

ГЕНЕТИЧЕСКИЕ ТЕХНОЛОГИИ
И СИСТЕМЫ ЖИЗНЕОБЕСПЕЧЕНИЯ

УДК 582.232; 519,688

ИСПОЛЬЗОВАНИЕ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ
ДЛЯ ПРЕДСКАЗАНИЯ СОДЕРЖАНИЯ КИСЛОРОДА
И ДИОКСИДА УГЛЕРОДА В ГАЗОВОЗДУШНОЙ СМЕСИ,
ВЫХОДЯЩЕЙ ИЗ ФОТОБИОРЕАКТОРА ПРИ КУЛЬТИВИРОВАНИИ
ARTHROSPIRA PLATENSIS

© 2023 г. П. С. Назин¹, Д. В. Сухинов¹, К. В. Горин^{1,*}

¹Национальный исследовательский центр «Курчатовский институт», Москва, Россия

*E-mail: Gorin_KV@nrcki.ru

Поступила в редакцию 10.10.2023 г.

После доработки 10.10.2023 г.

Принята к публикации 10.10.2023 г.

Сегодня фототрофные микроорганизмы рассматриваются не только как перспективный инструмент применения в различных отраслях для получения широкого ряда полезных веществ, но и встраиваются в системы по очистке воды, воздуха, а также в системы жизнеобеспечения при длительных космических полетах для утилизации продуктов жизнедеятельности экипажа, получения пищевых добавок, регенерации кислорода. Одним из инструментов, позволяющих предсказывать поведение фототрофных микроорганизмов в тех или иных условиях, являются искусственные нейронные сети. Для обучения системы использовались данные, полученные в ходе различных экспериментов, общий набор используемых данных составлял 11970 точек. Результаты показывают, что система способна предсказывать содержание кислорода и диоксида углерода в газовой смеси, выходящей из фотобиореактора, достигая R^2 от 0.8 до 0.98, что свидетельствует о высокой точности предсказания и перспективах для применения данного метода для разработки систем жизнеобеспечения и рециркуляции.

DOI: 10.56304/S2782375X23010102

ВВЕДЕНИЕ

Фототрофные микроорганизмы являются перспективным и активно развивающимся направлением в биоэнергетике, пищевой и фармацевтической промышленности, космической отрасли. Они могут использоваться в системах очистки воды, переработки отходов [1, 2], производства биотоплив, кормовой и пищевой биомассы и биологически активных веществ [3, 4], а также встраиваться в системы жизнеобеспечения космических аппаратов для обеспечения экипажа кислородом и пищевыми добавками, переработки отходов [5–7]. Темпы роста и параметры жизнедеятельности микроводорослей зависят от множества параметров среды и режима культивирования, таких как температура, интенсивность и спектр освещения, содержание различных питательных веществ в среде и т.д. Общий анализ различных проблем на пути промышленного культивирования микроводорослей проведен в [8].

Ввиду сложного метаболизма микроорганизмов создание классической модели, позволяющей предсказывать рост в различных условиях,

сильно затруднено. В качестве более подходящего метода предлагается использование методов машинного обучения, одним из которых являются искусственные нейронные сети (ИНС) [9]. Предсказание роста на основе ИНС позволит получить представление о поведении культуры без проведения длительных экспериментов.

ИНС – метод машинного обучения, использующий набор экспериментальных, полученных при наблюдении, или гипотетических данных, сопоставляющий входные параметры с известными корректными результатами. Обученная на этих данных система затем экстраполирует связи между ними на все пространство возможных входных параметров [10].

Примерами применения ИНС для решения предсказания и оптимизации культивации микроводорослей являются работы [9, 11–14] и работы, касающиеся производства различных веществ с их помощью [15–18]. Наиболее распространенным типом ИНС для решения таких задач являются ИНС с прямым распространением сигнала и обратным распространением ошибки. В качестве примеров

применения ИНС в работе с цианобактериями *Arthrospira (Spirulina) platensis* можно привести [19], где описана разработка системы на основе ИНС для наблюдения за ростом аквакультуры цианобактерии в открытых прудах для предотвращения возможной смерти культуры при резких изменениях погоды, где классические системы автоматического контроля показали себя недостаточно эффективными. Система принимала кислотность, температуру, электропроводимость среды и плотность культуры в качестве входных параметров, предсказывая эволюцию тех же параметров состояния среды и культуры с течением времени. Используемая ИНС включала в себя четыре физических процессора для эмуляции 20 слоев ИНС с 80 нейронами в сумме и требовала для обучения данные, полученные не менее чем за шесть недель наблюдения за ростом аквакультуры в различных условиях. В [20] проанализировано влияние использования различных систем аэрации и частоты сбора на продуктивность биомассы *A. platensis* в замкнутых фотобиореакторах. В дополнение к описанным выше параметрам в качестве входных параметров для ИНС изначально были использованы температура, кислотность среды, интенсивность освещения, содержание растворенного кислорода, выход кислорода, содержание бикарбонатов, фосфатов и нитратов в среде, однако тестирование показало, что в данном случае исключение интенсивности освещения, содержания бикарбонатов и фосфатов оказало положительное влияние на точность предсказания. Также можно упомянуть [21, 17], в которых внимание заострено на сложности эффективного обучения ИНС в условиях малого количества экспериментальных данных и применении ИНС для исследования роста микроводорослей в условиях открытого пруда, более сложных по сравнению с лабораторными условиями, соответственно.

Отметим, что цианобактерия *A. platensis* широко используется для получения съедобной биомассы и биологически активных веществ, находящихся применение в медицине и пищевой промышленности [22, 23].

В данной работе исследованы возможности применения алгоритма на основе ИНС для предсказания концентрации растворенного кислорода и диоксида углерода в выходящем газе из фотобиореактора при культивировании цианобактерии *A. platensis*.

МАТЕРИАЛЫ И МЕТОДЫ

Для обучения и тестирования программы предсказания на основе ИНС использовали набор данных, полученный при культивировании цианобактерии *A. platensis* В-12619 из коллекции ВКПМ. Цианобактерии культивировали в фотобиореакторе INFORS HT Labfors 5 в среде

Таблица 1. Состав компонента S1 среды Заррук

Вещество	Концентрация, г/л
NaHCO ₃	10.8
Na ₂ CO ₃	7.6
KH ₂ PO ₄	0.5

Таблица 2. Состав компонента S2 среды Заррук

Вещество	Концентрация, г/л
NaNO ₃	2.5
K ₂ SO ₄	1.0
NaCl	1.0
MgSO ₄ · 7H ₂ O	0.2
CaCl ₂ · 2H ₂ O	0.04
Микроэлементы	1 мл/л

объемом 1.8 л, в условиях постоянного освещения светодиодами 1500 Лк, перемешивания борботированием и при поддержании температуры среды на постоянном уровне 30°C. Анализ выходящего газа на содержание кислорода и диоксида углерода проводили при помощи газоанализатора BlueSens BlueInOne Cell. Цианобактерии выращивали на среде Заррук в модификации Корнета [24]. Отдельно готовили на дистиллированной воде компоненты среды S1 и S2 (табл. 1, 2) и смешивали в соотношении 1 : 1.

Среду использовали с вариациями различных параметров условий культивирования. Питательные вещества добавляли в среду при засеве культуры и их содержание не поддерживалось на постоянном уровне в течение роста. Интенсивность освещения и температура поддерживались на постоянном уровне в течение культивирования и не менялись между экспериментами.

В качестве инструмента обработки данных и моделирования была выбрана ИНС с прямым распространением данных и обратным распространением ошибки – методом машинного обучения, хорошо работающим со сложными нелинейными системами без необходимости создания аналитической модели. Для оптимизации архитектуры ИНС использовали модифицированный метод оптимизации гиперпараметров поиском по сетке, а также метод случайного поиска для получения начальных значений для дальнейшей работы и ручную подстройку в отдельных случаях и для набора сравнительной статистики.

ИНС с прямым распространением данных и обратным распространением ошибки, использованием дополнительного алгоритма предобработки обучающих данных и различными возможностями для оптимизации гиперпараметров раз-

Таблица 3. Список гиперпараметров ИНС

Гиперпараметр	Диапазон значений
Количество нейронов на первом скрытом слое	3–7
Количество нейронов на втором скрытом слое	3–7
Количество нейронов на третьем скрытом слое	3–7
Передающая функция 1	Сигмоидная, гиперболический тангенс, ReLu (усеченное линейное преобразование)
Передающая функция 2	
Передающая функция 3	
Размер пакета (при обучении)	50–500
Количество эпох обучения	1000–10000
Скорость обучения	0.01–0.1

Таблица 4. Расшифровка передаточных функций

Функция	Формула
Сигмоидная	$S(x) = \frac{1}{1 + e^x}$
Гиперболический тангенс	$\tanh(x) = \frac{e^{2x} - 1}{e^{2x} + 1}$
ReLu (усеченное линейное преобразование)	$F(x) = \max(0, x)$

Таблица 5. Входные параметры ИНС

Параметр	Диапазон значений	Значимость
Значение pH	9.5–11.0	Оказывает влияние на метаболизм клеток и рост культуры [28–31]
Концентрация растворенного кислорода в среде, %	80–100	Оказывает влияние на метаболизм клеток и рост культуры
Длительность культивирования, с	450–3265650	Определяет задачу предсказания
Мутность культурной среды	0.6–1.6	Пересчитывается в плотность биомассы культуры

работана на языке программирования Python 3.9 с использованием программного пакета TensorFlow 2.80. ИНС была затем обучена с использованием метода обучения с учителем и оптимизацией при помощи улучшенного метода градиентного спуска Adam на имеющемся наборе данных при помощи инструментария, предоставляемого пакетом TensorFlow. Выбор данной архитектуры обусловлен тем, что она менее требовательна к количеству обучающих данных, а также обучается быстрее, чем методы глубинного обучения [25]. Архитектуры типов PNN (*probabilistic neural network* – вероятностные нейронные сети) и CNN (*convolutional neural network* – сверточные нейронные сети) не подходят для решения задач этого класса [26, 27].

Используемая ИНС состоит из входного слоя, включающего в себя четыре нейрона по числу

входных параметров, трех скрытых слоев и выходного слоя, включающего для нейрона по числу выходных параметров. Выбор трех скрытых слоев обусловлен общими соображениями о сложности обрабатываемых данных и подтвержден сравнением с другими конфигурациями на ранних этапах разработки. Количество нейронов на скрытых слоях, типы передаточных функций и параметры скорости обучения задаются при оптимизации гиперпараметров.

В качестве входных параметров в фотобиореакторе выступают концентрация растворенного кислорода, значение pH, длительность культивирования и мутность культурной среды.

Используемый набор обучающих данных содержал 11970 точек данных, полученных в ходе пяти экспериментов культивирования.

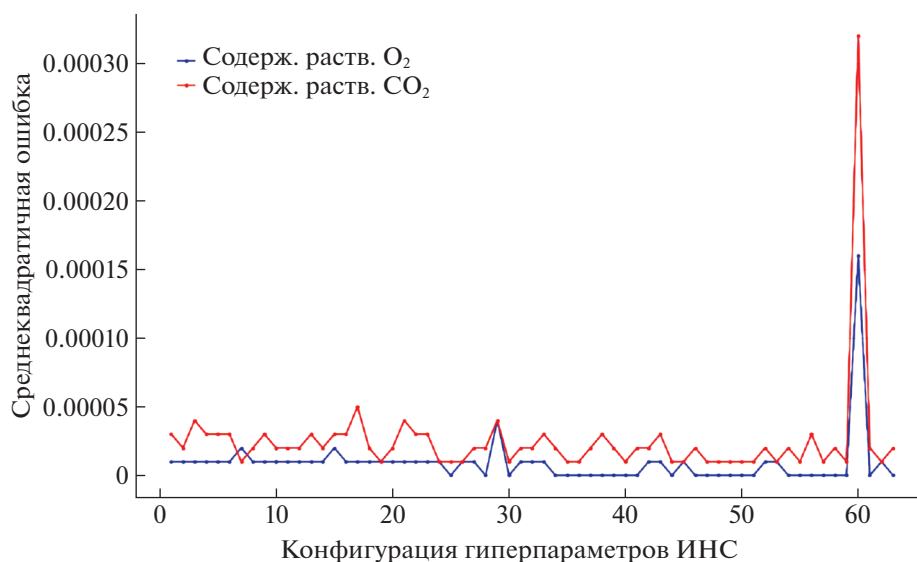


Рис. 1. Значения среднеквадратичной ошибки, полученные в процессе оптимизации гиперпараметров по сетке.

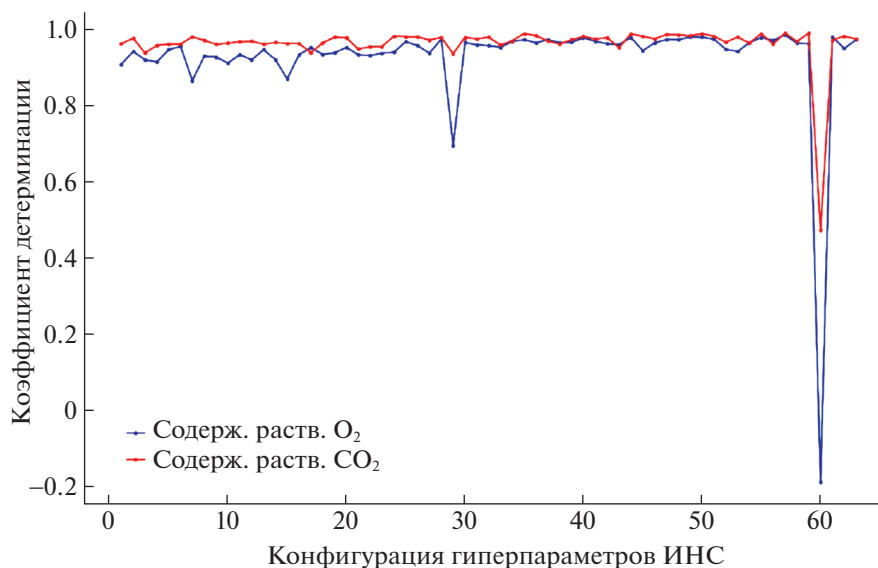


Рис. 2. Значения коэффициента детерминации, полученные в процессе оптимизации гиперпараметров по сетке.

РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

Гиперпараметры системы предсказания на основе ИНС были оптимизированы путем поиска по сетке для получения наилучшей конфигурации сети, отвечающей наилучшей точности предсказания в данных условиях. Для каждой конфигурации проводили десятикратную кросс-валидацию — имеющиеся данные разбивали на обучающий и валидационный набор в различных комбинациях, затем проводили обучение ИНС на обучающем наборе и проверку эффективности работы ИНС на валидационном наборе для раз-

личных комбинаций разбиения данных, покрывающих исключение всех точек данных из обучающего набора для проверки способности ИНС к генерализации. На иллюстрации процесса оптимизации (рис. 1, 2) показаны результаты тестирования различных конфигураций гиперпараметров ИНС. Можно заметить, что № 60 показывает anomalно плохой результат, в то время как большинство других исследуемых конфигураций показывают сравнительно высокие значения. Наилучшие результаты были получены при значениях гиперпараметров, представленных в табл. 7.

Таблица 6. Выходные параметры ИНС

Параметр	Значимость
Содержание кислорода в выходящем газе, %	Определяемый параметр
Содержание диоксида углерода в выходящем газе, %	Определяемый параметр

Таблица 7. Наилучшие значения гиперпараметров

Параметр	Значение
Количество нейронов на первом скрытом слое	6
Количество нейронов на втором скрытом слое	5
Количество нейронов на третьем скрытом слое	5
Передаточная функция 1	Сигмоидная
Передаточная функция 2	Сигмоидная
Передаточная функция 3	Сигмоидная
Размер пакета (при обучении)	250
Количество эпох обучения	5000
Коэффициент скорости обучения	0.025

Таблица 8. Результаты предсказаний для наилучшей конфигурации ИНС

Определяемый параметр	Наилучшее значение среднеквадратичной ошибки (MSE)	Наилучшее значение коэффициента детерминации (R^2)
O ₂ в выходящем газе	$6 \cdot 10^{-6}$	0.9755
CO ₂ в выходящем газе	$1 \cdot 10^{-5}$	0.97897

Таблица 9. Примеры результатов, полученных при оптимизации гиперпараметров методом случайного поиска

Среднеквадратическая ошибка		Коэффициент детерминации	
С. р. O ₂	С. р. CO ₂	С. р. O ₂	С. р. CO ₂
$2 \cdot 10^{-5}$	$3 \cdot 10^{-5}$	0.8311	0.95639
$3 \cdot 10^{-5}$	$4 \cdot 10^{-5}$	0.8013	0.93622
$6 \cdot 10^{-5}$	$1 \cdot 10^{-4}$	0.54444	0.86256
$1 \cdot 10^{-5}$	$1 \cdot 10^{-5}$	0.95139	0.94387
<u>$6 \cdot 10^{-6}$</u>	<u>$1 \cdot 10^{-5}$</u>	<u>0.9555</u>	<u>0.95897</u>
$1 \cdot 10^{-5}$	$2 \cdot 10^{-5}$	0.92087	0.94085
$5 \cdot 10^{-5}$	$4 \cdot 10^{-5}$	0.63606	0.93909
$1 \cdot 10^{-5}$	$2 \cdot 10^{-5}$	0.94692	0.94722
$4 \cdot 10^{-5}$	$9 \cdot 10^{-5}$	0.69421	0.87082
$2 \cdot 10^{-5}$	$4 \cdot 10^{-5}$	0.92987	0.94891

Тестирование обученной системы предсказания позволило достичь результатов для наилучшей конфигурации, показанных в табл. 8.

Алгоритм оптимизации гиперпараметров при помощи случайного поиска также показал спо-

собность к работе по данной задаче (табл. 9), но эффективность этого метода недостаточна из-за большого числа гиперпараметров, ввиду чего для конечной реализации выбор был сделан в пользу метода поиска по сетке (рис. 1, 2).

Таблица 10. Входные параметры для максимального количества выходящего кислорода и минимального количества диоксида углерода в газовой смеси из фотобиореактора

Входной параметр	Максимальное количества выходящего O ₂	Минимальное количества выходящего CO ₂
pH	10.67	10.45
Концентрация растворенного кислорода в среде, %	89.72	90.35
Длительность культивирования, с	1409850	749250
Мутность культуры	1.35	1.10

Таблица 11. Сравнение эталонных экспериментальных результатов и предсказаний ИНС для этих входных параметров

Определяемый параметр	max(O ₂), Эталон	min(CO ₂), Эталон	max(O ₂), ИНС	min(CO ₂), ИНС
O ₂ в выходящем газе	20.91	20.81	20.93	20.82
CO ₂ в выходящем газе	0.02	0.01	0.03	0.0149

Проверку работоспособности разработанной ИНС проводили путем сравнения полученных экспериментальных значений с максимальным количеством выходящего кислорода и минимальным количеством диоксида углерода в выходящей газовой смеси при культивировании в фотобиореакторе и предсказаний ИНС для экспериментальных точек, исключенных из обучающего набора. Полученные значения коэффициента детерминации R^2 составляли 0.9755 и 0.97897 для содержания кислорода и диоксида углерода соответственно (табл. 10, 11).

ЗАКЛЮЧЕНИЕ

Обучение искусственной нейронной сети на экспериментальных данных культивирования цианобактерий *Arthrospira platensis* позволило достичь высокой точности предсказания содержания кислорода и диоксида углерода для выходящей из фотобиореактора газовой смеси. Система показала высокую работоспособность на имеющемся наборе данных, достигая значений коэффициента детерминации R^2 , равных 0.9755 и 0.97897 для содержания кислорода и диоксида углерода соответственно, что свидетельствует о применимости разработанной ИНС для решения данной задачи.

Работа выполнена в рамках тематического плана НИЦ «Курчатовский институт» «Разработка платформенных технологий на основе фототрофных микроорганизмов для интенсивной утилизации CO₂ и производства биотоплив и других полезных продуктов».

СПИСОК ЛИТЕРАТУРЫ

1. Zhao B., Su Y., Zhang Y., Cui G. // Energy. 2015. V. 89. P. 347. <https://doi.org/10.1016/j.energy.2015.05.123>
2. Abdel-raouf N., Al-Homaidan A.A., Ibraheem I.B.M. // Saudi J. Biol. Sci. 2012. V. 19 (3). P. 257. <https://doi.org/10.1016/j.sjbs.2012.04.005>
3. Femando A.H., Mag A.B., Ubando A.T. // IEEE Region 10 Conference (TENCON). 2016. P. 3752.
4. Gorin K.V., Sergeeva Y.E., Butylin V.V. et al. // Biore-sour. Technol. 2015. V. 193. P. 178. <https://doi.org/10.1016/j.biortech.2015.06.097>
5. Menezes A.A., Cumbers J., Hogan J.A., Arkin A.P. // J. R. Soc. Interface. 2015. V. 12 (102). <https://doi.org/10.1098/rsif.2014.0715>
6. Belz S., Buchert M., Bretschneider J. et al. // Acta Astro-naut. 2014. V. 101. P. 170. <https://doi.org/10.1016/j.actaastro.2014.04.023>
7. Grobbelaar J.U., Soeder J., Groeneweg J. et al. // Water Res. 1988. V. 22 (11). P. 1459. [https://doi.org/10.1016/0043-1354\(88\)90105-4](https://doi.org/10.1016/0043-1354(88)90105-4)
8. Grobbelaar J.U. // Photosynth Res. 2010. V. 106 (1–2). P. 135. <https://doi.org/10.1007/s11120-010-9573-5>
9. García-Camacho F., López-Rosales L., Sánchez-Mirón A. et al. // Algal Res. 2016. V. 14. P. 58. <https://doi.org/10.1016/j.algal.2016.01.002>
10. Krushev S.S., Plyusnina T.Y., Antal T.K. et al. // Bio-phys. Rev. 2022. P. 821. <https://doi.org/10.1007/s12551-022-00982-2>
11. Vats S., Negi S. // Bioresour. Technol. 2013. V. 146. P. 682. <https://doi.org/10.1016/j.biortech.2013.07.141>
12. Chamilka V., Kankanamalage G., Hasara S. et al. // J. Biotechnol. 2020. V. 312. P. 44. <https://doi.org/10.1016/j.jbiotec.2020.02.010>
13. Noguchi R., Ahamed T., Saka D. et al. // Biosyst Eng. 2018. P. 1. <https://doi.org/10.1016/j.biosystemseng.2018.10.002>

14. *Franco B.M., Navas L.M., Gómez C. et al.* // *Algal Res.* 2019. V. 38. P. 101419.
<https://doi.org/10.1016/j.algal.2019.101419>
15. *Sivapathasekaran C., Mukherjee S., Ray A. et al.* // *Bioresour. Technol.* 2010. V. 101 (8). P. 2884.
<https://doi.org/10.1016/j.biortech.2009.09.093>
16. *Vats S., Negi S.* // *Bioresour. Technol.* 2013. V. 140. P. 392.
<https://doi.org/10.1016/j.biortech.2013.04.106>
17. *Pappu J.S.M., Vijayakumar G.K., Ramamurthy V.* // *Bioresour. Technol.* 2013. V. 130. P. 224.
<https://doi.org/10.1016/j.biortech.2012.12.082>
18. *Salam S., Verma T.N.* // *Energy Convers. Manag.* 2019. V. 180. P. 496.
<https://doi.org/10.1016/j.enconman.2018.11.014>
19. *Ponce E., Ponce C., Barraza B.* // *ICINCO-ICSO.* 2008. P. 289.
<https://doi.org/10.5220/0001482002890292>
20. *Susanna D., Dhanapal R., Mahalingam R., Ramamurthy V.* // *Biotechnol. Bioeng.* 2019.
<https://doi.org/10.1002/bit.27128>
21. *del Rio-Chanona E.A., Manirafasha E., Zhang D. et al.* // *Algal Res.* 2016. V. 13. P. 7.
<https://doi.org/10.1016/j.algal.2015.11.004>
22. *Ciferri O., Tiboni O.* // *Annu. Rev. Microbiol.* 1985. V. 39. P. 503.
<https://doi.org/10.1146/annurev.mi.39.100185.002443>
23. *Furmaniak M.A., Misztak A.E., Franczuk M.D., Wilmotte A.* // *Front Microbiol.* 2017. V. 8. P. 1.
<https://doi.org/10.3389/fmicb.2017.02541>
24. *Cornet J.F., Dussap C.G., Cluzel P., Dubertret G.* // *Biotechnol. Bioeng.* 1992. V. 40. P. 826.
<https://doi.org/10.1002/bit.260400710>
25. *Zhang X., Chen X., Yao L. et al.* // *International Conference on Neural Information Processing.* 2019.
<https://doi.org/10.13140/RG.2.2.16378.44481>
26. *Kokkinos Y.* // *IEEE 24th International Conference on Tools with Artificial Intelligence.* Athens: IEEE. 2012. P. 1094.
<https://doi.org/10.1109/ICTAI.2012.155>
27. *Albawi S., Mohammed T.A.* // *International Conference on Engineering and Technology (ICET).* 2017.
28. *Chakdar H., Pabbi S.* // *Frontier Discoveries and Innovations in Interdisciplinary Microbiology.* 2016. P. 45.
<https://doi.org/10.1007/978-81-322-2610-9>
29. *Mohite Y.S., Wakte P.S.* // *J. Algal Biomass. Util.* 2011. V. 2. P. 53.
30. *Dejsungkrant M., Chisti Y.* // *Bioprocess Biosyst Eng.* 2017.
<https://doi.org/10.1007/s00449-017-1778-y>
31. *Rio-chanona E.A., Zhang D., Xie Y. et al.* // *Ind. Eng. Chem. Res.* 2015. V. 54 (43).
<https://doi.org/10.1021/acs.iecr.5b03102>