

ПРИЛАДИ І МЕТОДИ КОНТРОЛЮ ТА ВИЗНАЧЕННЯ СКЛАДУ РЕЧОВИНИ

УДК 681.518.3: 535.243.22

С. М. Кватернюк

ОЦІНЮВАННЯ ДОСТОВІРНОСТІ КОНТРОЛЮ ТОКСИЧНОСТІ СТИЧНИХ ВОД МУЛЬТИСПЕКТРАЛЬНИМ МЕТОДОМ З ВИКОРИСТАННЯМ НЕЙРОМЕРЕЖІ

Вінницький національний технічний університет, м. Вінниця

Анотація. У роботі вдосконалено систему контролю індексу токсичності стічних вод за допомогою мультиспектрального методу з використанням біоіндикації по фітопланктону та нейромережі для обробки результатів. Розв'язання оберненої оптичної задачі для опосередкованого вимірювання параметрів середовища на основі мультиспектрального методу здійснено за допомогою нейромережі на основі багатшарового перцептрона з одним прихованим шаром, багатшарового перцептрона з двома прихованими шарами, мережі з радіальною базисною функцією, узагальненої регресійної нейромережі. Оцінено достовірність контролю токсичності та ймовірності помилок першого і другого роду.

Ключові слова: мультиспектральний контроль, токсичність, стічні води, біоіндикація.

Аннотация. В работе усовершенствована система контроля индекса токсичности сточных вод с помощью мультиспектрального метода с использованием биоиндикации по фитопланктону и нейросети для обработки результатов. Решение обратной оптической задачи для косвенного измерения параметров среды на основе мультиспектрального метода осуществлено с помощью нейросети на основе многослойного перцептрона с одним скрытым слоем, многослойного перцептрона с двумя скрытыми слоями, сети с радиальной базисной функцией, обобщенной регрессионной нейросети. Оценена достоверность контроля токсичности и вероятности ошибок первого и второго рода.

Ключевые слова: мультиспектральный контроль, токсичность, сточные воды, биоиндикация

Abstract. The system for monitoring the toxicity index of sewage with the help of the multispectral method with the use of bioindication on phytoplankton and neural network for processing the results is improved. The solution of the inverse optical problem for indirect measurement of the parameters of the medium on the basis of the multispectral method was carried out using neural networks based on a multilayer perceptron with one hidden layer, a multilayer perceptron with two hidden layers, a network with a radial basis function, a generalized regression neural network. The reliability of toxicity control and the probability of errors of the first and second kind are estimated.

Key words: multispectral control, toxicity, sewage, bioindication.

Вступ

Вдосконалений мультиспектральний метод контролю токсичності стічних вод включає відбір проб, визначення якісного та кількісного складу завислих частинок фітопланктону за допомогою проточного мультиспектрального телевізійного вимірювального аналізу частинок неперервної дії, порівняння отриманих даних з нормованими значеннями, а також формування мультиспектральних зображень водних середовищ з розсіювальними частинками фітопланктону на характеристичних довжинах хвиль від 300 до 1100 нм за допомогою ширококугової ПЗЗ-камери, перемикаемого вузькосмугового оптичного фільтра та джерела освітлення з рівномірним спектром випромінювання від 270 до 1200 нм, причому на основі отриманих даних за допомогою комп'ютера проводять аналіз мультиспектральних зображень та здійснюють опосередковане вимірювання концентрації частинок фітопланктону у кожному фрагменті зображення з використанням регресійного рівняння, яке пов'язує цей параметр з результатами мультиспектральних вимірювань. За рахунок використання ПЗЗ-камери з'являється можливість одночасного вимірювання концентрації частинок фітопланктону у великій кількості досліджуваних проб, які потрапляють у поле зору камери.

Актуальність

Актуальність роботи обумовлена необхідністю підвищення достовірності контролю індексу токсичності стічних вод за допомогою мультиспектрального методу з використанням біоіндикації по фітопланктону.

Мета

Метою дослідження є вдосконалення системи контролю індексу токсичності стічних вод за допомогою мультиспектрального методу з використанням біоіндикації по фітопланктону та нейромережі для обробки результатів, а також аналіз помилок першого та другого роду і достовірності контролю.

Задачі

1. Обернена задача для опосередкованого вимірювання параметрів середовища мультиспектральним методом.

2. Оцінювання достовірності контролю токсичності стічних вод мультиспектральним методом з використанням нейромережі.

1. Розв'язання оберненої задачі для опосередкованого вимірювання параметрів середовища мультиспектральним методом

Розв'язання оберненої оптичної задачі для опосередкованого вимірювання параметрів неоднорідного біологічного середовища за мультиспектральними параметрами можна провести декількома способами, зокрема, за допомогою множинної регресії, нечіткої логіки чи нейромережі.

Наприклад, опосередковане вимірювання концентрації частинок фітопланктону у кожному фрагменті зображення може здійснюватись з використанням регресійного рівняння [1], яке пов'язує цей параметр з результатами мультиспектральних вимірювань:

$$N = 1,4399 - 0,55 \cdot M_{540} + 3,225 \cdot M_{490} - 2,8 \cdot M_{440} - 4,5 \cdot M_{400} + 4,3 \cdot M_{315} + 0,4885 \cdot M_{670}, \quad (1.1)$$

де N – концентрація частинок фітопланктону, M_i – результати мультиспектральних вимірювань.

При цьому забезпечується методична похибка вимірювань 0,208%, що дозволяє отримати загальну похибку вимірювань 0,531% при використанні 6 спектральних каналів та ПЗЗ камери з розрядністю 12 біт і співвідношенням сигнал/шум 66 дБ.

Після опосередкованого вимірювання концентрації частинок фітопланктону у досліджуваній N_e і контрольній N_c пробах проводиться розрахунок індексу токсичності за формулою:

$$T = 100\% \frac{N_c - N_e}{N_c}. \quad (1.2)$$

Далі необхідно прийняти рішення про токсичність досліджуваної проби. Використовується норма індексу токсичності $T=20\%$ для контролю того, чи досліджувана проба є токсичною і норма індексу токсичності $T=50\%$ для контролю того, чи досліджувана проба є сильнотоксичною [2]:

– якщо $T < 20\%$, то досліджувана проба не є токсичною; якщо $T \geq 20\%$, то досліджувана проба токсична;
– якщо $T < 50\%$, то досліджувана проба не є сильнотоксичною; якщо $T \geq 50\%$, то досліджувана проба сильнотоксична.

Для розв'язання оберненої оптичної задачі опосередкованого вимірювання концентрації частинок фітопланктону параметрів неоднорідного біологічного середовища за мультиспектральними параметрами за допомогою нейромережі на її вхід необхідно подати результати мультиспектральних вимірювань у якості вхідних неперервних змінних, а на виході нейромережа повинна дати чисельне значення концентрації частинок фітопланктону у об'ємі досліджуваної проби. Для навчання нейромережі використаємо таблицю вихідних даних результатів мультиспектральних вимірювань та значень концентрації частинок фітопланктону визначених методом автоматизованої мікроскопії з використанням лічильної камери Гояєва (див. табл. 1).

Таблиця 1 – Результати мультиспектральних вимірювань концентрації частинок

n	M_{540}	M_{490}	M_{440}	M_{400}	M_{315}	M_{670}	N тис.кл./мл
1	11,604	18,723	20,925	19,336	23,899	15,081	19,970
2	11,814	18,081	21,724	18,856	24,671	15,226	21,094
3	11,556	18,610	21,478	19,383	24,670	14,902	21,096
4	11,951	18,626	21,560	18,856	23,843	15,808	19,964
...							
265	9,556	10,283	13,155	11,081	14,998	8,257	11,173

Для побудови нейромережі для системи контролю токсичності стічних вод мультиспектральні методом з використанням біоіндикації по фітопланктону використовуємо пакет STATISTACA 6.1 Neural Networks [3]. Перевіримо можливості застосування архітектури нейромережі на основі:

- багатошарового перцептрона з одним прихованим шаром – БШП1 (рис. 1, а);
- багатошарового перцептрона з двома прихованими шарами – БШП2 (рис. 1, б);
- мережі з радіальною базисною функцією – РБФ (рис. 1, в);
- узагальненої регресійної нейромережі – УРНМ (рис. 1, г).

Багатошаровий перцептрон (БШП) – одна з найбільш популярних типів мереж, і в багатьох класах задач він дає максимально можливу продуктивність. Багатошаровий перцептрон навчається за допомогою ітеративних алгоритмів, найбільш відомим з яких є зворотне поширення.

Нейромережа на радіальних базисних функціях (РБФ) поєднує одиночний радіальний прихований шар зі скалярним вихідним шаром. Нейрони прихованого шару діють як центри кластерів, групуючи схожі навчальні спостереження, а вихідний шар формує класифікуючу функцію регресії. Оскільки класифікуюче перетворення є нелінійним, лінійного вихідного шару достатньо, щоб створити в цілому нелінійну функцію.

Узагальнені регресійні нейронні мережі (УРНМ) будують ядерну оцінку регресійної поверхні. Вихідна змінна зазвичай числова. УРНМ має один прихований елемент на кожне навчальне спостереження. Однак, на відміну від імовірнісної нейромережі, можливо навчити УРНМ з меншим числом прихованих елементів, які є центрами тяжкості кластерів відомих даних [3].

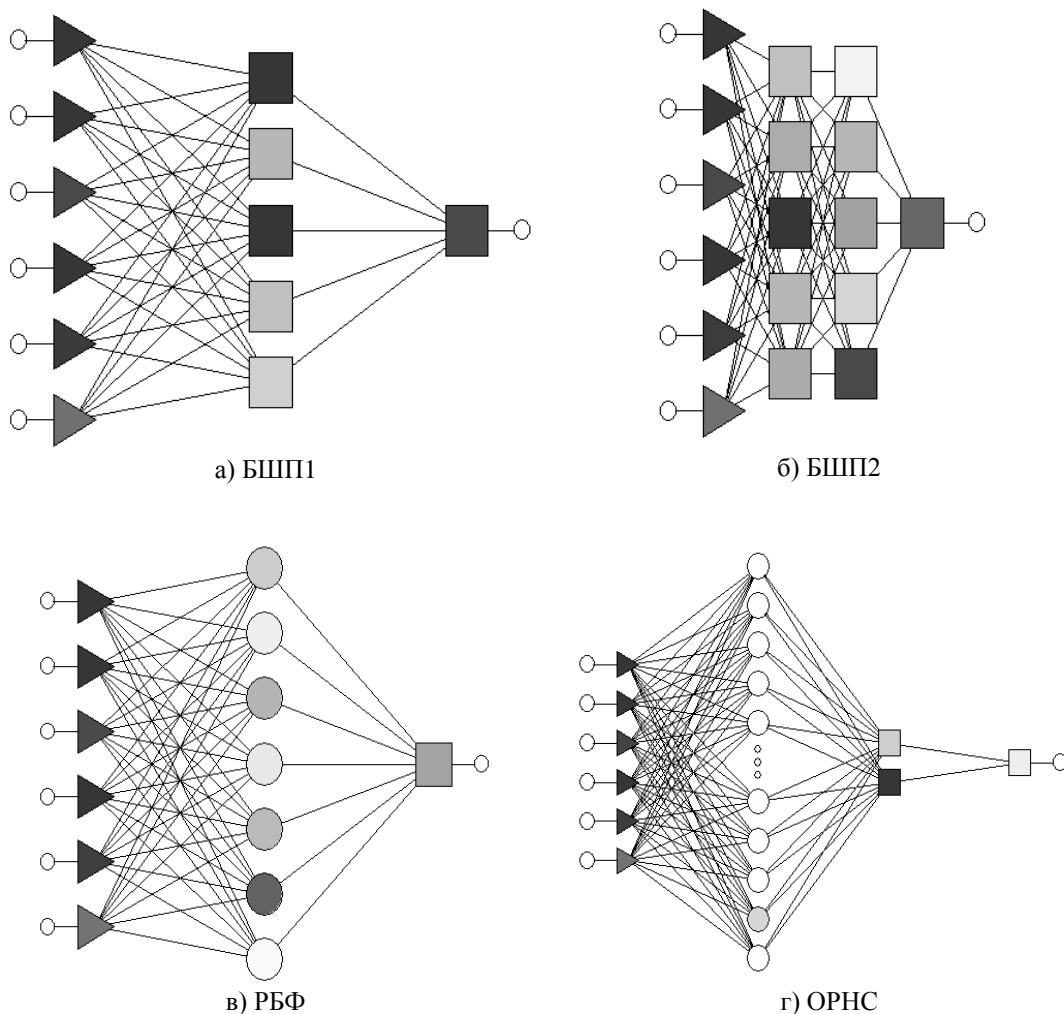


Рисунок 1 – Архітектура нейромереж для розв'язання оберненої задачі визначення параметрів неоднорідних біологічних середовищ на основі мультиспектральних параметрів

Після навчання нейромереж протягом 500 ітерацій на основі тестових послідовностей отримано такі результати продуктивності та помилок їх роботи (див. табл. 2). В подальшому використовуємо багатошаровий перцептрон, який дозволяє отримати найменшу тестову помилку.

Таблиця 2 – Результати продуктивності та помилок нейромереж

Архітектура	Продуктив. навч.	Продуктив. контр.	Продуктив. тест.	Помилка при навчанні, %	Контрольна помилка, %	Тестова помилка, %
БШП1	0,021271	0,018511	0,050634	0,005367	0,005287	0,013255
БШП2	0,022745	0,027138	0,027268	0,005506	0,005623	0,00496
РБФ	0,318117	0,282023	0,244629	0,09589	0,10716	0,073302
ОРНС	0,697079	0,662157	0,535462	0,193040	0,171507	0,141254

2. Оцінювання достовірності контролю токсичності стічних вод мультиспектральним методом з використанням нейромережі

Проаналізуємо достовірність контролю токсичності стічних вод мультиспектральним методом з використанням нейромережі використавши такий набір даних тестування: N_1 – концентрація частинок фітопланктону визначена методом автоматизованої мікроскопії з використанням лічильної камери Горяєва, N_2 – концентрація частинок фітопланктону визначена мультиспектральним методом з використанням нейромережі. Результати контролю токсичності з використанням автоматизованої мікроскопії вважаємо достовірними («золотий» стандарт). На основі отриманих значень концентрації частинок фітопланктону у контрольній і досліджуваній пробах за формулою (1.2) розраховуємо індекси токсичності, відповідно T_1, T_2 .

У результаті контролю токсичності необхідно визначити чи перевищує розраховане значення токсичності норму T_{norm} . У цьому випадку можливі такі варіанти:

$$\begin{cases} P_1 = 1 & \text{якщо } T_1 \geq T_{norm}, \text{ інакше } P_1 = 0, \\ P_2 = 1 & \text{якщо } T_1 < T_{norm}, \text{ інакше } P_2 = 0, \\ P_3 = 1 & \text{якщо } T_2 \geq T_{norm}, \text{ інакше } P_3 = 0, \\ P_4 = 1 & \text{якщо } T_2 < T_{norm}, \text{ інакше } P_4 = 0, \end{cases} \quad (2.1)$$

де P_1, P_2 – змінні множинного типу $\{0;1\}$, які відповідають дійсному значенню індексу токсичності визначеному за допомогою автоматизованої мікроскопії T_1 ; P_3, P_4 – змінні множинного типу $\{0;1\}$, які відповідають значенню індексу токсичності визначеному за допомогою нейромережі.

При цьому можливо розрахувати змінні A, B, C, D , які відповідають усім можливим подіям у процесі контролю:

$$\begin{cases} A = 1 & \text{якщо } ((P_1 = 1) \wedge (P_3 = 1)), \text{ інакше } A = 0, \\ B = 1 & \text{якщо } ((P_2 = 1) \wedge (P_3 = 1)), \text{ інакше } B = 0, \\ C = 1 & \text{якщо } ((P_1 = 1) \wedge (P_4 = 1)), \text{ інакше } C = 0, \\ D = 1 & \text{якщо } ((P_2 = 1) \wedge (P_4 = 1)), \text{ інакше } D = 0, \end{cases} \quad (2.2)$$

де A, B, C, D – змінні множинного типу $\{0;1\}$.

A – відповідає події, коли у процесі контролю отримано істинно-позитивний результат (у результаті контролю отримано позитивний результат – досліджувана проба токсична і дійсне значення теж було позитивним – проба токсична),

B – відповідає події, коли у процесі контролю отримано хибно-позитивний результат (у результаті контролю отримано позитивний результат – досліджувана проба токсична і дійсне значення ознаки було негативним – проба не токсична),

C – відповідає події, коли у процесі контролю отримано хибно-негативний результат (у результаті контролю отримано негативний результат – досліджувана проба не токсична і дійсне значення ознаки було позитивним – проба токсична),

D – відповідає події, коли у процесі контролю отримано істинно-негативний результат (у результаті контролю отримано негативний результат – досліджувана проба не токсична і дійсне значення ознаки теж було негативним – проба не токсична).

Розраховуємо змінні P_1, P_2, P_3, P_4 та, відповідно, A, B, C, D за формулами (2.1), (2.2) для набору даних тестування. Результати аналізу подій для контролю токсичності показано у табл. 3 для опрацювання результатів мультиспектральних вимірювань з використанням нейромережі.

Таблиця 3 – Аналіз подій при контролі токсичності при опрацюванні результатів мультиспектральних вимірювань з використанням нейромережі

N	T_1	T_2	Дійсне значення		Результати контролю		A	B	C	D
			$P_1(+)$	$P_2(-)$	$P_3(+)$	$P_4(-)$				
1	6,781	6,649	0	1	0	1	0	0	0	1
2	1,533	1,475	0	1	0	1	0	0	0	1
3	1,524	1,502	0	1	0	1	0	0	0	1
4	6,812	6,903	0	1	0	1	0	0	0	1
4	13,072	13,298	0	1	0	1	0	0	0	1
6	8,871	8,970	0	1	0	1	0	0	0	1
...										
93	47,846	48,853	1	0	1	0	1	0	0	0

Відповідно кількість подій кожного типу для контролю токсичності на основі мультиспектральних вимірювань буде $a = \sum_{i=0}^n A_i$, $b = \sum_{i=0}^n B_i$, $c = \sum_{i=0}^n C_i$, $d = \sum_{i=0}^n D_i$. Отримані значення a, b, c, d для порівняння двох груп за бінарною ознакою показано у табл. 4.

Таблиця 4 – Порівняння результатів контролю та дійсного значення при опрацюванні результатів мультиспектральних вимірювань з використанням нейромережі

Результати контролю	Дійсне значення ознаки («золотий» стандарт)		Разом
	Позитивне значення (+)	Негативне значення (-)	
Позитивний результат контролю (+)	істинно-позитивний результат (true positive (TP)) $a = 124$	хибно-позитивний результат (false positive (FP)) $b = 3$	$a + b = 127$
Негативний результат контролю (-)	хибно-негативний результат (false negative (FN)) $c = 2$	істинно-негативний результат (true negative (TN)) $d = 136$	$c + d = 138$
Разом	$a + c = 126$	$b + d = 139$	$a + b + c + d = 265$

На основі порівняльної таблиці визначаємо достовірність контролю токсичності при опрацюванні результатів мультиспектральних вимірювань з використанням нейромережі [4]:

Ймовірність помилки першого роду (α errors, type I error probability):

$$\alpha = b / (b + d) = 3 / (3 + 136) = 0,022. \quad (2.3)$$

Ймовірність помилки другого роду (β errors, type II errors probability):

$$\beta = c / (a + c) = 2 / (124 + 2) = 0,016. \quad (2.4)$$

Достовірність:

$$D = 1 - \alpha - \beta = 1 - 0,022 - 0,016 = 0,962 \quad (2.5)$$

Висновки

У роботі вдосконалено систему контролю індексу токсичності стічних вод за допомогою мультиспектрального методу з використанням біондикації по фітопланктону та нейромережі для обробки результатів. Перевірено можливості застосування архітектури нейромережі на основі багатошарового персептрона з одним прихованим шаром, багатошарового персептрона з двома прихованими шарами, мережі з радіальною базисною функцією, узагальненої регресійної нейромережі. Після навчання нейромереж протягом 500 ітерацій на основі тестових послідовностей досліджено результати продуктивності та помилок їх роботи. Найменшу тестову помилку забезпечила нейромережа на основі багатошарового персептрона з двома прихованими шарами. Оцінено достовірність контролю токсичності стічних вод мультиспектральним методом з використанням нейромережі та значення ймовірності помилки першого роду 0,022, ймовірність помилки другого роду 0,016, достовірність контролю токсичності 0,962, що достатні для використання розробленого засобу контролю у спеціалізованих лабораторіях природоохоронних закладів.

Список літератури

1. Кватернюк С. М. Метод та засоби мультиспектрального телевізійного вимірювального контролю стану неоднорідних біологічних середовищ / С. М. Кватернюк // Вісник Вінницького політехнічного інституту. – 2017. – № 1. – С. 15–22.
 2. Балтиев Ю.С. Методические указания по интегральной оценке качества окружающей среды (экологическая разведка местности) / Ю. С. Балтиев, Г. П. Усов. – Москва.- Военное издательств-во. – 2005. – 119 с.
 3. Нейронные сети. STATISTACA Neural Networks: Методология и технологии современного анализа данных / Под редакцией В.П. Боровикова. – 2-е изд., перераб. и доп. – М.: Горячая линия. – Телеком, 2008. – 392 с.
 4. ISO 3534-1:2006 Statistics – Vocabulary and symbols – Part 1: General statistical terms and terms used in probability. Committee ISO/TC 69/SC 1, Edition: 1, Publication date: 2010-10-25. – 105 p.
- Стаття надійшла: 20.11.2017.

Відомості про авторів

Кватернюк Сергій Михайлович – к.т.н., доцент, доцент кафедри екології і екологічної безпеки; м. Вінниця, Вінницький національний технічний університет, Хмельницьке шосе 95.

ДО ВІДОМА АВТОРІВ

Найновіші правила оформлення і подання статей знаходяться на сайті журналу <http://itce.vntu.edu.ua/>