BULLETIN OF THE POLISH ACADEMY OF SCIENCES
TECHNICAL SCIENCES, Vol. 71(2), 2023, Article number: e144603
DOI: 10.24425/bpasts.2023.144603

CONTROL AND INFORMATICS

The look-up algorithm of monitoring an object described by non-linear ordinary differential equations

Przemysław HAWRO[1]*, Tadeusz KWATER[1], Jacek BARTMAN[2], and Bogdan KWIATKOWSKI[3]

[1] Institute of Technical Engineering, The State University of Technology and Economics in Jaroslaw, Czarnieckiego 16, 37-500 Jarosław, Poland
[2] Institute of Computer Science, University of Rzeszow, Pigonia 1, 35-959 Rzeszów, Poland
[3] Faculty of Electrical and Computer Engineering, Rzeszow University of Technology, Pola 2, 35-959 Rzeszów, Poland

Abstract. The article proposes an adaptive algorithm that generates all object signals, including those for which measurements are not performed due to the difficulties associated with online measurements. The algorithm is modeled on the idea of the Kalman filter using its equation; however, the selection of gains is optimized differently, i.e. the constant values depend on the adopted ranges of adaptation errors. Moreover, the knowledge of the statistics of all noise signals is not imposed and there is no linearity constraint. This approach allowed us to reduce the complexity of calculations. This algorithm can be used in real-time systems to generate signals of objects described by non-linear differential equations and it is universal, which allows it to be used for various objects. In the conducted research, on the example of a biochemically contaminated river, only easily measurable signals were used to generate the object signals, and in addition, in the case of the absence of some measurements, the functioning of the algorithm did not destabilize.

Key words: look-up algorithm; monitoring online; adaptive estimation; polluted river; MATLAB.

1. INTRODUCTION

Issues related to ecology and production processes with the use of chemical processes play a significant role in the development of society in the world. It is related to the support of monitoring and control by means of real-time IT systems. In particular, it covers, among others, the monitoring of a polluted river or a continuous stirred-tank reactor (CSTR) [1].

The issues of online monitoring for polluted rivers are still present to solve as the development of agriculture, industry and cities has resulted in the deterioration of water resources of rivers and reservoirs in terms of their quality and quantity [2–4]. The purpose of monitoring water quality in rivers is to reduce the risk of pollution and ensure appropriate economic policies that take into account water protection [5]. The approaches to them vary, among others, there is the one which involves using a mobile measuring station [6]. The authors assure that this approach enables us to identify various situations of small and average river pollution. Another type is proposed in [7]. The adaptive algorithm of sampling is applied in order to improve the energetic efficiency of the automatic monitoring system with the simultaneous assurance of the precision of sampled data. As a measurement, there were used data of dissolved oxygen (DO) and water turbidity. The use of chemical sensors in the online monitoring of water quality is presented in [8]. It is indicated that systems based on chemical detection or a combination of them with other methods lead to the best results. The authors claim that the accuracy of monitoring certain water pollutants in real-time monitoring systems is best when using molecularly imprinted polymers.

Quite often in real-time monitoring systems, it is not possible to measure signals online. For a biochemically polluted river, such a signal is the biochemical oxygen demand indicator (BOD). The determination of BOD is usually carried out using laboratory service. Some studies use a method based on biosensors and correlation calculations to obtain representative information on this indicator [9, 10]. Another way is to use a modified observer by introducing additional coupling (for both adaptive and non-adaptive law) which facilitates a stable operation of the object [11].

In [12] the authors propose the use of optical sensors to estimate the BOD value, which allows for quick measurements for the water quality monitoring system, but the final result is obtained not directly but with a slight delay and a certain correlation coefficient. Machine learning methods, time series, statistical models, and environmental decision support systems are used to address monitoring issues in ecological systems to make water management more objective, dependable, and efficient [12–16]. Monitoring systems can also be targeted to meet the needs of different end-users, e.g. agriculture. The use of an Android-based monitoring system is presented in [17]. The pollution information obtained in this way is used by farmers.

In addition to the difficulties associated with estimating BOD for a river, there is also the impact of sudden weather anomalies, which can imply various sources of pollution with negative impacts on aquatic ecosystems [18]. Difficult tasks related to the influence of randomness, which are characterized by parametric uncertainty and chaotic responses to certain specified pa-

*e-mail: przemyslaw.hawro@pwste.edu.pl

Manuscript submitted 2022-07-08, revised 2022-11-29, initially accepted for publication 2023-01-13, published in April 2023.

rameter values in addition to ecosystems, include networks of permanent magnet synchronous motor systems [19]. The use of monitoring systems facilitates the immediate detection of unexpected changes in pollution, enabling us to take appropriate decision-making steps.

2. MATHEMATICAL DESCRIPTION OF THE OBJECT

Depending on the purpose of the water, its quality is determined by industry standards. In a standard analysis of water quality, indicators characterizing the state of water pollution are used. The relevant legal regulations for determining water quality use a description in the form of water class based on the values of the indicators used. One such indicator is BOD_5, the value of which for water taken for consumption purposes should not exceed 4 mg O_2/l [20]. The article assumes that the state of river pollution determines the appearance of non-zero values of the BOD indicator.

A general mathematical description of a biochemically polluted river takes into account the phenomena of advection-diffusion and self-purification [21–23]. This is captured by the Gauss-Ostrogradsky theorem given by the equation:

$$\frac{\partial x}{\partial t} + divq - \boldsymbol{A}x + \delta = 0, \tag{1}$$

in which: x represents in general the concentrations of various substances in the water (DO, BOD, chemical oxygen demand (COD) and other substances), q is the mass flux, δ is the internal source density, which determines the intensity of generation or decay of the transported medium, while div is the divergence operator.

In equation (1), the advective-diffusive flux q, when considering only the direction of the river flow, will cause the equation to take the form:

$$\frac{\partial x}{\partial t} + v\frac{\partial x}{\partial z_1} - D\frac{\partial^2 x}{\partial z_1^2} - \boldsymbol{A}x + \delta = 0. \tag{2}$$

Such simplifications involve the elimination of components from the model that do not significantly affect the accuracy of the results obtained. This concerns simplifications related to advection in directions other than in accordance with the direction of the river axis (z_1). Therefore, the description of advection in the plane perpendicular to the river axis can be omitted in the equations, as it does not cause a significant change in the concentration distribution. In the case of a river, the highest velocity of pollutant movement (v) occurs along its length (l), and movement in the other directions can be considered insignificant due to the much smaller object dimension in these directions relative to the river length.

Equation (2) describing the advection-diffusion transport of substances in the river can be simplified by omitting the contribution of diffusion to the distribution of pollutants. The condition, however, is a sufficiently high stream velocity. Elimination of the segment representing the diffusion phenomenon from the mathematical model considerably simplifies the complexity of numerical calculations. For further considerations, a conventional division of the river into sections with known flow velocity in that section is assumed, as presented in Fig. 1.

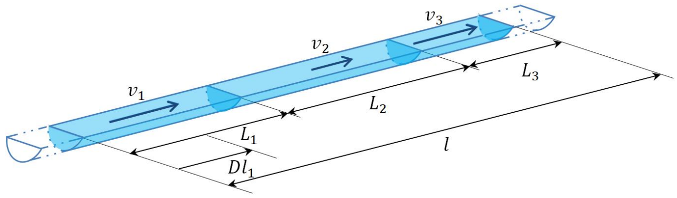

Fig. 1. Contractual division of the river into sections (v_i – speed ofmovement of pollutants in the river section L_i)

The lengths of the different river sections can be normalized by assuming that $Lz_i \in [0, 1]$, where $Lz_i = \dfrac{Dl_i}{L_i}$ normalized length of the i-th segment ($i = 1, 2, 3, \ldots$), where: Dl_i is the length in this river section from 0 to L_i [km], L_i – the length of this river section.

After dividing the river into sections and considering the above assumptions, the equation of the mathematical model for the i-th section of the polluted river can be expressed by a partial differential equation of the form:

$$\frac{\partial x_i}{\partial t}(Lz_i, t) + v(Lz_i t)\frac{\partial x_i}{\partial Lz_i}(Lz_i, t) = \boldsymbol{A}(Lz_i, t)\,x_i(Lz_i, t) + \boldsymbol{B}(Lz_i)\,w_i(t) \tag{3}$$

with boundary conditions (IC – initial condition, BC – boundary condition):

$$\text{IC: } x_i(Lz_i, t_0) = x_{i0}(Lz_i), \quad i = 1, 2, 3\ldots,$$

$$\text{BC: } x_i(0, t) = \boldsymbol{M}_{i-1}x_{i-1}(1, t) + v_{bi}(t), \ \ \boldsymbol{M}_0 = 0,$$

where: $w_i(t)$ – system disturbance vector in i-th section in time t, $\boldsymbol{M}_i$ – matrix taking into account the concentration and flow in the previous section and the current lateral inflow, v_{bi} – boundary disturbance in the i-th section.

Maintaining the accuracy of the description of phenomena described by differential equations with distributed parameters equation (3), an appropriate interpretation is made in order to obtain a description by differential equations with lumped parameters. This interpretation allows us to consider the model described by hyperbolic partial differential equations with the help of a set of easier-to-solve ordinary differential equations. The solution thus obtained is then used to construct a solution for the model with distributed parameters.

The idea of this method is to observe the distribution of pollutants in the river along the so-called characteristics (ch) in the spatial-temporal domain. These characteristics will be the lines determined by the known flow velocity as illustrated in Fig. 2, where one of such characteristics is shown (red color) taking into account pollution inflows distributed along the river length (influence of interference signals). The presented approach facilitates the observation of pollutant concentrations in the considered domain (time, length).

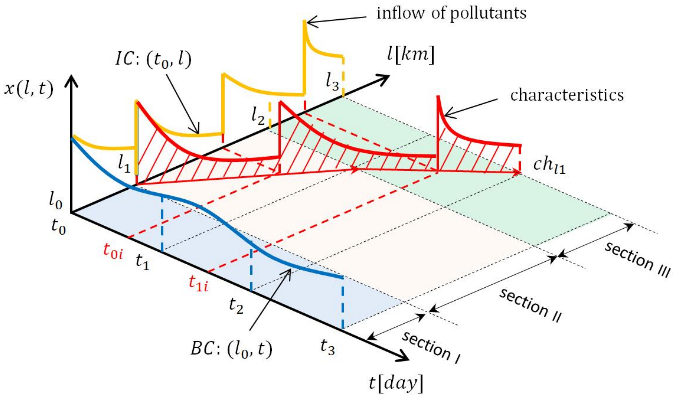

Fig. 2. Characteristics for partial differential equations

In Fig. 2, equal river current velocities are assumed in the color-coded areas. This approach is based on the interpretation of the description of the pollutant concentration distribution in the river observed for a freely moving volume of water. As a result of this interpretation, for flow velocity v_i in a river section the distribution of pollutants depends only on time, so the characteristics for the i-th river section are determined by the relation:

$$\frac{d}{dt} l_i(t) = v_i(l_i(t)) \tag{4}$$

whose limit points satisfy the condition:

$$\int_{t_{0,i}}^{t_{0,i+1}} v_i\left(l_i(t)\right) dt + l_{0i} = 1 \tag{5}$$

for $l_{0,i} \in [l_i, l_{i+1}]$.

In this way, a set of characteristics covering the entire solution domain is obtained, and the equation for each characteristic takes the form:

$$\frac{d}{dt} x(t) = \boldsymbol{A} x(t) + \boldsymbol{B} w(t), \tag{6}$$

where: x – state vector $x = col[x_1, x_2]$, $\boldsymbol{A} = \begin{bmatrix} k_1 & 0 \\ k_2 & k_3 \end{bmatrix}$ – coefficient matrix k_i, $i = 1,2,3$, $\boldsymbol{B}$ – interaction matrix of interfering signals, w – system disturbance vector $[w_1, w_2]$.

The coordinates of the vector x represent respectively: x_1 – concentration of organic pollutants expressed as BOD, with an initial condition $x_1(0)$ [mg/l], x_2 – the dissolved oxygen concentration deficit DO which is the difference $x_2 = x_{2S} - x_{2N}$ between the dissolved oxygen concentration x_{2S} and the oxygen content of the water at a saturation state x_{2N} with an initial condition $x_2(0)$ [mg/l].

The coefficients k_1, k_2 and k_3 in equation (6) describe the dynamics of the natural self-purification process of the river and depend primarily on temperature, in particular they mean: k_1 – coefficient of the rate of reaction of BOD [1/day], k_2 – coefficient of the influence of BOD on DO [1/day], k_3 – coefficient of the rate of change of DO [1/day]. In equation (6) the coordinates of the disturbance vector represent: w_1 – the intensity of pollutant inflow [mg/l/day], w_2 – the intensity of oxygen uptake/supply from/to the water [mg/l/day].The values of coefficients k_1, k_2 and k_3 depend mainly on temperature, which implies also a time dependence. The above assumptions make equation (6) a nonlinear ordinary differential equation, which will be used in further considerations.

For the purposes of online monitoring, such measurements are selected as can be performed directly and will not cause delays in obtaining them. For the issues under consideration, such a signal is the second coordinate of the vector x, i.e. DO – $x_2(t)$. Therefore, the general notation of the measurement equation takes the form:

$$y_i = \boldsymbol{C} x_i + v_{pi}, \tag{7}$$

where: $\boldsymbol{C} = [0\ 1]$ – measurement matrix, v_p – measurement disturbance.

It should be noted that it is treated that both the measurement and the x-signals are subject to a disturbance with a Gaussian distribution [24]. It is further assumed that the measurements may be infrequent or even random.

3. CONCEPTION AND IMPLEMENTATION OF THE LOOK-UP ALGORITHM

The idea of the presented adaptive look-up zonal algorithm is based on the structure of the Kalman filter, which takes the form:

$$\hat{x}_i = \boldsymbol{A}\hat{x}_i + K_i[y_i - \boldsymbol{C}\hat{x}_i]. \tag{8}$$

The proposed algorithm resigns from Kalman's optimal selection of gain and does not impose the necessity to know the statistics of forcing signals. It should be noted that the determination of the gain K_i in equation (8) is performed in a different way, i.e. an incremental method is used with fixed values of corrections depending on the magnitude of the adaptation error (see Fig. 3). The algorithm is characterized by lower computational complexity and the lack of limitations related to linearity and statistical knowledge of signals and acquires universal features. In the proposed algorithm, it has been assumed

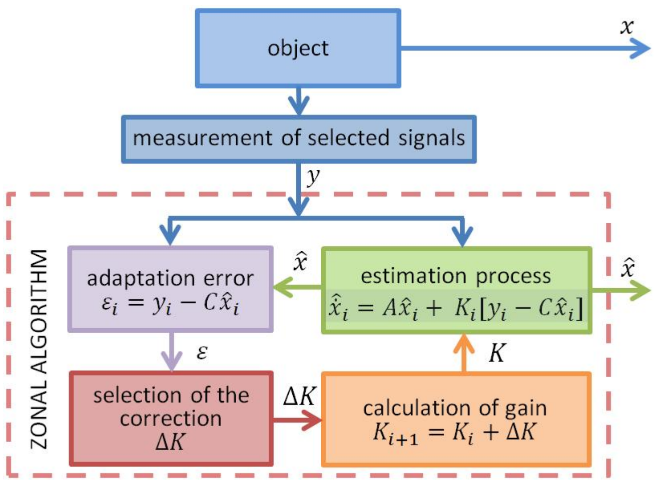

Fig. 3. Schematic of the look-up algorithm estimation idea

that the modification of the gain coefficient is performed after each measurement on the basis of the continuously determined adaptation error ε_i defined as the difference between the current measurement y_i and the corresponding coordinate of the generated estimation vector $\hat{x}_i$. The adaptation error is defined by the relation:

$$\varepsilon_i = y_i - \boldsymbol{C}\hat{x}_i\,. \tag{9}$$

The current value of the gain will be determined at each measurement step according to the rule defined by the expression:

$$K_{i+1} = K_i + \Delta K_j, \tag{10}$$

where: K_{i+1} gain coefficient for the next step, K_i – gain coefficient in the i-th step, ΔK_j – gain coefficient correction for the j-th zone.

The values of adjustments ΔK in general were assumed to depend on the value of the adaptation error $\varepsilon(t)$, i.e. $\Delta K = f(\varepsilon(t))$. The simplest solution is to assume that the function f is a step function, i.e. that ΔK values will be constant for certain ranges of $\varepsilon(t)$ changes.

The formal notation of the dependence of the correction ΔK on the estimation error for two error ranges $\varepsilon(t)$ takes the form:

$$\Delta K = \begin{cases} 0 & \text{for} \quad |\varepsilon(t)| < ER_1 \\ \Delta K_1 & \text{for} \quad |\varepsilon(t)| > ER_1 \end{cases}. \tag{11}$$

Relationship (11) represents the situation where the interval from $-ER_1$ to ER_1 can be called the insensitivity zone which means that for small adaptation errors, no gain corrections will be made.

For a more precise functioning of the algorithm it can be assumed that there can be more zones, so for N zones the ΔK values will be as follows:

$$\Delta K = \begin{cases} 0 & \text{for} \quad |\varepsilon(t)| < ER_1 \\ \Delta K_1 & \text{for} \quad ER_1 < |\varepsilon(t)| < ER_2 \\ \dots & \dots \quad \dots \\ \Delta K_{N-1} & \text{for} \quad ER_{N-1} < |\varepsilon(t)| < ER_N \\ \Delta K_N & \text{for} \quad |\varepsilon(t)| > ER_N \end{cases}. \tag{12}$$

It seems natural that a larger number of zones will produce more precise results.

For the sake of generality of considerations, in order to check the correctness of the algorithm, the worst situation was assumed regarding the knowledge of the initial gain and therefore the value of the coefficient K_0 (see equation (8)) in the first step is equal to zero. Graphical presentation of the zones of adaptation errors and the gain corrections assigned to them are shown in Fig. 4. In the variant of zone division presented in Fig. 4 three values of amplification corrections ΔK have been determined. It means that the ranges of values of the adaptation error ε obtained during the process of approximation of the monitored signals have been divided into three zones: zone I, i.e. the so-called insensitivity zone, in which no amplification correction is made, and two zones (II and III), in which the correction is made. In all cases, for each zone, a constant but different value

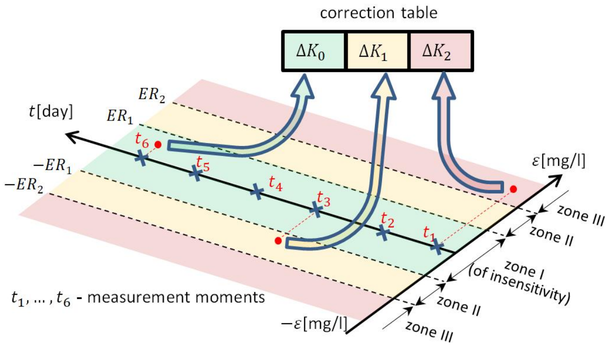

Fig. 4. Error zones and corresponding gain corrections (ΔK_i)

of the correction of filter gain coefficient ΔK was assumed. The magnitude of the correction ΔK in the proposed algorithm depends on the specifics of the object in question and is selected experimentally on the basis of expert knowledge. This correction takes different values depending on the zone to which it is assigned (see Fig. 4).

The detailed description of the algorithm is presented in the form of pseudo-code in Algorithm 1. The presented algorithm

ALGORITHM 1. Look-up algorithm

$x_0, K_0, \Delta K_1, \Delta K_2, \Delta K_3, ER_1, ER_2, ER_3, s$
//setting the initial conditions

Require: y_i //measurements
Ensure: $\hat{x}$
while $i < n$ **do**
  $\hat{x}_{i+1} \leftarrow A\hat{x} + K_i\varepsilon_i$;
  $\varepsilon_i \leftarrow y_i - C\hat{x}_i$
  if $i < s$ **or** $y_i \neq 0$ **then**
    if $\varepsilon_i \leq ER_1$ **then**
      $K_{i+1} \leftarrow K_i$
    end if
    if $ER_1 < \varepsilon_i \leq ER_2$ **then**
      $K_{i+1} \leftarrow K_i + \Delta K_1$
    end if
    if $ER_2 < \varepsilon_i \leq ER_3$ **then**
      $K_{i+1} \leftarrow K_i + \Delta K_2$
    end if
    if $\varepsilon_i > ER_3$ **then**
      $K_{i+1} \leftarrow K_i + \Delta K_3$
    end if
    if $i < s$ **then**
      ...//possibility to add more zones
    if $i < s$ **then**
      $s \leftarrow s - 1$;
    end if
  else
    $K_{i+1} \leftarrow K_i$
  end if
  $i \leftarrow i + 1$;
end while

selects the appropriate correction of the filter gain coefficient ΔK_j depending on the membership of the current adaptation error ε to the predetermined zone. If the value of the adaptation error ε is less than the error limit ER_1 then the coefficient correction is zero, i.e. the gain value is equal to the gain value from the previous step (lines 5–7 – zone I).

If the adaptation error is greater than the error limit ER_1 then the coefficient correction is equal to ΔK_1 (lines 8–10 – zone II). Subsequent zones correspond in Algorithm 1 to lines 11–13 zone III and 14–16 possible additional zone. The location of subsequent zones can be implemented from line 17.

The configuration of the algorithm includes a self-tuning process, which always occurs after the algorithm has been started and where the initial gain value is set based on measurements. The duration of the tuning period can be freely set but it cannot be shorter than the period with two measurements (in the algorithm the variable s means the number of measurements necessary to perform the tuning process). Then in this period, the initial values of the gain coefficients are determined. The actual operation of the algorithm starts after the self-tuning period.

3.1. Quality indicators for estimated signals

In the conducted research, there is a need to compare and evaluate the accuracy of the results generated by the proposed algorithm for the considered signals of the object, both measured and unmeasured. For this purpose, measures in the form of indices are used to assess the quality of the estimates [25, 26]. In research papers, various measures, both absolute and relative, can be used to analyze the errors in the obtained results.

In the research presented in this paper, two indices, i.e. root mean square error (RMSE) and mean absolute percentage error (MAPE), were used to assess the quality of the estimation process. The purpose of taking two different indices is intended to more accurately characterize the quality of the generated signals.

The first of the taken indicators RMSE in further considerations takes the form:

$$\mathrm{RMSE}_i = \sqrt{\frac{1}{n}\sum_{j=1}^{n} e_{i,j}^2}, \tag{13}$$

where: $e_{i,j} = x_{i,j} - \hat{x}_{i,j}$ – estimation error of the i-th coordinate of the signal vector x in the j-th time step, n – number of simulation time steps.

The second taken indicator MAPE takes the form:

$$\mathrm{MAPE}_i = \frac{1}{n}\sum_{j=1}^{n}\left|\frac{e_{i,j}}{x_{i,j}}\right| \cdot 100\%. \tag{14}$$

The RMSE indicator for an individual signal represents the absolute error, while the MAPE indicator is a measure of the percentage relative error. The RMSE indicator value captures a measure of error measured in units of a given signal and is not sensitive to small absolute errors. The MAPE indicator, on the other hand, represents percent error values and is sensitive to values close to or equal to zero.

4. EXPERIMENTAL STUDIES

The research of the proposed algorithm was carried out by simulating a mathematical model of a river described by ordinary differential equations. Both the simulation of the model and the implementation of the algorithm were carried out in the MATLAB environment. The simulation tests concerned in particular the time courses of BOD and DO signals with the participation of different values of system and measurement disturbances. The values of the noise signals affecting the model were obtained from a pseudorandom number generator with various parameters in the form of expected values and standard deviations determined from their assumed covariances. However, the designed monitoring algorithm does not use any information about the characteristics of these interfering signals. The quality of the generated signals was also tested at various frequencies of measurements, obtaining rare measurements. The obtained results were assessed using the RMSE and MAPE quality indicators.

The study was conducted for a river conventionally divided into three sections with side tributaries with higher pollution than the river. The average river current velocity in the sections is 35, 28 and 30 [km/day] respectively. The observation time is 36 days, which means that the river is 1116 km long. The values of BOD and DO indices vary within $(70\div 5)$ and $(-12\div -2)$ [mg/l], respectively. The lower values of these indices are similar to those of the Wisłok River near Rzeszów measured in summer.

The initial condition for the Kalman filter and adaptive algorithm estimations is for BOD: $\hat{x}_1 = 10$ [mg/l], and for DO: $\hat{x}_2 = -5$ [mg/l], while in the simulated facility the initial values for each section were assumed to be: $x_1 = 70$ [mg/l] and $x_2 = -10$ [mg/l], respectively. The values of the estimates at the beginning of subsequent sections (2 and 3) were not known but generated by the algorithm.

The following values for the parameters describing the dynamics of the self-purification process were used in the study: $k_1 = -0.23$ [1/day], $k_2 = -0.18$ [1/day], $k_3 = -0.72$ [1/day]. Furthermore, in the simulation experiments conducted, it was assumed that the river would be affected by different values of object and measurement disturbances expressed by their covariances W and V with values of $[3, -2; -2, 1]$, $[6, -4; -4, 2]$, $[9, -4; -4, 6]$ and $[18, -6; -6, 9]$, respectively, as well as $[0.1]$, $[0.3]$ and $[0.6]$. Experiments were performed for different combinations of these covariances in order to verify the influence of changes in the intensity of disturbances affecting the object on the correct functioning of the algorithm proposed in this paper, which means that the tests were performed for low and high disturbance intensities.

Figure 5 shows that the quality of the generated signals with the proposed look-up algorithm is significantly better compared to the Kalman filter. This difference can be seen in the case of sudden changes in pollution caused by lateral inflows, where there is a rapid change in the pollution values. Then, the look-up algorithm generates signals closer to the signals from the object. The results obtained apply to the case of measurements taken infrequently, i.e. in the period every 20-time steps.

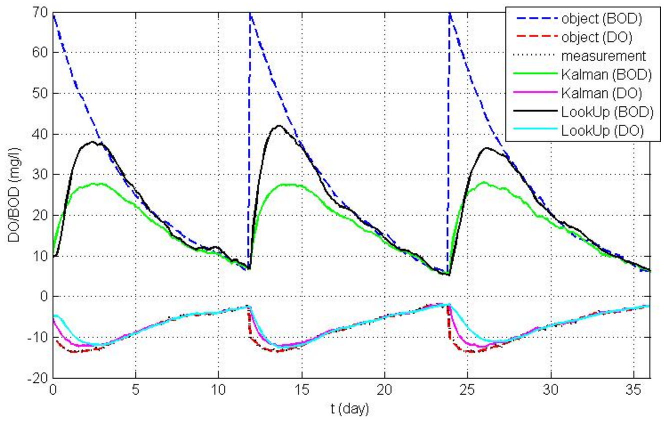

Fig. 5. Time waveforms of BOD and DO signals for a three-section river (measurements taken every 20-time steps)

The values of RMSE and MAPE indices determining the quality of algorithm performance at different measurement frequencies are presented in Table 1. The large differences between the values of indicators of the quality of monitoring for BOD and DO are due to the measurements made only for DO. The water quality indicator BOD is not measured, so the monitoring system is not directly provided with information about this signal, therefore the RMSE and MAPE indicators for it take on greater values than those for the DO signal. MAPE indicator is characterized by the fact that it reaches high values for real signals close to zero. Therefore, at the stage when the state of the river reaches low concentrations of pollutants, the index values naturally assume higher values. In simulation experiments, unusual situations with very large and sudden inflows of pollutants were researched, among others, to test the correct reaction of the algorithm to such cases. These abrupt changes in BOD cause the largest errors which have a large share in the values of the quality indicators used. In a situation where the volume of sudden pollutants inflows was limited, thus making the natural conditions for a river more realistic, a significant reduction in the value of the RMSE monitoring quality indicator is achieved. This is illustrated by the obtained time courses from Fig. 6 for which the RMSE index values were obtained, respectively, for BOD and DO with the values 4.5195, 0.91603. The quality of monitoring, measured by the MAPE indicator, also turns out to be slightly better.

Table 1
Values of monitoring quality indicators

Quality indicators	RMSE		MAPE	
Signal	BOD	DO	BOD	DO
frequent	14.1301	1.8992	12.1924	8.346
rare (10 steps)	13.4262	1.7067	13.154	7.82
rare (20 steps)	14.8757	2.1752	13.7328	10.0051
rare (80 steps)	14.9047	2.1491	13.892	10.2211
Kalman	16.6546	1.3296	21.8441	7.3888
Parameters:	look-up algorithm: $ER = 0.05$, $\Delta K = [0.85; -0.2]$, Kalman filter gain: $K_F = [-5.425; 2.835]$ covariance of system and measurement disturbances: $W = [3, -2; -2, 1]$, $V = [0.1]$			

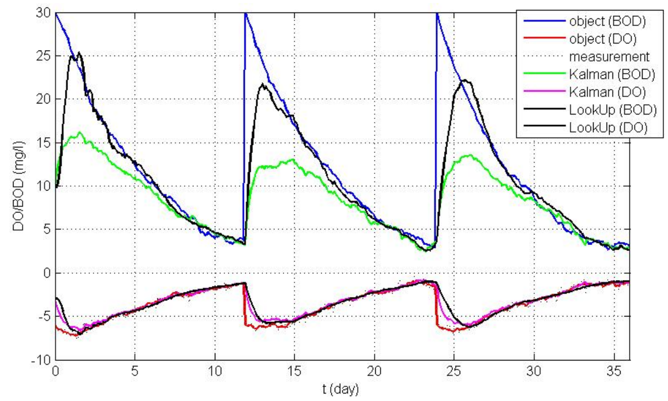

Fig. 6. Time waveforms of BOD and DO signals for a three-section river with reduced sudden pollutant inflows

In all investigated cases the look-up algorithm with frequent and less frequent measurements is superior to the Kalman filter, but for very rare measurements its quality decreases. The choice of the correction of filter gain coefficient K is a decisive element affecting the quality of the generated signals. Example waveforms of such amplifications are presented in Fig. 7.

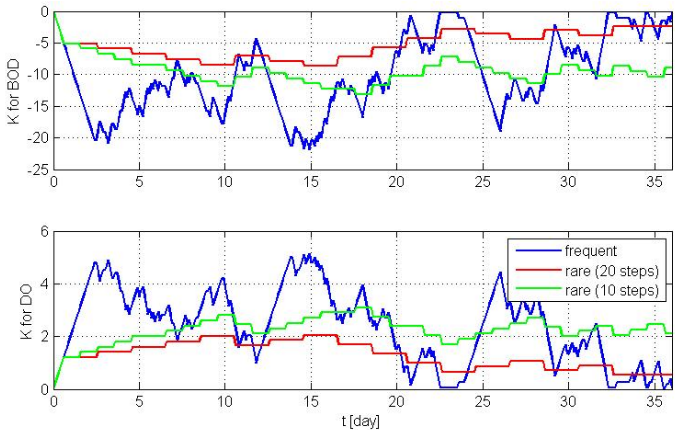

Fig. 7. Look-up algorithm gain coefficients (K) for different measurement realisations

In all cases tested, the initial gain value was zero and the algorithm always generated stable gain values. The amplification waveforms depend on the changes in the BOD and DO indices, with amplification values varying by several for DO and minus a dozen for BOD. With less frequent measurements, smaller variations in gain values were obtained.

The research carried out included the impact of changes in the system and measurement disturbances on the quality of monitoring. In a series of simulation experiments, the values of RMSE and MAPE indices were obtained for different sizes of system and measurement disturbances. Figure 8 presents the averaged values of the quality indices from 10 runs for the BOD signals generated with the look-up algorithm in comparison with the results obtained with the Kalman filter algorithm. The observations show that for the BOD signal, both monitoring quality indices always took lower values for the look-up algorithm even with rare measurements. This indicates better monitoring quality of the signals, regardless of the combination of system and measurement disturbances.

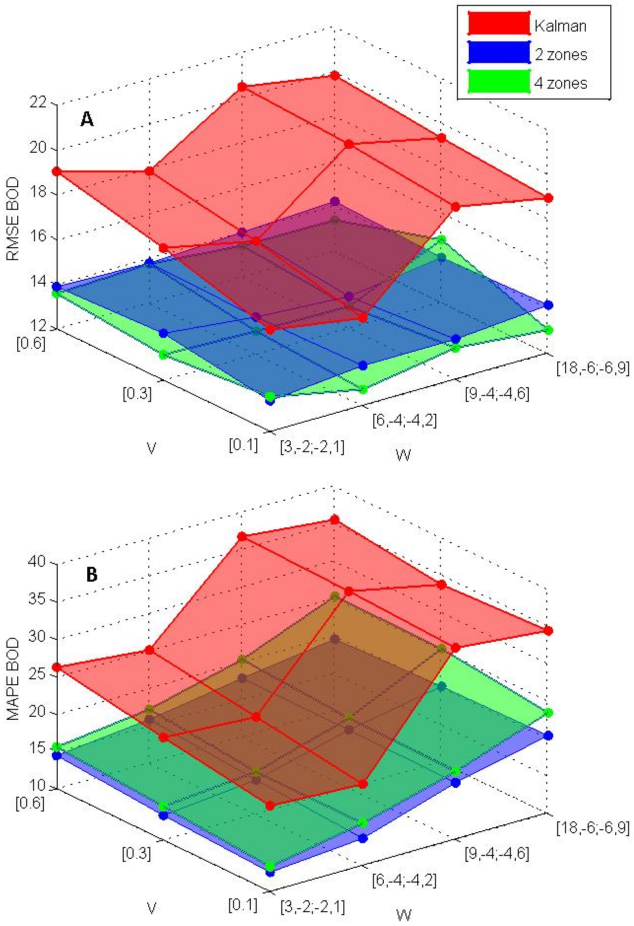

Fig. 8. RMSE (A) and MAPE (B) quality indicators for BOD signal (rare measurements – 20 steps), W, V – covariance of system and measurement disturbances

Figure 9 shows the quality indicators RMSE (A) and MAPE (B) obtained for the waveforms of the variable DO. In the performed calculations, the monitoring quality indicators for DO deficit always had smaller values in comparison with BOD. This was due to the measurements of this signal, i.e. the current information about the object. Comparing the results obtained for the Kalman filter and the look-up algorithm, it can be concluded that they are comparable for both MAPE and RMSE indicators for DO signals. A slight difference in favour of the proposed algorithm can be seen with small system disturbances and larger measurement disturbances. The Kalman filter, on the other hand, was slightly better at generating DO signals at higher values of measurement and system disturbances.

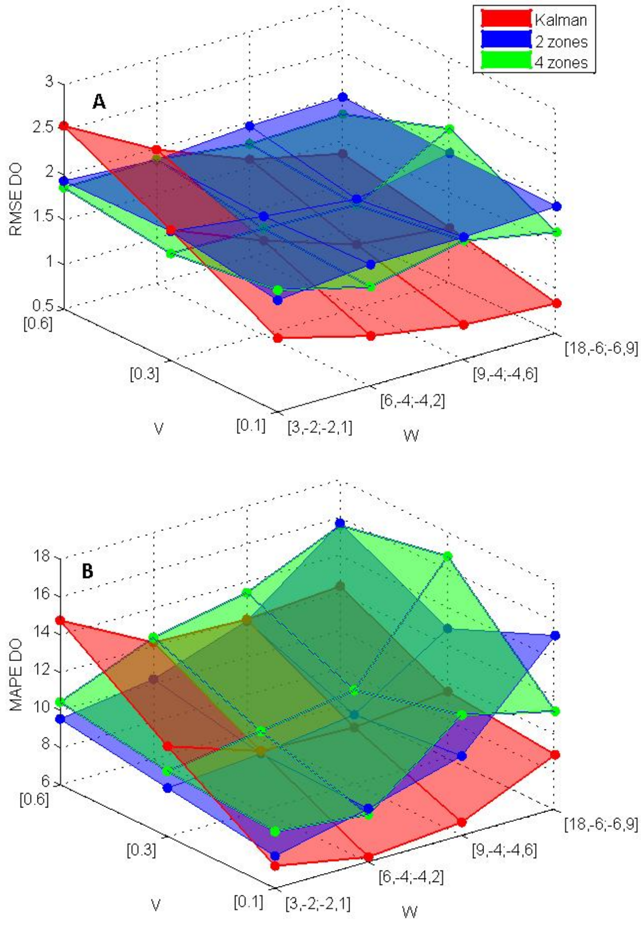

Fig. 9. RMSE (A) and MAPE (B) quality indicators for DO signal (rare measurements – 20 steps), W, V – covariance of system and measurement disturbances

5. CONCLUSIONS

This paper presents an adaptive look-up algorithm that generates online monitoring signals for an object. The object is a biochemically polluted river whose water quality is represented by water quality indicators BOD and DO deficit. The algorithm reconstructs the signals based only on the DO measurements and the mathematical model of the object. The essence of the algorithm is the adaptive selection of the gain and the correction of the filter gain coefficient in the filtration equation.

The quality of the monitored signals was determined by the RMSE and MAPE indices. The RMSE indicator values are higher for BOD compared to DO regardless of the algorithm used. For the MAPE quality indicator there is a similar variation for BOD and DO, but for BOD the look-up algorithm obtains better quality than the Kalman filter. The presented concept of online monitoring for monitoring an object described by ordinary differential equations representing two water quality indicators, i.e. BOD and DO works well.

The authors believe that the approach to river water quality monitoring presented in this article differs from the propositions considered in other scientific studies. The most similar approach is presented in the article [23] in which the monitoring quality indicator has values comparable to those presented.

Similar values of the RMSE index for BOD were also obtained in the article [27] where they are, however, the results of statistical research.

Stable operation of the algorithm was obtained in all the conducted studies with variable disturbances and rare measurements. In addition, acceptable results were obtained using an algorithm with low computational complexity, which is important for real-time systems. The presented approach works properly without the need to know the characteristics of the noise signals and any of their distributions. Correct results are also obtained with a non-linear description of the object [1]. Satisfactory results should also be expected in the further extension of the mathematical model to other water quality indicators. The proposed algorithm has low computational complexity, less than in Kalman algorithms, and can also be used for other objects.

The envisaged continuation of research work will be related to testing the presented algorithm in real conditions, e.g. sewage treatment plants or river monitoring in the Podkarpackie region. Further work will concern the adaptation of the algorithm for the needs of diagnostic monitoring and a possible control system. In order to determine the quality of monitoring, the use of other quality indicators is also envisaged.

REFERENCES

[1] P. Hawro, "Adaptive algorithms for signal estimation in the monitoring system of objects described by non-linear ordinary differential equations," Ph.D. dissertation, Faculty of Computer Science, Electronics and Telecommunications, Kraków, 2020.

[2] M. Marselina, A. Sabar, and N. Fahimah, "Spatial and temporal assessment of surface water quality using water quality index the Saguling Reservoir, Indonesia," *J. Water Land Dev.*, vol. 49, pp. 111–120, 2021.

[3] E. Novita, H.A. Pradana, B.H. Purnomo, and A.I. Puspitasari, "River water quality assessment in East Java, Indonesia," *J. Water Land Dev.*, vol. 47, no. 1, pp. 135–141, 2020.

[4] M.C. Obeta, U.P. Okafor, and C.F. Nwankwo, "Influence of discharged industrial effluents on the parameters of surface water in Onitsha urban area, southeastern Nigeria," *J. Water Land Dev.*, vol. 42, no. 1, pp. 136–142, 2019.

[5] A. Steinhoff-Wrześniewska, M. Strzelczyk, and M. Helis, "Identification of catchment areas with nitrogen pollution risk for lowland river water quality," vol. 48, no. 2, pp. 53–64, 2022.

[6] A.M. Meyer, C. Klein, E. Fünfrocken, R. Kautenburger, and H.P. Beck, "Real-time monitoring of water quality to identify pollution pathways in small and middle scale rivers," *Sci. Total Environ.*, vol. 651, pp. 2323–2333, Feb. 2019.

[7] T. Shu, M. Xia, J. Chen, and C. De Silva, "An energy efficient adaptive sampling algorithm in a sensor network for automated water quality monitoring," *Sensors*, vol. 17, no. 11, p. 2551, 2017.

[8] I. Yaroshenko *et al.*, "Real-Time Water Quality Monitoring with Chemical Sensors," *Sensors*, vol. 20, no. 12, p. 3432, Jun. 2020.

[9] P. Villalobos *et al.*, "A BOD monitoring disposable reactor with alginate-entrapped bacteria," *Bioprocess Biosyst. Eng.*, vol. 33, no. 8, pp. 961–970, 2010.

[10] Ł. Górski, K.F. Trzebuniak, and E. Malinowska, "Low BOD determination methods: The state-of-the-art," *Chem. Process Eng. – Inz. Chem. i Proces.*, vol. 33, no. 4, pp. 629–637, 2012.

[11] M. Morawiec and P. Kroplewski, "Nonadaptive estimation of the rotor speed in an adaptive full order observer of induction machine," *Bull. Pol. Acad. Sci. Tech. Sci.*, vol. 68, no. 4, pp. 973–981, 2020.

[12] A.U. Alam, D. Clyne, and M.J. Deen, "A Low-Cost Multi-Parameter Water Quality Monitoring System," *Sensors*, vol. 21, no. 11, p. 3775, May 2021.

[13] Z. Di, M. Chang, and P. Guo, "Water Quality Evaluation of the Yangtze River in China Using Machine Learning Techniques and Data Monitoring on Different Time Scales," *Water*, vol. 11, no. 2, p. 339, 2019.

[14] P. Hawro, T. Kwater, R. Pękala, and B. Twaróg, "Soft sensor with adaptive algorithm for filter gain correction in the online monitoring system of a polluted river," *Appl. Sci.*, vol. 9, no. 9, p. 1883, 2019.

[15] T. Kwater, P. Hawro, J. Bartman, and D. Mazur, "The algorithm of adaptive determination of amplification of the PD filter estimating object state on the basis of signal measurable on-line," *Arch. Control Sci.*, vol. 31, no. 1, pp. 129–143, 2021.

[16] P. Yu, J. Cao, V. Jegatheesan, and X. Du, "A Real-time BOD Estimation Method in Wastewater Treatment Process Based on an Optimized Extreme Learning Machine," *Appl. Sci.*, vol. 9, no. 3, p. 523, Feb. 2019.

[17] R. Sulistyowati, A. Suryowinoto, A. Fahruzi, and M. Faisal, "Prototype of the Monitoring System and Prevention of River Water Pollution Based on Android," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 462, p. 012028, Jan. 2019.

[18] S. Smoroń and S. Twardy, "Concentrations and loads of N-NO3, N-NH4, PO 4 and BOD5 in waters of the upper Dunajec (in the years 1985-1998)," *J. Water Land Dev.*, vol. 10, no. 10, pp. 151–162, 2006.

[19] N. Siddique and F.U. Rehman, "Hybrid synchronization and parameter estimation of a complex chaotic network of permanent magnet synchronous motors using adaptive integral sliding mode control," *Bull. Pol. Acad. Sci. Tech. Sci.*, vol. 69, no. 3, pp. 1–9, 2021.

[20] A.L. Kowal and M. Świderska-Bróż, *Oczyszczanie wody*. Warszawa: Wydawnictwo Naukowe PWN, 2009.

[21] Z. Gomolka, B. Twarog, and E. Zeslawska, "State Analysis of the Water Quality in Rivers in Consideration of Diffusion Phenomenon," *Appl. Sci.*, vol. 12, no. 3, p. 1549, 2022.

[22] R. Szymkiewicz, *Mathematical modeling of flows in rivers and canals*. Warszawa: PWN, 2000.

[23] Z. Gomolka, B. Twarog, E. Zeslawska, A. Lewicki, and T. Kwater, "Using Artificial Neural Networks to Solve the Problem Represented by BOD and DO Indicators," *Water*, vol. 10, no. 1, p. 4, Dec. 2017.

[24] K.E. Rudolph, "A New Approach to Linear Filtering and Prediction Problems," *Trans. ASME – J. Basic Eng.*, vol. 82, pp. 35–45, 1960.

[25] T. Chai and R.R. Draxler, "Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature," *Geosci. Model Dev.*, vol. 7, no. 3, pp. 1247–1250, 2014.

[26] S. Kim and H. Kim, "A new metric of absolute percentage error for intermittent demand forecasts," *Int. J. Forecast.*, vol. 32, no. 3, pp. 669–679, 2016.

[27] L.C. Ke, H.T. Van Trang, V.H. Liem, T.N. Tuong, and P.T. Duyen, "Assessment of surface water pollutant models of estuaries and coastal zone of Quang Ninh – Hai Phong using Spot-5 images," *Geod. Cartogr.*, vol. 64, no. 1, pp. 29–42, 2015.