

УДК 519.237.5

ДОСЛІДЖЕННЯ ІНДУКТИВНИМИ МЕТОДАМИ ВПЛИВУ РІЗНИХ ФАКТОРІВ НА ВМІСТ ХЛОРОФІЛУ *a* У ФІТОПЛАНКТОНІ КАХОВСЬКОГО ВОДОСХОВИЩА

Г.А. Піднебесна¹, А.В. Курейшевич²,

¹ *Міжнародний науково-навчальний центр інформаційних технологій
та систем НАН та МОН України*

² *Інститут гідробіології НАН України,*

pidnebesna@ukr.net, alischuk@rambler.ru

В роботі досліджується вплив різних природних і антропогенних чинників на вміст хлорофілу *a* – основного фотосинтезуючого пігмента фітопланктону, який є показником рівня його вегетації і величини первинної продукції. Дослідження проведено за допомогою множинного регресійного аналізу даних Каховського водосховища та з застосуванням МГУА.

Ключові слова: вміст хлорофілу, фітопланктон, регресійний аналіз, Каховське водосховище, МГУА

The paper studies the impact of various natural and anthropogenic factors on the content of chlorophyll *a* being the main photosynthetic pigment of phytoplankton, which is a measure of its vegetation and volume of primary production. The study was conducted using multiple regression analysis of Kakhovka reservoir data and GMDH.

Keywords: chlorophyll content, phytoplankton, regression analysis, Kakhovka reservoir, GMDH

В работе исследуется влияние различных природных и антропогенных факторов на содержание хлорофилла *a* – основного фотосинтезирующего пигмента фитопланктона, который является показателем уровня его вегетации и величины первичной продукции. Исследование проведено с помощью множественного регрессионного анализа данных Каховского водохранилища, а также с использованием МГУА.

Ключевые слова: содержание хлорофилла, фитопланктон, регрессионный анализ, Каховское водохранилище, МГУА

Вступ.

Водні ресурси мають винятково важливе значення для суспільства. Натомість забезпечення потреб населення і народного господарства водою можливі лише при збереженні та покращенні якості води, раціональному використанні її підприємствами всіх галузей та відтворенні водних ресурсів. Тому вивчення закономірностей функціонування водних екосистем і формування їх біологічної продуктивності та якості води залишається наразі вельми актуальним.

Як відомо, первинною автотрофною ланкою водойм є фітопланктон (водорості різних систематичних груп, що населяють товщу води) Асимілюючи

сонячну радіацію та перетворюючи її у органічну речовину в процесі фотосинтезу, він створює первинну продукцію, за рахунок якої через ланцюг живлення існують інші гідробіоти (організми, що живуть у воді). Однак підвищення рівня первинної продукції водойм внаслідок евтрофування – збільшення у воді концентрації біогенних елементів, насамперед азоту і фосфору з часом призводить до негативних наслідків. Проявом їх є «цвітіння» води синьозеленими водоростями, накопичення в ній органічних речовин та зменшення вмісту розчиненого кисню, що призводить до замору риби, заростання прибережної зони вищими водними рослинами, замулення водойм.

Одним з напрямків досліджень сучасної екології є необхідність з'ясування відгуку фітопланктону на зміни навколишнього середовища. Дослідження фундаментальних основ формування фітопланктону і глобального впливу природних чинників на цей процес в умовах антропогенного навантаження важливі для розробки теорії біологічної продуктивності водойм, науково обґрунтованих методів управління функціонуванням екосистем і прогнозування якості води.

Для водних ресурсів характерна сильна мінливість режиму, починаючи від добових і закінчуючи віковими коливаннями водо насиченості кожного джерела. Складна взаємодія великої кількості факторів додає коливанням характер випадкового процесу. Тому розрахунки, що відносяться до вивчення закономірностей функціонування водних ресурсів, носять імовірнісний, статистичний характер. Тож у дослідженні екологічних процесів дніпровських водосховищ, зокрема Каховського, було використано методи регресійного аналізу та методу групового урахування аргументів. Було проведено моделювання за статистичними даними багаторічних спостережень.

Постановка задачі

Реакцію планктонних водоростей на дію природних і антропогенних чинників можна досліджувати за вмістом хлорофілу a – основного фотосинтезуючого пігмента фітопланктону, оскільки він є показником рівня його вегетації та первинної продукції. За концентрацією хлорофілу a в одиниці об'єму води оцінюють також ступінь трофності водойм, їх екологічний стан і біомасу планктонних водоростей. Вміст хлорофілу a фітопланктону в одиниці об'єму води залежить від багатьох чинників: гідрологічних та гідрометеорологічних умов, світлового та температурного режимів, сезону року, часу доби, видового складу водоростей, їх забезпеченості біогенними елементами, фізіологічного стану популяцій тощо. [1, 2]

Мета роботи: для дослідження екологічних процесів дніпровських водосховищ за статистичними даними багаторічних спостережень побудувати математичні моделі залежності концентрації хлорофілу a від різних чинників для Каховського водосховища.

Опис дослідження

В основі досліджень залежності концентрації хлорофілу *a* у фітопланктоні від низки різних факторів лежить індуктивний підхід до моделювання і прогнозування, спрямований на максимальне видобування інформації, прихованої в даних. Одним з найефективніших методів, заснованих на індуктивному підході, є метод групового урахування аргументів (МГУА), автором якого є академік О.Г. Івахненко [3]. Метод базується на принципах самоорганізації моделей та автоматичному відборі кращих з них за зовнішнім критерієм.

Для побудови моделей було використано два підходи: багатфакторний регресійний аналіз та комбінаторний алгоритм МГУА. Виконано аналіз та порівняння результатів.

Моделювання проводилось за даними багаторічних спостережень (таблиця 1) середніх величин вмісту хлорофілу *a* фітопланктону в одиниці об'єму води.

До розгляду брались наступні фактори:

x_1 – сумарний вміст розчиненого неорганічного азоту, мг N / л
(фактор 1);

x_2 – вміст розчиненого неорганічного фосфору, мг P / л (фактор 2);

x_3 – N/P – відношення вмісту азоту до вмісту фосфору, відносні одиниці
(фактор 3);

x_4 – температура води, t°C (фактор 4);

x_5 – об'єм стоку води, м³×10⁹ / місяць (фактор 5);

x_6 – сумарна сонячна радіація, МДж / м²×місяць (фактор 6).

Таблиця даних багаторічних спостережень Каховського водосховища

N	Рік (серпень місяць)	Вміст хлоро- філу <i>a</i> , мкг/л	Міне- раль- на форма N, мг/л	Міне- раль- на форма P, мг/л	N/P	Темпе- ратура води <i>t</i> , °C	Стік, м ³ ×10 ⁹ / місяць	Освітле- ність, МДж/м ² × місяць
		Y	x1	x2	x3	x4	x5	x6
1	1980	58,1	3,224	0,065	49,60	19,4	5,08	552,7
2	1981	27,4	1,2312	0,076	16,20	21,1	3,20	703,8
3	1982	16,3	1,3108	0,058	22,60	24,0	4,57	620,7
4	1983	41,4	1,4348	0,068	21,10	21,0	2,54	606,8
5	1984	14,1	0,617	0,095	6,49	21,4	3,39	673,3
6	1985	47,8	1,065	0,036	29,58	23,5	4,62	645,1
7	1986	29,1	1,156	0,027	42,81	23,4	4,15	497,9
8	1987	17,7	0,081	0,026	3,12	20,8	2,83	724,3
9	1988	19,5	0,291	0,094	3,10	23,4	3,57	653,0
10	1989	11,09	0,406	0,026	15,62	22,7	4,26	753,0
11	1990	9,8	0,369	0,030	12,30	21,0	2,95	750,0
12	1991	17,9	0,578	0,123	4,70	21,0	3,36	693,0
13	1993	14,3	0,438	0,111	3,95	22,9	2,84	625,0

Моделювання за допомогою регресійного аналізу

У процесі першого етапу дослідження було побудовано регресійні моделі з різною кількістю факторів (від 1 до 6) за статистичними даними за 1980–1993 роки і розраховано для них коефіцієнти кореляції та детермінації [4]. З кожної групи моделей з однаковою кількістю факторів для подальшого аналізу вибиралась модель з найбільшим значенням коефіцієнта множинної кореляції R . Якщо значення R наближається до 1, це означає, що побудована модель пояснює майже всю мінливість відповідних змінних. Адекватними вважаються моделі, коефіцієнт детермінації яких більше ніж 0,5.

В таблиці 2 представлені отримані моделі з найбільшими значенням коефіцієнта множинної кореляції. Значення R у проміжку [0,7 – 0,89] говорить про те, що ці моделі мають сильний зв'язок між незалежними (вхідними) та залежним факторами.

Таблиця 2

Моделі, отримані за допомогою регресійного аналізу

Структура	Модель	R
$y=f(x_1, x_3)$	$y=9,86+11,6*x_1+0,24*x_3$	0,831
$y=f(x_1, x_3, x_5)$	$y=18,69+11,47*x_1+0,36*x_3-0,01*x_5$	0,838
$y=f(x_1, x_3, x_5, x_6)$	$y=23,05+11,41*x_1+0,33*x_3-2,84*x_5-0,007*x_6$	0,839
$y=f(x_1, x_3, x_4, x_5, x_6)$	$y=27,01+11,14*x_1+0,33*x_3-0,14*x_4-2,66*x_5-0,008*x_6$	0,839
$y=f(x_1, x_2, x_3, x_5, x_6)$	$y=25,08+11,72*x_1-7,33*x_2+0,3*x_3-2,77*x_5-0,009*x_6$	0,839
$y=f(x_1, x_2, x_3, x_4, x_5, x_6)$	$y=29,53+11,45*x_1-7,88*x_2+0,29*x_3-0,16*x_4-2,58*x_5-0,01*x_6$	0,839

З таблиці 2 видно, що чотири моделі мають однакові значення коефіцієнта множинної кореляції $R=0,839$. Це означає, що абсолютно достатнім є використання моделі з найменшою кількістю чинників (чотири), тобто подальше збільшення кількості враховуваних факторів буде означати зайве переускладнення моделі.

Побудова моделей за допомогою методу групового урахування аргументів (МГУА)

В другому етапі дослідження для побудови моделей було застосовано комбінаторний алгоритм МГУА. МГУА – оригінальний метод побудови моделей за експериментальними даними в умовах невизначеності. Метод ґрунтується на різних процедурах перебору множини моделей-кандидатів за зовнішніми критеріями селекції і є ефективним засобом автоматичного розв’язання задач структурно-параметричної ідентифікації. Отримані за цим методом моделі оптимальної складності враховують невідомі закономірності функціонування досліджуваного об’єкта, інформація про які неявно міститься у вибірці даних [5, 6].

Задача індуктивного моделювання зводиться до формування за вибіркою експериментальних даних деякої множини моделей-кандидатів F та пошуку оптимальної моделі з цієї множини за заданим критерієм селекції CR :

$$f^* = \arg \min_{f \in \mathfrak{F}} CR(y, f(X, \hat{\theta}_f)), \quad (1)$$

де оцінки параметрів $\hat{\theta}_f$ для кожної $f \in \mathfrak{Z}$ є розв'язком задачі виду:

$$\hat{\theta}_f = \arg \min_{\theta_f \in R^{s_f}} Q(y, X, \theta_f), \quad (2)$$

де $Q(\cdot) \neq CR(\cdot)$ – критерій якості розв'язання задачі параметричної ідентифікації кожної окремої моделі, що генерується в процесі структурної ідентифікації.

Основний принцип методу полягає у знаходженні мінімуму зовнішнього критерію при послідовному ускладненні моделі. Знайдений мінімум критерію вказує на модель оптимальної складності. Під складністю моделі розуміють кількість оцінюваних параметрів або/та степінь нелінійності.

Для пошуку оптимальної моделі використовуються зовнішні критерії, які базуються на розбитті вибірки на дві й більш частин. Оцінювання параметрів і перевірка якості моделей виконується на різних підвибірках. Поділ вибірки дозволяє неявно (автоматично) врахувати різні види апріорної невизначеності при побудові моделі.

Якість моделі визначається за допомогою критерію регулярності $AR(s)$ (3), який будується на розбитті вибірки вхідних даних X на дві підвибірки: навчальну X_A та перевірочну X_B . На навчальній підвибірці X_A оцінюємо параметри моделі, а на перевірочній підвибірці X_B обчислюємо помилку:

$$AR(s) = \|y_B - X_B \theta_A\|^2, \quad (3)$$

де θ_A – вектор параметрів, визначений за МНК на підвибірці X_A .

Якість моделі вважається найкращою s^* , коли критерій $AR(s)$ буде найменшим, тобто:

$$s^* = \arg \min_{s=1,m} AR(s), \quad (4)$$

При використанні комбінаторного алгоритму виконується перебір всіх можливих моделей з вибором найкращої з них за критерієм селекції (4).

Для оцінювання коефіцієнтів моделей [7] було розглянуто три варіанти поділу вибірки на навчальну X_A та перевірочну X_B :

- 7/6 – навчальна вибірка X_A містить 7 точок даних за 1980-1986 роки (таблиця 1), перевірочна вибірка X_B містить 6 точок даних з 1987 по 1993 роки,
- 8/5, – X_A містить 8 точок даних за 1980-1987 роки, X_B – 5 точок даних з 1988 по 1993 роки,

- 9/4 – X_A , містить 9 точок даних за 1980-1988 роки, X_B – 4 точок даних з 1989 по 1993 роки.

Точка даних – це всі дані вимірювань за серпень місяць конкретного року (рядок в таблиці 1).

Вибір кращих моделей при моделюванні залежності вмісту хлорофілу *a* у фітопланктоні Каховського водосховища відбувався за кращим значенням критерію регулярності *AR* для кожного з варіантів поділу вибірки даних. В результаті розрахунків було обрано такі три кращих моделі (таблиця 3):

Таблиця 3.

Моделі, отримані за комбінаторним алгоритмом МГУА

Структура	Модель	Коефіцієнт множинної кореляції <i>R</i>	Поділ вибірки X_A / X_B
$y=f(x_1)$	$y_1 = 21.59x_1$	0,68	9/4
$y=f(x_1, x_2)$	$y_2 = 18.23x_1 + 96.25x_2$	0,74	8/5
$y=f(x_1, x_4, x_5, x_6)$	$y_3 = 16.1x_1 + 0.81x_4 - 1.56x_5 - 0.002x_6$	0,82	7/6

Значення коефіцієнту множинної кореляції говорить про сильний зв'язок між вхідними факторами та незалежною змінною, тобто моделі є адекватними.

В результаті моделювання за комбінаторним алгоритмом МГУА вибрано модель структури $y = f(x_1, x_4, x_5, x_6)$, в яку входять такі фактори:

x_1 – сумарний вміст розчиненого неорганічного азоту, мг N / л (фактор 1);

x_4 – температура води, t°C (фактор 4);

x_5 – об'єм стоку води, $m^3 \times 10^9$ / місяць (фактор 5);

x_6 – сумарна сонячна радіація, МДж / $m^2 \times$ місяць (фактор 6).

Фактори x_5 (об'єм стоку води) та x_6 (сумарна сонячна радіація) входять до моделі зі знаком «мінус», що свідчить про зворотний зв'язок між вмістом хлорофілу *a* фітопланктону у одиниці об'єму води та цими чинниками, що узгоджується з отриманими нами раніше результатами досліджень (2, 4).

На рисунку 1 показаний графік, на якому зображені реальні дані (ряд 1) та отримані кращі регресійна модель (ряд 2) та модель, побудована за комбінаторним алгоритмом МГУА (ряд 3). На графіку видно, що отримані моделі є близькими за значеннями вмісту хлорофілу *a*.

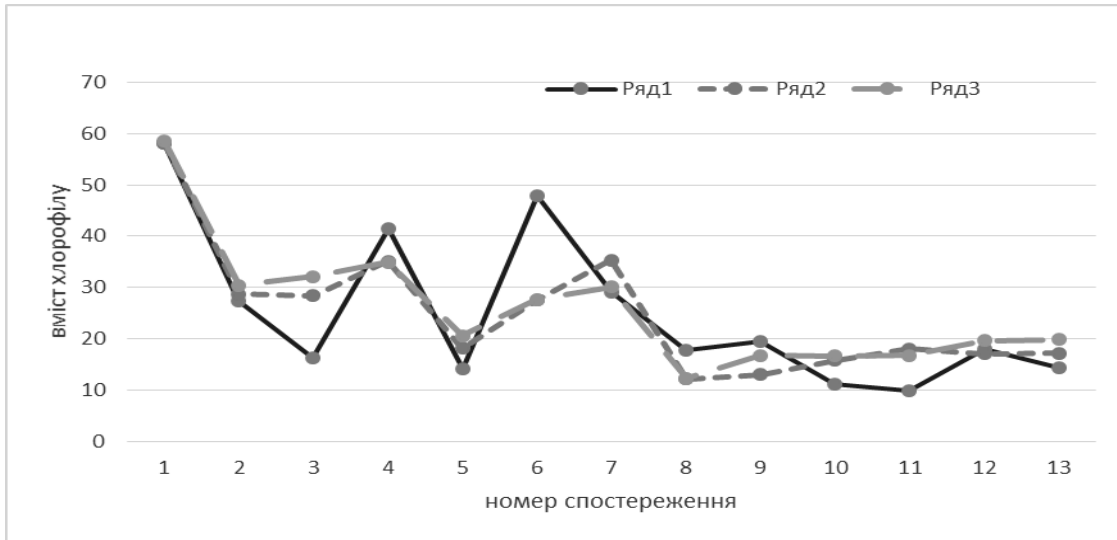


Рис. 1 Порівняння реальних та модельних значень:
 ряд 1 – реальні дані, ряд 2 – регресійна модель,
 ряд 3 – модель за комбінаторним алгоритмом МГУА

Отримані двома різними методами моделі залежності вмісту хлорофілу *a* у фітопланктоні від різних факторів виявились близькими за значеннями та структурою. Вони містять три однакових фактори: концентрація азоту у воді x_1 , стік води через турбіни ГЕС x_5 та сумарна сонячна радіація x_6 . Отримані моделі відрізняються тільки присутністю фактора x_3 (відношення вмісту азоту до вмісту фосфору) в регресійній моделі та фактора x_4 (температура води) в моделі МГУА. Враховуючи, що фактор x_1 та x_3 відображають вміст азоту, а температура води x_4 є важливим чинником впливу на вміст хлорофілу, можна вважати, що модель, побудована за МГУА, є більш адекватною.

Висновки

Аналіз отриманих результатів показав, що в усі без винятку відібрані моделі обов'язково входить змінна x_1 – мінеральна форма азоту. Отже, найважливішим чинником, який впливав на вміст хлорофілу *a* у фітопланктоні Каховського водосховища в літній сезон (у серпні, що характеризується інтенсивним «цвітінням» води синьозеленими водоростями), є саме x_1 . Це закономірно, враховуючи, що при евтрофуванні водойм вміст фосфору у воді зростає швидше, ніж азоту, і він (азот) стає тим чинником, що лімітує розвиток фітопланктону, що спостерігалось у водосховищах Дніпра (і найбільш чітко у Каховському, яке замикає каскад дніпровських водосховищ) у 1980–1993 рр. Також до складу кращих отриманих моделей входять показники об'єму стоку води та сумарної сонячної радіації. Але між вмістом хлорофілу *a* фітопланктону в одиниці об'єму води та цими чинниками спостерігається зворотна залежність.

Результати дослідження показали, що для Каховського водосховища абсолютно достатнім є використання моделі з чотирма чинниками. Моделі, отримані за двома порівнюваними методами, відрізняються тільки присутністю факторів x_3 в регресійній моделі та x_4 в моделі МГУА.

Схожість отриманих результатів дозволяє зробити висновок про ефективність застосування методу групового урахування аргументів в порівнянні з регресійним аналізом, оскільки при цьому немає потреби вручну перебирати варіанти моделей і оцінювати їх адекватність – МГУА виконує перебір і порівняння ефективності моделей автоматично.

Сказане стосується саме Каховського водосховища, наявної вибірки статистичних даних та літнього сезону, що характеризується інтенсивним «цвітінням» води синьозеленими водоростями.

Література

1. Ліщук А.В. Еколого-фізіологічні основи формування фітопланктону прісноводних екосистем// Автореф. дис. докт. біол. наук: 03.00.17 / Ін-т гідробіології НАН України. – Київ, 2007. – 38 с.
2. Курейшевич А.В. Зависимость многолетней динамики содержания хлорофилла *a* в планктоне днепровских водохранилищ от суммарной солнечной радиации и активности Солнца - Гидробиол. журн. –2004. – Т. 40, № 3. – С.16–29.
3. Ивахненко А.Г. Метод группового учета аргументов - конкурент методу стохастической аппроксимации // Автоматика. - 1968. - № 3. - С. 58–72.
4. Курейшевич А.В., Піднебесна Г.А., Степашко В.С. Аналіз залежності вмісту хлорофілу *a* у фітопланктоні Каховського водосховища від сукупної дії факторів засобами множинної регресії. // Індуктивне моделювання складних систем. Зб. наук. праць. - К. : МННЦІТС, 2015, вип. 7. – С. 147–153.
5. Степашко В.С., Єфіменко С.М., Савченко Є.А. Комп'ютерний експеримент в індуктивному моделюванні. – Київ: Наукова думка, 2014. – 222 с.
6. <http://www.mgua.irtc.org.ua/ukr/index.php?page=gmdh>.
7. Піднебесна Г.А., Курейшевич А.В. Застосування МГУА для моделювання впливу різних факторів на вміст хлорофілу *a* у фітопланктоні Каховського водосховища. // Інтелектуальні системи прийняття рішень і проблеми обчислювального інтелекту. Матеріали міжнародної наукової конференції ISDMCI-2016. – Херсон: видавництво ПП Вишемирський В.С., 2016. – С.303–305.