UDC 519.8, 519.7
doi:10.31799/1684-8853-2023-2-51-60
EDN: WUHAWQ

Articles

Dynamic strategies for monitoring quality control at a complex bioengineering facility

S. I. Kolesnikova[a]**,** *Dr. Sc., Tech., Professor, orcid.org/0000-0001-7158-2747, skolesnikova@yandex.ru*
A. A. Fomenkova[a]**,** *Senior Lecturer, orcid.org/0000-0002-5570-6743*
[a]*Saint-Petersburg State University of Aerospace Instrumentation, 67, B. Morskaia St., 190000, Saint-Petersburg, Russian Federation*

Introduction: *Stream processing of the data generated by bioengineering facilities is still an unresolved issue during real time identification of significant patterns in them. As a result, the lack of a unified approach to diagnose the macrostates of a water purification system has led to a significant complication of correct predictive analytics and untimely precursor detection of undesirable situations.* ***Purpose:*** *To develop a dynamic procedure for a real-time selection of the most preferred set of diagnostic features from a variety of all possible sets that are obtained during the training and that are replenished during the operation of the system, the above-mentioned procedure providing a flexible optimization strategy for its monitoring and control.* ***Methods:*** *We use a new procedure to form preference relations for a set of alternatives by group accounting for their relative superiority for the huge volume of continuously updated data.* ***Results:*** *Using the dynamic decision rule for switching between sets of criteria, which depends on the actual data, we obtain correct quantitative estimates of the state for a certain type of a purification system. The key idea of a dynamic decision rule is to take into account the dependence of the number of state classification errors obtained at the historical data on the criteria set used at the current control moment. The new algorithm provides a gain in the quality of macrostate diagnosing (by 15–20%) and an exponential decrease in decision-making time compared with the classical stationary model.* ***Practical relevance:*** *The results of this study are used to develop a monitoring plan for anaerobic treatment systems and to create and maintain a data base with an essential reduction of the time of historical data processing and of computational resources for an industrial hybrid bioreactor.* ***Discussion:*** *We can assume the possibility of full automation of self-regulation of biotechnical objects when we consider the contribution to the quality of the recognition system (object states). It consequently makes a certain set of criteria in the diagnosis of various states. This may be of particular importance because of the natural instability of interrelated biophysical and chemical processes and the possibility to design a stabilizing regulator.*

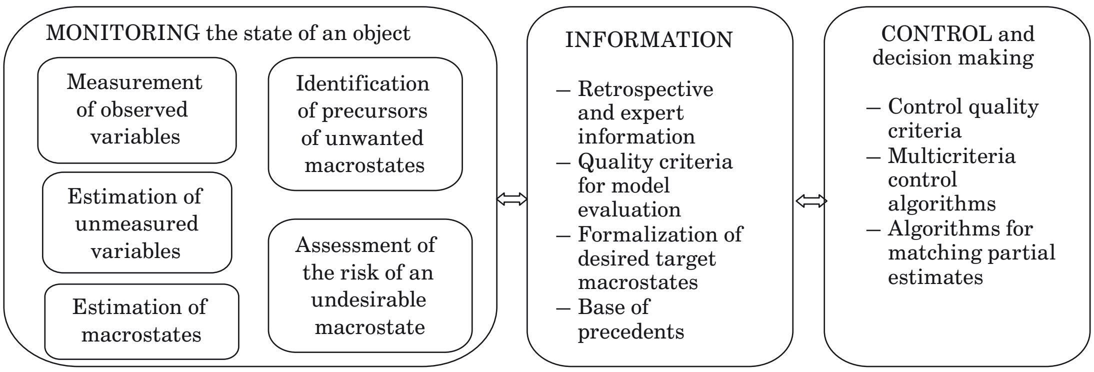
Keywords *— identification system for complex object macrostates, macrostate, dynamic rule, wastewater treatment system, dynamic decision rule.*

For citation: Kolesnikova S. I., Fomenkova A. A. Dynamic strategies for monitoring quality control at a complex bioengineering facility. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2023, no. 2, pp. 51–60. doi:10.31799/1684-8853-2023-2-51-60, EDN: WUHAWQ

Introduction to the problem

Control over complex chemical and biotechnological systems under conditions of multiple benchmarks and control quality criteria [1–4] is often implemented based on the multiple targets and peer recommendations, frequently without using the methods of mathematical programming and decision making, despite the increasing number of publications concerning the issue of monitoring such systems (e.g., [5–8]) (Fig. 1).

The problem of monitoring and control over a bioengineering system — a multidimensional and multi-loop controlled object with unstable dynamic modes in an open-loop condition is exceptionally important and has all attributes of a complex problem [9], since the object under study is a symbiosis of certain physical, technical and biochemical param-



■ ***Fig. 1.*** Conceptual scheme of monitoring of a complex system eters nonlinearly varying in time, and it is the biological component which provides a non-stationary dynamics of the predator-prey type [10].

One of the sources of type I–II errors in rendering a decision on the item-related state of a complex bioengineering object is a natural behavioral nonstationarity inherent in a (nonlinear) object, in every time interval requiring a set of indices with the best discriminating properties for the reliable differentiation of one state from the other, where the best set of indices for such an identification of one state could differ from that of the other.

Let us define the macrostates of a complex object as a subset of its phase space with certain functional constraints on the controlled (target) variables.

The *dynamic strategies of control over monitoring quality* will be here understood as an adaptive rule of observation and evaluation of the main indices of the process under study, which ensures:

– readjustment of the observer parameters in dependence of the dynamics of a process using the relevant period of its prehistory;

– changing of the set of indicators (criteria, benchmarks) in the course of the process evolution with the best properties in terms of the problem of state identification;

– real-time time of the response in the form of a state under big data conditions (multiple alternatives and criteria).

For the sake of readability and better understanding of the research results presented below, we are going to analyze a concrete complex system of anaerobic biological wastewater treatment (ABWT) [11, 12] (without loss of generality) and the problem of monitoring and control for this system [13].

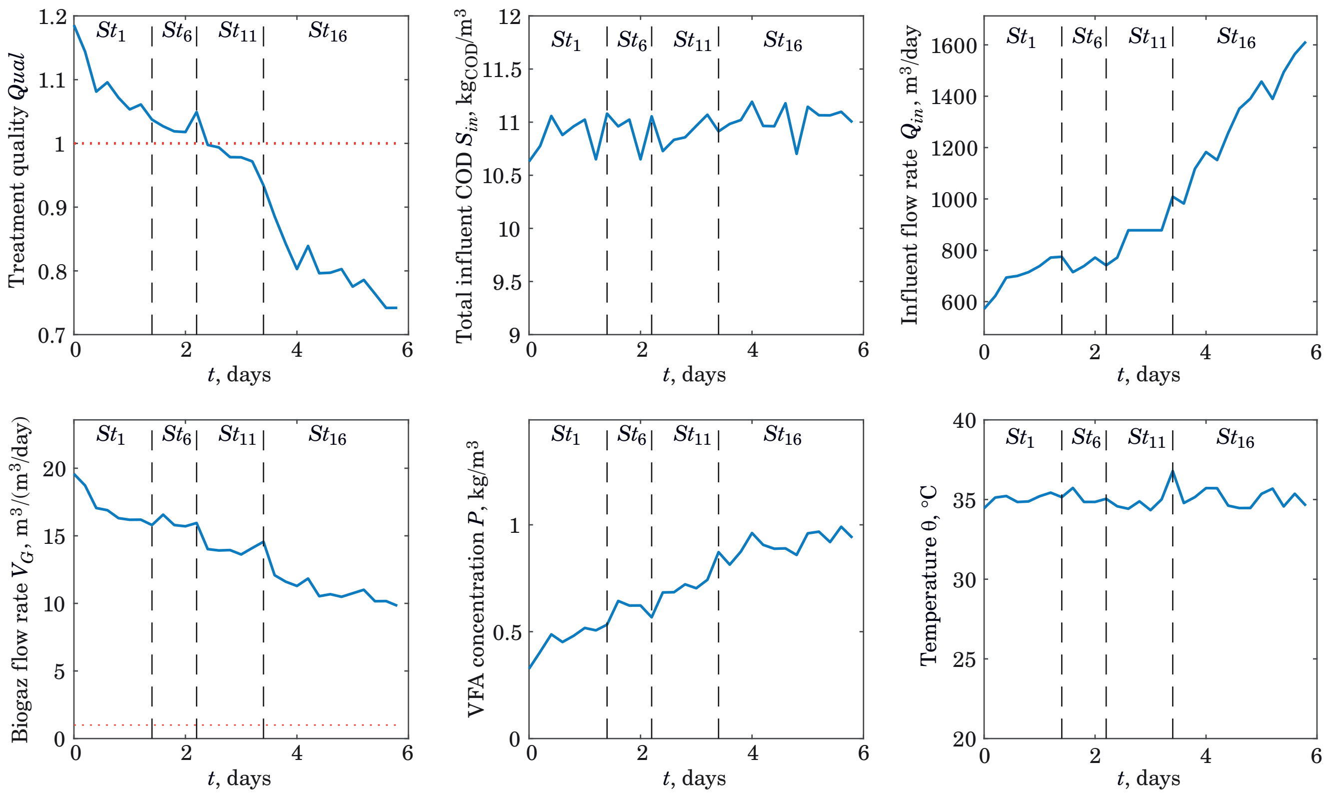
Controlled object description and motivation of a necessary use of dynamic strategies for controlling monitoring quality

Let us turn to an object reported earlier (e.g., [13, 14]), whose mathematical description represents a system on nonlinear differential equations.

The state of ABWT is a symbiosis of certain physical, technical and biochemical parameters representing, in a formalized form, a set of dynamic variables $\left\{x_i\left(t\right),\ i=\overline{1,\ n}\right\}$, which comprehensively define the object's position in a given space of states and form a state vector $\mathbf{x}(t)$.

Let us take that the entire phase state of state vector $\mathbf{x}(t)$ is conventionally divided into subsets $\Omega_i \in St,\ St=\{\Omega_i \mid i=\overline{1,N_\Omega}\}$ — macrostates with characteristic properties making it possible to distinguish between them (Fig. 2 and Table 1).

A macrostate can be prescribed analytically [14] by a limiting equality $\psi(\mathbf{x})=0$, $\mathbf{x}=\mathbf{x}(t)$, $t\to\infty$, where



■ ***Fig. 2.*** Results of simulation modeling of the ABWT state change $St_1 \to St_6 \to St_{11} \to St_{16}$

$\psi(\mathbf{x}(t))$ is the prescribed function and, as a manifold, it would possess asymptotic stability.

Remark 1. The technical states accepted in the diagnostic problems [] are partial cases of macrostates.

One of the important monitoring problems — to identify and maintain the desired states — is based on the formation of a minimal set of diagnostic attributes for the analysis of ABWT processes, taking into account the biomass degradation. The table of ABWT states formed relying on the results of simulation modeling (see Table 1) for every diagnostic attribute $\pi_j \in \Pi,\ \Pi = \left\{\pi_j \middle| j = \overline{1, N_\Pi}\right\}$ contains the ranges $l_{ij} \in L,\ L = \left\{l_{ij} \middle| i = \overline{1,\ N_\Omega};\ j = \overline{1,\ N_\Pi}\right\}$ of possible values of the types of system's states $\Omega_i \in St,\ St = \left\{\Omega_i \middle| i = \overline{1,\ N_\Omega}\right\}$.

The ranges of values of one and the same attribute in different types of states can coincide or partially overlap (see Fig. 2). In the case where for two different states the ranges of values of an attribute coincide or overlap, these types are thought to be indistinguishable by this attribute.

Table 2 presents some performance indicators of water treatment systems. There is also a number of information indices not presented here due to their awkwardness.

■ ***Table 1.*** Types of ABWT states

State code	State group	State type
Ω_1	Up state with normal biomass activity	Up state with normal biomass activity
Ω_2	Up state with signs of beginning destabilization of anaerobic biomass	Impact of high substrate concentrations with attributes of bioreactor acidification
Ω_3		Insufficient feed
Ω_4		Destruction of biomass structures
Ω_5		Fall-out from temperature regime
Ω_6		Increased wastewater load
Ω_7	Partial up state with developing degradation of anaerobic biomass	Impact of high substrate concentrations with attributes of bioreactor acidification
Ω_8		Insufficient feed
Ω_9		Destruction of biomass structures
Ω_{10}		Fall-out from temperature regime
Ω_{11}		Increased wastewater load
Ω_{12}	Down state with considerable degradation of anaerobic mass	Impact of high substrate concentrations with attributes of bioreactor acidification
Ω_{13}		Insufficient feed
Ω_{14}		Destruction of biomass structures
Ω_{15}		Fall-out from temperature regime
Ω_{16}		Increased wastewater load

■ ***Table 2.*** Criteria for formulation of classification attributes

Criteria name / measurement unit	Designation
Hydraulic retention time (HRT), days	$HRT = V/Q \to \min$, V — bioreactor volume, Q — influent flow rate
Content of volatile fatty acids (VFAs), kg_{COD}/m^3	$P \to \min$
Organics load rate (OLR), $kg_{COD}/(m^3day)$	$OLR = S_{in}/HRT = S_{in}Q/V \to \max$
Commulative biogas production, m^3	$G \to \max$
Biogas formation rate, $m^3/(m^3day)$	$I_G = dG/dt/V \to \max$
Energy efficiency	$Eff = \dfrac{E_G - E_{heating}}{E_{heating}} \to \max$
Organics removal rate, $kg_{COD}/(m^3day)$	$I_{COD} = \dfrac{Q_{in}S_{in} - Q_{out}S_{out}}{V} \to \max$
Degree of wastewater treatment, %	$\eta = \dfrac{S_{in} - (S_{out} + P_{out})}{S_{in}} 100\% \to \eta^*$, $\eta^* \in [0.7;\ 0.9]$

Formulation of problem for determination of dynamic control strategies of monitoring quality

In the general formulation, the problem of searching for optimal dynamic strategies for controlling the monitoring quality of a complex biotechnical object (multidimensional, multi-loop, nonlinear) with unstable dynamic modes in an open-loop condition can be realized via solving the following local tasks:

1) formalize a dynamic rule of selecting the controlling strategies for choosing a most "suit-

able" algorithm using the minimum error criterion; the outcome is the current macrostate code (number);

2) develop an algorithm of correct upgrading of the dynamic rule in accordance with the update information incoming as the ABWT process develops;

3) obtain all possible sets of classification (diagnostic) attributes providing a correct (error-free or with an admissible number of ambiguities on a training sample) identification of the peer-selected ABWT macrostates.

Without the loss of generality and for the sake of readability of the utilized control design technology, stabilizing the CO in the neighborhood of the desired state, let us look at a generalized mathematical model of an anaerobic bioreactor with a suspended-deposited biomass [13, 14].

This class of bioreactors is used in the cases where the waste water treatment is not subject to stringent requirements and even at a small capacity demonstrates good economic performance; these reactors are highly needed in the municipal wastewater treatment facilities.

In a mixer-bioreactor, both the organic substrate and the biomass are homogeneously distributed over the working volume of the purification apparatus. A substrate is fed to a bioreactor of volume V at flow rate $Q_{in}(t)$, concentration of nutritious organic substances $S_{in}(t)$ and biomass $B_{in}(t)$, and an equal volume of the working liquid is drained off. Due to stirring of the medium, the convective flows of substances prevail over the diffusion-induced flows. The medium homogenization allows describing the convectional component of the mass transfer process in a bioreactor according to the balance ratio (Fig. 3).

The functions

$$f_1(t)=\frac{Q_{in}(t)}{V}\left(S_{in}(t)-S(t)\right),$$

$$f_2(t)=\frac{Q_{in}(t)}{V}\left(B_{in}(t)-B(t)\right),\ f_3(t)=-\frac{Q_{in}(t)}{V}P(t)$$

characterize the inflow and outflow of the substrate, biomass and the biochemical reaction products. During the bioreactor operation without a recirculation flow, neither the biomass nor the substance — reaction product, is introduced into the reactor and the temperature is sustained at a constant value. Given these assumptions, the ABWT system will transform from Fig. 3 into

$$\frac{\mathrm{d}S(t)}{\mathrm{d}t}=\frac{Q_{in}}{V}\left(S_{in}-S(t)\right)-\left(\frac{\mu_{\max1}}{Y_{X1}}+K_{mX1}\right)\times$$
$$\times\frac{S(t)B_1(t)}{K_{S1}+S(t)}-B_1(t)K_{SX1};\ \frac{\mathrm{d}B_1(t)}{\mathrm{d}t}=-\frac{Q_{in}}{V}B_1(t)+$$
$$+\mu_{\max1}\frac{S(t)B_1(t)}{K_{S1}+S(t)+K_i^{-1}S^2(t)}-k_{d1}B_1(t);$$

$$\frac{\mathrm{d}P(t)}{\mathrm{d}t}=-\frac{Q_{in}}{V}P(t)+Y_{S1}B_1(t)\times$$
$$\times\left(K_{SX1}+K_{mX1}\frac{S(t)}{K_{S1}+S(t)}\right)-\left(\frac{\mu_{\max2}}{Y_{X2}}+K_{mX2}\right)\times$$
$$\times\frac{P(t)B_2(t)}{K_{S2}+P(t)+K_i^{-1}P^2(t)}-B_2(t)K_{SX2};$$

$$\frac{\mathrm{d}B_2(t)}{\mathrm{d}t}=-\frac{Q_{in}(t)}{V}B_2(t)+\mu_{\max2}\frac{P(t)B_2(t)}{K_{S2}+P(t)}-k_{d2}B_2(t);$$

$$\frac{\mathrm{d}G(t)}{\mathrm{d}t}=V_{m\max}B_2(t)\frac{P(t)}{K_m+P(t)}\cdot\frac{K_{im}}{K_{im}+P(t)}+$$
$$+\left(\mu_{\max1}\frac{S(t)B_1(t)}{K_{S1}+S(t)}Y_{CO_2S}+\right.$$
$$\left.+\mu_{\max2}\frac{P(t)B_2(t)}{K_{S2}+P(t)}Y_{CO_2P}\right)\frac{M_{CO_2}}{M_B}.$$

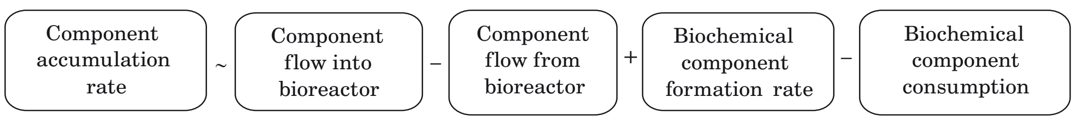
The peculiar life activity of the acidogenic and methanogenic microbial populations is characterized by a set of kinetic parameters: μ_i, Y_{Xi}, $K_{SX\,i}$, $K_{mX\,i}$, K_{Si}, Y_{Si} and $V_{m\,\max}$, K_m, K_{im}, respectively.

Problem solution

Stage 1. Rule of dynamic strategies

Let at the time point $t=\overline{1,\ 2,}\ \ldots,\ T$ of checking the macrostate $\Omega_i(t)$, $i=\overline{1,\ N_\Omega}$, $t=1,\ 2,\ \ldots,\ T$ for the sake of identification of its type b *basic sets* of classification attributes Π_l, $\Pi^b=\left\{\Pi_l \middle| l=\overline{1,\ b}\right\}$ be formed.

The algorithm for identification ABWT macrostates using set Π_l, $l=\overline{1,\ b}$ will be denoted as $a_l=a(\Pi_l)$, $l=1,\ b$. The outcome of the algorithm $a_l(t)$, $t=\overline{1,\ T}$ is the code (number) of a macrostate at time t (as the time point of inspecting the system) relying on the available information obtained in the interval [1, t – 1].



■ ***Fig. 3.*** Relationships between the rates of change in the concentration of the components with time

A linear combination of basic algorithms (algorithmic composition [15]) $a_l(t),\ t=\overline{1,\ T}$:

$$a(t)=\sum_{k=1}^{b} w_k(t)a_k(t),\ t=1,\ 2,\ \ldots,\ T;$$
$$\sum_{k=1}^{b} w_k(t)=1,\ w_k(t)=\begin{cases} k=l^*(t),\\ k\neq l^*(t),\end{cases} \quad (1)$$

with the time-updated weight coefficients $w_k(t)$ will be termed as the dynamic strategy of the quality monitoring control of a complex bioengineering object. The weights are updated at every time point $t = 1, 2, \ldots, T$ immediately prior to calculating the outcome on the state type.

In (1), $l^*(t)$ is the number of the 'best' model selected by the criterion of the average (r.m.s.) prediction error

$$l^*(t)=\arg\min_{k=1,\ldots,b}\sum_{\tau=1}^{t-1}\gamma^{t-\tau-1}\left(a_k(\tau)-true(\tau)\right)^2, \quad (2)$$

where $true(\tau)$ is the true state number; $\gamma \in [0, 1]$ is the algorithmic composition parameter determining the update information volume (prehistory), which is taken into account in decision making according to the rule (1), (2).

The prediction accuracy of the ABWT state type according to (1) will be estimated using the cross-validation functional within a time interval of length $[t_0,\ t_1]$ $\sum_{t=t_0}^{t_1}\left(a(t)-true(t)\right)^2$. The value of this interval is an additional optimization parameter.

In order to obtain the dynamic estimates of $w_k(t)$ we will apply a modification of the analytic hierarchy process (AHP) (let us term it as AHP+) under the conditions of high-dimensional set of alternatives and criteria used in the monitoring and control of complex objects.

Stage 2. Rule of updating weight coefficients of controlling strategy selection

For a better understanding of the approach proposed for a correct control over processing of the dynamically added large-volume information, let us present one by one three algorithms at the level of basic concepts, where the first key point (historically AHP [16]) remains the framework of the following two concepts but is free from its limitations.

Algorithms 1 and 2 considered below are a preamble to the main algorithm 3, on which the dynamic rule (1) and (2) for recognizing the states of a complex system is based [17, 18].

Algorithm 1. AHP

Input: sets of alternatives $\left\{A_i \middle| i=\overline{1,\ N_A}\right\}$ and criteria $\left\{C_j \middle| j=\overline{1,\ N_C}\right\}$.

1. Define the expert ratings on the scale 1÷9 for the initial sets $w_{ij}=w\left(A_i/C_j\right),\ i=\overline{1,\ N_A},\ j=\overline{1,\ N_C}$ (assessment of alternative A_i with reference to criteria C_j) and $v\left(C_j\right),\ j=\overline{1,\ N_C}$.
2. Form the pairwise comparison matrices (PCMs) for the criteria $\mathbf{C}=\left\|v\left(C_i\right)/v\left(C_j\right)\right\|_{N_C\times N_C}$ and alternatives $\mathbf{A}_j=\left\|w\left(A_i/C_j\right)\right\|_{N_A\times N_A}$ for every criterion.
3. Check all PCMs in terms of their consistency and repeat Step 2 in case it is distorted. Consistency here is understood as a transitivity of priorities in using the expert ratings (1÷9).
4. Normalize (to unity) the eigenvectors of PCMs for criteria $\mathbf{C}$ and alternatives $\mathbf{A}_j$ for every criterion, which are denoted as

$$\mathbf{C}^*=\left\|v_j^*\right\|_{1\times N_C},\ v_j^*=v_j^{norm},\ \mathbf{A}_j^*=\left(w_{1j}^*,\ \ldots,\ w_{N_A j}^*\right)^{\mathrm{T}},$$
$$w_{ij}^*=w_{ij}^{norm},\ i=\overline{1,\ N_A},\ j=\overline{1,\ N_C},$$

respectively.

5. Obtain the weight coefficients of alternatives as a linear criteria convolution via formula $u_i=\sum_{j=1}^{N_C}v_j^* w_{ij}^*,\ i=\overline{1,\ N_A}$, followed by the resulting vector normalization $\mathbf{u}^*=\left\|u_i^*\right\|_{1\times N_A},\ u_i^*=u_i^{norm}$.

Rank the set of alternatives in descending order of the weight coefficients.

Output: a set of normalized estimates of eigenvectors for every alternative for each criterion:

$$\mathbf{W}=\begin{pmatrix} w_{11}^* & \ldots & w_{1N_C}^* \\ \ldots & \ldots & \ldots \\ w_{N_A 1}^* & \ldots & w_{N_A N_C}^* \end{pmatrix}, \quad (3)$$

where w_{ij}^* is the assessment of the i-th alternative for C_j criterion $i=\overline{1,\ N_A};\ j=\overline{1,\ N_C}$, $\sum_{i=1}^{N_A}w_{ij}^*=1,\ j=\overline{1,\ N_C}$.

Comment 1. In the matrix notation the result of algorithm 1 is the product of $\mathbf{u} = \mathbf{W}\mathbf{C}^{*\mathrm{T}}$.

A generalization of AHP+ allows correctly comparing the numerical indicators (without involving the Saaty rating scale) and increasing the set of alternatives during the assessment without breaking the earlier achieved relations between the alternatives.

Then follows a pairwise comparison of the normalized components of the alternative weight coefficients for all of the criteria with normalization (to unity) *in every pair* and a subsequent linear convolution of the criteria estimation with the assessment of alternatives *in every pair*.

Algorithm 2. AHP+. Correct generalization of AHP

Input: matrix **W** and vector of criteria-based assessment $\mathbf{C}^*$.

1. Form a set of matrices for pairwise comparisons $\mathbf{W}_j,\ j=\overline{1,\ N_C}$, corresponding to criteria C_j with the vector elements

$$\mathbf{W}_j=\left\|\mathbf{w}_j\left(i,k\right)\right\|_{N_A\times N_A},$$

$$\mathbf{w}_j\left(i,k\right)=\left(w_{ij}^*(i),\ w_{kj}^*(k)\right),\ i,\ k=\overline{1,\ N_A},\ j=\overline{1,\ N_C},$$

where the element of matrix $\mathbf{w}_j(i, k)$ has the following structure:

$$w_{ij}^*(i)=\frac{w_{ij}^*}{w_{ij}^*+w_{kj}^*},\ w_{kj}^*(k)=\frac{w_{kj}^*}{w_{ij}^*+w_{kj}^*},$$

$$w_{ij}^*(i)+w_{kj}^*(k)=1,\ i,\ k=\overline{1,\ N_A},\ j=\overline{1,\ N_C}. \quad (4)$$

Comment 2. Indices of element $\mathbf{w}_j(i,\ k)$ stand for: $(i,\ k)$ is the address of the vector element of matrix $\mathbf{W}_j$ of the relative weights of alternatives A_i and A_k only in this pair in terms of criterion C_j.

2. For a matrix of the vector elements, whose components are the linear criteria-based convolutions given by the following:

$$\bar{\mathbf{W}}=\left\|\bar{\mathbf{w}}\left(i,k\right)\right\|_{N_A\times N_A},\ \bar{\mathbf{w}}\left(i,\ k\right)=\left(\bar{w}(i),\ \bar{w}(k)\right),$$

$$\bar{w}(i)=\sum_{j=1}^{N_C}v_j^*w_{ij}^*(i),\ \bar{w}(k)=\sum_{j=1}^{N_C}v_j^*w_{kj}^*(k),\ i,\ k=\overline{1,\ N_A}.$$

3. Normalize the vector elements of matrix $\bar{\mathbf{W}}$ via the rule (4) [17]:

$$\bar{w}^*(i)=\frac{\bar{w}(i)}{\bar{w}(i)+\bar{w}(k)},\ \bar{w}^*(k)=\frac{\bar{w}(k)}{\bar{w}(i)+\bar{w}(k)},$$

$$\bar{w}^*(i)+\bar{w}^*(k)=1,\ i,\ k=\overline{1,\ N_A}.$$

4. Determine the resulting weight coefficient values of alternatives $w\left(A_i\right),\ i=\overline{1,\ N_A}$ based on the summation of the first components of the vector elements of matrix

$$\bar{\mathbf{W}}^*=\left\|\bar{\mathbf{w}}^*\left(i,\ k\right)\right\|_{N_A\times N_A},\ \bar{\mathbf{w}}^*\left(i,k\right)=\left(\bar{w}^*(i),\ \bar{w}^*(k)\right),$$

$$w\left(A_i\right)=\sum_{l=1}^{N_A}\bar{w}^*(i). \quad (5)$$

Output: a set of normalized estimates of weight coefficients of alternatives.

Comment 3. The maximum effect on the computational complexity is exerted by the volume of alternatives, since the computational intensity ratio of algorithm 2 is $O\left(N_A^2N_C\right)$.

Traditionally, a multi-criteria ranking problem is solved in two steps: first of all the alternatives are put in correspondence with the ‘weights’ according to some rule and then the list of alternatives is sorted out.

Algorithm 3 is based on the violation of this sequence, specifically: a sequence of selection problems is formed *on a pair of alternatives* (from the prescribed set of alternatives), where the results of solution via algorithm AHP+ are used to compare two weight coefficients. Further, using the principle of sorting a numerical array, the elements (weight coefficients) are rearranged. Running through the entire array, the synchrony of operations is implemented: comparison with new alternatives and ordering of the weight coefficients of alternatives (here algorithms $a_l\left(t\right),\ t=\overline{1,\ T}$).

Algorithm 3. AHP+SORT. Generalization of AHP+ for big data

Input: matrix **W** and vector of criteria-based assessment $\mathbf{C}^*$.

1. Initialize algorithm 3. Take $i = 1, j = i + 1, i, j \in \{1, \ldots, N_A\}$. Conventionally order the set of alternatives and select alternative A_1 for comparison.

2. Select the next element A_j from the ranked set (list of alternatives) A and calculate the weights for the pair A_i, A_j, using algorithm 2 (AHP+).

3. Verify the sorting condition in the descending order of estimates of weight coefficients of alternatives:

if $w(A_j) > w(A_i)$, then rearrange the elements $w(A_i), w(A_j)$ in the data series $w(A_1), w(A_2), \ldots, w(A_i)$, otherwise take $j := j + 1$ and proceed to Step 2.

Outcome: a set of normalized estimates of weight coefficients of alternatives.

Comment 4. Algorithm AHP+SORT uses sequential comparisons of the pairs of elements of the set of alternatives for the sake of calculating the entire set of alternatives due to modification of AHP+.

Comment 5. Sorting can be based a combination of different approaches, such as a fast sorting modifications [19, 20]:

– division of the input array into sub-arrays using a special algorithm;

– sorting by merge insertions from every subarray;

– combination of the sorted sub-arrays into the final array by the modified merged sorting.

Implementation conditions for Algorithm 3 and its properties.

1. Algorithm AHP+ is not applied to the entire list of alternatives but rather to the pairs of alternatives during the sorting procedure, thus solving the problem of the curse of dimensionality.

2. The strongest influence on the increasing computational complexity is exerted by the volume of alternatives; the algorithmic complexities of algorithms AHP+ and AHP+SORT are estimated (without considering the number of criteria) as $O\left(N_A^2\right)$ and $O(N_A \log(N_A))$, respectively.

3. Obeying the Arrow's axiom [21]: obtain a correct ranking of the dynamically added list of alternatives, while maintaining the earlier achieved relations between the alternatives.

4. A possibility of using numerical values of criteria-based assessments of alternatives, by-passing the 1÷9 scale of the classical AHP.

5. While considering the process in time, it becomes *possible* to determine in real time a set of attributes that "best of all" characterize the separating capacity of the algorithm for identification of the bioengineering object's states.

Paper [18] presents an example of calculating the weight coefficients of alternatives according to AHP+ in order to provide a better understanding of Algorithm 3.

At the same time, Algorithm 2 is the basis for determination of various sets of classification attributes for identification of the ABWT macrostate.

Solution of applied problem. Simulation modeling of dynamic strategies

A determination of the minimal set of diagnostic attributes, which is correct for the fixed time t of ABWT inspection, according to which all types of ABWT states (see Table 1) are pairwise-distinguishable, using a set of criteria (see Table 1) is a non-trivial problem even with the delayed (trained) data, to say nothing of the real-time mode.

Specifically, a minimal set of attributes which is optimal in terms of verifications, can be determined using the concise Yablonskii algorithm [22], according to which the stub (shortest) coverings of the discrimination matrix (a binary table where every element is equal to zero, if two states are indiscriminable by the relevant attribute, and to unity in the opposite case) are determined.

Remark 2. The presence of disturbances and unstable states of the object in question would lead to a certain random set of binary tables and the next task would be related to a correct processing of the resulting set for formulating a reliable (meaningful under the conditions of this sample) final decision on the type of ABWT's state.

The cost of verification of the attributes is calculated in arbitrary units, including the costs of the instrumentation and control systems and the reagents for carrying out a single inspection, the number of check-up points where the parameters are measured in the apparatus, and the necessary laboratory tests.

The set of attributes providing the maximum information content of this analysis will contain those attributes from the set $\Pi = \left\{\pi_j \middle| j = \overline{1, N_\Pi}\right\}$, which during regular inspections had to the largest extent decreased the residual entropy in the course of analysis of the system's state.

The minimal sets of diagnostic attributes obtained in the sequential two-criteria optimization based on the minimum costs of verifications (C_1) and the Shannon data capacity (C_2) are given in Table 3.

Here *TSS* — total suspended solids, mg/L. The problem of a correct formation of diagnostic attributes even using only two criteria turned out to be nontrivial, and its complexity had been due to the following factors:

1) exhaustive enumeration and comparison of the effectiveness of the sets of attributes form a large volume of the domain sets of attributes is limited by the computational performance of the algorithms available for the solution of multi-criteria selection problem;

2) sets of diagnostic attributes can differ in terms of their composition (Table 3 is an example of differences between two sets by a single parameter);

3) selection of diagnostic sets by individual attributes (or one principal attribute) does not guaranty the exhaustiveness of characteristics of the real observed state of the system (pH value in certain bioreactors cannot indirectly indicate the concentration of volatile fatty acids, VFA);

4) sets of diagnostic attributes *is a function of time* in the sense that in different stages of the process different groups of attributes can possess discriminating properties, since a decreased content of methane CH_4 in biogas and the volumes of generated biogas V_G manifest themselves in the already

■ ***Table 3.*** Minimum sets of diagnostic attributes for identification of a certain ABWT state under biomass degradation

Criteria	Physical-technical parameters	Biochemical parameters
C_1	Q_{in}, θ	V_G, pH, S_{out}, CH_4, TSS
C_2	Q_{in}, θ	V_G, VFA_{total}, S_{out}, CH_4, TSS

developed process of biomass degradation and are not early indicators of destabilization of the anaerobic fermentation;

5) parameter VFA_{total}, in combination with the calculated OLR allows revealing the initial states of destabilization of the anaerobic fermentation but requires that additional current samples be selected during the process development over time.

Comments on applicability of multi-criteria methods for formation of sets of performance indicators depending on object's state

According to [1], the well-known methods of solution of multi-criteria problems could be roughly divided into four following groups:

1) ranking of optimality criteria and alternatives, including the criteria-based convolutions with the weight coefficients prescribed by the decision maker (expert);

2) selection of one basic criterion and conversion of other targets into constraints (which could appear to be incompatible);

3) construction of a certain generalized criterion, based on physical considerations and peer opinion (e.g. maintaining the desired invariant of the target system);

4) use of a concept of a normalized criteria space and the search for a solution providing (in a certain metric) the minimum distance between the target functions and the optimal values.

In [23], this principle is used to search for an optimal trade-off solution on the basis of a game-theory model. However, the trade-off solution is selected not from the entire domain of determination of variables, but rather from a certain subset formed by a linear combination of the solutions optimal in terms of individual criteria. Thereby, a possibility of losing the 'best solutions' relative to the found solutions is ruled out.

In [24], the latter shortcoming was eliminated, and an optimal solution was searched for in the entire region of feasibility, not going beyond the convex programming, which is exceptionally convenient from the computational standpoint. However, there is a danger of obtaining an incorrect solution of the multi-objective problem, because in the general case, the solutions obtained upon normalization of the spaces of criteria and alternatives can disagree.

In the frames of the above-mentioned, the classification of AHP and its correct modification AHP+ are the most mathematically sound methods and represent a hybrid process involving the advantages of all approaches 1)–4).

The AHP+ modification is free from the following shortcomings of the classical AHP method [16] and essentially extends its application:

1) matrix representation of the pairwise-compared data (criteria and alternative) gives rise to an increased computational awkwardness in the case of a high power of a set of alternatives and criteria;

2) non-observance of the axiom of indifference in terms of the Arrow's alternatives: the achieved priorities can change as new candidates (alternatives) are added;

3) disturbance of the consistency of the pairwise-comparison matrix in the case where not the initial criteria and alternatives are used for comparison, but rather a well-known scale of Saaty's (AHP author) assessment ratings (1÷9), indisputably introducing an ambiguity into the assessment of the degree of preference between the compared indicators.

Remark 3. In [25], the authors provide an example of inconsistency of the main assumption in AHP: the scales, in which the preferred alternatives of each criterion are assessed, are taken to be neither related to each other or to the criterion priorities.

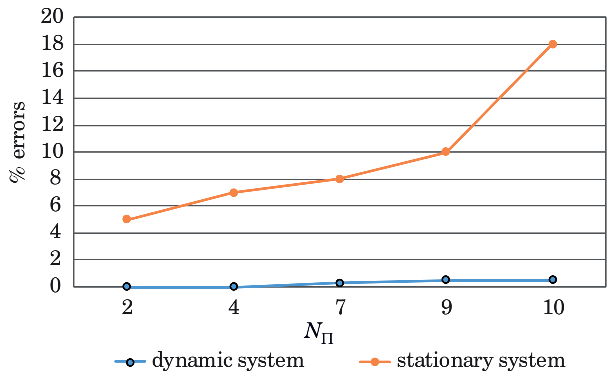
The amount of possible macrostates and sets of criteria (performance indices) in bioengineering systems can be quite large (on the order of 100 and more) and their formulation — resource- and time-consuming. Therefore, traditionally a statistical set of attributes is selected during monitoring and used in the course of the entire treatment process.

The accuracy of the rule (1) and (2) was estimated from the simulation modeling data involving the use of ABWT's mathematical model from [13, 14]. The basic algorithms used in the estimation were based on 10 sets of attributes.

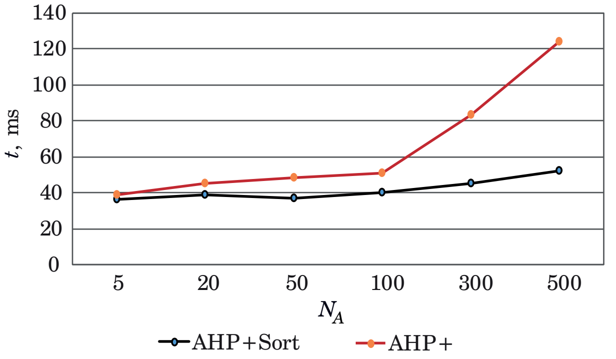
Algorithm AHP+SORT allows making a real-time decision on the most effective set of attributes and the relevant classification algorithm. The use of this algorithm allowed:

— essentially reducing the time of in-line processing of the incoming information and the time of detection of undesirable states of the system (Figs. 4 and 5);

— decreasing the number of classification errors (on model data) by 10–20%;



■ ***Fig. 4.*** Comparative graphs of the errors in decision-making system on solution of ranking problems versus the dimension of initial sets of criteria



■ ***Fig. 5.*** Comparative graphs of the time spent on solution of ranking problem versus the dimension of initial sets of alternatives

— providing a possibility of formation of correct managerial solutions.

Conclusion

A new mathematically sound method for making correct (consistent and error-free, based on a set of baseline cases) managerial decisions has been proposed as applied to a bioengineering object containing unstable states, among its stationary states, due to the presence of a biological sub-system.

It has been shown that the accuracy of state identification for various sets of classification attributes depends on the process evolution character and, from this standpoint, in every monitoring point it is reasonable to place reliance on that set of attributes which possesses the highest predictive accuracy based on the earliest process prehistory (baseline cases). These monitoring control approaches have been termed here as dynamic strategies.

The proposed mathematical tool for monitoring control has been tested using an example of a complex system of anaerobic biological water treatment, aiming the solution of the problem of formation of diagnostic attributes for identification of certain phase-space subsets of the dynamic ABWT model, referred to as ABWT macrostates.

The proposed algorithm of correct solution of the multi-objective selection problem can be applied to a wide range of industrial problems solved using a number of business indices (maximum product yield, profit, labor productivity, introduced innovation volumes, etc.). The software program implementing the AHP+SORT generalization algorithm makes offers real-time solutions of multicriterion problems containing on the order of 10^3 and more criteria with the order of matrix of alternatives 10^5 and over.

Financial support

This research was supported by the Russian Science Foundation (project no. 23-29-00336).

References

1. Kafarov V. V., Lazarev G. B., Avdeev V. I. A method for solving multicriteria control problems for a complex chemical-technological system. *Reports of the USSR Academy of Sciences*, 1971, vol. 198, no. 1, pp. 62–63 (In Russian).
2. Kafarov V. V. Dorokhov I. N. *Sistemnyj analiz himiko-tekhnologicheskih processov: osnovy strategii* [System Analysis of Chemical Technology Processes: Fundamentals of Strategy]. Moscow, YUrajt Publ., 2018. 499 p. (In Russian).
3. Kalyuzhnyi S. V. Batch anaerobic digestion of glucose and its mathematical modeling. II. Description, verification and application of model. *Bioresource Technology*, 1997, vol. 59, pp. 249–258. doi:10.1016/S0960-8524(96)00125-3
4. Vavilin V. V. Dynamics of anaerobic microorganisms community: experience in mathematical modeling. *Proc. of the Samara Scientific Center of the Russ. Acad. of Sci.*, 2009, vol. 11, no. 1(7), pp. 1615–1619 (In Russian).
5. Schügerl K. Progress in monitoring, modeling and control of bioprocesses during the last 20 years. *Journal of Biotechnology*, 2001, vol. 85, pp. 149–173. doi:10.1016/S0168-1656(00)00361-8
6. Sonnleitner B. Instrumentation of biotechnological processes. *Advances in Biochemical Engineering/Biotechnology*, 2000, vol. 66, pp. 1–64.
7. Fehrenbach R., Comberbach M., Petre J. O. On-line biomass monitoring by capacitance measurement. *Journal of Biotechnology*, 1992, vol. 23(3), pp. 303–314. doi:10.1016/0168-1656(92)90077-M
8. Lyberatos G. and Skiadas I. V. Modelling of anaerobic digestion — a review. *Global Nest Journal*, 1999, vol. 1(2), pp. 63–76.
9. Rastrigin L. A. *Sovremennye principy upravleniya slozhnymi ob”ektami* [Modern Principles of Control for Complicated Object]. Moscow, Sovetskoe radio Publ., 1980. 232 p. (In Russian).
10. Buyvolova A. G. On the issue of managing the “predator-prey” system. *Automatic Remote Control*, 1993, vol. 54, iss. 4, pp. 546–554 (In Russian).
11. Batstone D. J., Keller J., Angelidaki I., Kalyuzhnyi S. V., Pavlostathis S. G., Rozzi A., Sanders W. T. M., Siegrist H., Vavilin V. A. The IWA Anaerobic Digestion Model No 1 (ADM1). *Water Science and Technology*, 2002, vol. 45, no. 10, pp. 65–73.
12. Kiely G., Tayfur G., Dolan C., and Tanjf K. Physical and mathematical modelling of anaerobic digestion of organic wastes. *Water Research*, 1997, vol. 31, no. 3, pp. 534–540.

13. Kolesnikova S., Fomenkova A. Formalization of data exchange between the systems of monitoring and control over an anaerobic bioreactor. *Proc. 4th Intern. Conf. on Control Systems, Mathematical Modeling, Automation and Energy Efficiency (SUMMA)*, 2022, pp. 550–555. doi:10.1109/SUMMA57301.2022.9973923
14. Fomenkova A. A., Klucharev A. A., and Kolesnikova S. I. Formalization of target invariants and designing an adaptive control system for the model of anaerobic biological wastewater treatment. *Journal of Physics Conference Series*, 2021, 2099 012035. doi: 10.1088/1742-6596/2099/1/012035
15. Vorontsov K. V., Egorova E. V. Dynamically adaptable compositions of forecasting algorithms. *Artificial Intelligence*, 2006, no. 2, pp. 277–280 (In Russian).
16. Saaty T. L. Decision making with the analytic hierarchy process. *International Journal of Services Sciences*, 2008, vol. 1(1), 83. doi:10.1504/ ijssci.2008.017590
17. Samokhvalov Y. Y. Group accounting of relative alternative superiority in decision-making problems. *Cybernetics and Systems Analysis*, 2003, vol. 39, pp. 897–900. https://doi.org/10.1023/B:CASA.0000020231.09571.33
18. Kolesnikova S. I., Karavanova S. A. Optimization of the multi-criteria choice algorithm with a dynamically replenished large set of alternatives. *Information Technology*, 2021, no. 5, pp. 235–241 (In Russian). doi:10.17587/it.27.235-241
19. Knuth D. E. *The Art of Computer Programming*. First ed. 1973. Vol. 3. 1997. 723 p.
20. Hoare C. Quicksort. *The Computer Journal*, 1962, vol. 5, no. 1, pp. 10–16.
21. Kenneth J. Arrow. *Social Choice and Individual Values*. 3rd ed. Yale University Press, 2012.
22. Zhuravlev Yu. I. *Algorithms for constructing minimal disjunctive normal forms for functions of the algebra of logic*. In: *Discrete Mathematics and Mathematical Problems of Cybernetics*. Moscow, Nauka Publ., 1974. Pp. 67–98. (In Russian).
23. Juttler H. Linear model with several objective functions. *Economics and the Mathematical Methods*, 1967, vol. 3, no. 3, pp. 397–406.
24. Ermolenko B. V., Kafarov V. V., Kazbekova N. V. Construction of a mathematical model for finding the optimal production program for a chemical enterprise. *Theoretical Foundations of Chemical Engineering*, 1969, vol. 3, no. 6, p. 903 (In Russian).
25. Podinovsky V. V., Podinovskaya O. V. On the incorrectness of the hierarchy analysis method. *Control Sciences*, 2011, no. 1, pp. 8–13 (In Russian).

УДК 519.8, 519.7
doi:10.31799/1684-8853-2023-2-51-60
EDN: WUHAWQ

Динамические стратегии управления качеством мониторинга сложного биоинженерного объекта

С. И. Колесникова[a], доктор техн. наук, профессор, orcid.org/0000-0001-7158-2747, skolesnikova@yandex.ru
А. А. Фоменкова[a], старший преподаватель, orcid.org/0000-0002-5570-6743
[a]Санкт-Петербургский государственный университет аэрокосмического приборостроения, Б. Морская ул., 67, Санкт-Петербург, 190000, РФ

Введение: обработка больших потоков данных, порождаемых биоинженерными объектами, является по-прежнему нерешенной проблемой при выявлении в них значимых закономерностей в реальном времени. Отсутствие единого подхода к диагностированию макросостояний очистительной системы привело к существенному усложнению проведения корректной предиктивной ее аналитики и несвоевременности обнаружения предвестников нежелательных ситуаций. **Цель:** разработка динамической процедуры выбора в реальном времени наиболее предпочтительного набора диагностических признаков из множества полученных на обучении и пополняемых в процессе эксплуатации системы, обеспечивающей гибкую стратегию оптимизации ее мониторинга и управления. **Методы:** применяется новая процедура формирования отношений предпочтения по набору альтернатив посредством группового учета их относительного превосходства для большого объема непрерывно пополняемых данных. **Результаты:** алгоритм оценивания макросостояния системы биоочистки, основанный на правиле переключения между наборами критериев в зависимости от актуальных данных, позволил получать корректные количественные оценки работоспособности системы определенного типа. Суть идеи динамической процедуры заключается в учете зависимости числа ошибок классификации состояний, полученных на исторических данных, от используемого критериального множества в текущий момент контроля. Совместное использование алгоритма группового учета предпочтений и метода сортировки обеспечило выигрыш по качеству диагностирования макросостояний (на 15–20 %) и экспоненциальному уменьшению времени принятия решений относительно ранее разработанной стационарной классической модели парных сравнений анализируемых альтернатив. **Практическая значимость:** результаты исследований использованы при разработке мониторинга системы анаэробной очистки, создании и сопровождении базы данных и знаний с существенным сокращением времени обработки поступающих данных и вычислительных ресурсов для промышленного секционного гибридного биореактора. **Обсуждение:** учитывая вклад в качество распознающей системы состояний объекта, который вносят динамические стратегии, можно предположить возможность полной автоматизации саморегулирования биотехнических объектов. Это может иметь особое значение в связи с природной неустойчивостью взаимосвязанных биофизических и химических процессов и возможностью конструирования стабилизирующего регулятора.

Ключевые слова — система распознавания макросостояний сложного объекта, макросостояние, динамическое правило, система очистки сточных вод, динамическое правило принятия решения.

Для цитирования: Kolesnikova S. I., Fomenkova A. A. Dynamic strategies for monitoring quality control at a complex bioengineering facility. *Информационно-управляющие системы*, 2023, № 2, с. 51–60. doi:10.31799/1684-8853-2023-2-51-60, EDN: WUHAWQ
For citation: Kolesnikova S. I., Fomenkova A. A. Dynamic strategies for monitoring quality control at a complex bioengineering facility. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2023, no. 2, pp. 51–60. doi:10.31799/1684-8853-2023-2-51-60, EDN: WUHAWQ