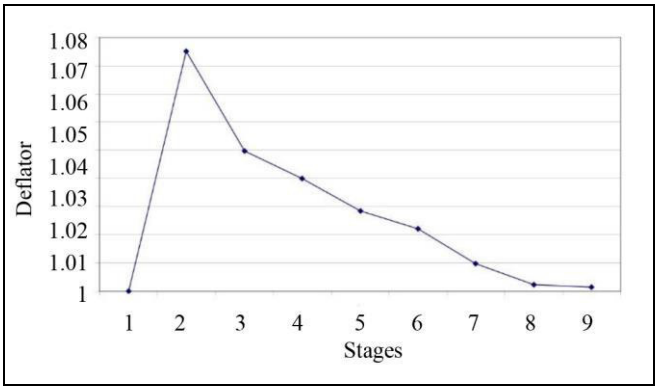

Fig. 6. The deflator under multistage output and price balancing.

The price growth constraint in the initial stages can be toughened to reduce the deflator peak. As a result, the growth of productivity at these stages will decrease.

CONCLUSIONS

This paper has considered analysis methods for control mechanisms of developing systems in crisis situations. Methods for assessing productivity indicators in natural and value terms have been developed within the economic system operation model in the autonomous mode. The procedure for calculating the

indicative dynamics of price and output proportions in the autonomous mode has been described. An illustrative example has been given: calculation results for the structure of Russia's intersectoral balance.

Despite the existing high potential for development, Russia's economy faces crisis phenomena. These factors cause the incomplete realization of the potential of its technological core [5]. The results of this paper demonstrate the possibilities of increasing the efficiency of the economy through the systematic structural modification of the technological core. Proceeding from the Rosstat data, the possible growth of productivity of the economy can be exceeded more than 2 times. Such possibilities should be implemented by elaborating strategic economic development plans. Along with choosing priority directions of development of the technological core, it is necessary to apply adequate forecasting methods for multisectoral dynamics considering the main aspects of economic activity: asset formation, accumulation, final consumption of the state and households, and export-import flows [14, 15]. New-level planning also involves appropriate organizational and institutional mechanisms [16].

Unlike a directive plan, an indicative plan is nonbinding. If any sectors deviate from the plan, it will be recalculated to reflect the changed circumstances. The new plan and estimates of under-received benefits are reported to economic agents. Within fixed technological outputs, the stage of controlled (indicative plan-based) restructuring takes a limited time. The additional funding for asset formation and wages can be provided either by the sectors' own value added or by redistributing the funds between different sectors. In the latter case, the restructuring period may be shorter. The strategic plan is considered fulfilled after the restructuring period. However, the emergence of new technologies and refined data, as well as the adjustment of structural constraints on the outputs and prices, may require elaborating a new strategic plan. Therefore, the indicative plan should be corrected on a regular basis.

Adequate assessments of the socio-economic state of national and global systems and the consequences of control impacts are required at the strategic decision level (the state and interstate affairs). The following elements should be assessed:

– the imbalance of the national economy,
– the current and potential state of intersectoral interaction,
– strategic decision-making under instability,
– structural innovation planning.

The results above illustrate the specifics of the proposed methodology with relevant calculation and analysis tools. This class of problems has several peculiarities. Here, high-dimensional models are of real interest. (The analysis may cover hundreds of economic activities.) Besides, the practical application of technological core models requires effective algorithms for solving mathematical programming problems [10, 12, 17] and linguistic calculation management means, particularly their integration into the working environment [11, 13]. Free access is needed to actual verified data and modern information technology, including an appropriate computing environment and interface devices. The paper [18] described the application of a similar open-access toolkit (Thread Pool Executor of Akka) for the processing of high-dimensional problems.

APPENDIX

P r o o f of Proposition 1.

Indeed, the maximum values of the criteria γ and r are the eigenvalues of the matrices $\mathbf{A}$ and $\mathbf{A}^{\mathrm{T}}$, respectively. Let $\mathbf{V}$ and $\mathbf{P}$ be the solutions of problems (2) and (8) with constraints (3) and (9), respectively. Since these constraints are homogeneous, the solutions under consideration are determined within the multiplier and, hence, are ambiguous.

The characteristic polynomials for both matrices coincide:

$$L(\gamma) = (1-\gamma a_{11})(1-\gamma a_{22})...(1-\gamma a_{nn}) -$$
$$a_{12}a_{21}(1-\gamma a_{33})... - a_{13}a_{31}(1-\gamma a_{22}) - ...,$$
$$L(r) = (1-ra_{11})(1-ra_{22})...(1-ra_{nn}) -$$
$$a_{12}a_{21}(1-ra_{33})... - a_{13}a_{31}(1-ra_{22}) - ...,$$

where γ is the eigenvalue of the matrix $\mathbf{A}$ and r is the eigenvalue of the matrix $\mathbf{A}^{\mathrm{T}}$. ♦

P r o o f of Proposition 2.

Indeed, consider the transformation $\mathbf{y} = \mathbf{Ax}/\|\mathbf{x}\|$. We decompose the vector $\mathbf{x}$ with respect to the eigenvectors $\mathbf{s}^i$, $i = 1,.., n$, of the matrix $\mathbf{A}$ corresponding to the eigenvalues λ_i:

$$\mathbf{x} = \sum_{i=1}^{n} b_i \mathbf{s}^i.$$

Then

$$\mathbf{Ax} = \sum_{i=1}^{n} b_i \mathbf{A}\mathbf{s}^i = \sum_{i=1}^{n} b_i \lambda_i \mathbf{s}^i$$

and

$$\mathbf{Ax}/\|\mathbf{x}\| = \frac{\sum_{i=1}^{n} b_i \lambda_i \mathbf{s}^i}{\left\| \sum_{i=1}^{n} b_i \mathbf{s}^i \right\|}.$$

By assumption, the matrix spectrum satisfies the condition $|\lambda_i| < 1$, $i = 1,..., n$. Hence, this transformation is a contraction, and $\mathbf{s}^1$ is a fixed attraction point of this transformation.

If the initial approximation has nonnegative components, all subsequent iterations will give the same result as the obtained eigenvector. The eigenvectors $\mathbf{s}^i$ correspond-

ing to the other eigenvalues λ_i, $i \neq 1$, contain negative components because they are orthogonal to $\mathbf{s}^1$. Therefore, due to the nonnegativity of the matrix $\mathbf{A}$,

$$|\lambda_1| = \|\mathbf{A}\mathbf{s}^1\| / \|\mathbf{s}^1\| > \|\mathbf{A}\mathbf{s}^i\| / \|\mathbf{s}^i\| = |\lambda_i|, \; i \neq 1.$$

Thus, deviations from the equilibrium vector $\mathbf{s}^1$ will decrease the estimate $\|\mathbf{A}\mathbf{x}\| / \|\mathbf{x}\|$ for $\mathbf{x} \neq \mathbf{s}^1$; due to the contraction property of the operator, this estimate will increase at the subsequent iterations. Therefore, the iterative process will converge to the vector $\mathbf{s}^1$ and the value λ_1. ♦

P r o o f of Proposition 3.

Indeed, the gradient of $a(\mathbf{x})$ has the form

$$\nabla_x a(\mathbf{x}^k) = -(a(\mathbf{x}^k)\mathrm{I} - \mathbf{A}^{\mathrm{T}}\mathbf{I}) / \|\mathbf{x}^k\|,$$

$\mathbf{z}^k$ is the projection of the point $\mathbf{y}^k$ on the technological constraint

$$\mathbf{y}^k = a\mathbf{A}\mathbf{y}^k,$$

and $\mathbf{x}^{k+1}$ is the projection of the point $\mathbf{z}^k$ on the constraint

$$\mu \leq x_i \leq \theta, \; i = 1, \ldots, n.$$

Since $\nabla_x a(\mathbf{x}^k) \neq 0$, the projection operators are monotonic, and there exists a step $h > 0$ such that the sequence $a(\mathbf{x}^k)$ will monotonically decrease. Moreover, because the admissible domain is a compact set, the limit point $a(\mathbf{x}^*)$ will be bounded and represent a constrained optimum. ♦

REFERENCES

1. Uzyakov, M.N., Problems of Economic Measurements and Possibilities of Structural Analysis, *Studies on Russian Economic Development*, 2020, vol. 31, no 1, pp. 3–4. (In Russian.)
2. Gusev, V.B., Equilibrium Models of Multi-Resource Self-Developing Systems, *Control Sciences*, 2007, no. 3, pp. 18–24. (In Russian.)
3. Leontief, W.W., *Essays in Economics. Theories, Theorizing, Facts, and Policies*, New York: Oxford University Press, 1966.
4. *Indikativnoe planirovanie i provedenie regional'noi politiki* (Indicative Planning and Regional Policy Implementation), Abdikeev, M.N., et al., Moscow: Finansy i Statistika, 2007. (In Russian.)
5. Gusev, V.B., Models of Autonomous Control in the Developing Systems, *Control Sciences*, 2018, no. 6, pp. 2–17. (In Russian.)
6. Polyak, B.T., Khlebnikov, M.V., and Rapoport, L.B., *Mathematicheskaya teoriya avtomaticheskogo upravleniya* (Mathematical Theory of Automatic Control), Moscow: URSS, 2019. (In Russian.)
7. Gusev, V.B., The Technological Core Model of a Large-Scale Economic System: Optimal Characteristics, *Control Sciences*, 2021, no. 6, pp. 25–33.
8. *Supply and Use Tables of the Russian Federation for 2019* (*in Current Prices, mil. rub.*). Published by the Federal State Statistics Service of the Russian Federation, January 26, 2022. https://rosstat.gov.ru/statistics/accounts. (In Russian.)
9. Dorfman, R., Samuelson, P.A., and Solow, R.M., *Linear Programming and Economic Analysis*, New York: McGraw-Hill, 1958.
10. Fox, W.P. and Burks, R., Mathematical Programming: Linear, Integer, and Nonlinear Optimization in Military Decision-Making, in *Applications of Operations Research and Management Science for Military Decision Making*, New York: Springer, 2019, pp. 137–191.
11. Mason, A.J., OpenSolver - An Open Source Add-in to Solve Linear and Integer Progammes in Excel, *Operations Research Proceedings 2011*, Klatte, D., Lüthi, H.-J., and Schmedders, K., Eds., Berlin-Heidelberg: Springer, 2012, pp. 401–406. http://dx.doi.org/10.1007/978-3-642-29210-1_64.
12. Bergstra, J., Bardenet, R., Bengio, Y., and Kégl, B., Algorithms for Hyper-parameter Optimization, *Proceedings of the 24th International Conference on Neural Information Processing Systems (NIPS'11)*, Red Hook, NY, USA: Curran Associates Inc., 2011, pp. 2546–2554.
13. Doumic, M., Perthame, B., Ribes, E., et al., Toward an Integrated Workforce Planning Framework Using Structured Equations, *European Journal of Operational Research*, 2017, vol. 262, iss. 1, pp. 217–230.
14. Samuelson, P.A., *Economics*, New York: McGraw-Hill, 1989.
15. Petrov, A.A., Pospelov, I.G., and Shananin, A.A., *Opyt matematicheskogo modelirovaniya ekonomiki* (An Experience of Mathematical Modeling of an Economy), Moscow, Energoatomizdat, 1996. (In Russian.)
16. Antipov, V.I., *GOSPLAN. Vchera, Segodnya, zavtra* (State Planning: Yesterday, Today, Tomorrow), Moscow: Kontseptual, 2019. (In Russian.)
17. Polyak, B.T., *Introduction to Optimization*, Optimization Software, 1987.
18. Hai, T.N., Tien, V.D., and Csaba, R., Optimizing the Resource Usage of Actor-based Systems, *Journal of Network and Computer Applications*, 2021, vol. 190:103143. DOI: https://doi.org/10.1016/j.jnca.2021.103143.

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

Received August 13, 2022, and revised November 6, 2022. Accepted November 23, 2022

Author information

Gusev, Vladislav Borisovich. Cand. Sci. (Phys.–Math.), Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia
✉ gusvbr@ipu.ru

Cite this paper

Gusev, V.B., A Strategic Management Model for Restructuring the Technological Core of an Economy, *Control Sciences* **6,** 11–20 (2022). http://doi.org/10.25728/cs.2022.6.2

Original Russian Text © Gusev, V.B., 2022, published in *Problemy Upravleniya*, 2022, no. 6, pp. 14–25.

Translated into English by *Alexander Yu. Mazurov*,
Cand. Sci. (Phys.–Math.),
Trapeznikov Institute of Control Sciences,
Russian Academy of Sciences, Moscow, Russia
✉ alexander.mazurov08@gmail.com

DOI: http://doi.org/10.25728/cs.2022.6.3

FUZZY VOLATILITY MODELS WITH APPLICATION TO THE RUSSIAN STOCK MARKET

V.A. Sviyazov

National Research University Higher School of Economics, Moscow, Russia

✉ v.sviyazov.96@gmail.com

Abstract. Volatility modeling and forecasting is a topical problem both in scientific circles and in the practice. This paper develops an approach combining the GARCH model and fuzzy logic. The Takagi–Sugeno fuzzy inference scheme is adopted to fuzzify an original autoregression model (the conditional heteroskedasticity model). As a result, several different local GARCH models can be used in different input data domains with soft switching between them. This approach allows considering such phenomena as volatility clustering and asymmetric volatility (the properties of real financial markets). The proposed algorithm is applied to the historical values of the RTS Index and compared with the classical GARCH model. As demonstrated below, in several cases, fuzzy models have advantages over traditional ones, namely, higher forecasting accuracy. Thus, the proposed method should be considered among others when modeling the volatility of the Russian financial market instruments: it demonstrates qualities superior to the conventional counterparts.

Keywords: fuzzy systems, forecasting, time series, volatility.

INTRODUCTION

The translation of several well-known econometric models from the probabilistic to fuzzy formulation has recently attracted increasing attention; for example, see the papers [1–9]. This also applies to volatility modeling, an important problem from both scientific and applied points of view. Many models imply a single dependence of output values on input data, but the factual dependence may vary in different domains of the input variables. Introducing fuzziness into the system eliminates this drawback by using several local models and aggregating them through fuzzy membership functions.

Volatility modeling somewhat differs from other problems of financial econometrics: unlike, e.g., the prices of financial instruments or interest rates, volatility is unobservable. Moreover, there is no consensus on what to call volatility. However, the capability to calculate the current volatility of an instrument or portfolio and forecast its future value correctly is crucial for financial institutions: it underlies market risk assessment. In turn, a sufficiently accurate assessment of market risks improves the stability of a financial institution and allows avoiding fatal losses in turbulent markets.

One of the most well-known and widespread methods for measuring volatility is the implied volatility within the Black–Scholes model [10, 11]. It is calculated from the market prices of European options (financial derivatives giving the right to buy or sell an underlying asset, e.g., a stock, at a definite price on a definite date). This approach has several disadvantages. First of all, a liquid options market is needed; otherwise, the option price with a high probability can be unfair, and the entire procedure will make no sense. In addition, the Black–Scholes model contains several strong assumptions, which are often not satisfied in reality. For example, we mention the assumption of a time-constant risk-free interest rate to borrow and lend money, or the assumption of a time-constant volatility of the underlying asset price. Due to the latter assumption, implied volatility may vary for different option exercise prices. For stocks, this effect is known as the volatility smile or volatility smirk. Thus, the method is somewhat contradictory and, in addition, it requires rather strict preconditions for application.

Another well-known approach is econometric modeling based on historical volatility values. As a rule, volatility here is understood as the rate of return of an asset. In this direction, we note GARCH (*generalized autoregressive conditional heteroskedasticity*) [12, 13], a generally recognized model pioneered in [14] and based on the ARCH model. However, the classical GARCH model neglects an effect demonstrated by modern financial markets. It consists in the skewed distribution of asset returns in the markets: negative external shocks cause a sharper fall and higher volatility, whereas positive external shocks cause less sharp growth and less high volatility. Different researchers proposed quite a wide range of modifications for the classical GARCH model to take asymmetric volatility into account: NAGARCH (nonlinear asymmetric GARCH) [15], EGARCH (exponential GARCH) [13], QGARCH (quadratic GARCH) [16], GJR-GARCH (Glosten, Jagannathan, and Runkle GARCH) [17], TGARCH (threshold GARCH) [18], VSGARCH (volatility-switching GARCH) [19], and others. However, these approaches disregard, e.g., the presence of four volatility clusters.

Due to some shortcomings of conventional methods, we propose a model incorporating elements of fuzzy logic. The fuzzy system used in this paper, the Takagi–Sugeno fuzzy inference scheme, stems from [20]. The authors [21] introduced the fuzzy model parameters estimation procedure via the least squares method. These two papers conditioned the wide application of such models in various fields and the fuzzy formulation added to classical econometric models. Fuzzy systems were described, e.g., in [22, 23].

Fuzzy models are adopted to forecast stock index volatility as well. According to the works on this subject, there is a wide range of algorithms and input data. For example, the study [1] presented an asymmetric fuzzy GARCH model with a fuzzy inference scheme to determine the switching threshold. The paper demonstrated the effectiveness of the proposed method on the returns of several stock indices: NASDAQ (the USA), Nikkei 225 (Japan), the Taiwan Weighted Index, and the Hang Seng Index (Hong Kong). The original idea of [1] was further developed in [2]. The same asymmetric fuzzy GARCH model was used (as before, to determine the switching threshold), but fuzziness was also introduced into the characteristic function. (In the modified model, it can take any value in the interval [0, 1].) Three variants were proposed to fuzzify the characteristic function. The authors compared the effectiveness of the presented methods with the GJR-GARCH model and the model [1]. The MOEX Russia Index (formerly the MICEX Index) and the RTS Index were studied as time series.

The authors [3] proposed an *adaptive fuzzy inference* system (AdaFIS), which dynamically determines the required number of fuzzy rules and their parameters. The model was applied to the Bovespa Index (Brazil's main stock index), the BRL/USD exchange rate, and the Petrobras preferred stock prices. The paper [4] described *evolving participatory learning* (ePL), a dynamic estimation method for model parameters. The researchers tested their method on historical values of the S&P 500 and Bovespa stock indices. Note that this method is an extension of the *evolving Takagi–Sugeno* (eTS) model [24]. The authors [5] continued studying the evolving fuzzy GARCH model in the paper [4], but they applied another fuzzy clustering method (the eClustering algorithm). The values of the S&P 500 and Bovespa indices were taken as the real series as well. The fuzzy GARCH model was also presented in [6]. In the cited work, asymmetry was considered by using known returns as explanatory variables: since the return (not its square) is employed in fuzzy clustering, the sign is taken into account. The authors applied the model to the Dow Jones Industrial Average Index.

In this paper, we propose a model combining the standard GARCH model and fuzzy logic. The model can be briefly described as follows. The input data are divided into several fuzzy clusters, and a different local GARCH model is applied within each cluster. The outputs of each local model are then aggregated into one via a pre-selected membership function. We conducted the empirical part of this research on historical values of the RTS Index, one of the main stock indices of the Russian market. Two models (benchmarks) are used to compare the forecasting properties of the proposed model: GARCH without recalculation and GARCH with recalculation. (The latter model will be defined below.) According to the calculation results, there exist fuzzy GARCH models with a higher forecasting accuracy than their classical counterparts.

This paper is organized as follows. In Section 1, we describe theoretically the proposed model and the approach to input data clustering. Section 2 presents the initial data, the problem statement, and the results of the empirical study. The Conclusions section summarizes the outcomes and outlines possible lines for further research.

1. METHODOLOGY

1.1. The Fuzzy GARCH Model

The classical GARCH model was described, e.g., in [13]. The proposed fuzzy model is based on the GARCH model but also includes soft switching between fuzzy rules. Each rule corresponds to a fuzzy cluster in the input data space. With C clusters, the fuzzy GARCH model can be represented as a set of C fuzzy IF-THEN rules:

$$\begin{gathered}\text{IF } x_t \in F_k \text{ THEN} \\ h_t^{(k)} = \alpha_{k0} + \sum_{i=1}^{q} \alpha_{ki} y_{t-i}^2 + \sum_{j=1}^{p} \beta_{kj} h_{t-j}, \\ \alpha_{k0} > 0, \\ \alpha_{ki} > 0 \forall k, i, \\ \beta_{kj} \geq 0 \forall k, j.\end{gathered} \tag{1}$$

The notations are as follows: k is the cluster number ($k = 1,..., C$); F_k denotes the kth fuzzy cluster; $x_t = (x_t^1,..., x_t^n)'$ is the variable vector to determine the membership function at a time instant t. In addition, y_t gives the time series under consideration; $h_t^{(k)}$ is the conditional variance corresponding to the kth fuzzy rule at a time instant t, and h_t is the conditional variance at a time instant t (see below); α_{k0}, α_{ki}, and β_{kj} are the estimated model parameters. From this point onwards, the symbol ' means transposition. Generally speaking, the vector x_t may have an arbitrary dimension n depending on t. The parameters α_{ki} and β_{kj} will be called the consequent parameters.

The expression $x_t \in F_k$ is understood in a fuzzy sense (the degree of membership of the vector x_t to a cluster F_k, a real number from the interval [0, 1]). The degree of membership to a certain cluster may have different functional forms. In this paper, we use the Gaussian membership function, which analytically coincides with the density of the multivariate Gaussian distribution:

$$\mu_k(x_t) = \frac{1}{(2\pi)^{n/2} (\det(\Sigma_k))^{1/2}} e^{-\frac{1}{2}(x_t - c_k)' \sum_k^{-1} (x_t - c_k)}.$$

Here, $\mu_k : \mathbb{R}^n \to \mathbb{R}$ is the membership function of a vector to the kth cluster, whereas $c_k \in \mathbb{R}^n$ and $\Sigma_k^* \in \mathbb{R}^n$ are the center and covariance matrix of the kth cluster. This paper considers diagonal positive definite covariance matrices. The matrix Σ_k and the vector c_k, the parameters of the kth cluster, completely define it. For all k together, they will be called the antecedent parameters.

The degrees of membership are normalized so that at each point, their sum over all clusters is 1:

$$\mu_k^*(x_t) = \frac{\mu_k(x_t)}{\sum_{k=1}^{C} \mu_k(x_t)}.$$

All variances $h_t^{(k)}$, $k = 1,..., C$, are aggregated into a single value h_t using the membership functions:

$$h_t = \sum_{k=1}^{C} \mu_k^*(x_t) h_t^{(k)},$$

where $h_t^{(k)}$ are given by (1).

The value y_t^2 is forecasted. In this paper, the series y_t consists of the returns on some financial instrument. The forecast has the form

$$\hat{y}_t^2 = h_t.$$

Let T be the number of elements in an aggregate sample (including training and test samples). The consequent parameters are estimated using the least squares method: we choose the parameters α_{ki} and β_{kj} by minimizing the sum

$$\sum_{t=1}^{T} (y_t^2 - \hat{y}_t^2)^2,$$

where y_t are known values from the sample.

Hereinafter, y_t describes the logarithmic return on some asset or index, expressed as a percentage. For a given series z_t of some values (the prices of a financial instrument or index values),

$$y_t = \ln \frac{z_t}{z_{t-1}} \cdot 100. \tag{2}$$

In all calculations below, $C = 4$.

1.2. Clustering and Antecedent Parameter Estimation

In this study, we cluster the entire initial series of known returns y_t at time instants t, i.e., $x_t = (y_1,..., y_t)'$ $\forall t$. Within the proposed approach, the dimension n of the vector x_t is a function of time: $n = n(t) = t$. Thus, we build a family of fuzzy systems: at each time instant, when a new value of the return becomes known, the entire series is clustered again with this new value, generating a new fuzzy system.

The input data are divided into two parts: training and test samples.

There exist different approaches to data clustering, which essentially means estimating the antecedent pa-

rameters. In this paper, the grid search method is used. A detailed description of the grid is provided in subsection 2.1.

Let T_{train} and T_{test} denote the sizes of the training and test samples, respectively.

On the same training sample used for the fuzzy model, we construct the classical GARCH(p, q) model with the same values of the parameters p and q as in the fuzzy model. The forecasting accuracy of the classical model is estimated using two approaches as follows. The first approach is producing the usual model forecast for T_{test} days ahead. The second approach consists in reestimating the GARCH model parameters daily for T_{test} days and obtaining the forecast for the next day only. (For convenience, the latter approach will be called GARCH with recalculation). Note that the recalculation procedure allows reducing the forecasting error; see the empirical results below (Tables 1 and 2 in Section 2). For each approach mentioned, we find the mean square errors: $MSE_{n/r}$ for GARCH without recalculation and $MSE_{w/r}$ for GARCH with recalculation. The two resulting errors are used as benchmarks for the fuzzy model. For comparing the fuzzy model with the classical one, we introduce $ratio_{w/r} = \frac{MSE_{w/r}}{MSE_{fuzzy}}$ and $ratio_{n/r} = \frac{MSE_{n/r}}{MSE_{fuzzy}}$. If $ratio_i > 1$, the error of the fuzzy method is smaller than that of the classical ith method. (Here, i = "*w/r*" or i = "*n/r*"). In this case, the higher value $ratio_t$ takes, the "better" the fuzzy model will be compared to the ith classical model.

To define the fuzzy model, we use the set $\Theta = (\Theta, ..., \Theta_C) \in \mathbb{R}^{n \times n \times C}$ of possible values of the antecedent parameters. Each $\Theta_k = \{c_k, \Sigma_k \,|\, c_k \in \mathbb{R}^n, \Sigma_k \in \mathbb{R}^n\}$ represents all possible combinations of the values c_k and Σ_k. The procedure of estimating the antecedent parameters begins with constructing a particular fuzzy model for each element Θ and calculating the value MSE_{fuzzy}. To find the best antecedent parameters c_k and Σ_k, we maximize $ratio_{w/r} = ratio_{w/r}(c_1, ..., c_C, \Sigma_1, ..., \Sigma_C)$.

Thus, these steps yield the best set of antecedent parameters, i.e., the best fuzzy model compared to the classical GARCH model.

2. EMPIRICAL STUDY

2.1. The Set of Antecedent Parameters

Let $p = 1$ and $q = 1$. Assume that the number of clusters is fixed: $C = 4$. The center of each cluster $c_k \in \mathbb{R}^n$ is a vector with the same components c_k^*:

$$c_k = \begin{pmatrix} c_k^* \\ \dots \\ c_k^* \end{pmatrix}.$$

The covariance matrix is a diagonal real-valued matrix of dimensions $n \times n$, with the same element $\Sigma_k^* > 0$ on the main diagonal:

$$\Sigma_k = \begin{pmatrix} \Sigma_k^* & 0 & \dots & 0 \\ 0 & \Sigma_k^* & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \Sigma_k^* \end{pmatrix}.$$

This element can be treated as the variance of the kth cluster.

Let $c^* = (c_1^*, c_2^*, c_3^*, c_4^*)$ and $\Sigma^* (\Sigma_1^*, \Sigma_2^*, \Sigma_3^*, \Sigma_4^*)$. These four-dimensional vectors completely define all four clusters. We employ them to parameterize the admissible vector space instead of the vector Θ.

In this paper, the set of centers is not varied but set expertly: in all calculations, c^* = (–7.5; –1.5; 1.5; 4). Intuitively, the centers c^* can be interpreted as follows: center 1.5 corresponds to small positive returns whereas center 4 to large returns; centers –1.5 and –7.5 correspond to small and large negative returns, respectively (in absolute terms). The fact that |–7.5| > |4| reflects a characteristic feature of capital markets: the growth under positive external shocks is smoother than the drop under external negative shocks.

The variances are estimated by the grid search method. The grid was constructed with the following ranges: for Σ_1^*, from 4 to 12; for Σ_2^* and Σ_3^*, from 1 to 6; for Σ_4^*, from 2 to 10. The grid step was set equal to 1 for all Σ_k^*. Thus, 2916 grid nodes were considered in total.

2.2. Results

The daily logarithmic returns of the RTS Index were used as the initial series. Figure 1 shows the closing index data for a long historical period.

The beginning of the training sample coincides with the first trading day of 2014 (January 6, 2014). Figure 2 presents the values of the RTS Index for the historical period used in the calculations (the series z_t), about 3 years from the beginning of the training sample. The corresponding logarithmic returns (the series y_t (2)) are demonstrated in Fig. 3.

Fig. 1. Daily values of the RTS Index since 2001.

Fig. 2. Daily values of the RTS Index: January 2014–December 2016.

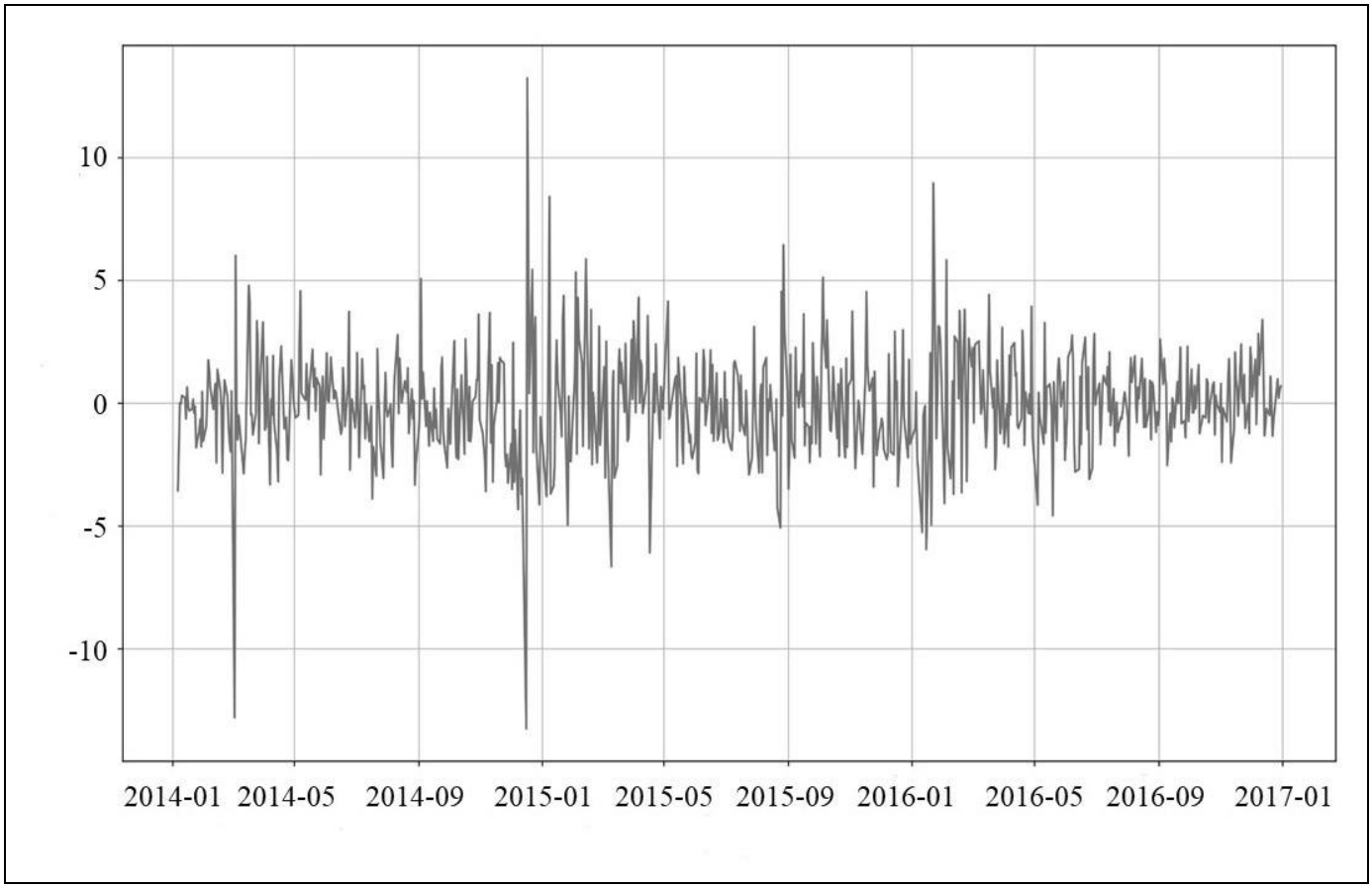

Fig. 3. Daily logarithmic returns of the RTS Index: January 2014–December 2016.

The antecedent parameters (i.e., the variances under fixed centers) were estimated using 100 elements in the training sample and 10 elements in the test sample. The best set of the variances was (7, 6, 3, 5). The model with these parameters was then applied to the samples of other sizes; see the corresponding results in Table 2. The sample sizes were selected from the following considerations: 252 is the approximate number of trading days in one year (504 in two years, 21 in one month, 42 in two months, and 63 in three months).

The results of the empirical study are tabulated below. Table 1 combines the characteristics of the systems with the best results compared to the classical model. Note that $ratio_{w/r} > 1$ in all these models. The first row of this table contains information about the best fuzzy model. Next, Table 2 shows the results of applying the best system from Table 1 to the samples of other sizes. According to the calculations, these systems have a higher error than the classical model ($ratio_{w/r} < 1$).

Good results were achieved for 100 elements in the training sample and 10 elements in the test sample. The quality of the fuzzy model gradually declines when increasing the size of either sample. This phenomenon can be explained as follows: the antecedent parameters need to be reestimated on other sample sizes. (This was not done due to high computational costs.) Moreover, adding variability to the cluster centers may contribute to finding a model with higher accuracy.

Thus, for sufficiently short time series, there exist fuzzy GARCH systems superior to the classical GARCH model in their forecasting properties. With increasing the time series length, the classical GARCH model becomes a "stronger" competitor. Most probably, the reason lies in the advantages of the maximum likelihood method, which show up with increasing the time series length. However, it seems that fuzzy GARCH systems superior to the classical GARCH model can be obtained even for longer time series using broader classes of membership functions.

Table 1

The best models on the training and testing samples of 100 and 10 elements, respectively

T_{train}	T_{test}	$\sum^*$	MSE_{fuzzy}	$MSE_{n/r}$	$MSE_{w/r}$	$ratio_{n/r}$	$ratio_{w/r}$
100	10	[7, 6, 3, 5]	9.66	17.89	12.75	1.85	1.32
100	10	[6, 4, 4, 6]	9.72	17.89	12.75	1.84	1.31
100	10	[8, 6, 4, 6]	10.18	17.89	12.75	1.76	1.25
100	10	[8, 1, 4, 6]	11.31	17.89	12.75	1.58	1.13

Table 2

The best model among the ones applied to the samples of other sizes

T_{train}	T_{test}	$\sum^*$	MSE_{fuzzy}	$MSE_{n/r}$	$MSE_{w/r}$	$ratio_{n/r}$	$ratio_{w/r}$
252	21	[7, 6, 3, 5]	353.14	273.58	189.09	0.77	0.51
252	42	[7, 6, 3, 5]	251.84	201.34	137.60	0.80	0.55
252	63	[7, 6, 3, 5]	190.30	141.74	99.21	0.74	0.52
504	21	[7, 6, 3, 5]	542.75	382.64	300.49	0.70	0.55
504	42	[7, 6, 3, 5]	290.84	203.56	160.35	0.70	0.55
504	63	[7, 6, 3, 5]	202.47	143.12	112.76	0.70	0.56

CONCLUSIONS

In this paper, a fuzzy GARCH system has been applied to model volatility. The classical GARCH models with and without recalculation were considered as alternatives to the proposed model (benchmarks). The effectiveness of the fuzzy model has been tested on the main index of the Russian stock market (the RTS Index). As shown above, there exist fuzzy systems producing more accurate forecasts compared to traditional models.

Nevertheless, the forecasting properties of the fuzzy system obviously tend to deteriorate (compared to the benchmark models) with increasing the sample size. This can be due to the parameter estimation methods used in the fuzzy model and in the classical GARCH model.

A possible line for further research is practical methods for finding such fuzzy systems. Of particular interest are other forms of membership functions as well as more universal methods for estimating the antecedent parameters.

Acknowledgments. *This research was supported in part through computational resources of HPC facilities at HSE University* [25].

REFERENCES

1. Hung, J.C., A Fuzzy Asymmetric GARCH Model Applied to Stock Markets, *Information Sciences*, 2009, vol. 179, no. 22, pp. 3930–3943.
2. Lepskiy, A. and Suevalov, A., Application of Fuzzy Asymmetric GARCH-Models to Forecasting of Volatility of Russian Stock Market, *Advances in Intelligent Systems and Computing*, 2018, vol. 679, pp. 286–294.
3. Luna, I. and Ballini, R., Adaptive Fuzzy System to Forecast Financial Time Series Volatility, *Journal of Intelligent and Fuzzy Systems*, 2012, vol. 23, pp. 27–38. DOI: 10.3233/IFS-2012-0491.
4. Maciel, L., Gomide, F., and Ballini, R., Enhanced Evolving Participatory Learning Fuzzy Modeling: An Application for Asset Returns Volatility Forecasting, *Evolving Systems*, 2013, vol. 5, pp. 1–14. DOI: 10.1007/s12530-013-9099-0.
5. Maciel, L., Gomide, F., and Ballini, R., Evolving Fuzzy-GARCH Approach for Financial Volatility Modeling and Forecasting, *Computational Economics*, 2016, vol. 48, no 3, pp. 379–398.
6. Popov, A.A. and Bykhanov, K.V., Modeling Volatility of Time Series Using Fuzzy GARCH Models, *Proceedings of the 9th Russian-Korean International Symposium on Science and Technology* (*KORUS-2005*), 2005, vol. 1, pp. 687–692.
7. Tan, L., Wang, S., and Wang, K., A New Adaptive Network-Based Fuzzy Inference System with Adaptive Adjustment Rules for Stock Market Volatility Forecasting, *Information Processing Letters*, 2017, vol. 127, pp. 32–36. DOI: 10.1016/j.ipl.2017.06.012.
8. Troiano, L., Mejuto, E., and Kriplani, P., An Alternative Estimation of Market Volatility Based on Fuzzy Transform, *Proceedings of the Joint 17th World Congress of the International Fuzzy Systems Association and 9th International Conference on Soft Computing and Intelligent Systems* (*IFSA-SCIS 2017*), Otsu, 2017, pp. 1–6. DOI: 10.1109/IFSA-SCIS.2017.8023316.
9. Thavaneswaran, A., Liang, Y., Zhu, Z., et al., Novel Data-Driven Fuzzy Algorithmic Volatility Forecasting Models with Applications to Algorithmic Trading, *Proceedings of the 2020 IEEE International Conference on Fuzzy Systems* (*FUZZ-IEEE*), Glasgow, 2020, pp. 1–8. DOI: 10.1109/FUZZ48607.2020.9177735.

10.Black, F. and Scholes, M., The Pricing of Options and Corporate Liabilities, *Journal of Political Economy*, 1973, vol. 81, no. 3, pp. 637–657.
11.Hull, J.C., *Options, Futures, and Other Derivatives*, 9th ed., London: Pearson, 2014.
12.Bollerslev, T., Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics*, 1986, vol. 31, no. 3, pp. 307–327. DOI: https://doi.org/10.1016/0304-4076(86)90063-1.
13.Tsay, R.S., *Analysis of Financial Time Series*, 3rd ed., Hoboken: John Wiley & Sons, 2010.
14.Engle, R.F., Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation, *Econometrica*, 1982, vol. 50, no. 4, pp. 987–1007.
15.Engle, R.F. and Ng, V.K., Measuring and Testing the Impact of News on Volatility, *The Journal of Finance*, 1993, vol. 48, no. 5, pp. 1749–1778.
16.Sentana, E., Quadratic ARCH Models, *The Review of Economic Studies*, 1995, vol. 62, no. 4, pp. 639–661.
17.Jagannathan, R., Glosten, L., and Runkle, D., On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks, *Journal of Finance*, 1993, vol. 48, pp. 1779–1801.
18.Zakoian, J.-M., Threshold Heteroskedastic Models, *Journal of Economic Dynamics and Control*, 1994, vol. 18, no. 5, pp. 931–955.
19.Fornari, F. and Mele, A., Modeling the Changing Asymmetry of Conditional Variances, *Economics Letters*, 1996, vol. 50, no. 2, pp. 197–203.
20.Takagi, T. and Sugeno, M., Fuzzy Identification of Systems and Its Applications to Modeling and Control, *IEEE Transactions on Systems, Man, and Cybernetics*, 1985, vol. SMC-15, no. 1, pp. 116–132. DOI: 10.1109/TSMC.1985.6313399.
21.Sugeno, M. and Kang, G.T., Structure Identification of Fuzzy Model, *Fuzzy Sets and Systems*, 1988, vol. 28, no. 1, pp. 15–33.
22.Piegat, A., *Fuzzy Modeling and Control*, Physica Heidelberg, 2001.
23.Baczyński, M. and Jayaram, B., *Fuzzy Implications*, Berlin–Heidelberg: Springer, 2008.
24.Angelov, P.P. and Filev, D.P., An Approach to Online Identification of Takagi–Sugeno Fuzzy Models, *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 2004, vol. 34, no. 1, pp. 484–498. DOI: 10.1109/TSMCB.2003.817053.
25.Kostenetskiy, P.S., Chulkevich, R.A., and Kozyrev, V.I., HPC Resources of the Higher School of Economics, *Journal of Physics: Conference Series*, 2021, vol. 1740, no. 1, p. 012050. DOI: 10.1088/1742-6596/1740/1/012050.

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

Received June 10, 2022, and revised December 10, 2022. Accepted December 13, 2022

Author information

Sviyazov, Vladimir Andreevich. Postgraduate, National Research University Higher School of Economics, Moscow, Russia
✉ v.sviyazov.96@gmail.com

Cite this paper

Sviyazov, V.A., Fuzzy Volatility Models with Application to the Russian Stock Market, *Control Sciences* **6,** 21–28 (2022). http://doi.org/10.25728/cs.2022.6.3

Original Russian Text © Sviyazov, V.A., 2022, published in *Problemy Upravleniya*, 2022, no. 6, pp. 26–34.

Translated into English by *Alexander Yu. Mazurov,*
Cand. Sci. (Phys.–Math.),
Trapeznikov Institute of Control Sciences,
Russian Academy of Sciences, Moscow, Russia
✉ alexander.mazurov08@gmail.com

DOI: http://doi.org/10.25728/cs.2022.6.4

ESTIMATING INDUSTRIAL PROCESS STABILITY BY WHITNEY'S SINGULARITY THEORY WHEN CHOOSING A SUFFICIENT TIME-SAMPLING FREQUENCY OF THE CONTROL SIGNAL

M.A. Rabotnikov[1], B.G. Stafeichuk[2], and A.G. Shumikhin[3]

Perm National Research Polytechnic University, Perm, Russia

[1]✉ rabotnikovma@gmail.com, [2]✉ bgstaf@mail.ru, [3]✉ shumichin@gmail.com

Abstract. In this paper, we estimate the stability of continuous-type automated industrial processes and choose a sufficient time-sampling frequency of the control signal using Whitney's singularity theory. The proposed stability analysis approach is based on constructing typical bifurcations for the historical data of a technological object under different time-sampling frequencies of its control signal. The singularity equation serves for obtaining the equation of the equilibrium state curves of the system and a sufficient time-sampling frequency of the control signal corresponding to the vertex of the resulting curve. As an illustrative example, the developed method is applied to the control system of the mass balance stripping section in the purification process of a styrene distillation column of the ethylbenzene, styrene, and polystyrene plants. Based on the quantitative analysis results, we construct a bifurcation and determine a sufficient time-sampling frequency of the control signal to ensure system stability.

Keywords: catastrophe theory, bifurcation, dynamical systems, stability criteria.

INTRODUCTION

Catastrophe theory has been widely applied to analyze the behavior of economic and social systems, assess the design properties of structures and apparatuses over time, and study the properties and quantitative characteristics of dynamical systems [1–6]. The existing mathematical framework numerically describes jump-like transients of the output variable due to a smooth change in the input parameters of a dynamical system [7]. In particular, this approach can be used to estimate the stability of continuous-type automated industrial processes [8].

The problem under consideration—the stable operation of industrial processes—consists in complying with the regulatory limits of the detected technological parameters to minimize the risk of stopping production and any violations of the quality composition of commercial products (on the one hand) and maximize the technological efficiency under various disturbances, often of sporadic nature (on the other hand); for details, see [9]. The modern technical base and the existing mathematical approaches and algorithmic solutions used in automated control allow reducing this problem (technological mode stabilization and compliance with the regulatory limits) to multidimensional model predictive control [10]. For systems of this class, the time-sampling frequency of the control device is one of the indicators characterizing the inertia of the controlled process and the effect of disturbing factors on the controlled technological parameters. Under fixed values of other characteristics, a sufficient value of this frequency ensures the stable operation of a technological object within the specified regulatory limits of the corresponding industrial process.

1. CONSTRUCTING BIFURCATIONS OF CONTROL SYSTEMS

How does the stability of an industrial process depend on the control signal frequency? What is the range of frequencies ensuring system operation within given constraints? To answer these questions, we consider the dynamics of a controlled process parameter as a typical bifurcation.

By conditions, the bifurcation depends on the time-sampling frequency of the control device and the effect of disturbing factors on the system, possessing a codimension of 2. Due to the continuity and natural origin of the processes under study, the bifurcation dependence must be a continuously differentiable function on the entire definitional domain.

For the problem under consideration, it is convenient to represent the bifurcation as a second Whitney singularity, the cusp [11]. Its functional mapping is given by

$$f(x, g(\omega)) = x^3 + g(\omega)x,$$

with the following notations: $f(x)$ is the normalized distribution of the controlled parameter; $x \in [-1, 1]$ is the normalized spectrum of the equivalent disturbance characterizing the resulting effect of all external factors of the system on the controlled variable; $g(\omega)$ is the frequency function; finally, ω is the time-sampling frequency characterizing the periodicity of the control signal.

According to the bifurcation properties, the control system is in an unstable state if $g(\omega) > 0$: the function $f(x, g(\omega))$ has no nondegenerate singularities and increases continuously over the entire definitional domain [12]. For $\omega = \omega_s$, we obtain $g(\omega) = 0$, where ω_s is a sufficient time-sampling frequency of the control device algorithms: the system passes through a bifurcation point S, being in an equilibrium state. In the case $g(\omega) < 0$, the system comes to a stable equilibrium, forming a bifurcation with two stationary points.

Note that the frequency function $g(\omega)$ has no strict mathematical formalization. For the quantitative description of observed real processes, it can be constructed by piecewise linear approximation based on the historical operation data of the controlled object. Considering the inverse dependence of the stability margin on the frequency ω ensuring system stability within the range $\omega \in (\omega_s, +\infty)$, the frequency function for this system operation range takes the form

$$g(\omega) = -\omega + \omega_s .$$

In this case, the bifurcation is transformed to

$$f(x, \omega) = x^3 + (\omega_s - \omega)x.$$

Let us reduce the distribution $f(x, \omega)$ from the normalized variables to those expressed in the original units of measurement:

$$y(x, \omega) = \alpha f(x, \omega) + M_y,$$

where α is the normalization coefficient and M_y denotes the expectation of $y(x, \omega)$. As a result, the surface is given by

$$y(x,\omega) = \alpha\left(x^3 + (\omega_s - \omega)x\right) + M_y. \qquad (1)$$

The projection of the cusp $y(x, \omega)$ onto the plane (y, ω) yields the set of singularities y_s with the vertex at the initial bifurcation point S. The bifurcation set under study forms the boundaries of the branches of the equilibrium state curves. According to the definition [13], these curves are obtained by equating the derivative $y'_x(x, \omega)$ to zero:

$$3x^2 + \omega_s - \omega = 0 .$$

Substituting this condition into formula (1), we determine the general equation of the bifurcation set of the control system:

$$y_s(\omega) = M_y \pm 2\alpha\left(\frac{\omega - \omega_s}{3}\right)^{3/2}, \ \omega \ge \omega_s . \qquad (2)$$

The graph of the resulting bifurcation is shown in Fig. 1.

The parameters ω_s and α, characterizing the stability of the control system, are determined empirically by the historical data of the closed-loop dynamics under different frequencies ω. According to the correspondence principle for the degrees of freedom and the necessary number of equations to identify the parameters ω_s and α, a statistical sample must contain historical data sets for two operation modes under different frequencies ω_1 and ω_2 of the control system: $y_1(x_1, \omega_1)$ and $y_2(x_2, \omega_2)$, where y_1 and y_2 are the detected states of the controlled parameter under external disturbances x_1 and x_2, respectively. In this case, we have the system of equations

$$\begin{cases} y_1 = \alpha\left(x_1^3 + (\omega_s - \omega_1)x_1\right) + M_y, \\ y_2 = \alpha\left(x_2^3 + (\omega_s - \omega_2)x_2\right) + M_y. \end{cases}$$

After trivial transformations, this system yields expressions for the parameters ω_s and α:

$$\omega_s = \frac{(y_1 - M_y)(x_2\omega_2 - x_2^3) - (y_2 - M_y)(x_1\omega_1 - x_1^3)}{(y_1 - M_y)x_2 - (y_2 - M_y)x_1}, \qquad (3)$$

$$\alpha = \frac{(y_1 - M_y)x_2 - (y_2 - M_y)x_1}{x_1x_2\left(x_1^2 - x_2^2 - \omega_1 + \omega_2\right)}. \qquad (4)$$

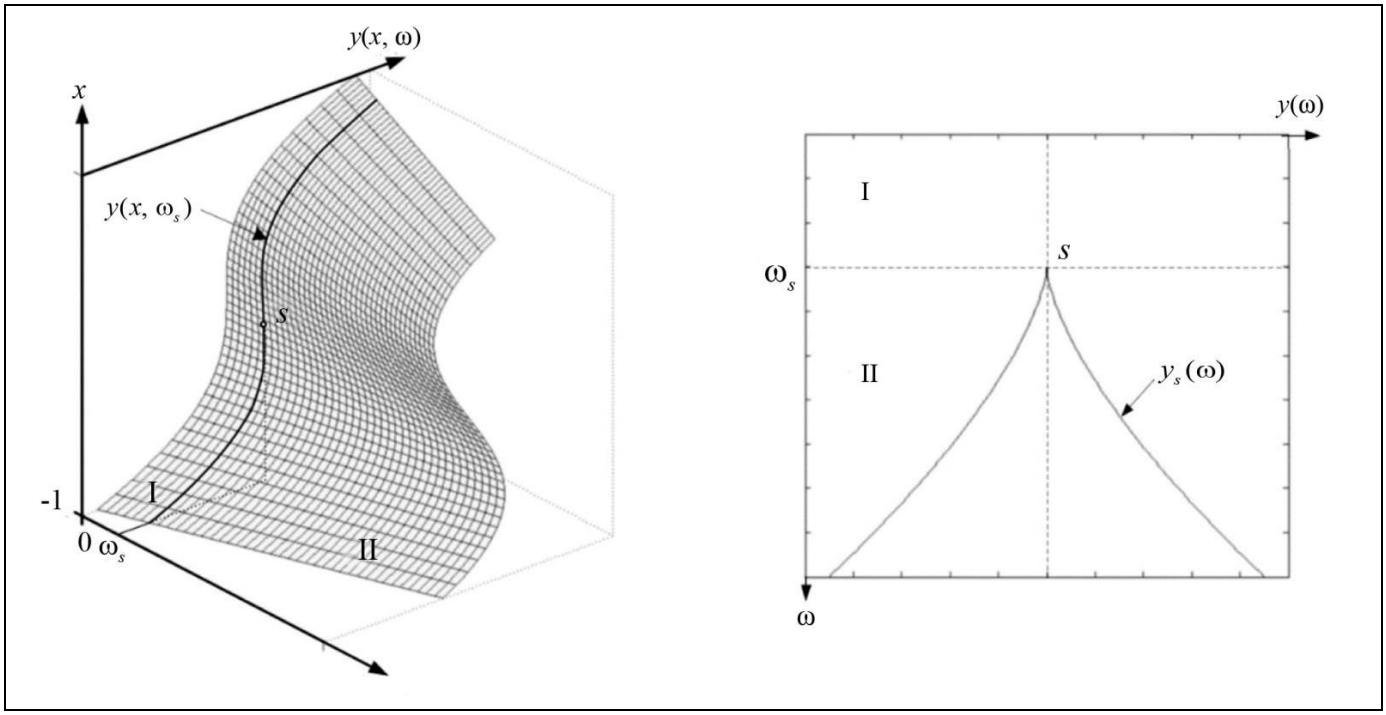

Fig. 1. Control system bifurcation: *s*—bifurcation point, I—unstable state domain, and II—stable state domain.

It is convenient to calculate ω_s and α at the maximum spike points of the controller parameter: y_1 and y_2 for the frequencies ω_1 and ω_2, respectively. These responses of the output signal correspond to the limits of the undetected spectrum of the resulting disturbance of the system in the normalized form: $|x_{1,2}| = 1$. The sign of $x_{1,2}$ is determined depending on the location of the spike point of the controlled variable relative to the sampling mean M_y.

2. BIFURCATIONS OF REAL TECHNOLOGICAL OBJECTS: ONE EXAMPLE OF CONSTRUCTION

As an industrial process example, we consider the stripping section of the distillation column C-2 in the separation unit of ethylbenzene, styrene, and polystyrene production (PESP): construct a bifurcation and choose a sufficient time-sampling frequency for the control signal of the automated process control system. This unit consists of three columns (C-1, C-2, and C-3) with the following functions: separating hydrocarbon condensate supplied from the ethylbenzene dehydrogenation block into raw styrene and the benzene-toluene-ethylbenzene fraction; purifying commercial styrene from heavier fractions (residue of styrene rectification); separating the benzene-toluene-ethylbenzene fraction into recycled ethylbenzene and the benzene-toluene fraction (benthol). Figure 2 shows the general diagram of the main material flows of the distillation unit for ethylbenzene, styrene, and polystyrene production.

The technological mode of the stripping section of column C-2 is maintained by multidimensional control: it is required to stabilize the level in the column (*LI001*) under a given constraint (the limit value of the column temperature profile (*TI001*)) due to beginning the styrene polymerization reaction. The main control action is the superheated steam supply into the heat exchanger H-2 (*FIC001*), intended for heating the bottom fraction. The evacuation of high boiling components from the column (*FIC002*) is fixed at the minimum value due to commercial styrene losses in the mixture. The key disturbance for this automated process control system is the quantitatively undetected change in the composition of the incoming hydrocarbon feedstock.

The dynamics of the technological object are described by the system of differential equations

$$\begin{cases} \sum\limits_{i=0}^{n_1} a_{1i} \dfrac{\partial^i \Delta L(t)}{\partial t^i} \\ \quad = \sum\limits_{j=0}^{m_1} b_{1j} \dfrac{\partial^j \Delta F(t-\tau_{11})}{\partial t^j} + \sum\limits_{k=0}^{q_1} c_{1k} \dfrac{\partial^k \Delta Q(t-\tau_{12})}{\partial t^k}, \\ \sum\limits_{i=0}^{n_2} a_{2i} \dfrac{\partial^i \Delta T(t)}{\partial t^i} \\ \quad = \sum\limits_{j=0}^{m_2} b_{2j} \dfrac{\partial^j \Delta F(t-\tau_{21})}{\partial t^j} + \sum\limits_{k=0}^{q_2} c_{2k} \dfrac{\partial^k \Delta Q(t-\tau_{22})}{\partial t^k}, \end{cases} \quad (5)$$

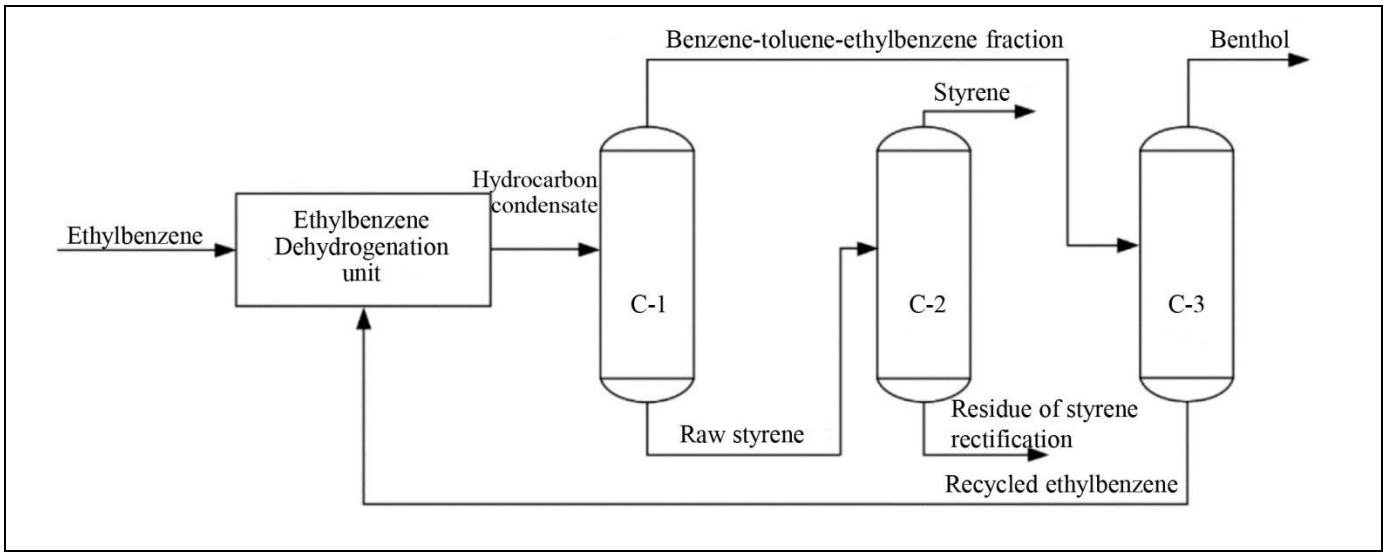

Fig. 2. Distillation unit diagram.

with the following notations: L is the liquid phase level in column C-2; F is the steam flow rate in the heat exchanger H-2; Q is the quantitative composition of column feeding; T is the bottom temperature in column C-2; a, b, and c are the coefficients of differential equations; n, m, and k are the orders of polynomials ($n > m$, $n > k$); finally, τ is the time lag of dynamic channels.

The technological mode is maintained within the specified regulatory limits by an automated *model predictive control* (MPC) system [14]. The model is presented as a matrix transfer function approximating the real behavior of the industrial process according to the empirical transient characteristics of the transmission channels (5). Figure 3 shows the process control diagram for the stripping section of column C-2.

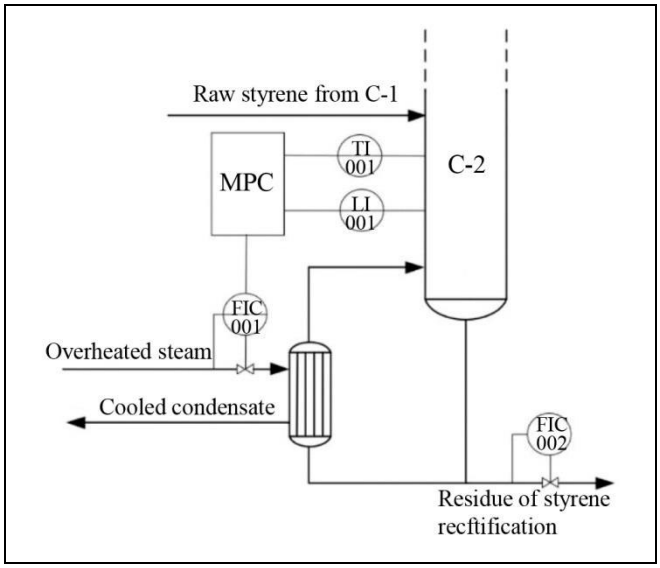

Fig. 3. Process control for the stripping section of column C-2.

Due to the relatively small dimensions of the column and the high rate of mass exchange processes, the distillation column C-2 has low inertia. Figure 4 demonstrates the level in the column under different time-sampling frequencies of the control signal. For $\omega_1 = 0.1$/min, we observe jump changes of this level with a root-mean-square (RMS) deviation of 7.29%, which determine the possibility of an emergency shutdown of the plant: the pumping equipment will switch off under the full release of the bottom fraction from the column. For $\omega_2 = 1$/min (all other system parameters remain unchanged), the liquid phase level in the column has admissible fluctuations with an RMS deviation of 4.33%: the industrial process is in a steady state.

To construct a bifurcation of the system on a given historical data set (a representative sample), we selected necessary initial data. For this purpose, we considered the maximum deviations of the liquid phase level in column C-2 relative to the mean $M_y = 88.75\%$ for the system operating with the frequencies $\omega_1 = 0.1$/min and $\omega_2 = 1$/min: $y(x_1, \omega_1) = 55.29\%$, $x_1 = -1$; $y(x_2, \omega_2) = 82.32\%$, $x_2 = -1$.

According to formulas (3) and (4), the quantitative properties of this bifurcation are given by the parameters $\omega_s = 0.22$/min and $\alpha = 30.03$. Then the bifurcation of the automated process control system for the stripping section of the distillation column C-2 and the corresponding function of the bifurcation take the following form (Fig. 5):

$$y(x, \omega) = 30.03\left(x^3 + (0.22 - \omega)x\right) + 88.75,$$

$$y_s(\omega) = 88.75 \pm 60.07\left(\frac{0.22 - \omega}{3}\right)^{3/2}, \ \omega \geq 0.22.$$

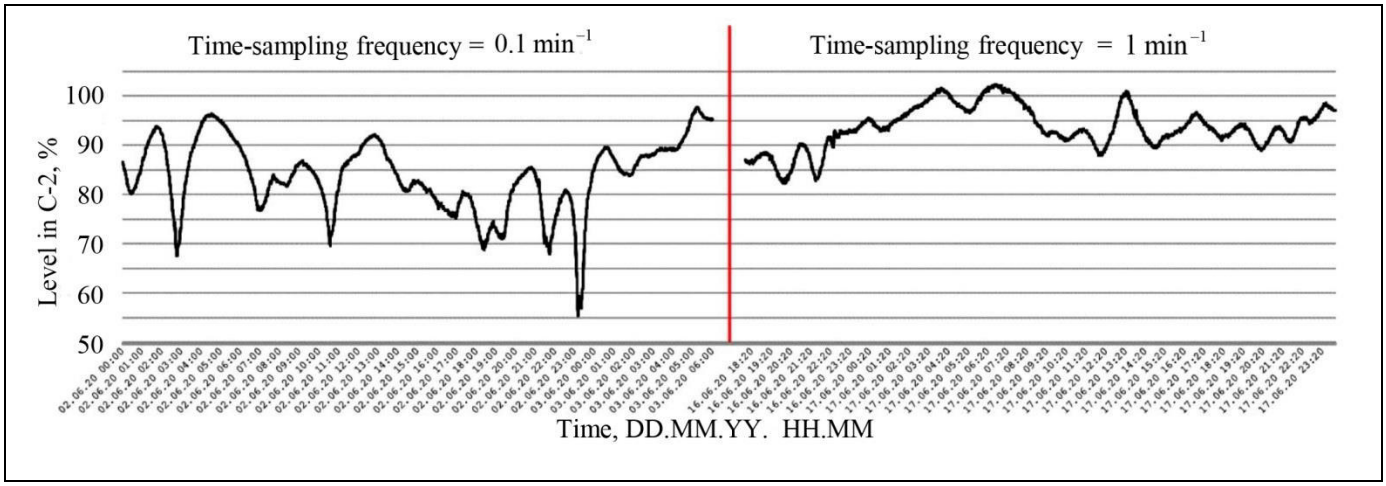

Fig. 4. The level in column C-2

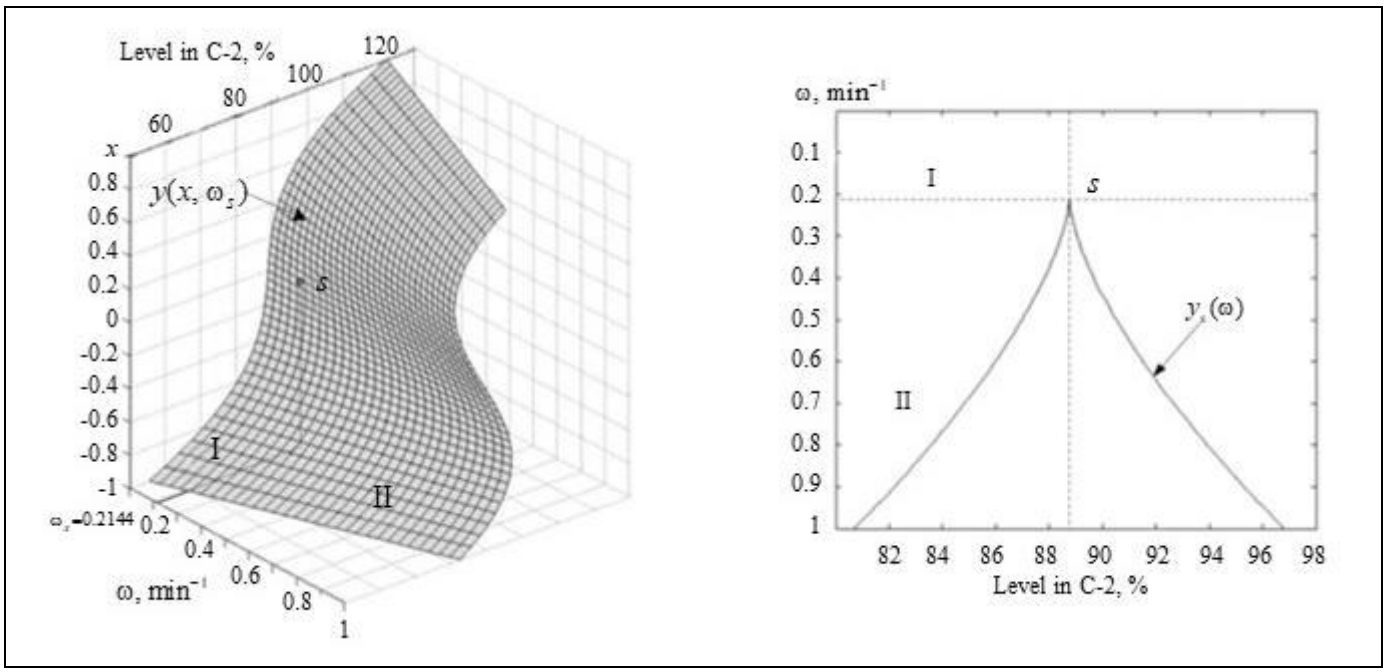

Fig. 5. Bifurcation of the stripping section of column C-2: *s*—bifurcation point, I—unstable state domain, and II—stable state domain.

The constructed bifurcation for choosing the sufficient time-sampling frequency of the control signal is adequate: it agrees with the dynamical characteristics of the real material balance stabilization system for the stripping section of the distillation column C-2. The control action channel (Fig. 3) has the delay $\tau_e \approx 5$ min; therefore, the admissible time-sampling frequency is empirically defined by $\omega_e = \tau_e^{-1} \approx 0.2$ min^{-1}. The analytical solution obtained by the method of bifurcation diagrams gives the sufficient time-sampling frequency $\omega_s \geq 0.22$ min^{-1}.

CONCLUSIONS

This paper has presented a method for constructing bifurcations of dynamical systems based on historical data. It has been applied to estimate the stability of the control system of the mass balance stripping section in the purification process of a styrene distillation column of the ethylbenzene, styrene, and polystyrene plants. According to the analysis results, the control signal ensures stable system operation within the admissible regulatory limits of the technological process

parameters under a sufficient time-sampling frequency of 0.22 min^{-1}.

This approach to constructing bifurcations of control systems and calculating the stable state domain of the chemical engineering object is estimative: it serves for preliminarily determining the optimal time step of the control system considering the dynamical properties of the controlled object.

REFERENCES

1. Zinenko, A.V., Catastrophe Theory and Price Dynamics, *Economics and Mathematical Methods*, 2018, vol. 54, no. 4, pp. 116–123. (In Russian.)
2. Nedel'ko, N.S., Catastrophe Theory and Analysis of Economic Systems' Behavior, *Vestnik of MSTU*, 2010, vol. 13, no. 1, pp. 223–227. (In Russian.)
3. Schmitt, N., Tramontana, F., and Westerhoff, F., Nonlinear Asset-price Dynamics and Stabilization Policies, *Nonlinear Dynamics*, 2020, no. 102, pp. 1045–1070.
4. Muzhikov, G.P. and Gilev, M.A., Analysis of Dynamic Systems Using Elements of the Theory of Bifurcations, in *Sovremennye tendentsii razvitiya i perspektivy vnedreniya innovatsionny tekhnologii v mashinostroenii, obrazovanii i ekonomike* (Current Trends and Prospects for Introducing Innovative Technologies in Engineering, Education, and Economics), 2017, vol. 3, no. 1 (2), pp. 35–37. (In Russian.)
5. Qin, L., Qin, H., and Xing, J., Energy Flow Characteristics of Friction-Induced Nonlinear Vibrations in a Water-Lubricated Bearing-Shaft Coupled System, *Acta Mechanica Sinica*, 2021, no. 37, pp. 679–704.
6. Skorobogatov, S.M., *Catastrophes and Serviceability of Reinforced Concrete Structures (Classification and Elements of Theory)*, Yekaterinburg: Ural State University of Railway Transport, 2020.
7. Hassard, B., Kazarinov, D., and Wan, Y., *Theory and Applications of Hopf Bifurcation*, Cambridge: Cambridge University Press, 1982.
8. Ostreikovskii, V.A., *Analiz ustoichivosti i upravlyaemosti dinamicheskikh sistem metodami teorii katastrof* (Analysis of Stability and Controllability of Dynamic Systems by Catastrophe Theory Methods), Moscow: Vysshaya Shkola, 2005. (In Russian.)
9. Chereshko, A.A. and Shunderyuk, M.M., Applicability Limits of Model-Based Predictive Control Algorithms under Uncertain Control Object Dynamics, *Control Sciences*, 2020, no. 1, pp. 17–23. (In Russian.)
10. Bakhtadze, N.N. and Lototskii, V.A., Contemporary Methods of Production Process Control, *Control Sciences*, 2009, no. 3, pp. 56–63. (In Russian.)
11. Arnold, V.I., *Catastrophe Theory*, Berlin–Heidelberg: Springer, 1984.
12. Iooss, G. and Joseph, D., *Elementary Stability and Bifurcation Theory*, New York: Springer, 2014.
13. Moskalenko, A.V., Tetuev, R.K., and Makhortykh, S.A., On Studies of Bifurcation Phenomena Such as Memory and Delay, *KIAM Preprint No. 109*, Moscow: Keldysh Institute of Applied Mathematics RAS, 2019, pp. 1–44. (In Russian.)
14. Rabotnikov, M.A., Strategy of Material Balance Control Implementation of the Styrene Rectifying Column Bottom, *Trudy Vserossiiskoi nauchno-prakticheskoi konferentsii "Khimiya. Ekologiya. Urbanistika"* (Proceedings of the All-Russian Scientific and Practical Conference on Chemistry, Ecology and Urbanism), Perm, 2021, vol. 4, pp. 242–246. (In Russian.)

This paper was recommended for publication by B.G. Il'yasov, a member of the Editorial Board.

Received March 1, 2022, and revised September 11, 2022. Accepted December 16, 2022.

Author information

Rabotnikov, Mikhail Alekseevich. Assistant, Perm National Research Polytechnic University, Perm, Russia
✉ rabotnikovma@gmail. com

Stafeichuk, Boris Grigor'evich. Cand. Sci. (Eng.), Perm National Research Polytechnic University, Perm, Russia
✉ bgstaf@mail.ru

Shumikhin, Aleksandr Georgievich. Dr. Sci. (Eng.), Perm National Research Polytechnic University, Perm, Russia
✉ shumichin@gmail.com

Cite this paper

Rabotnikov, M.A., Stafeichuk, B.G., Shumikhin, A.G., Estimating Industrial Process Stability by Whitney's Singularity Theory When Choosing a Sufficient Time-Sampling Frequency of the Control Signal. *Control Sciences* **6,** 29–34 (2022). http://doi.org/10.25728/cs.2022.6.4

Translated into English by *Alexander Yu. Mazurov*,
Cand. Sci. (Phys.–Math.),
Trapeznikov Institute of Control Sciences,
Russian Academy of Sciences, Moscow, Russia
✉ alexander.mazurov08@gmail.com

DOI: http://doi.org/10.25728/cs.2022.6.5

ANALYSIS OF STRESS EXPOSURES ON AUTONOMOUS NAVIGATION CONDITIONS IN SEARCH CORRELATION-EXTREME NAVIGATION SYSTEMS

A.I. Alchinov[1] and I.N. Gorokhovsky[2]

[1]Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia
[2]Research Center of Topographic and Navigational Support, Central Research Institute No. 27, Moscow, Russia

✉ alchinov46@mail.ru, ✉ gin_box@mail.ru

Abstract. This paper further develops the concept of an applied geographic information system (AGIS) for modeling search correlation-extreme navigation systems (CENSs), which was presented in [4]. As shown below, the AGIS can be configured to perform computational experiments with computer models of the existing CENSs and those undergoing various development stages without programming in universal languages. Strict reliability requirements for CENSs increase the role of testing their computer models under stress exposures. During stress testing, the negative effects of different exposures on autonomous navigation conditions are assessed in application areas. Such exposures are not considered at the CENS design stage (reference point masking, distortion of terrain objects borders, etc.). The exposures that prevent CENSs from performing their tasks effectively (critical exposures) are described. Stability to critical exposures is a strong motivation for improving all CENS elements: sensors of geophysical fields, onboard algorithms, and CENS preparation procedures for performing particular tasks in application areas. The mathematical model of approximation by generalized step functions [4] is used to analyze critical exposures on CENS operation. Computer simulation models of different shooting systems are considered as the most important sources of initial data on the approximated function. The mathematical model of stress exposures on CENSs that match images by the mutual correlation criterion is developed further.

Keywords: search correlation-extreme navigation system, shooting system, computer simulation model, stress exposure, stress testing, approximation of functions, generalized step function, reference image, current image.

INTRODUCTION

An inertial navigation system (INS) is the main means of controlling the spatial position of ground, air, sea, and space objects. This system controls the coordinates, velocity, and angular position of vehicles relative to the vertical position. Autonomy is one of the main advantages of INSs. However, errors in determining navigation parameters increase over time, and INS data must be corrected, e.g., using satellite navigation systems (SNSs). In this case, autonomy, a very important quality of any moving object, actually disappears [1, 2]. For example, search correlation-extreme navigation systems (CENSs) have wider capabilities in this regard, as they use map-matching autonomous navigation methods. Such systems serve to refine off-line information about the location, orientation, and other parameters of a moving object coming from the main navigation system. A control system uses this information to compensate the deviations in the object's motion parameters to follow a given route. Search CENSs check hypotheses about the values of motion parameters by matching the current terrain sector image received by the onboard shooting system with fragments of a reference image of the application area. The reference images are prepared in advance

and stored in the memory of the onboard computer. When searching for a reference image fragment close by content to the current image (in the sense of a closeness function in the onboard algorithm), a regular shift grid of the frame selecting the next fragment of the reference image is used. The hypotheses that the sought parameters have values equal to those at the grid nodes are checked. The hypothesis for which the closeness function achieves maximum is accepted. Global search schemes, gradient methods from the arsenal of numerical optimization methods, and their combinations are often used [3].

Correlation-extreme navigation systems provide:

– orientation relative to the real terrestrial surface without binding to the Earth reference system, occupying a small area "on the air," in air, land, sea, and outer space;

– high-precision autonomous navigation, whose accuracy depends mainly on the accuracy of the maps used and their processing methods;

– high or complete noise immunity and hidden operation without any maintenance or recovery cost for navigational fields;

– no negative impact on the environment.

Correlation-extreme navigation systems belong to the class of the most stable and reliable systems.

In CENSs, image matching involves a set of features describing the reference and current images that are independent of flight routes and the altitudes and angles of approaching the correction area of INSs. These features must be invariant with respect to possible mutual transformations: shifting, rotation, and scaling of images.

According to the analysis of CENSs, methods for determining coordinates based on correlation image matching are used in a pre-known set of directions of motion to a given point, and the target is a terminal point in the form of some zone. The reliable operation of CENSs requires correlation functions sensitive to mutual rotation, shifting, and scaling of compared images. To form a reference image, CENSs have special technologies for processing preliminarily obtained images and maps.

Presently, to form control actions, necessary data are acquired in modern vision systems to correct the current coordinates of aircraft. As a result, the crew avoids several functions on information processing and aircraft control: CENSs use information from many sensors (shooting systems), namely, a television camera, a thermal imager, a lidar (detection of low-visibility objects in bad weather and at night), and a radar. The information coming from these sensors is subjected to filtering and segmentation. Then contours are identified, the most significant objects of the terrain are classified, and the resulting information is processed jointly with cartographic information. Joint processing yields the most detailed information about the operating environment of CENSs; due to combining heterogeneous information, objects are detected and identified with considerably higher reliability [2].

This paper further develops the concept of an applied geographic information system (AGIS) for modeling search correlation-extreme navigation systems (CENSs), which was presented in [4]. With this system, experts on the autonomous navigation of moving objects by map-matching methods will receive an information technology (a) to create computer models of the existing CENSs and those undergoing various development stages without programming in universal languages, (b) adjust their operation in application areas, and (c) perform necessary computational experiments with these models. With a special language implemented in the AGIS CENSs, which is close to the professional language for the subject domain, and a friendly application programming interface, an expert adjusts the system for a particular modeling task, getting access to the ready-made software components and geospatial data he needs.

According to the paper [4], strict reliability requirements for CENSs increase the role of testing their computer models under stress exposures. During stress testing, the negative effects of different exposures on autonomous navigation conditions are assessed in application areas. Such exposures are not considered at the CENS design stage. Critical exposures are the exposures that prevent CENSs from performing their tasks effectively. Revealing critical exposures motivates improving all CENS elements: sensors of geophysical fields, onboard algorithms, and CENS preparation procedures for performing particular tasks in application areas. Two problems, i.e., revealing critical stress exposures on autonomous navigation conditions in a given application area of a CENS and preparing this CENS for operation in this application area, can be formulated in general form within the mathematical model of approximation by generalized step functions proposed in [4]. For details, see Section 1 below. Also, we justify the requirements for software components necessary to test the solution methods for these problems.

As demonstrated in [4], computer simulation models of CENS shooting systems are the source of the initial data on the approximated function in these problems. Section 2 below introduces approaches to building such computer models. The mathematical model of stress exposures on autonomous navigation

conditions proposed in [4] covers a particular case of search CENSs, in which onboard algorithms implement the procedure of matching the reference image (created when preparing a CENS for operation in an application area) and the current image received by a CENS sensor at the application instant. In the mathematical model mentioned, the closeness of two images is measured by their correlation coefficient. In Section 3 of this paper, we build another mathematical model of significantly more effective stress exposures on autonomous navigation conditions in application areas of such CENSs.

1. PREPARING SEARCH CENSS FOR OPERATION IN AN APPLICATION AREA. CRITICAL EXPOSURES ON AUTONOMOUS NAVIGATION CONDITIONS

In this paper, we continue conceptualizing the AGIS CENSs, resting on the mathematical model of adjusting a CENS to operate in a given application area [4]. Let us extend this model considering its relations with critical exposures.

Like in the paper [4], for illustrative clarity and easy comprehension of the main features, we choose the following class of search CENSs. The shooting system captures a scene image S on a terrain section, and the onboard algorithm refines the planned coordinates $d=(X,Y)$ of the aircraft at the shooting instant. These limitations will affect neither the set of CENS variants (and the procedures of their adjustment to perform a particular task in a given application area) covered by the model, nor the generality of the analysis results and their practical importance for the conceptual AGIS CENSs with critical exposures.

According to the traditional approach, the adjustment problem is to prepare appropriate a priori (reference) information of a reference image and record it into the CENS onboard memory. "Appropriate" means that the onboard computer will determine the navigation parameters of the moving object with a given accuracy by comparing such information with the current information from the onboard shooting system upon arrival in the application area. On the other hand, the problem of critical exposures is preventing from achieving this goal. We denote these problems by $\boldsymbol{Z}^+$ and $\boldsymbol{Z}^-$, respectively.

The mathematical statement of the problem $\boldsymbol{Z}^+$ involves the following basic notions of map-matching autonomous navigation:

• The autonomous navigation conditions of the aircraft in an application area of a CENS are defined by a function $f(S):M\to D$, where M is the set of possible images S coming from the shooting system to the onboard algorithm input in a given application area, and D is the set of possible locations $d\in D$ of the aircraft at the shooting instant. This function describes an objectively existing relationship between the image content (the "value" of the "variable" $S\in M$) and the aircraft location $d=(X,Y)\in D$ when the shooting system "fixes the value" S. The properties of this relationship can facilitate or hinder using map-matching autonomous navigation.

• A computer model of the shooting system can be the source of initial information $I_0\{f(S):M\to D\}$ about this function when adjusting a CENS for operation in a given application area. After adaptation to the application area, this model simulates shooting system operation in this area: $I_0\{f(S):M\to D\} = \hat{f}^{-1}(\theta)(d,p)\colon D\times P\to M$, where $\theta\in\Theta$ is a generalized parameter with a constant value for all $d\in D$ and $p\in P$. Adaptation consists in setting this particular "value." Its physical meaning will become clear from further presentation. The generalized parameter $p\in P$, where P is the set of admissible values, corresponds to the disturbing factors considered in the shooting system modeling. Selecting an application area means choosing the sets D and P of admissible values of the refined and disturbing parameters, respectively. Note that the function $\hat{f}^{-1}$ is the inverse of $f(S):M\to D$, and the computer simulation model of the shooting system implements a parametric family of such functions, $\{\hat{f}^{-1}(\theta)\}_{\theta\in\Theta}$.

• The CENS onboard computer can be treated as a technical realization of the parametric family of single-valued functions $\{\hat{f}(\alpha)(S)\}_{\alpha\in A}$, where $\hat{f}(\alpha)(S):M\to\hat{D}$ is a particular function from this family. During the CENS adjustment procedure, this function is uniquely chosen depending on the value of the generalized parameter $\alpha\in A$.

As shown in [4], the elementary piecewise constant (step) functions of one variable can be generalized and used to approximate the functions $f(S):M\to D$. In this case, the partitions of the real axis into segments correspond to the partitions of the image set M into classes. It has been found that it is reasonable to solve some actual problems of CENS by applying hierarchical partitions of the set into classes and subclasses. For generalized step functions with the same number

of subclasses in each class at the same hierarchical partition level, we have [4] the analytical expression

$$\hat{f}(S)=\sum_{i_1=1}^{l}\chi_{i_1}\left(\pi(S)\right)\sum_{i_2=1}^{l_1}\chi_{i_1i_2}\left(\pi_{i_1}(S)\right)\ldots$$

$$\sum_{i_r=1}^{l_r}\chi_{i_1i_2\ldots i_r}\left(\pi_{i_1i_2\ldots i_{r-1}}(S)\right)\hat{d}_{i_1i_2\ldots i_r}.$$

Removing the constraint on the number of subclasses yields

$$\hat{f}(S)=\sum_{i_1=1}^{l}\chi_{i_1}\left(\pi(S)\right)\sum_{i_2=1}^{l_{i_1}}\chi_{i_1i_2}\left(\pi_{i_1}(S)\right)\ldots$$

$$\sum_{i_r=1}^{l_{i_1i_2\ldots i_r}}\chi_{i_1i_2\ldots i_r}\left(\pi_{i_1i_2\ldots i_{r-1}}(S)\right)\hat{d}_{i_1i_2\ldots i_r}. \quad (1)$$

Then the problem $\boldsymbol{Z}^+$ can be formulated as a function approximation problem.

Statement of the problem $\boldsymbol{Z}^+$. *The function $f(S):M\to D$ is defined by the inverse $\hat{f}^{-1}(\theta)(d,p):D\times P\to M$, where the parameter $\theta\in\Theta$ has a constant value for all $d\in D$ and $p\in P$, being uniquely determined by choosing the sets D and P (all possible values of the refined and disturbing parameters, respectively). The parametric family of generalized step functions $\left\{\hat{f}(\alpha)(S)\right\}_{\alpha\in A}$ (2) and an admissible CENS error $\varepsilon>0$ are given. It is required to find $\alpha^*\in A$ such that $\rho_M\left(\hat{f}\left(\alpha^*\right)(S),f(S)\right)\le\varepsilon$, where ρ_M denotes a metric in the space of functions with the domain M and the codomain D.*

Recall that the analysis is temporarily restricted to the set $D\subset R^2$. Therefore, for a fixed image *S*, the system response $\hat{d}$ is correct if $\rho\left(\hat{d},d\right)\le\varepsilon$, where ρ denotes the distance function of two points in the plane R^2.

Critical exposures are planned and implemented due to the following reasons. Suppose that the problem $\boldsymbol{Z}^+$ is solved. Upon arriving at the application area and receiving the image *S*, the system will calculate the value of the function $\hat{f}(\alpha^*)(S)=\hat{d}$ and return it as the response. Let the correct response be *d*; the stress exposure will be critical if its realization ensures $\rho\left(\hat{d},d\right)>\varepsilon$ for all possible input images *S* in this application area. Within map-matching autonomous navigation, the function $f(S):M\to D$ is subjected to the stress exposure; see the representation above. The function $\hat{f}^{-1}(\theta)(d,p):D\times P\to M$ is the data source for choosing the values of the stress exposure parameters. The exposure has the following effect: when solving the problem $\boldsymbol{Z}^+$, we use the "obsolete" data on the autonomous navigation conditions (before the exposure) instead of the current data. We should have used not $\hat{f}^{-1}(\theta)(d,p)$ but $(R\cdot\hat{f}^{-1}(\theta))(d,p)$, where *R* is an operator transforming the obsolete function $\hat{f}^{-1}(\theta)$ into the current one considering the exposure.

Therefore, the problem $\boldsymbol{Z}^-$ can be generally formulated as follows.

Statement of the problem $\boldsymbol{Z}^-$. *The function $f(S):M\to D$ is given by the inverse $\hat{f}^{-1}(\theta)(d,p):D\times P\to M$, where the parameter $\theta\in\Theta$ has a known value. The solution $\hat{f}\left(\alpha^*\right)(S):M\to\hat{D}$ of the corresponding problem $\boldsymbol{Z}^+$ is also known. It is required to find an operator R such that*

$$\rho\left(\hat{f}\left((\alpha^*)(\hat{f}^{-1}(\theta)(d,p))\right),\hat{f}\left(\left(\alpha^*\right)(R\cdot\hat{f}^{-1}(\theta))(d,p)\right)\right)>\varepsilon$$

$$\forall\,(d,p)\in D\times P.$$

Consider a particular case of this problem to clarify the meaning of general expressions and further develop the concept of the AGIS CENSs. In formula (1), let *r* = 2 (the levels of hierarchical partitioning into classes and subclasses) and the preliminary image transformations be excluded from the analysis, i.e., $\pi(S)=S$ and $\pi_{i_1i_2\ldots i_{r-1}}(S)=S$. Then we have:

$$M=\bigcup_{i=1}^{l}K_i,\ \text{where}\ K_m\bigcap K_n=\varnothing\ \forall m,n\in[1,l],\ m\ne n;$$

$$K_i=\bigcup_{j=1}^{l_i}K_{ij},\ \text{where}\ K_{im}\bigcap K_{in}=\varnothing$$

$$\forall i=1,\ldots,l\ \text{ and }\ m,n\in[1,l_i],\ m\ne n;$$

$$\hat{f}(S)=\sum_{i=1}^{l}\chi_i(S)\sum_{j=1}^{l_i}\chi_{ij}(S)\hat{d}_{ij}. \quad (2)$$

The expression (2) can be written in the vector form

$$\hat{f}(S)=\left\langle\boldsymbol{\chi}(S),\hat{\boldsymbol{f}}(S)\right\rangle, \quad (3)$$

where angle brackets denote the scalar product of generalized vectors and

$$\boldsymbol{\chi}(S)=(\chi_1(S),\chi_2(S),\ldots,\chi_l(S)),$$

$$\hat{\mathbf{f}}(S)=\left(\hat{f}_1(S),\hat{f}_2(S),\ldots,\hat{f}_l(S)\right),$$

$$\hat{f}_i(S)=\sum_{j=1}^{l_i}\chi_{ij}(S)\hat{d}_{ij}=\left\langle\boldsymbol{\chi}_i(S),\hat{\mathbf{d}}_i\right\rangle,$$

$$\boldsymbol{\chi}_i(S)=\left(\chi_{i1}(S),\chi_{i2}(S),\ldots,\chi_{il_i}(S)\right),$$

$$\hat{\mathbf{d}}_i = \left(\hat{d}_{i1}, \hat{d}_{i2}, \ldots, \hat{d}_{il_i}\right),$$
$$i = 1, \ldots, l.$$

Note that $\hat{f}_i(S)$ are approximating functions for the contractions of the function $f(S): M \to D$ into the subsets $K_i \in M$.

The vector function $\hat{\mathbf{f}}(S)$ has the coordinatewise representation

$$\hat{\mathbf{f}}(S) = \left(\left\langle \boldsymbol{\chi}_1(S), \hat{\mathbf{d}}_1 \right\rangle, \left\langle \boldsymbol{\chi}_2(S), \hat{\mathbf{d}}_2 \right\rangle, \ldots, \left\langle \boldsymbol{\chi}_l(S), \hat{\mathbf{d}}_l \right\rangle\right).$$

As a result, the expression (3) takes the form

$$\hat{f}(S) = \left\langle \boldsymbol{\chi}(S), \left(\left\langle \boldsymbol{\chi}_1(S), \hat{\mathbf{d}}_1 \right\rangle, \left\langle \boldsymbol{\chi}_2(S), \hat{\mathbf{d}}_2 \right\rangle, \ldots, \left\langle \boldsymbol{\chi}_l(S), \hat{\mathbf{d}}_l \right\rangle\right)\right\rangle. \quad (4)$$

According to (4), only the vector functions $\boldsymbol{\chi}(S)$ and $\boldsymbol{\chi}_i(S)$ depend on the input image. By definition, however, any algorithm for calculating such functions is a recognition algorithm [5]. As proved therein, any recognition algorithm can be represented as applying the recognition operator **B**(*S*) to the image *S* and the decision rule **C**(**B**(*S*)) to the result. Note that recognition algorithms use feature sets of the recognized object, extracting them from the images *S* coming from the shooting system through their prior transformations. Now let us analyze such transformations. In the expression (4), we have

$$\boldsymbol{\chi}(S) = \mathbf{C}\left(\mathbf{B}\left(\pi(S)\right)\right),$$

where:

$\mathbf{B}\left(\pi(S)\right) = \left(b_1(\pi(S), b_2(\pi(S), \ldots, b_l(\pi(S))\right)$, $b_i\left(\pi(S)\right)$ is a numerical measure of the closeness of a given image to a class $K_i \in M$;

$C\left(b_1\left(\pi(S)\right), b_2\left(\pi(S)\right), \ldots, b_l\left(\pi(S)\right)\right) = (c_1, c_2, \ldots, c_l)$, where $c_i \in \{0, 1\}$, $i = 1, \ldots, l$;

$\boldsymbol{\chi}_i(S) = \mathbf{C}_i\left(\mathbf{B}_i\left(\pi(S)\right)\right)$, $\mathbf{B}_i$, and $\mathbf{C}_i$, where $i = 1, \ldots, l$, are described by analogy.

Hence, the AGIS CENSs should include libraries of software components implementing these operators and the preliminary transformations of images at different levels: from refinement and filtering to the scene description on the terrain section shot [6].

We adopt the obtained structure of vector characteristic functions to reveal the structure of the parameters $\alpha \in A$ in the parametric families of generalized step functions $\left\{\hat{f}(\alpha)(S)\right\}_{\alpha \in A}$. The analysis below is restricted to the decision rules **C** and $\mathbf{C}_i$ without parameters (the most common case according to [5]). Then the only parametric families are those of recognizing operators. For $\left\{\left(\mathbf{B}\left(\alpha^B\right)\left(\pi\left(\alpha^\pi\right)(S)\right)\right)\right\}_{\alpha^B \in A^B}$, we have

$$\boldsymbol{\chi}(S) = \mathbf{C}\left(\mathbf{B}\left(\alpha^B\right)\left(\pi\left(\alpha^\pi\right)(S)\right)\right),$$

where:

$$\mathbf{B}\left(\alpha^B\right)\left(\pi\left(\alpha^\pi\right)(S)\right) = \left(b\left(\alpha_1^B\right)(\pi\left(\alpha^\pi\right)(S),$$
$$b\left(\alpha_2^B\right)\left(\pi\left(\alpha^\pi\right)(S)\right), \ldots, b\left(\alpha_l^B\right)\left(\pi\left(\alpha^\pi\right)(S)\right)\right),$$

$b\left(\alpha_i^B\right)\left(\pi\left(\alpha^\pi\right)(S)\right)$ is a numerical parametric measure of the closeness of a given image to a class $K_i \in M$;

$$\mathbf{C}\left(b\left(\alpha_1^B\right)\left(\pi\left(\alpha^\pi\right)(S)\right), b\left(\alpha_2^B\right)\left(\pi\left(\alpha^\pi\right)(S)\right), \ldots,$$
$$b\left(\alpha_l^B\right)\left(\pi\left(\alpha^\pi\right)(S)\right)\right) = (c_1, c_2, \ldots, c_l),$$

where $c_i \in \{0, 1\}$, $i = 1, \ldots, l$.

The parametric families $\left\{\mathbf{B}_i\left(\alpha_i^{B_i}\right)\left(\pi\left(\alpha^\pi\right)(S)\right)\right\}_{\alpha_i^{B_i} \in A_i^{B_i}}$ are described by analogy.

Then the parameter α can be represented by the set

$$\alpha = (\alpha^\pi; \alpha_1^B, \ldots, \alpha_l^B; \alpha_1^{B_1}, \ldots, \alpha_{l_1}^{B_1}; \alpha_1^{B_2}, \ldots, \alpha_{l_2}^{B_2}; \ldots;$$
$$\alpha_1^{B_l}, \ldots, \alpha_{l_l}^{B_l}; \hat{\mathbf{d}}_1, \ldots, \hat{\mathbf{d}}_l). \quad (5)$$

The CENS is adjusted for operation in an application area by determining the particular values of the preliminary transformation parameters α^π and recognizing operators α^B and α_i^B and the set of vectors $\hat{\mathbf{d}}_i$. (The component $\hat{d}_{ij} = \left(\hat{X}_{ij}, \hat{Y}_{ij}\right)$ of some vector $\hat{d}_i$ will be returned by the system as a response for $S \in K_{ij}$.)

Section 2 considers image modeling issues for different CENS shooting systems and different preliminary transformations $\pi\left(\alpha^\pi\right)(S)$ with the parameters α^π as well as the corresponding parametric families of functions $\hat{f}^{-1}(\theta)(d, p): D \times P \to M$, where D and P are the value sets of the navigation parameter refined by CENSs and the disturbing parameter, respectively.

Concluding this section, we analyze the other parameters of the set (5) and the parameters $\theta \in \Theta$ used to adjust the shooting system model for a given application area in the problems $\boldsymbol{Z}^+$ and $\boldsymbol{Z}^-$. Consider the example of single-level partitions into classes and CENSs with image matching under no disturbing parameters p (Fig. 1).

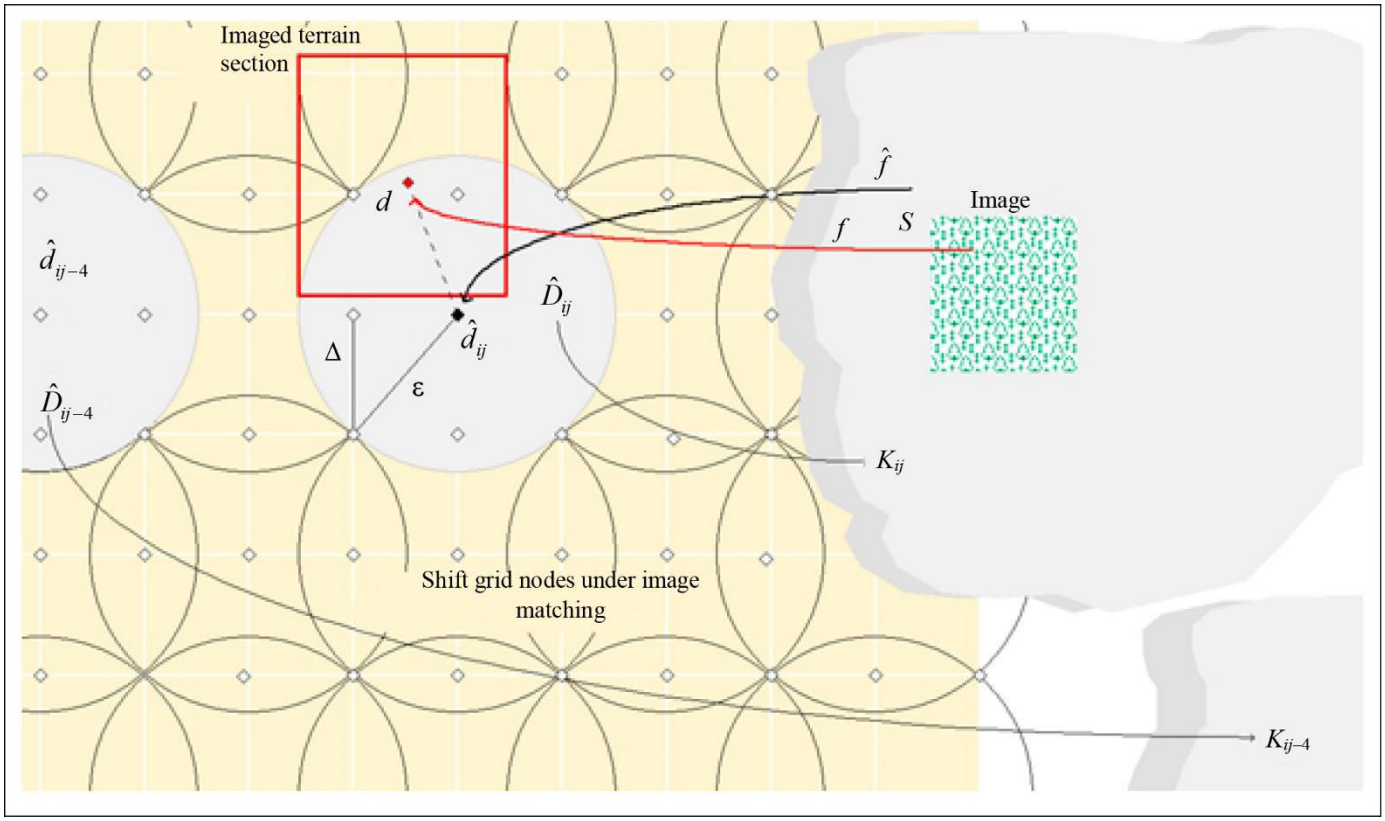

Fig. 1. The notions of the proposed mathematical model.

In this case, the set of possible CENS responses—the nodes of a rectangular grid of shifts—is described by the matrix $\left\|\hat{d}_{ij}\right\| = \left\|\left(\hat{X}_{ij}, \hat{Y}_{ij}\right)\right\|$, where $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$. The one-level partitioning of the image set M and the expression for the step function take the form

$$M = \bigcup_{i=1}^{m}\bigcup_{j=1}^{n} K_{ij}, \text{ where } K_{ij} \bigcap K_{tq} = \varnothing \ \ \forall i \neq t \text{ and } j \neq q,$$

and

$$\hat{f}(S) = \sum_{i=1}^{m}\sum_{j=1}^{n} \chi_{ij}\left(\pi\left(\alpha^{\pi}\right)(S)\right)\hat{d}_{ij} = \sum_{i=1}^{m}\left\langle \boldsymbol{\chi}_i(S), \hat{\mathbf{d}}_i \right\rangle,$$

where $\boldsymbol{\chi}_i(S) = \left(\chi_{i1}(S), \chi_{i2}(S), \ldots, \chi_{in}(S)\right)$,

$\hat{\mathbf{d}}_i = \left(\hat{d}_{i1}, \hat{d}_{i2}, \ldots, \hat{d}_{in}\right), \ i = 1, \ldots, m.$

With $\boldsymbol{\chi}_i$ written through the superposition of the recognizing operator and the decision rule, we obtain

$$\boldsymbol{\chi}_i(S) = \mathbf{C}_i\left(\mathbf{B}_i\left(\alpha^{B_i}\right)\left(\pi\left(\alpha^{\pi}\right)(S)\right)\right),$$

where:

$$\mathbf{B}_i\left(\alpha^{B_i}\right)\left(\pi\left(\alpha^{\pi}\right)(S)\right) = \left(b_1\left(\alpha_1^{B_i}\right)\left(\pi\left(\alpha^{\pi}\right)(S)\right),\right.$$
$$\left. b_2\left(\alpha_2^{B_i}\right)\left(\pi\left(\alpha^{\pi}\right)(S)\right), \ldots, b_n\left(\alpha_n^{B_i}\right)\left(\pi\left(\alpha^{\pi}\right)(S)\right)\right),$$

$b\left(\alpha_j^{B_i}\right)\left(\pi\left(\alpha^{\pi}\right)(S)\right)$ is a numerical parametric measure of the closeness of a given image to a class $K_{ij} \in M$;

$$\mathbf{C}_i\left(b\left(\alpha_1^{B_i}\right)\left(\pi\left(\alpha^{\pi}\right)(S)\right), \ldots, b\left(\alpha_m^{B_i}\right)\left(\pi\left(\alpha^{\pi}\right)(S)\right)\right) = \left(c_1^i, c_2^i, \ldots, c_m^i\right), \text{ where } c_j^i \in \{0, 1\}, \ j = 1, \ldots, m.$$

Then the parameter α is represented by the set

$$\alpha = (\alpha^{\pi}; \alpha_1^{B_1}, \ldots, \alpha_m^{B_1}; \alpha_1^{B_2}, \ldots, \alpha_m^{B_2}; \ldots; \alpha_1^{B_m}, \ldots, \alpha_m^{B_m}; \hat{\mathbf{d}}_1, \ldots, \hat{\mathbf{d}}_m). \quad (6)$$

A domain is called a shooting area if it combines all terrain sections in the shooting system frame in an application area of a CENS. For a fixed shooting system in the case under consideration, the boundaries of such sections depend only on the planned coordinates of the shooting points.

When solving the problems $\boldsymbol{Z}^+$ and $\boldsymbol{Z}^-$, the partition of the set of possible input images M into classes K_{ij} is supposed to satisfy the following condition: if S belongs to a class K_{ij}, it is obtained in a small neighborhood $\hat{D}_{ij}$ of a node $\hat{d}_{ij}$ of the shift grid $\left\|\hat{d}_{ij}\right\| = \left\|\left(\hat{X}_{ij}, \hat{Y}_{ij}\right)\right\|$, where $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$. In this case, the CENS error will not exceed ε:

$\rho\left(\hat{d}_{ij},\ d\right) \le \varepsilon = \Delta\sqrt{2}$, where Δ denotes the distance between the grid nodes. Moreover, it is assumed possible to obtain an appropriate reference image (RI) of the shooting area. "Appropriate" means that this image can be used as the parameter $\theta = RI$ of the simulation model of the shooting system $\hat{f}^{-1}(RI)(d): D \to M$ to find the solution (6) of the problem $\boldsymbol{Z}^{+}$ (and solve the problem $\boldsymbol{Z}^{-}$ as well). The first problem consists in calculating the values of the parameters in formula (6) using $\hat{f}^{-1}(RI)(d)$. By assumption, with some preliminary transformation $(\pi(\alpha^{\pi})(S))$, the solutions are $\alpha_j^{B_i} = \hat{f}^{-1}(RI)(\hat{d}_{ij})$ and $\hat{\mathbf{d}}_i = (\hat{d}_{i1},\ldots,\hat{d}_{in})$. The decision rule searches for the maximum value of $b\left(\alpha_j^{B_i}\right)(\pi(\alpha^{\pi})(S)$ over all i and j and assigns c_j^i a value of 1; for details, see [6, 7].

The formation of reference images satisfying these assumptions is described in Section 2. Section 3 considers the case $b\left(\alpha_j^{B_i}\right)(\pi(\alpha^{\pi})(S) = \mathbf{r}\left(\hat{f}^{-1}(RI)(\hat{d}_{ij}), S\right)$, where $\mathbf{r}$ is the mutual correlation of the image and the reference image fragment corresponding to the shift of the image frame to the node $\hat{d}_{ij}$. For this case, we find an effective operator R solving the problem $\boldsymbol{Z}^{-}$.

2. COMPUTER SIMULATION OF DIFFERENT-TYPE SHOOTING SYSTEMS USED IN CENSS

The correlation image processing algorithm in CENSs is based on maximizing the mutual correlation function C of the current and reference images to decide that at the determination instant the aircraft's coordinates coincide with the reference image center.

Consider control of the value C_1^s during S cycles of stress exposures sequentially introduced in pixels (i, j) of the reference image:

$$C_1^s = \frac{\sum_{i=1}^{Mx}\sum_{j=1}^{My} a_{ij}\left(a_{ij} + \sum_{k=1}^{N} b_k\left(i - x_k, j - y_k\right)\right)}{\sqrt{\sum_{i=1}^{Mx}\sum_{j=1}^{My} a_{ij}^2}\cdot\sqrt{\sum_{i=1}^{Mx}\sum_{j=1}^{My}\left(a_{ij} + \sum_{k=1}^{N} b_k\left(i - x_k, j - y_k\right)\right)^2}}. \quad (7)$$

Here, we use the following notations: a_{ij} is the brightness of the reference image; $a_{ij}\left(a_{ij} + \sum_{k=1}^{N} b_k\left(i - x_k, j - y_k\right)\right)$ is the brightness of the current image (changed by the means of stress exposure by the value b_k in the corresponding pixel (i, j) of the reference image); N is the number of the means of stress exposure; finally, Mx and My are the dimensions of the reference and current images, respectively.

Formula (7) presents the reference and current images. The peculiarities of forming these images and preparing information for the operation of some types of CENS sensors are described below.

The following types of onboard sensors are most developed and widespread.

- Optical sensors in the visible wavelength range of electromagnetic radiation (EMR). The wavelength range of optical radiation is from 100 nm to 1 mm. It is divided into ultraviolet (100–400 nm), visible (400–700 nm), and infrared (700 nm–1 mm); see Fig. 2.
- Thermal sensors in spectral ranges in transparency windows of 0.7–0.9 μm, 0.9–2.5 μm, 3–5 μm, and 8–14 μm (Fig. 2). Thermal sensors have stable operation in the ranges of 3–5 μm and 7–14 μm. CENSs with such sensors work at ranges from a hundred meters to several kilometers.
- Radar sensors in decimeter, centimeter, and millimeter radio bands. They are used mainly to detect ground objects. Active radar sensors can operate at several tens of kilometers but are almost not used for imaging in CENSs. However, there exist many different means with active radar navigation systems, and they can be subjected to stress exposure [8–11].
- Hyper-spectrometers. The development of CENSs with hyper-spectrometers is a priority: multiple information acquisition channels provide real opportunities related to permeability in any environment, reliability of object classification, and selection of interference of natural and artificial origin. In practice, CENSs can implement intelligent positioning of aircraft using modern signal processing methods [12, 13].

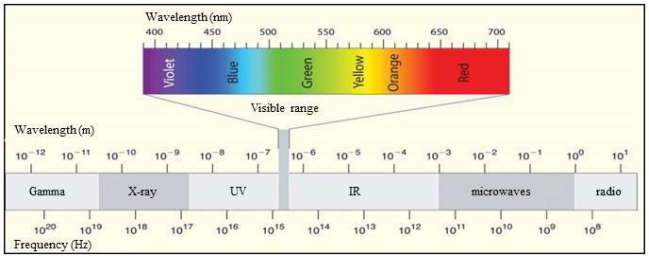

Fig. 2. Radiation ranges.

When creating the unified means of reconnaissance and CENSs, adaptation to a particular task can be carried out by changing the mathematical apparatus of signal processing. Hyperspectral images allow detecting buried and caved objects, minefields, and underground communications due to the significant difference in their reflection spectra. Hyperspectral images display continuous spectral bands, unlike multiband images with separated spectral bands.

CENS sensors of visible and thermal ranges of EMR wavelengths have high spatial and temperature resolution. Radio-thermal CENS sensors provide the resolution necessary for recognizing many areal terrain objects [13].

An important issue is determining the number of the means of stress exposure in the INS correction area. For this purpose, we estimate the accuracy of aircraft navigation using CENSs [5, 8, 9, 14–16]:

$$P_{\text{obj}} = 1 - \exp\left(-\frac{R_{\text{obj}}^2}{2\sigma_{\text{nav}}^2}\right),$$

where P_{obj} is the probability of determining the object's coordinates, R_{obj} is the object's radius on the terrain (in m), and σ_{nav} is the navigation accuracy of the inertial system (in m).

By assumption, an object is detected if the CENS sensor's field of view at least touches the outer boundary of the terrain section where the object is located. The navigation point obeys the Gaussian distribution law, and the lateral deviation and range deviation coincide with one another (Fig. 3).

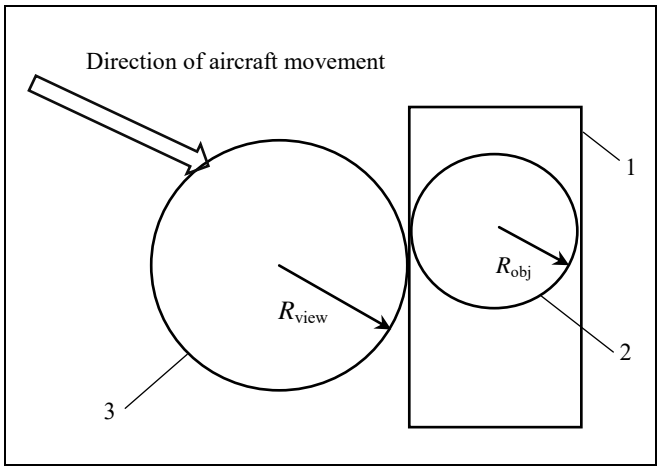

Fig. 3. The sighting scheme for aircraft equipped with CENS at a given point: 1—object location area, 2—the largest radius of the circle fully inscribed in the outer contours of the object, 3—CENS field of view.

When the CENS sensor's field of view (R_{view}) touches the object's location area, particularly the circle of radius R_{obj}, the probability of CENS activation can be supposed 1.

Under the maximum possible errors of the inertial navigation system, the dimensions of the reference image of the terrain must be larger than those of the current image. This condition has not been justified in the available literature. According to [8, 9, 17], the dimensions of the current image obtained by the CENS sensor during navigation can be half the dimensions of the reference image stored in the onboard navigation system.

We emphasize another important aspect. The approach of a moving object to the CENS operation area is guided by the INS and therefore depends on its accuracy. The CENS sensor's field of view on the terrain from the activation altitude must have a radius exceeding three times the mean square error of the INS (σ_{INS}). This requirement ensures position correction by the CENS with a probability close to 1, in accordance with the Gaussian distribution of errors when approaching the object area (under the failure-free operation of the CENS).

If the CENS fails, the accuracy of aircraft navigation will be determined by the accuracy of the inertial navigation system, $\sigma_{\text{nav.}}$ For this purpose, the generated current image should be appropriately modified at the position correction instant by the ground means of stress exposure. Also, they can be placed in the cone of space between the CENS sensor and the object (the circle of radius R_{obj}).

Analysis shows that the expression (7) is sensitive to image distortions and the value C is controllable. Hence, we can propose an optimal method of finding the number and power of the means of stress exposure on the CENS sensor. This method sequentially extracts the brightest or dimmest pixels in the reference image depending on the types of such means.

Methods of purposeful impact on the operation of CENS sensors can reduce their efficiency in modern navigation systems.

Consider the issues of preparing images for use when implementing stress exposure methods for CENSs.

For the areas of stress exposures on the CENS operation, it is necessary to prepare special images by processing photos. A digital image should be partitioned into larger pixels [18]. For example, an original color image of the terrain (Fig. 4) was divided into pixels of dimensions 587×441, and the side length of each pixel of the black-and-white image was 7 m on the terrain (Fig. 5). The following objects were represented in pixel format: a road network (Fig. 6), vegetation (Fig. 7), hydrography (Fig. 8), and a settlement (Fig. 9). These objects are the most informative in the visible range of EMR wavelengths.

The pixel images in Figs. 4–9 have dimensions of 587×441 elements. Each pixel was assigned an average brightness of several pixels of the original image. For preliminary processing, we used a terrain photo consisting of 4241×3769 elements. The side sizes of one pixel were 0.85 m (the original color image on the terrain) and about 7 m (the transformed black-and-white image). These figures are important for further calculations, especially when processing images of large dimensions.

Fig. 4. An original image.

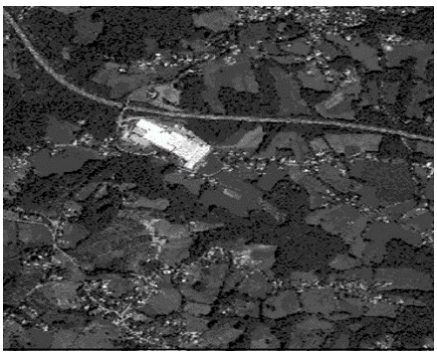

Fig. 5. A black and white image.

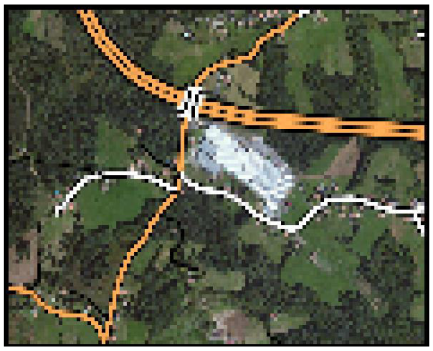

Fig. 6. A road network.

Fig. 7. Vegetation.

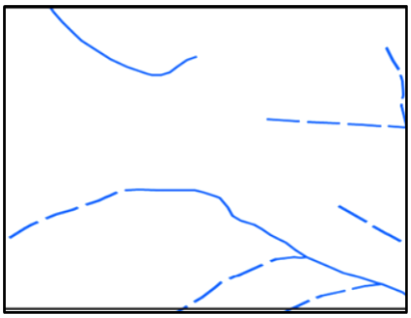

Fig. 8. Hydrography.

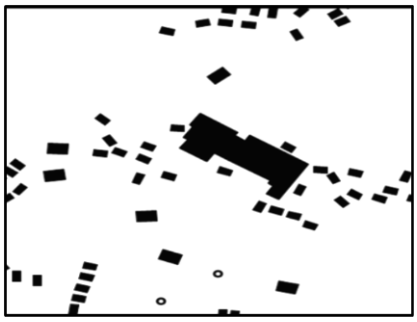

Fig. 9. A settlement.

Figure 5 is a black-and-white image where each pixel shows an optical density value.

This image contains optical density values as the brightness of the reference image (formula (7)). As a rule, the minimum optical density limit does not equal 0: usually, its value is 0.1 conventional units and greater. On the other hand, the maximum limit never equals 4.0: usually, it is less than 3.8. Such ranges can be processed by scanners (e.g., when inputting a digital image). On terrain images, it is possible to display objects with optical densities of 0.1–2.2 and greater. The practice-based scale of these values is presented in the table below.

Tone scale on a black-and-white terrain image

Image phototone	Separation principle	Optical density D
White	Visually distinguishable tone in the image	0.1 or less
Almost white	Optical density of the veil	0.2–0.3
Light gray	Minimum optical density	0.4–0.6
Gray	Average optical density	0.7–1.1
Dark gray	Maximum density	1.2–1.6
Almost black	Tone exceeding the maximum density	1.7–2.2
Black	Visually distinguishable tone of scale	2.2 or more

Figure 10 shows a monochrome image; shades of gray are more likely, but other combinations with tones of one color, such as green-white or green-red, as well as tones from light brown to dark brown, are also possible.

Under digital processing, monochrome has only two values:

– only one color (either on or off), i.e., a binary image;

– shades of this color.

Figure 11 demonstrates a posterized image, specifying the number of tonal levels (brightnesses) of the image. This is necessary, e.g., to create large monotone areas when forming reference images for the subsequent modeling of effective stress exposure conditions for the CENS operation. In all digital images, the color levels are discrete, and the smooth continuous transition between them is achieved by the number of these levels. The black-and-white image in Fig. 12 serves for comparison with the posterized one.

Let us discuss some features of using radar images.

Getting detail at less than 50 cm per pixel is not a challenge for modern satellite radars. However, this resolution applies to only one axis: the pixel will be not square. On the second axis, the resolution could be around 100 cm. For example, 100×25 cm is one of the best current results for the reference image [10]. A radar image is obtained monochrome, with various

Fig. 10. A monochrome image.

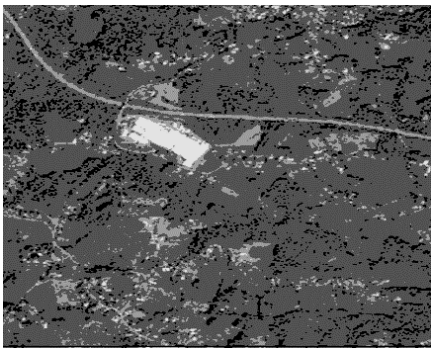

Fig. 11. A posterized image.

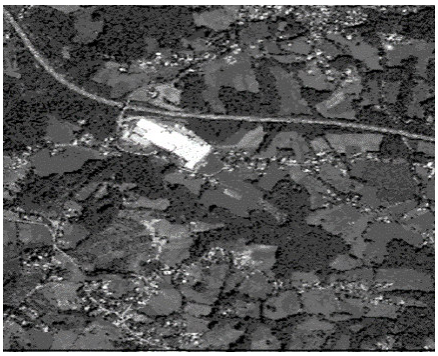

Fig. 12. A black-and-white image.

features of radio waves reflection from different objects, including the display of some details hidden on conventional optical images.

For map-matching navigation systems, we consider the methods and technologies of obtaining and processing satellite photos: other sources are unlikely [3, 11].

When modeling a navigational sensor, a monochrome image most clearly reflects the available reference objects as shades of black or white image fragments, whereas a color image is used only to represent perceived brightness by combining several channels (usually, red, blue, and green). The individual channels can be weighted to achieve the desired result. For example, the green and blue channels can be combined and the red channel can be turned off.

Under hypercube information processing, various representations are possible for the final expressive results in the image, associated with separating the necessary reference objects for CENSs. The reference points should be selected using SURF, one of the most effective modern pattern recognition algorithms [12, 19].

This algorithm includes the following stages:

– performing the scale-space representation,
– calculating the Hessian values,
– searching for local maxima,
– determining the true maximum,
– determining the reference point orientation,
– forming the reference point descriptor.

SURF searches for the key points of the image using the Hessian matrix and creates their descriptors invariant to scale and rotation. The Hessian achieves maximum at the maximum change points of the brightness gradient. It is good at detecting spots, corners, and line edges. The Hessian is invariant to image brightness shift but not to scale. This problem is solved by testing different scales and filters, applying them one by one to a single pixel. The method splits the entire set of scales into octaves.

Different types of images allow choosing objects for managing the computer simulation of CENS performance improvement conditions and CENS failure conditions when the means of stress exposure are applied to onboard motion control programs.

A binary (black and white) image involves only two quantization levels and represents only white and black colors. A grayscale image uses 256 quantization levels, with 8 bits (1 byte) reserved for the description of each image element. Black always corresponds to level 0, whereas white corresponds to level 1 of the binary image and level 255 of the grayscale image.

A color image is formed using a particular palette (RGB, CMYK, etc.). In each palette, colors and their tones are created by mixing the three primary colors in proportions corresponding to their quantization levels.

The RGB palette uses three primary colors: Red, Green, and Blue.

To represent one image element, we need:

– 1 bit for a binary image,
– 8 bits (1 byte) for a grayscale image with 256 quantization levels,
– 24 bits (3 bytes) for a color image with the same number of quantization levels.

When forming an image, the capabilities of onboard CENS computers and their information support are taken into account.

3. OPTIMIZATION OF STRESS EXPOSURES ON AUTONOMOUS NAVIGATION CONDITIONS OF CENSS WITH IMAGE MATCHING BY MUTUAL CORRELATION

Mathematical expressions for calculating automatically the optimal powers of the means of stress exposure located in calculated coordinates were obtained in the paper [4]. As established therein, the means of stress exposure arranged on the terrain (object) have different impacts on the current image formed by the CENS sensor, affecting the similarity of the reference

and current images. By moving the means of stress exposure in a certain installation domain, we can minimize the correlation function of the reference and current images.

Problem statement. *The physical field brightness distribution a_{ij} is known. There are N means of stress exposure on a navigation system sensor, and means k induces a given additional field brightness $b_k(i, j)$ when placed at the origin. Consider a known variant of placing means k at a point with coordinates (x_k, y_k), $k = 1,..., N$. After automatically calculating the optimal powers A_k of each means of stress exposure, it is required to find the displacements $(\Delta x_k, \Delta y_k)$, $k = 1,..., N$, ensuring the reduced correlation between the current and reference images* [10].

Problem solution. In the elementary case, the correlation is calculated on the window $1 \le i \le Mx$, $1 \le j \le My$. Under a given arrangement of the means of stress exposure, the physical field brightness at a point with coordinates (i, j) will be changed means k with power A_k to the value

$$a_{ij} + \sum\nolimits_{k=1}^{N} A_k b_k \left(i - x_k - \Delta x_k, j - y_k - \Delta y_k\right),$$

where: i and j are the pixel coordinates on the reference image; x_k and y_k are the coordinates of the means of stress exposure; Δx_k and Δy_k are the displacements of the means of stress exposure.

The correlation between the current and reference images is given by

$$C = \frac{\sum\limits_{i=1}^{Mx}\sum\limits_{j=1}^{My} a_{ij}\left(a_{ij} + \sum\limits_{k=1}^{N} A_k b_k\left(i - x_k - \Delta x_k,\ j - y_k - \Delta y_k\right)\right)}{\sqrt{\sum\limits_{i=1}^{Mx}\sum\limits_{j=1}^{My} a_{ij}^2} \times \sqrt{\sum\limits_{i=1}^{Mx}\sum\limits_{j=1}^{My}\left(a_{ij} + \sum\limits_{k=1}^{N} A_k b_k\left(i - x_k - \Delta x_k,\ j - y_k - \Delta y_k\right)\right)^2}}. \quad (8)$$

To solve the problem, we find the components of the displacement vector $(\Delta x_1,..., \Delta x_N, \Delta y_1,..., \Delta y_N)$ that decrease the correlation value C. For this purpose, the displacement vector must be opposite to the gradient vector of the function C (the partial derivatives of C in the direction $(\Delta x_1,..., \Delta x_N, \Delta y_1,..., \Delta y_N)$. These partial derivatives are:

$$\frac{\partial C}{\partial \Delta x_s} = \frac{1}{\sqrt{\sum\limits_{i=1}^{Mx}\sum\limits_{j=1}^{My} a_{ij}^2}} \times \left\{ \frac{-\sum\limits_{i=1}^{Mx}\sum\limits_{j=1}^{My} a_{ij} A_s \frac{\partial b_s}{\partial x}\left(i - x_s - \Delta x_s,\ j - y_s - \Delta y_s\right)}{\sqrt{\sum\limits_{i=1}^{Mx}\sum\limits_{j=1}^{My}\left(a_{ij} + \sum\limits_{k=1}^{N} A_k b_k\left(i - x_k - \Delta x_k,\ j - y_k - \Delta y_k\right)\right)^2}} + \right.$$

$$\left. \frac{\left[\sum\limits_{i=1}^{Mx}\sum\limits_{j=1}^{My} a_{ij}\left(a_{ij} + \sum\limits_{k=1}^{N} A_k b_k\left(i - x_k - \Delta x_k,\ j - y_k - \Delta y_k\right)\right) \times \right.}{\sqrt{\left(\sum\limits_{i=1}^{Mx}\sum\limits_{j=1}^{My}\left(a_{ij} + \sum\limits_{k=1}^{N} A_k b_k\left(i - x_k - \Delta x_k,\ j - y_k - \Delta y_k\right)\right)^2\right)^3}} \right.$$

$$\left. \left. \times \sum\limits_{i=1}^{Mx}\sum\limits_{j=1}^{My}\left(A_s \frac{\partial b_s}{\partial x}\left(i - x_s - \Delta x_s,\ j - y_s - \Delta y_s\right) \times \left(a_{ij} + \sum\limits_{k=1}^{N} A_k b_k\left(i - x_k - \Delta x_k,\ j - y_k - \Delta y_k\right)\right)\right)\right] \right\} =$$

$$\frac{1}{\sqrt{\sum\limits_{i=1}^{Mx}\sum\limits_{j=1}^{My} a_{ij}^2} \times \sqrt{\left(\sum\limits_{i=1}^{Mx}\sum\limits_{j=1}^{My}\left(a_{ij} + \sum\limits_{k=1}^{N} A_k b_k\left(i - x_k - \Delta x_k, j - y_k - \Delta y_k\right)\right)^2\right)^3}} \times$$

$$\left\{ -\sum\limits_{i=1}^{Mx}\sum\limits_{j=1}^{My} a_{ij} A_s \frac{\partial b_s}{\partial x}\left(i - x_s - \Delta x_s,\ j - y_s - \Delta y_s\right) \times \sum\limits_{i=1}^{Mx}\sum\limits_{j=1}^{My}\left(a_{ij} + \sum\limits_{k=1}^{N} A_k b_k\left(i - x_k - \Delta x_k,\ j - y_k - \Delta y_k\right)\right)^2 + \right.$$

$$\left. \sum\limits_{i=1}^{Mx}\sum\limits_{j=1}^{My} a_{ij}\left(a_{ij} + \sum\limits_{k=1}^{N} A_k b_k\left(i - x_k - \Delta x_k,\ j - y_k - \Delta y_k\right)\right) \times \right.$$

$$\sum_{i=1}^{Mx}\sum_{j=1}^{My}\left(A_s \frac{\partial b_s}{\partial x}\left(i-x_s-\Delta x_s,\, j-y_s-\Delta y_s\right)\times\left(a_{ij}+\sum_{k=1}^{N} A_k b_k\left(i-x_k-\Delta x_k,\, j-y_k-\Delta y_k\right)\right)\right)\Bigg\}=$$

$$\frac{1}{\sqrt{\sum_{i=1}^{Mx}\sum_{j=1}^{My} a_{ij}^2}\times\sqrt{\left(\sum_{i=1}^{Mx}\sum_{j=1}^{My}\left(a_{ij}+\sum_{k=1}^{N} A_k b_k\left(i-x_k-\Delta x_k,\, j-y_k-\Delta y_k\right)\right)^2\right)^3}}\times$$

$$\Bigg\{\underline{-\sum_{i=1}^{Mx}\sum_{j=1}^{My} a_{ij} A_s \frac{\partial b_s}{\partial x}\left(i-x_s-\Delta x_s,\, j-y_s-\Delta y_s\right)\times\sum_{i=1}^{Mx}\sum_{j=1}^{My} a_{ij}^2}-$$

$$2\sum_{i=1}^{Mx}\sum_{j=1}^{My} a_{ij} A_s \frac{\partial b_s}{\partial x}\left(i-x_s-\Delta x_s,\, j-y_s-\Delta y_s\right)\times\sum_{i=1}^{Mx}\sum_{j=1}^{My} a_{ij}\sum_{k=1}^{N} A_k b_k\left(i-x_k-\Delta x_k,\, j-y_k-\Delta y_k\right)-$$

$$\sum_{i=1}^{Mx}\sum_{j=1}^{My} a_{ij} A_s \frac{\partial b_s}{\partial x}\left(i-x_s-\Delta x_s,\, j-y_s-\Delta y_s\right)\times\sum_{i=1}^{Mx}\sum_{j=1}^{My}\left(\sum_{k=1}^{N} A_k b_k\left(i-x_k-\Delta x_k,\, j-y_k-\Delta y_k\right)\right)^2+$$

$$\underline{\sum_{i=1}^{Mx}\sum_{j=1}^{My} a_{ij}^2\times\sum_{i=1}^{Mx}\sum_{j=1}^{My} a_{ij} A_s \frac{\partial b_s}{\partial x}\left(i-x_s-\Delta x_s,\, j-y_s-\Delta y_s\right)}+$$

$$\sum_{i=1}^{Mx}\sum_{j=1}^{My} a_{ij}^2\times\sum_{i=1}^{Mx}\sum_{j=1}^{My}\left(A_s \frac{\partial b_s}{\partial x}\left(i-x_s-\Delta x_s,\, j-y_s-\Delta y_s\right)\times\sum_{k=1}^{N} A_k b_k\left(i-x_k-\Delta x_k,\, j-y_k-\Delta y_k\right)\right)+$$

$$\sum_{i=1}^{Mx}\sum_{j=1}^{My}\left(a_{ij}\sum_{k=1}^{N} A_k b_k\left(i-x_k-\Delta x_k,\, j-y_k-\Delta y_k\right)\right)\times\sum_{i=1}^{Mx}\sum_{j=1}^{My} a_{ij} A_s \frac{\partial b_s}{\partial x}\left(i-x_s-\Delta x_s,\, j-y_s-\Delta y_s\right)+$$

$$\sum_{i=1}^{Mx}\sum_{j=1}^{My}\left(a_{ij}\sum_{k=1}^{N} A_k b_k\left(i-x_k-\Delta x_k,\, j-y_k-\Delta y_k\right)\right)\times$$

$$\sum_{i=1}^{Mx}\sum_{j=1}^{My}\left(A_s \frac{\partial b_s}{\partial x}\left(i-x_s-\Delta x_s,\, j-y_s-\Delta y_s\right)\times\sum_{k=1}^{N} A_k b_k\left(i-x_k-\Delta x_k,\, j-y_k-\Delta y_k\right)\right)\Bigg\}=$$

(*By analogy with the paper* [4], *the underlined terms are mutually reduced.*)

$$=\frac{1}{\sqrt{\sum_{i=1}^{Mx}\sum_{j=1}^{My} a_{ij}^2}\times\sqrt{\left(\sum_{i=1}^{Mx}\sum_{j=1}^{My}\left(a_{ij}+\sum_{k=1}^{N} A_k b_k\left(i-x_k-\Delta x_k,\, j-y_k-\Delta y_k\right)\right)^2\right)^3}}\times$$

$$\Bigg\{-2\sum_{i=1}^{Mx}\sum_{j=1}^{My} a_{ij} A_s \frac{\partial b_s}{\partial x}\left(i-x_s-\Delta x_s,\, j-y_s-\Delta y_s\right)\times\sum_{i=1}^{Mx}\sum_{j=1}^{My} a_{ij}\sum_{k=1}^{N} A_k b_k\left(i-x_k-\Delta x_k,\, j-y_k-\Delta y_k\right)-$$

$$\sum_{i=1}^{Mx}\sum_{j=1}^{My} a_{ij} A_s \frac{\partial b_s}{\partial x}\left(i-x_s-\Delta x_s,\, j-y_s-\Delta y_s\right)\times\sum_{i=1}^{Mx}\sum_{j=1}^{My}\left(\sum_{k=1}^{N} A_k b_k\left(i-x_k-\Delta x_k,\, j-y_k-\Delta y_k\right)\right)^2+$$

$$\sum_{i=1}^{Mx}\sum_{j=1}^{My} a_{ij}^2\times\sum_{i=1}^{Mx}\sum_{j=1}^{My}\left(A_s \frac{\partial b_s}{\partial x}\left(i-x_s-\Delta x_s,\, j-y_s-\Delta y_s\right)\times\sum_{k=1}^{N} A_k b_k\left(i-x_k-\Delta x_k,\, j-y_k-\Delta y_k\right)\right)+$$

$$\sum_{i=1}^{Mx}\sum_{j=1}^{My}\left(a_{ij}\sum_{k=1}^{N} A_k b_k\left(i-x_k-\Delta x_k,\, j-y_k-\Delta y_k\right)\right)\times\sum_{i=1}^{Mx}\sum_{j=1}^{My} a_{ij} A_s \frac{\partial b_s}{\partial x}\left(i-x_s-\Delta x_s,\, j-y_s-\Delta y_s\right)+$$

$$\sum_{i=1}^{Mx}\sum_{j=1}^{My}\left(a_{ij}\sum_{k=1}^{N}A_k b_k\left(i-x_k-\Delta x_k,\, j-y_k-\Delta y_k\right)\right)\times$$

$$\sum_{i=1}^{Mx}\sum_{j=1}^{My}\left(A_s\frac{\partial b_s}{\partial x}\left(i-x_s-\Delta x_s,\, j-y_s-\Delta y_s\right)\times\sum_{k=1}^{N}A_k b_k\left(i-x_k-\Delta x_k,\, j-y_k-\Delta y_k\right)\right)\Bigg\}.$$

A similar expression for $\frac{\partial C}{\partial \Delta y_s}$ is obtained by replacing $\frac{\partial b_s}{\partial x}$ with $\frac{\partial b_s}{\partial y}$. The gradient vector in the direction (Δx_1,..., Δx_N, Δy_1,..., Δy_N) gives the displacement vector of the means of stress exposure that reduces the correlation. The displacement value can be calculated by minimizing a function of one variable using the standard bisection algorithm. The iterative application of such displacements allows reducing the correlation at each step. The proposed bisection method is a sequential minimization method for function (8). It is possible to construct nested segments, each containing at least one of the optima (the minimum of the correlation function). There are no such applications of the bisection method, as well as other optimization methods and algorithms for correlation functions, in the available literature. This process continues until achieving a given rate of convergence. Its value is obtained experimentally for each type of CENS sensors during simulation modeling.

The described method can serve to optimize the arrangement of the means of stress exposure in optical, thermal, and radio-thermal wavelength EMR ranges EMI and when using the hypercube of information obtained by hyper-spectrometers in transparency windows. In the radar wavelength range, however, an additional problem arises: the placed means of stress exposure must be oriented in space relative to the field of view of the navigation system sensor so that the reflected signal will return to the system sensor. The means of stress exposure in the radar range are aimed at reducing or increasing the effective scattering surfaces and the distortion of the backward secondary radiation diagram based on reflecting and absorbing composite materials.

CONCLUSIONS

This paper continues studies to justify the project of an applied geographic information system (AGIS) for modeling search correlation-extreme navigation systems (CENSs). The corresponding concept was presented in [4]. Let us summarize the results:

- The mathematical model of approximation of the functions describing autonomous navigation conditions in application areas of search CENSs by generalized step functions has been further developed. This model has been employed to analyze critical stress exposures on the CENS operation. Based on the analysis results, the requirements for the applied geographic information system (AGIS) for the developers of CENSs have been specified.
- Possible dimensions have been determined for the reference image formation area to implement stress testing. Also, the impact of stress testing on the location of the carrier of the map-matching autonomous navigation system has been assessed.
- New procedures have been formulated for the technology of preparing special images with a certain structure. These procedures allow constructing interference zones for onboard sensors, including purposeful impacts to violate the system.
- The arrangement of the means of stress exposure on map-matching autonomous navigation systems with image matching by mutual correlation has been improved.
- Attention has been paid to the use of different physical fields of the Earth in navigation systems. They significantly affect the approaches to stress exposures on the operation of navigation systems.
- The obtained results can be used in technologies for creating modern navigation systems and their testing in complex operating conditions of information sensors in the AGIS CENSs.

REFERENCES

1. Avgustov, L.I., Orientation by Geophysical Fields Provides Autonomous Navigation of Combat Aircraft, *Kommersant-nauka*, 2015, no. 2, pp. 34–35 (In Russian.)
2. Karshakov, E.V., Pavlov, B.V., and Tkhorenko, M.Yu., Promising Map-Aided Aircraft Navigation Systems, *Gyroscopy Navig.*, 2021, vol. 12, no. 1, pp. 38–49. https://doi.org/10.1134/S2075108721010077.
3. Krasovskii, A.A., Beloglazov, I.N., and Chigin, G.P., *Teoriya korrelyatsionno-ekstremal'nykh navigatsionnykh system* (Theory of Correlation-Extreme Navigation Systems), Moscow: Nauka, 1979. (In Russian.)
4. Alchinov, A.I. and Gorokhovsky, I.N., A Conceptual Applied Geographic Information System for Modeling Search Autonomous Correlation-Extreme Navigation Systems, *Control Sciences*, 2022, no. 1, pp. 43–54.

5. Zhuravlev, Yu.I., Zenkin, A.A., Zenkin, A.I., et al., Recognition and Classification Problems with Standard Training Information, *USSR Computational Mathematics and Mathematical Physics*, 1980, vol. 20, no. 5, pp. 195–211.
6. Duda, R.O. and Hart, P.E., *Pattern Classification and Scene Analysis*, Wiley, 1973.
7. Volkovitskiy, A.K., Gladyshev, A.I., Goldin, D.A., et al., A Computer Simulation Complex for Analysis of Magnetic Gradiometry Systems, *Control Sciences*, 2021, no. 3, pp. 57–65.
8. Bolkunov, A.A., Ryazantsev, L.B., and Sidorenko, S.V., Assessment of the Radar Visibility of Armaments, Military and Special Equipment with Using Unmanned Aerial Vehicles, *Voennaya Mysl'*, 2017, no. 9, pp. 70–73. (In Russian.)
9. Likhachev, V.P., Pantyukhin, M.A., and Sidorenko, S.V., An Algorithm for Morphological Processing of Radar Images and Automatic Detection of Objects by Radar Shadow, *Proceedings of Voronezh State Univ. Ser. Syst. Anal. Inform. Tech.*, 2018, no. 2, pp. 150–161.
10. Kupryashkin, I.F., Likhachev, V.P., Seleznev, D.A., and Usov, N.A., A Method to Distort the Radar Image in a Space Radar Station with Synthetic Aperture Antenna, *RF Patent 2622904 RU*, 2017. (In Russian.)
11. *Tekhnicheskie sredstva razvedsluzhb kapitalisticheskikh gosudarstv* (Technical Means of Intelligence Services of Capitalist States), Information Bulletin of All-Russian Institute for Scientific and Technical Information RAS (VINITI), 2009–2015. (In Russian.)
12. Umale, P., Patil, A., Sahani, C., et al., Planer Object Detection Using SURF and SIFT Method, *International Journal of Engineering Applied Sciences and Technology*, 2022, vol. 6, iss. 11, pp. 36–39.
13. Świeżewski, J., YOLO Algorithm and YOLO Object Detection, 2020. URL: https://appsilon.com/object-detection-yolo-algorithm/. (Accessed November 11, 2022.)
14. Abezgauz, G.G., Tron', A.P., Kopenkin, Yu.R., and Korovina, I.A., *Spravochnik po veroyatnostnym raschetam* (Handbook of Probabilistic Calculations), Moscow: Voenizdat, 1970. (In Russian.)
15. Kutakhov, V.P. and Meshcheryakov, R.V., Group Control of Unmanned Aerial Vehicles: A Generalized Problem Statement of Applying Artificial Intelligence Technologies, *Control Sciences*, 2022, no. 1, pp. 55–60.
16. Yuyukin, I.V., Correlation-Extreme Navigation Through Geophysical Fields Based on the Use of Spline Technology, *Vestnik Gos. Univ. Morsk. Rechn. Flota im. Adm. S.O. Makarova*, 2021, vol. 13, no. 4, pp. 505–517. DOI: 10.21821/2309-5180-2021-13-4-505-517. (In Russian.)
17. Antyufeev, V.I., Bykov, V.N., Grichanyuk, A.M., et al., *Matrichnye radiometricheskie korrelyatsionno-ekstremal'nye sistemy navigatsii letatel'nykh apparatov* (Matrix Radiometric Correlation-Extreme Navigation Systems for Aircraft), Kharkov: Shchedraya Usad'ba Plyus, 2014. (In Russian.)
18. Biryukov, V.S., Digital Images in Photogrammetry, *Geodesy and Cartography*, 2000, no. 10, pp. 33–36. (In Russian.)
19. Dyshlyuk, V. O., A Study of Quality and Speed Indicators When Searching for Reference Points on Images by the SURF Method, *Molodoi Uchenyi*, 2018, no. 27 (213), pp. 23–26. (In Russian.)

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

Received July 12, 2022, and revised November 14, 2022.
Accepted November 29, 2022.

Author information

Alchinov, Alexander Ivanovich. Dr. Sci. (Eng.), Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia
✉ alchinov46@mail.ru

Gorokhovsky, Igor Nikolaevich. Dr. Sci. (Eng.), Research Center of Topographic and Navigational Support, Central Research Institute No. 27, Moscow, Russia
✉ gin_box@mail.ru

Cite this paper

Alchinov, A.I., Gorokhovsky, I.N., Analysis of Stress Exposures on Autonomous Navigation Conditions in Search Correlation-Extreme Navigation Systems. *Control Sciences* **6,** 35–48 (2022). http://doi.org/10.25728/cs.2022.6.5

Translated into English by *Alexander Yu. Mazurov*,
Cand. Sci. (Phys.–Math.),
Trapeznikov Institute of Control Sciences,
Russian Academy of Sciences, Moscow, Russia
✉ alexander.mazurov08@gmail.com

DOI: http://doi.org/10.25728/cs.2022.6.6

15TH INTERNATIONAL CONFERENCE ON MANAGEMENT OF LARGE-SCALE SYSTEM DEVELOPMENT (MLSD'2022)

The 15th International Conference on Management of Large-Scale System Development (MLSD'2022) was held on September 26–28, 2022. This conference is organized annually by Trapeznikov Institute of Control Sciences, Russian Academy of Sciences (ICS RAS), with the support of the IEEE Russia Section. The conference aims to promote R&D cooperation on various managerial aspects of large-scale system development at sectoral, regional, national, and transnational levels.

Leading scientists from academia, research institutes, universities, and governmental and commercial organizations, professionally involved in the theory and practice of management in the modern era of the information society, are traditional participants of the conference.

Following the publication policy, MLSD conferences broadly present to the scientific community new approaches, principles, and capabilities of cybernetic large-scale management based on mathematical modeling and modern information and communication technologies.

Every year the conference proceedings are published in Russian; since 2017, selected conference papers have been placed in the IEEE *Xplore* digital repository. About 150 MLSD papers annually contribute to the world's scientific collection.

The MLSD'2022 program included 18 plenary and 199 sectional papers of leading experts from 30 cities of Belarus, Kazakhstan, China, the USA, and Russia. Amongst them, 155 papers were extended and published electronically in IEEE *Xplore*.[1]

Traditionally, the conference has a plenary session and 16 sections in the following areas.

Section 1. Management problems of 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. Management of development of a digital economy. Design offices both situational and expected analytical centers, institutes of development of large-scale systems.

Section 4. Simulation and optimization in the problems of development management of large-scale systems.

Section 5. Nonlinear processes and computing methods in the problems of management of large-scale systems.

Section 6. Management of 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 objects of nuclear power and other objects of increased danger.

Section 12. Information support and software management systems for large-scale production.

Section 13. Methodology, methods, software, and algorithmic support of intellectual processing of large volumes of information.

Section 14. Monitoring in the management of large-scale systems.

Section 15. Management of large-scale systems advancement in healthcare, medico-biological systems, and technologies.

Section 16. Managing the development of social systems.

MLSD'2022 aimed to cover big data management issues, including big data use in various areas of management, as well as the standardization of methods, models, and tools for big data processing.

The main theme of the conference was theoretical foundations for the strategic management of large-scale system development in the context of national security.

[1] https://ieeexplore.ieee.org/xpl/conhome/9933724/proceeding

This range of problems is topical, as was clearly justified by Dr. Sci. (Eng.), Prof. *A.D. Tsvirkun* (ICS RAS) in his plenary paper "Managing the Development of Large-Scale Systems in the New Conditions of Sanctions." The author defined the essence and content of the modern geopolitical situation as the confrontation of three alternative scenarios of world order evolution (mondialism, unipolarity, and multipolarity of the world, respectively). Mondialism is the elimination of national states and the transition of full power into the hands of transnational corporations. Unipolarity characterizes a world where power and control of resources are concentrated in a chosen territory, and other countries and peoples become its colonies. Multipolarity is the freedom of creation and competition between different civilizations. This context of the modern world order is determinative to understanding sanctions policy.

The paper described TEO-INVEST, a digital platform (software package) to elaborate investment projects for a group of enterprises or enterprises with a complex internal structure. This platform has high potential due to possible integration with Capital Investments, the state information system to monitor the implementation of agreements, reimbursements of investors, and tax deductions. In the future, the Federal Tax Service should become a single complex for attracting private capital to new investment projects and a navigator of state support measures for business.

The complexity of the current geopolitical situation was also addressed by Dr. Sci. (Eng.), Prof. *V.V. Tsyganov* (ICS RAS). In the plenary paper "Methods and Models for Adapting the Russian Transport Infrastructure under Sanctions," he systemically analyzed the destructive effects of anti-Russian sanctions on Russia's transport sector. Obviously, the Western countries and their allies in Asia are attempting a global blockade of the Russian Federation in trade and transport spheres. Due to unprecedented sanctions, the economic situation in the country is very difficult. From the author's point of view, the socio-economic development of Russia under sanctions is impossible without an advanced geographical transformation of the national transport infrastructure. Turning logistics from the West to the East is a transport development strategy for improving social stability and defensive potential. As discovered, this strategy is complicated to develop: in practice, it required a new "theory of large transport systems" and a complex set of transport infrastructure management models under sanctions. The proposed digital platform includes a dynamic real-time model of government operation in uncertain conditions and five functional blocks (management of transport infrastructure development for socio-economic systems; selection and expertise of transport infrastructure development projects; training and restructuring of transport infrastructure; formation of transport corridors; ensuring security). In conclusion, Tsyganov dwelled on some issues of implementing the prototype of such a set of models in the projects of new transport infrastructure to improve living standards as well as the economic efficiency and security of the Russian Federation.

Academician of RAS *S.P. Filippov* and the co-authors, Cand. Sci. (Econ.) *F.V. Veselov* and Drs. Sci. (Eng.) *A.V. Keiko* and *T.G. Pankrushina* (Energy Research Institute RAS) presented the plenary paper "Information and Model Support for Decarbonization Management in the Russia Energy Sector." It was devoted to the key direction in the transformation of the modern economy. The authors considered the production structure of the country's energy economy as an objective source of greenhouse gas emissions, acting as the center of strategies and measures of decarbonization programs and the low-carbon restructuring of the strategic planning system. The *National Climate and Energy Plans* (NCEP) of EU countries were analyzed. As applied to Russian reality, the matter concerns the close integration of the three most important documents: the Energy Strategy, the Forecast of the Scientific and Technological Development of the Fuel and Energy Sectors, and the Forecast of the Long-Term Socio-economic Development of the country. In the new conditions, only this approach can become an effective tool for managing decarbonization in almost all spheres of the economy with fossil fuels (the harmonization of the pace of technological re-equipment and the commensurability of required costs and achieved environmental effects along all fuel and energy resource production and consumption chains). For the practical implementation of such a strategy, the paper proposed SCANER, a digital platform (a modeling and information package) to study the fundamental interconnections of the Russian fuel and energy complex with the global energy system (at the level of Russian energy resources exports) and the national economy (considering sectoral production capabilities of the fuel and energy complex and solvent domestic demand).

The plenary paper "The Interaction of Federal District Economies of the Russian Federation (Coalition Analysis Results)" by Corresponding Member of RAS *V.I. Suslov* and co-authors *Yu.S. Ershov* (Institute of Economics and Industrial Engineering, SB RAS, Novosibirsk) and Cand. Sci. (Econ.) *N.M. Ibragimov*

(Novosibirsk State University) was devoted to regional economic integrity. In terms of strategic planning, this paper logically agrees with the directions of energy and transport development mentioned above. The research was based on applied calculations on the interregional multisector optimization model for 40 sectors and 8 federal districts of Russia for the year 2030. The proposed digital platform involves calculations by the coalition analysis methodology (assessing the autonomous development capabilities of all possible coalitions of federal districts). The coalition analysis rested on input-output tables and the theory of cooperative games; the semi-dynamic version of the interregional multisector optimization model with the nonlinear setting (8 federal districts and 40 types of economic activity) was used for implementation. According to the calculations, the rupture of ties between regions and their ties with the outer world caused changes in almost all regional development indicators: the final product, gross output, investment, and employment. The authors presented several negative impacts of the potential rupture on various regional characteristics. As demonstrated by the calculations, the national economy has a high degree of interregional integration and involvement in the global economy. The greatest effect is expected from the coalition interaction of the eastern regions: they "work" for the all-Russian market, provide exports (oil and gas, non-ferrous metals, timber), and have a positive balance of commodity exchanges and a smaller share in the total final consumption (as compared to the share in the total GRP).

Note the paper "Scenario Analysis of Problems in the Security Management of Complex Socio-Economic Systems" by Corresponding Member of RAS *V.L. Schultz* (Center for Security Studies RAS), Dr. Sci. (Eng.), Prof. *V.V. Kul'ba*, and Cands. Sci. (Eng.) *I.V. Chernov* and *A.B. Shelkov* (ICS RAS). It considered the problem of increasing the efficiency of security management through the formal target forecasting of the behavior of the controlled object and its environment. The authors summarized the experience accumulated in improving the security management of socio-economic systems based on the scenario approach, including theoretical problems and applications. The digital platform proposed in the paper is represented by a software-analytical complex. Methodologically, the complex is based on a mathematical model of signed, weighted signed, and functional signed digraphs. This model extends the classical graph model with additional components. In particular, each vertex is assigned its parameter; each arc is assigned either a sign, or a weight, or a function (i.e., an arc transformation functional is introduced). In a practical interpretation, the parameters of graph vertices are the key indicators or factors describing the state and dynamics of the situation, and the graph structure reflects the cause-and-effect relations between them. A totality of vertex parameter values in the graph model describes a particular state of the situation at a given time. A change in these values generates an impulse (a transition of the system from one state to another). Management of situation development is modeled by changing the structure and impulses applied to certain graph vertices. The authors performed scenario studies of the problems of regional security, information security, public security, and technogenic security in the Russian Federation.

An important life support condition in an unstable environment was considered by Dr. Sci. (Eng.), Prof. *V.N. Burkov*, Dr. Sci. (Eng.), Prof. *A.Yu. Zalozhnev*, and *A.D. Kostyreva* (ICS RAS) in their paper "Consumer Market Stabilization Modeling" (Section 16). The authors presented a macro description of the operation of an economic system. The main variables in the model are the free price index, money savings of the population, commodity stocks of consumer goods in monetary terms, and the regulated price index. The value of money acts as an auxiliary variable. This economic and mathematical model is described by a system of ordinary differential equations. In the model, all production is aggregated into one sector with the gross output divided into consumption goods and all other products (fixed and current assets, non-productive goods not representing consumer demand items). The paper introduced approaches to the development of anti-inflationary economic policy and considered economic policies to stabilize the consumer market. As shown, four major stabilization policies can be chosen within the model. Also, the modeling results and their qualitative assessment were given. The model neglects budget revenues and their balancedness. To study consumer market stabilization, it suffices to represent the budget only by government expenditures for a certain period (the wages in the non-productive sphere, social security payments, and social insurance payments).

On the background of modern globalization and the informatization of public life, information wars and confrontations in social networks become harsh tools for violating social stability. In this regard, R&D works focused on mechanisms to counter information attacks, recognize initiators, block transmitted messages, and launch a reverse information wave are crucial for national security. At the plenary session, this subject was considered by Dr. Sci. (Eng.) *D.A.*

Gubanov and Dr. Sci. (Phys.–Math.) *A.G. Chkhartishvili* (ICS RAS) in the paper "Forming Opinions in Social Networks: The Confrontation of Several Information Sources." The authors studied a model in which agents' opinions (or preferences) are unobservable and the observed actions do not fully reflect their opinions. For such a model, the following problems of information control and confrontation were stated and solved: the state of agents is subjected to mass influence (e.g., through controlled media) to obtain network actions beneficial for the Principal. The Principal's strategy is choosing the degree of information influence. Analytical solutions were obtained in the case of a conditionally "atomized" network (in which the agents do not trust each other).

An important area of the country's development under sanctions was addressed by Cand. Sci. (Eng.) *M.V. Smirnov* (Financial University under the Government of the Russian Federation) in the paper "Methods and Models of Decision-Making to Achieve Russia's Technological Sovereignty" (Section 6). The author posed the topical problem of choosing priority parameters in the process of decision-making to achieve Russia's technological sovereignty. Based on matrix convolution and prioritization, he proposed an approach to organizing a binary tree of financial and economic parameters involved in the integral assessment of managerial decisions and their results.

The paper "Model of Dynamics the Inflation Macro Indicators Taking into Account External Influences" by Cand. Sci. (Phys.–Math.) *V.B. Gusev* (ICS RAS) analyzed macroeconomic indicators of Russia's economy (the consumer price index, the ruble exchange rate, and the dynamics of the annual GDP in the medium term). The model involves expert hypotheses, patterns, and statistical data. The so-called phenomenological approach was used: the properties of the economy's observed response to external influences were formalized, and the resulting model was verified.

The paper "Single-Factor Stress Testing of the System-Forming Enterprises" by Dr. Sci. (Eng.) *O.I. Dranko*, Corresponding Member of RAS *A.F. Rezchikov* (ICS RAS), Dr. Sci. (Eng.) *A.S. Bogomolov* (Institute for Precision Mechanics and Control Problems RAS, Saratov), and Cand. Sci. (Econ.) *M.M. Dvoryashina* (ICS RAS) was focused on the control of recession risks for real-sector enterprises. The authors presented the results of calculations of the single-factor stress tests (the critical reduction in revenue) for a list of system-forming organizations. According to the calculations, many Russian organizations have a good margin of safety for 2022.

Cybersecurity of the technosphere plays a significant role in the national security system. The main problems and approaches in this field were considered in the plenary paper "Risk Assessment and Cybersecurity of Nuclear Power Plants" by Cand Sci. (Phys.–Math.) *V.G. Promyslov*, *E.A. Abdulova*, Cands. Sci. (Eng.) *E.F. Jharko* and *A.Yu. Iskhakov*, Dr. Sci. (Eng.), Prof. *R.V. Meshcheryakov*, Dr. Sci. (Eng.) *A.G. Poletykin*, Cand. Sci. (Phys.–Math.) K.V. Semenkov (ICS RAS) and *N.N. Akimov*, *P.A. Golubev*, and *I.Yu. Lepekhin* (Sedakov Research Institute of Measuring Systems, Branch of All-Russian Research Institute of Experimental Physics, Nizhny Novgorod). The paper considered cybersecurity risk assessment for automated process control systems (APCSs) of critical facilities. The internal and external contexts of risk assessment were discussed for APCSs of nuclear power plants (NPPs). Two methodologies were proposed: the ones for the R&D life cycle stage and the operation stage of APCSs. The authors formulated the main tasks of ensuring the cybersecurity of APCSs of NPPs and outlined the principles of security architecture. The connection between the classical principle of defense-in-depth, applicable for nuclear safety, and its projection to information security in APCSs of NPPs was described. As justified by the authors, the main goal of cyber protection in APCSs of NPPs is to prevent violation of nuclear safety standards. Due to this feature, there is a strong connection between the classifications of APCSs of NPPs by nuclear security and cybersecurity.

Methodological grounds for developing digital decision support platforms were presented in the paper "Hierarchical Structures in Strategic Planning and Control" by Dr. Sci. (Eng.), Prof. *F.I. Ereshko* (Federal Research Center "Computer Science and Control" RAS). The author proposed mathematical models of controlled systems with a hierarchical organization.

The papers mentioned above are only a small part of the conference contributions, methodologically verified and implemented in the form of services, analytical applications, and software systems. They provide a toolkit for analytics and the choice of management strategies for the development of large-scale systems in complex macroeconomic and geopolitical conditions.

Thus, the main result of MLSD'2022 was proposals on implementing a digital strategic planning platform for transport, energy, economic, and social systems and investment processes under foreign sanctions.

The results obtained allow optimizing the choice of state-supported enterprises by the criterion of pre-

serving the basic socio-economic processes in complex macroeconomic conditions and under high uncertainty. One example is choosing the set of enterprises supported to maintain the vitality of industries in economic mobilization conditions and under emerging constraints (sanctions, epidemics, etc.).

The modern growing challenges and threats to Russia's economy require urgent and adequate preventive measures based on the risk-oriented approach to economic management. According to the Russian and foreign experience, the risk-oriented approach should be applied in a broad sense: as a mechanism of forming financial resources to cover losses and minimize risks and as a mechanism of forming a wide range of financial measures and coordinated actions in addressing strategic development objectives of Russia's economy.

Chair of the Organizing Committee
A.D. Tsvirkun

Secretary of the Organizing Committee
I.A. Stepanovskaya

Author information

Tsvirkun, Anatoly 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 paper
Tsvirkun, A.D., Stepanovskaya, I.A., 15th International Conference on Management of Large-Scale System Development (MLSD'2022). *Control Sciences* **6**, 49–53 (2022). http://doi.org/10.25728/cs.2022.6.6

Original Russian Text © Tsvirkun, A.D., Stepanovskaya, I.A., 2022, published in *Problemy Upravleniya*, 2022, no. 6, pp. 59–64.

Translated into English by *Alexander Yu. Mazurov*, Cand. Sci. (Phys.–Math.), Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia
✉ alexander.mazurov08@gmail.com