

УДК 004.94 004.89 664.1

# АНАЛІЗ СИНГУЛЯРНОГО СПЕКТРА ВИТРАТИ ВОДИ ДИФУЗІЙНОЇ УСТАНОВКИ ЦУКРОВОГО ВИРОБНИЦТВА ДЛЯ ЗАДАЧ АВТОМАТИЗАЦІЇ

**А.С. Горпинченко***Department of Automation and Computer Technologies for Control Systems named after Prof. A.P. Ladanuk, National University of Food Technologies, Kyiv, Ukraine**ORCID <https://orcid.org/0009-0005-6190-7348>**E-mail: [antongorpinchenkodra@gmail.com](mailto:antongorpinchenkodra@gmail.com)*

## АНОТАЦІЯ

У роботі досліджено можливості застосування методу аналізу сингулярного спектра для обробки й аналізу часових рядів витрати води дифузійної установки в технологічному процесі виробництва цукру. Стабільність витрати води є важливим фактором ефективності дифузійного процесу, оскільки вона впливає на гідродинамічні умови екстракції сахарози, енергетичні витрати та загальну ефективність роботи обладнання. У реальних виробничих умовах сигнали технологічних параметрів містять значну кількість шумів, випадкових збурень і короточасних аномалій, що ускладнює їх використання в системах автоматичного керування.

Під час дослідження використано метод аналізу сингулярного спектру, який належить до класу непараметричних методів аналізу часових рядів і дає можливість виділяти трендові, сезонні та шумові складові сигналу без необхідності попереднього припущення щодо математичної моделі процесу. Запропоновано послідовність практичного застосування методу, що містить етапи формування траєкторної матриці, виконання сингулярного розкладу, групування компонент і реконструкції сигналу за допомогою діагонального усереднення.

Проведено аналіз часових рядів витрати води з використанням вікна вкладення, що відповідає характерній періодичності процесу. На основі сингулярного спектра визначено основні компоненти, які формують структуру сигналу. Показано, що перші компоненти описують тренд та сезонні коливання технологічного процесу, тоді як інші компоненти характеризують випадкові шумові збурення.

Результати дослідження демонструють, що реконструкція сигналу на основі перших компонент дає змогу ефективно фільтрувати шум, зберігаючи основні закономірності зміни витрати води. Оцінювання якості реконструкції виконано за допомогою статистичних показників точності, зокрема середньої абсолютної похибки, середньоквадратичної похибки та середньої відносної похибки. Отримані результати підтверджують можливість використання методу аналізу сингулярного спектра для підвищення достовірності вимірювальних сигналів та їх подальшого використання в системах автоматичного керування, діагностики технологічного обладнання та прогнозування режимів роботи дифузійної установки.

Практична цінність роботи полягає у можливості інтеграції запропонованого підходу в програмно-технічні комплекси автоматизації цукрового виробництва для підвищення стабільності технологічних режимів та ефективності використання ресурсів.

**Ключові слова:** аналіз сингулярного спектра, метод гусениці, часові ряди, дифузійна установка, витрата води, автоматизація цукрового виробництва, обробка сигналів, промислова автоматизація, статистичний аналіз даних, фільтрація сигналів, реконструкція часових рядів.

## Вступ

Цукрове виробництво належить до енерго- та ресурсоємних галузей харчової промисловості. Стабільність технологічних режимів на дифузійній ділянці суттєво впливає на вихід сахарози та витрати пари, води й електроенергії. Витрата води на дифузю є одним із ключових параметрів, що визначає

гідродинамічні умови, інтенсивність масообміну та якість дифузійного соку. Для автоматизованих систем керування важливо мати сигнал витрати води з мінімальним впливом випадкових збурень, датчикового шуму та короточасних аномалій.

Типові промислові дані мають складну структуру, що поєднує трендові компоненти, сезонні коливання

та випадкові збурення. Для моделювання й аналізу таких сигналів широко використовуються методи аналізу часових рядів, статистичні підходи та методи головних компонент [5], [6]. Вони поєднують повільний тренд, добові коливання, вплив режимів пуску та зупинки, а також нерегулярні відхилення, зумовлені зміною сировини та діями персоналу. Через це використання простих лінійних фільтрів або класичних авторегресійних моделей не завжди дає надійний результат. Перспективним підходом до аналізу складних часових рядів є аналіз сингулярного спектра, який у прикладній літературі часто називають методом гусениці. Цей метод поєднує ідеї спектрального аналізу та статистичного розкладу сигналів і широко використовується для виділення трендів, періодичних складових та шуму у часових рядах [1], [2].

### Аналіз останніх досліджень і постановка задачі

Методи обробки сигналів у системах автоматизації традиційно передбачають згладжування, спектральний аналіз, фільтрацію Калмана й ідентифікацію моделей у просторі станів. Для процесів харчової промисловості характерні нелінійності, зміна параметрів у часі та періодичні впливи, що ускладнює побудову єдиної моделі. Аналіз сингулярного спектра розглядається як інструмент, який поєднує переваги статистичного підходу та спектрального розкладу, забезпечуючи адаптивне виділення компонент сигналу без попереднього припущення щодо моделі процесу [1], [3], [4].

**Мета роботи** полягає у розробленні та демонстрації інженерної методики застосування методу аналізу сингулярного спектра для часових рядів витрати води на дифузійну установку з отриманням числових показників якості реконструкції та рекомендацій для подальшого використання результатів у задачах автоматичного керування та діагностики, базується на перетворенні одновимірного часового ряду в багатовимірний за допомогою формування траєкторної матриці.

### Матеріали та методи

Об'єктом дослідження є часовий ряд витрати води на дифузійну установку, отриманий із засобів вимірювання та реєстрації технологічних параметрів. У цьому матеріалі наведено демонстраційний фрагмент для ілюстрації розрахунків. Під час підготовки до подання в конкретний журнал таблицю та графік легко замінити на ряд із вашої системи Supervisory Control and Data Acquisition без зміни методики.

Вибрано дискретизацію один вимір на годину та тривалість спостереження три доби. Позначимо початковий ряд як  $x$  з індексом часу  $t$ .

Модельне подання сигналу в задачі розкладу має вигляд суми корисних складових та шуму [6], [7]:

$$x_t = y_t + e_t. \quad (1)$$

Тут  $y$  є сукупністю тренду та сезонних складових, а  $e$  – шумом і нерегулярними збуреннями. Метод аналізу сингулярного спектра виконується в чотири етапи: вбудовування, сингулярний розклад, групування компонент і діагональне усереднення.

#### Вбудовування та траєкторна матриця

Першим і фундаментальним етапом алгоритму аналізу сингулярного спектра є процедура вбудовування, яка полягає у відображенні вихідного одновимірного часового ряду в багатовимірний простір. Нехай маємо часовий ряд витрати води на дифузійну установку (2) загальною довжиною  $N$ , для реалізації процедури вбудовування потрібно визначити ціле число  $L$ , яке називається довжиною вікна (window length) (3) [7]:

$$F = (f_0, f_1, \dots, f_{N-1}), \quad (2)$$

$$1 < L < N. \quad (3)$$

Процес вбудовування перетворює часовий ряд у послідовність (4) векторів вбудовування (5) розмірності  $L$ :

$$K = N - L + 1, \quad (4)$$

$$F = (f_{0-1}, \dots, f_{i+L-2})^T. \quad (5)$$

Результатом цього кроку є формування траєкторної матриці  $X$  розміром  $L \times K$ , яка має таку структуру (6):

$$X = [X_1, \dots, X_k] = \begin{bmatrix} f_0 & f_1 & \dots & f_{k-1} \\ f_1 & f_2 & \dots & f_k \\ \vdots & \vdots & \ddots & \vdots \\ f_{L-1} & f_L & \dots & f_{N-1} \end{bmatrix}. \quad (6)$$

Траєкторна матриця  $X$  за своєю будовою є ганкелевою матрицею, оскільки вона має однакові елементи на діагоналях, що йдуть справа наліво. Це означає, що всі елементи  $x_{ij}$  для яких  $i + j$ , є ідентичними.

Вибір довжини вікна  $L$  є найбільш критичним моментом на етапі вбудовування, оскільки цей параметр визначає здатність методу розділяти різні компоненти сигналу. З теоретичного погляду  $L$  має бути достатньо великим, щоб охопити основну динаміку процесу, але не перевищувати  $N$ . У контексті автоматизації цукрового виробництва, де процеси мають виражену циклічність, вікно  $L$  доцільно вибирати кратно очікуваній сезонності або періоду домінуючих коливань.

Для добового циклу спостережень із кроком дискретизації один вимір на годину оптимальними значеннями є  $L = 24$  або  $L = 48$ . Такий вибір дає можливість матриці «акумулювати» інформацію про добові ритми споживання води та технологічні зміни протягом зміни. У цій науковій роботі прийнято значення  $L = 24$ , що забезпечує адекватне розбиття сигналу на

низькочастотний тренд, добову періодику та високо-частотний шум. Формування такої матриці дає змогу на наступних етапах застосувати апарат сингулярного розкладу для виявлення прихованих закономірностей у динаміці витрати води, які є невидимими за стандартного візуального аналізу часового ряду (7):

$$X = [x_1 \dots x_K]. \quad (7)$$

#### Сингулярний розклад

На другому етапі алгоритму виконується сингулярний розклад (Singular Value Decomposition) сформованої траєкторної матриці  $X$ . Ця процедура є еквівалентною розкладу вихідного сигналу на систему ортогональних компонент у просторі вкладених векторів [5]. Математично цей процес подається у вигляді добутку трьох матриць (8):

$$X = U \Sigma V^T. \quad (8)$$

Відповідно  $U$  – ортогональна матриця розміром  $L$ , стовпці якої є лівими сингулярними векторами (власними векторами матриці  $XX^T$ );  $V$  – ортогональна матриця розміром  $K$ , що містить праві сингулярні вектори;  $\Sigma$  – діагональна матриця розміром  $L \times K$ , на головній діагоналі якої розташовані невід'ємні сингулярні значення  $s_j$ , впорядковані за спаданням (9):

$$(s_1 \geq s_2 \geq \dots \geq s_L \geq 0). \quad (9)$$

Діагональні елементи матриці відіграють ключову роль у структурному аналізі часового ряду витрати води. Кожне сингулярне значення характеризує масштаб відповідної компоненти розкладу. Квадрати сингулярних значень, які часто позначають як власні числа (10)

$$\lambda_i = s_i^2, \quad (10)$$

є прямо пропорційними внеску відповідних компонент у загальну дисперсію досліджуваного сигналу. Аналіз сингулярного спектра дає змогу ідентифікувати структуру ряду: великі сингулярні значення зазвичай відповідають головним компонентам – низькочастотному тренду та періодичним гармонікам, тоді як довга ланка малих значень свідчить про наявність випадкового шуму та збурень.

Для потреби в автоматизації технологічного процесу такий розклад є корисним, оскільки він дає можливість перейти від спотвореного шумом часового ряду до його енергетичного представлення [8], [9]. Власні вектори  $U_i$  визначають форму коливань, а головні компоненти відображають динаміку цих коливань у часі. Таким чином, сингулярний розклад виступає адаптивним фільтром, який самостійно підлаштовується під специфіку даних конкретної дифузійної установки, не вимагаючи від розробника точного знання частотних характеристик процесу.

Внесок  $i$  компоненти в загальну структуру сигналу в процентах обчислюється (11):

$$p_i = \frac{s_i}{\sum_{j=1}^s s_j} * 100\%. \quad (11)$$

Такий підхід забезпечує високу роздільну здатність методу та дає можливість ефективно відокремити детерміновану частину витрати води від стохастичних збурень, що є критичним для стабілізації роботи контурів регулювання.

Компоненти об'єднують у групи, які відповідають тренду, сезонності та шуму. Практичне правило полягає у виборі першої компоненти як тренду та парних компонент як сезонних гармонік, якщо в сингулярному спектрі спостерігаються близькі значення та характерні коливальні власні вектори.

Для кожної групи формується матриця відновлення як сума елементарних матриць. Після цього застосовується діагональне усереднення, яке перетворює матрицю назад у часовий ряд. Підсумковий реконструйований сигнал дорівнює сумі відновлених груп тренду та сезонності (12):

$$x^\circ = x\_trend + x\_season. \quad (12)$$

Якість реконструкції оцінено за середньою абсолютною похибкою MAE (13), коренем середньоквадратичної похибки RMSE (14) та середньою відносною похибкою MAPE (15). Для уникнення неоднозначностей у записі введемо допоміжні величини [11], [13]:

$$a_t = x_t - x_t^\circ, \quad (13)$$

$$MAE = \frac{1}{N} \sum_{t=1}^N a_t,$$

$$b_t = x_t - x_t^{\circ 2}, \quad (14)$$

$$RMSE = \frac{1}{N} \sum_{t=1}^N b_t,$$

$$c_t = \frac{a_t}{x_t}, \quad (15)$$

$$MAPE = 100 \frac{1}{N} \sum_{t=1}^N c_t.$$

#### Результати та їх аналіз

Після побудови траєкторної матриці з параметром  $L$ , що дорівнює 24, виконано сингулярний розклад та обчислено внесок перших компонент у сумарну енергію сигналу. Для практичного використання важливо контролювати, щоб перші компоненти відповідали фізичним закономірностям процесу, а не випадковим пікам. У демонстраційному прикладі перша компонента відображає тренд, друга та третя

компоненти утворюють сезонну пару, інші компоненти переважно описують шум.

Таблиця 1 містить значення часток дисперсії для перших десяти компонент, а також накопичену частку. Це дає можливість обґрунтувати вибір числа компонент для реконструкції.

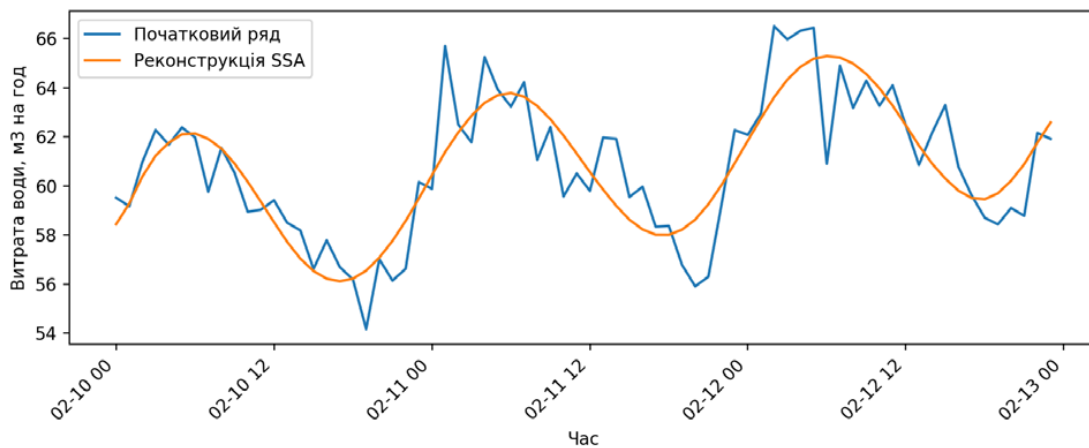
За даними таблиці 1, перші три компоненти описують основну структуру сигналу. Реконструкцію виконано як суму тренду та сезонності, що відповідає компонентам 1, 2 та 3.

За вибраною реконструкцією отримано такі показники точності: середня абсолютна похибка дорівнює 1,128 м<sup>3</sup> на год, корінь середньоквадратичної похибки дорівнює 1,490 м<sup>3</sup> на год, середня відносна похибка – 1,859 відсотка. Невелика відносна похибка свідчить, що відновлений сигнал зберігає фізично значущі коливання та є придатним для використання в автоматичних контурах керування.

На рис. 1 показано початковий часовий ряд і реконструкцію сигналу. Видно, що реконструкція

**Табл. 1. Внесок перших компонент методу аналізу сингулярного спектра в дисперсію сигналу витрати води**

Номер компоненти	Сингулярне значення $s$	Частка дисперсії $p$ , відсотків	Накопичено, відсотків
1	2085,346	99,78	99,78
2	57,780	0,08	99,86
3	56,686	0,07	99,93
4	27,648	0,02	99,95
5	27,324	0,02	99,96
6	13,774	0,00	99,97
7	12,782	0,00	99,97
8	10,922	0,00	99,97
9	10,527	0,00	99,98
10	10,282	0,00	99,98



**Рис. 1. Початковий часовий ряд витрати води та реконструкція методом аналізу сингулярного спектра**

згладжує випадкові коливання та локальні викиди, зберігаючи тренд і добову сезонність. Це зменшує ризик помилкових спрацьовувань регуляторів та підвищує стабільність технологічного режиму.

#### **Практичні рекомендації для впровадження у системі автоматизації**

Для інтеграції методу аналізу сингулярного спектра в програмно-технічний комплекс рекомендовано реалізувати алгоритм як окремий модуль

обробки сигналів. Вхідними даними є масив значень витрати води з буфера архіву, виходом є очищений сигнал і показники якості. Оновлення доцільно виконувати ковзним вікном із кроком один вимір.

Довжину вікна  $L$  слід узгоджувати із характерним періодом коливань. Для добової сезонності за годинної дискретизації рекомендовано  $L$ , що дорівнює 24. Для більш складних режимів можна застосовувати  $L$ , що дорівнює 48. Занадто мале  $L$

приводить до змішування сезонності із шумом, занадто велике  $L$  знижує стійкість оцінок у разі коротких вибірок.

Групування компонент варто виконувати на основі двох критеріїв: частка дисперсії та форма власних векторів. Для промислової експлуатації зручно фіксувати правило: перша компонента відповідає тренду, друга та третя компоненти відповідають сезонності, інші компоненти відкидаються як шум. Далі правило уточнюють на основі архівних даних за декілька змін і для різних типів навантаження.

Отриманий очищений сигнал може бути використаний для підвищення енергоефективності виробництва й оптимізації режимів роботи технологічного обладнання відповідно до сучасних стандартів управління енергоспоживанням [15].

### Висновки

1. Виконано обґрунтування використання аналізу сингулярного спектру для обробки часових рядів витрати води дифузійної установки цукрового виробництва.

2. Запропоновано інженерну методику налаштування методу аналізу сингулярного спектра з вибором довжини вікна, сингулярним розкладом і групуванням компонент для виділення тренду та сезонності.

3. Для демонстраційного фрагмента даних, якщо  $L$  дорівнює 24, реконструкція за першими трьома компонентами забезпечила середню абсолютну похибку 1,128 м<sup>3</sup> на год, корінь середньоквадратичної похибки – дорівнює 1,490 м<sup>3</sup> на год, середня відносна похибка – 1,859 відсотка.

4. Показано, що очищений сигнал придатний для застосування в контурах автоматичного керування та діагностики, а підхід може бути перенесений на суміжні задачі, зокрема на керування активною вентиляцією цукрового буряка.

### Конфлікт інтересів

Автор декларує, що не має конфлікту інтересів стосовно цього дослідження, у тому числі фінансового, особистісного характеру, авторства чи іншого характеру, який міг би вплинути на дослідження та його результати, представлені в цій статті.

### Фінансування

Дослідження проводилося без фінансової підтримки.

### Доступність даних

Рукопис не має пов'язаних даних.

### ЛІТЕРАТУРА

- [1] N. Golyandina, Analysis of Time Series Structure. SSA and Related Techniques. Boca Raton, FL, USA: Chapman and Hall/CRC, 2001. DOI: 10.1201/9781420035841.
- [2] N. Golyandina and A. Zhigljavsky, Singular Spectrum Analysis for Time Series. Berlin, Heidelberg: Springer, 2013. DOI: 10.1007/978-3-642-34913-3.
- [3] H. Hassani, “Singular Spectrum Analysis: Methodology and Comparison,” Journal of Data Science, vol. 5, no. 2, pp. 239–257, 2007. DOI: 10.6339/JDS.2007.05(2).396.
- [4] R. Vautard and M. Ghil, “Singular spectrum analysis in nonlinear dynamics,” Physica D, vol. 35, no. 3, pp. 395–424, 1989. DOI: 10.1016/0167-2789(89)90077-8.
- [5] T. Jolliffe, Principal Component Analysis. New York, NY, USA: Springer, 2002. DOI: 10.1007/b98835.
- [6] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, Time Series Analysis: Forecasting and Control. Hoboken, NJ, USA: Wiley, 2015.
- [7] D. C. Montgomery, Introduction to Statistical Quality Control. Hoboken, NJ, USA: Wiley, 2019.
- [8] K. Ogata, Modern Control Engineering. Upper Saddle River, NJ, USA: Pearson, 2010.
- [9] D. E. Seborg, T. F. Edgar, and D. A. Mellichamp, Process Dynamics and Control. Hoboken, NJ, USA: Wiley, 2011.
- [10] ДСТУ 8302:2015 Інформація та документація. Бібліографічне посилання. Загальні положення та правила складання. Київ: Держспоживстандарт України, 2015.
- [11] В. Крупко, Автоматизація технологічних процесів харчових виробництв. Навчальний посібник. Київ: Видавництво, 2018.
- [12] О. Якубенко, Технологія цукру. Дифузійні процеси та режими. Харків: Видавництво, 2016.
- [13] С. Бойко, Системи автоматичного керування. Теорія і практика. Львів: Видавництво, 2017.
- [14] І. Левченко, Методи фільтрації та обробки сигналів у системах керування. Одеса: Видавництво, 2020.
- [15] ISO 50001 Energy management systems. Requirements with guidance for use. International Organization for Standardization, 2018.

### ANALYSIS OF THE SINGULAR SPECTRUM OF WATER CONSUMPTION IN A DIFFUSION UNIT OF A SUGAR PRODUCTION PLANT FOR AUTOMATION APPLICATIONS

Anton Horpynchenko

*This paper investigates the application of Singular Spectrum Analysis for processing time series of water flow rate in a diffusion unit of sugar production. Stable water flow is an important factor affecting the efficiency of the diffusion process, since it influences hydrodynamic conditions, sucrose extraction efficiency, and energy consumption of the technological equipment. In industrial environments, measurement signals often contain noise, random disturbances, and short term anomalies, which complicates their direct use in automatic control systems.*

*The study applies Singular Spectrum Analysis as a non parametric method for time series decomposition that allows separating trend, periodic components, and*

noise without assuming a predefined mathematical model of the process. A practical implementation procedure is proposed, including trajectory matrix construction, singular value decomposition, component grouping, and signal reconstruction using diagonal averaging.

Experimental analysis of water flow rate data demonstrates that the leading components represent the main physical structure of the signal, including trend and seasonal variations, while higher order components correspond to noise. Reconstruction based on the dominant components significantly reduces random disturbances while preserving the informative dynamics of the technological process.

The quality of reconstruction is evaluated using statistical accuracy indicators such as Mean Absolute Error, Root Mean Square Error, and Mean Absolute Percentage Error. The obtained results confirm that Singular Spectrum Analysis can effectively improve the reliability of measurement signals and may be applied in industrial automation systems for control, diagnostics, and forecasting of diffusion unit operating modes.

**Keywords:** singular spectrum analysis, the Gusevitsa method, time series, diffusion unit, water flow rate, sugar production automation, signal processing, industrial automation, statistical data analysis, signal filtering, time series reconstruction.

## REFERENCES

- [1] N. Golyandina, Analysis of Time Series Structure. SSA and Related Techniques. Boca Raton, FL, USA: Chapman and Hall/CRC, 2001. DOI: 10.1201/9781420035841.
- [2] N. Golyandina and A. Zhigljavsky, Singular Spectrum Analysis for Time Series. Berlin, Heidelberg: Springer, 2013. DOI: 10.1007/978-3-642-34913-3.
- [3] H. Hassani, "Singular Spectrum Analysis: Methodology and Comparison," Journal of Data Science, vol. 5, no. 2, pp. 239–257, 2007. DOI: 10.6339/JDS.2007.05(2).396.
- [4] R. Vautard and M. Ghil, "Singular spectrum analysis in nonlinear dynamics," Physica D, vol. 35, no. 3, pp. 395–424, 1989. DOI: 10.1016/0167-2789(89)90077-8.
- [5] T. Jolliffe, Principal Component Analysis. New York, NY, USA: Springer, 2002. DOI: 10.1007/b98835.
- [6] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, Time Series Analysis: Forecasting and Control. Hoboken, NJ, USA: Wiley, 2015.
- [7] D. C. Montgomery, Introduction to Statistical Quality Control. Hoboken, NJ, USA: Wiley, 2019.
- [8] K. Ogata, Modern Control Engineering. Upper Saddle River, NJ, USA: Pearson, 2010.
- [9] D. E. Seborg, T. F. Edgar, and D. A. Mellichamp, Process Dynamics and Control. Hoboken, NJ, USA: Wiley, 2011.
- [10] DSTU 8302:2015 Information and documentation. Bibliographic reference. General provisions and rules of composition. Kyiv: Derzhspozhyvstandard of Ukraine, 2015.
- [11] V. Krupko, Automation of technological processes in food production. Textbook. Kyiv: Vydavnytstvo, 2018.
- [12] O. Yakubenko, Sugar technology. Diffusion processes and modes. Kharkiv: Vydavnytstvo, 2016.
- [13] S. Boiko, Automatic control systems. Theory and practice. Lviv: Vydavnytstvo, 2017.
- [14] I. Levchenko, Methods of filtering and signal processing in control systems. Odesa: Vydavnytstvo, 2020.
- [15] ISO 50001:2018, Energy management systems – Requirements with guidance for use. International Organization for Standardization, 2018.

Дата першого надходження статті до видання:

12.02.2026

Дата прийняття статті до друку

після рецензування: 05.03.2026

Дата публікації (оприлюднення) статті:

12.05.2026



Стаття поширюється на умовах ліцензії відкритого доступу CC BY 4.0