МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Київський національний університет будівництва і архітектури

МОДЕЛЮВАННЯ ТА ПРОГНОЗУВАННЯ СТАНУ ДОВКІЛЛЯ

Методичні вказівки
до виконання практичних робіт
для здобувачів першого (бакалаврського)
рівня вищої освіти за спеціальностями
101«Екологія», 183 «Технології захисту
навколишнього середовища»

Київ 2025

УДК 504.064.2.001.18:519.711
М74

Укладачі: О. А. Котовенко, канд. техн. наук, доцент
О. Ю. Мірошниченко, старш. викладач

Рецензент Л.О. Василенко, канд. техн. наук, доцент

Відповідальний за випуск Т.М. Ткаченко, д-р техн. наук, професор

Затверджено на засіданні кафедри охорони праці і навколишнього середовища, протокол № 10 від 04 березня 2025 року

В авторській редакції

Моделювання і прогнозування стану довкілля [Електронний]:
М74 методичні вказівки до виконання практичних робіт/ уклад.: О. А. Котовенко, О. Ю. Мірошниченко. – Київ : КНУБА, 2025. – 35 с.

Містять методичні вказівки до виконання практичних робіт та необхідний для їх виконання стислий теоретичний матеріал з курсу «Моделювання та прогнозування стану довкілля».

Призначені для здобувачів першого (бакалаврського) рівня вищої освіти за спеціальностями 101«Екологія», 183 «Технології захисту навколишнього середовища».

ЗМІСТ

ЗАГАЛЬНІ ПОЛОЖЕННЯ

В наш час методи математичного моделювання застосовуються в різних областях науки, в тому числі і в екології. Сьогодні неможливо вирішувати складні проблеми, що виникають при дослідженні стану біосфери та її компонентів, а також процесів, що проходять в цій складній системі, без сучасних методів інформатики, системного аналізу, методів фізичного і математичного моделювання і на їх основі прогнозування станів цих складних систем.

Методичні вказівки призначені для надання допомоги студентам-екологам при виконанні практичних робіт з курсу «Моделювання і прогнозування стану довкілля». Це один з найважливіших курсів для спеціаліста-еколога, який визначає рівень його професійної підготовки і дає йому потужний інструмент як в дослідницькій (науковій), так і в практичній діяльності.

Метою практичних робіт є одержання студентами практичних навичок застосування методів математичного моделювання та прогнозування в екологічних дослідженнях та охороні навколишнього середовища.

ЧАСТИНА І. Моделювання стану довкілля у випадку детермінованої базової інформації

ПРАКТИЧНА РОБОТА № 1

ЕКСПЕРТНА ОЦІНКА, НА ОСНОВІ ІМІТАЦІЙНОГО МОДЕЛЮВАННЯ, ЕКОЛОГІЧНОЇ ПРИДАТНОСТІ ДІЛЯНКИ ДЛЯ БУДІВНИЦТВА

МЕТА РОБОТИ : Ознайомлення студентів з принципом імітаційного моделювання з застосуванням комп'ютерної техніки для вирішення екологічних задач у будівництві.

ЗАВДАННЯ ДО РОБОТИ: За даними, наданими викладачем, провести екологічну оцінку будівельного майданчика, використовуючи програму RANK-KOR.

ТЕОРЕТИЧНІ ВІДОМОСТІ

Імітаційна модель експертної оцінки основа на упорядкуванні альтернатив у відповідності із зменшенням чи збільшенням будь-якої кількісно вимірюваної ознаки. Якщо запропоновано впорядкувати n альтернативних варіантів $A_j, j = 1, n$, кожен з яких має K ознак $f_q, q = 1, k.$ кожен з m експертів ранжує альтернативні варіанти, присвоюючи за кожною ознакою ранг x_i. В результаті буде отримана матриця X розміром $K \times n$, яка вміщує сумарне ранжування за всіма ознаками (показниками). Суми у рядках отриманої таблиці завжди рівні, оскільки вміщують натуральний ряд чисел (ранги), що розташовані у довільному порядку. Середнє натуральне число дорівнює $(n+1)/2$ і тоді середнє для всієї таблиці $a = \dfrac{m(n+1)}{2}$. За даними таблиці розраховується коефіцієнт узгодження W, що характеризує узгодженість поглядів експертів. Повна узгодженість поглядів експертів існує в тому випадку, якщо $W = 1$, а повна її відсутність при $W = 0$. Ступінь узгодженості визначається шляхом порівняння критерію $\chi^2{}_{\text{розр}} = m(n-1)W$ і $\chi^2{}_{\text{теор}}$ для кількості ступенів свободи $(n-1)$. Якщо розрахункова величина більше теоретичної, то погляди експертів узгоджені. В тому випадку, якщо якісні показники мають різну цінність, вони також узгоджуються експертами.

Присвоєння вагових коефіцієнтів для кожного сумарного рангу якісних показників здійснюється за таким правилом: якщо позначити вагу найбільш важливого показника як V_s, а найменш важливого V_0, то можна присвоїти іншим якісним показникам вагу V_q, пропорційно їх сумарному рангу:

$$V_q = V_0 + \frac{y_q - y_0}{(y_s - y_0)(V_s - V_0)},$$

де y_0 – сумарний ранг найменш важливого показника, y_s – сумарний ранг найбільш важливого показника. Сумарне зважене ранжування альтернативних варіантів визначається із матриці $V \times X$. А оптимальним буде елемент найменший за величиною у одержаному після множення векторі рядку [1,2].

Контрольні запитання

1. Який метод закладено в основу імітаційної моделі RANK-KOR?
2. В яких випадках застосовується метод узгодження експертних оцінок?
3. Яким чином кожний експерт ранжує альтернативні варіанти?

ПРАКТИЧНА РОБОТА № 2

МОДЕЛЮВАННЯ ПРОЦЕСІВ ВЗАЄМОДІЇ ДВОХ ПОПУЛЯЦІЙ (МОДЕЛЬ ХИЖАК-ЖЕРТВА)

МЕТА РОБОТИ : Ознайомлення студентів з класичною моделлю Віто Вольтера «хижак-жертва» та її можливим використанням в екологічних дослідженнях.

ЗАВДАННЯ ДО РОБОТИ: Використовуючи програму DYN-MID виконати моделювання біологічної рівноваги в природі за заданими викладачем даними

ТЕОРЕТИЧНІ ВІДОМОСТІ

Математична модель, що описує взаємодії між популяціями типу "хижак-жертва", вперше були одержані відомим італійським математиком Віто Вольтерра, якого вважають батьком математичної біології і математичної екології. Рівняння зазвичай складаються відносно чисельності популяцій або відносно біомаси організмів, що її складають.

При побудові моделі використовують прийняту в екології класифікацію взаємодій між видами (при описі неживої системи необхідно знати, які сили врівноважують систему) – конкуренція, симбіоз, хижацтво. Базовими рівняннями для побудови моделі є диференційні рівняння моделей росту популяцій (найпростішої – експоненціальної моделі росту популяцій та логістичної моделі росту популяцій, що враховує вплив на популяцію зовнішнього середовища) [1].

В наш час для опису системи "хижак-жертва" або "споживач-ресурс" широко використовується така система:

$$\frac{dR}{dt} = Q - V(R,N)N$$

$$\frac{dN}{dt} = -mN + kV(R,N)N$$

де R – кількість ресурсу; N – чисельність споживачів (популяції); Q – швидкість надходження ресурсу в систему; $V(R,N)$ – швидкість споживання ресурсу однією особиною популяції; $k<1$– частина ресурсу, що витрачається на відтворення (виробництво потомства); m – коефіцієнт смертності, який обернено пропорційний середній тривалості життя особин в даних умовах середовища. Найпростішим випадком є такий, коли Q, k, m – сталі величини. Відносно вибору трофічної функції $V(R,N)$ приймаються різні припущення, зокрема припускають, що вона залежить тільки від R, причому в нулі вона дорівнює нулю, а при збільшенні R вона зростає, причому має асимптоту $V=A$.

Вихідними даними для програми DYN-MID є: -початкове значення часу (t_0); – кінцеве значення часу $(t_к)$; – початкова середня чисельність хижаків; – початкова середня чисельність рослиноїдних; – початковий крок інтегрування (0.1 – 1). Результатами моделювання є: – прогнозна середня чисельність хижаків; – середня чисельність рослиноїдних наприкінці заданого часового періоду. За отриманими табличними даними необхідно графічно побудувати залежність середньої чисельності хижаків і рослинноїдних від часу t.

Контрольні запитання

1. Яка модель є базовою для побудови моделі «хижак-жертва»?
2. Чим відрізняються моделі експоненціального і логістичного росту популяцій?
3. Які класи взаємодій у живій і неживій природі, що розглядаються у екології застосовуються при моделюванні взаємодії популяцій ?

ПРАКТИЧНА РОБОТА № 3

МОДЕЛЮВАННЯ ПРОЦЕСУ БІОЛОГІЧНОГО ОЧИЩЕННЯ СТІЧНИХ ВОД

МЕТА РОБОТИ: Ознайомити студентів з імітаційним моделюванням процесу біологічного очищення стічних вод у біоставках.

ЗАВДАННЯ ДО РОБОТИ: Використовуючи комп'ютерну програму bio_prud виконати моделювання біоставків для очищення стічної води за даними, що надаються викладачем.

ТЕОРЕТИЧНІ ВІДОМОСТІ

Одним із найбільш розповсюджених методів біологічної очистки води є застосування біоставків(при БСК$_{повн}$.до 500 мг /л). При цьому використовують природну аерацію (БСК$_{повн}$ не перевищує 200мг/л) і штучну (в інших випадках). Моделювання процесу біоочищення у ставках також передбачає визначення наступних параметрів: часу перебування води в ставку; площі дзеркала води у ставку, за якими встановлюються розміри конструкції і обладнання споруд.

Час перебування стічних вод в біоставках з природною аерацією визначають за формулою:

$$\tau = (1/K_1)\lg[(L_a - L_c)/(L_\tau - L_c)],$$

де K_1 – константа швидкості споживання кисню, 1/доб; L_a – БСК$_{повн}$ води, що надходить на очищення, мг(O_2)/л; L_τ – БСК$_{повн}$ води, що виходить з ставка мг(O_2)/л; L_c – БСК$_{повн}$ води, обумовлена внутрішньоводоймищними процесами мг(O_2)/л.

Активну поверхню біологічного ставка визначають за формулою:

$$F_a = aQ(L_a - L_\tau)/[(a-b)r_1],$$

де Q – витрати стічної води, м^3/доб; L_a, L_τ – БСК$_{повн}$ відповідно стічних вод, що надходять і виходять із ставка, мг(O_2)/л; a, b – розчинність кисню у воді відповідно на початку процесу і через час τ, мг/л; r_1 – атмосферна реаерація, г/(м^2·доб).

Фактична площа ставка (площа дзеркала) дорівнює $F = F_a / \alpha'$, для ставків з сильно порізаними берегами $\alpha' = 0,5 - 0,6$; для ставків з рівними берегами $\alpha' = 0,8 - 0,9$.

Для підвищення швидкості розчинення кисню використовують ставки з аерацією. Аерація дозволяє в 3 – 3,5 рази підвищити навантаження за забрудненнями і збільшити глибину ставка до 3,5 м [5] .

Контрольні запитання

1. Який фактор найбільш суттєво впливає на процесу якості води та стану водної екосистеми?
2. Що таке біоставки і в яких випадках їх застосовують?
3. Коли доцільно застосовувати біоставки з природною аерацією, а коли з штучною?

ПРАКТИЧНА РОБОТА № 4

МОДЕЛЮВАННЯ ТА ПРОГНОЗУВАННЯ ЗАБРУДНЕННЯ ВОДОТОКУ В РАЙОНІ ВОДОЗАБОРУ

МЕТА РОБОТИ : Ознайомлення студентів з ідентифікацією моделі динаміки забруднень у водотоках.

ЗАВДАННЯ ДО РОБОТИ: Дані до роботи надаються викладачем у вигляді таблиці вимірювання гідрологічних характеристик і показника біхроматної окиснюваності на постах спостереження. Необхідно розрахувати концентрації забруднення в наступні моменти часу.

ТЕОРЕТИЧНІ ВІДОМОСТІ

Розв’язання водогосподарських проблем, у тому числі визначення допустимого навантаження на водотоки, потребує складання водогосподарського балансу водних ресурсів, що дає змогу прогнозувати якість води річок і водоймищ на перспективу за встановлених початкових і граничних умов. Екологічні вимоги визначаються за різницею нормативного й фактичного навантаження в різних створах. Обмеження зображується у вигляді:

$$S_i < [S_i] \qquad (4.1.)$$

де $[S_i]$ – нормативна концентрація речовини для заданої категорії водоспоживання й водокористування.

У балансових розрахунках навантаження обмежується у вигляді імовірнісної характеристики :

$$P\{S_i Q_i \le [G_i]\} \ge P^*, \qquad (4.2.)$$

де P, P^* – імовірнісні характеристики відповідно водних ресурсів і водоспоспоживання; G_i – розрахункові (мінімальні) витрати води в i-му створі. Необхідно запобігти зниження якості питної води нижче за санітарно-гігієнічні нормативи (гранично допустимі концентрації – ГДК).

Підприємство розташовано в точці А (місце випуску стічних вод) (рис. 3.1). За умов незадовільного режиму роботи очисних споруд, коли порушується технологія або режим випуску стічних вод, у районі водозабору B може скластися екологічно небезпечна ситуація.

Схема спостережень має бути постійною (для забезпечення початкових значень) на основі невеликого числа точок і періодичною за ширшою мережею спостережень (для ідентифікації моделей).

Періодична схема спостережень, яка організується для збирання даних натурних спостережень за концентрацією забруднюючих речовин з метою ідентифікації моделей прогнозу, здійснюється в такий спосіб. Здійснюється спостереження за скидом стічних вод у точці A безпосередньо на підприємстві (об'єм скиду, час, концентрація стічних вод тощо). В точці 1 (повторного змішування стічних вод), а також у точках 2…6 здійснюються синхронні вимірювання концентрацій забруднюючих речовин з інтервалом часу τ. Відстань Δx між точками краще брати постійною (рівномірна мережа спостережень). Ідентифікацію параметрів моделі можна здійснювати і в разі нерівномірної мережі спостережень, коли немає деяких даних, однак вирішення таких питань потребує спеціальних досліджень.

Виходячи з одновимірного рівняння процесів дифузії, переносу й самоочищення [3,4], структуру різницевого оператора можна подати у вигляді

$$U_{k,n+1} = a_{-1}U_{k-1,n} + a_0 U_{k,n} + a_1 U_{k+1,n}\ , \tag{4.3}$$

де
$$a_{-1} = \left(\frac{a}{(\Delta x)^2} + \frac{V}{2\Delta x}\right)\tau\ ; \tag{4.4.}$$

$$a_0 = \left(\frac{1}{\tau} - \frac{2a}{(\Delta x)^2} + \lambda\right)\tau\,; \tag{4.5.}$$

$$a_1 = \left(\frac{a}{(\Delta x)^2} - \frac{V}{2\Delta x}\right)\tau\ ; \tag{4.6.}$$

Якщо швидкість течії на відрізку $A-B$ постійна, а умови самоочищення ідентичні, то синхронними вимірюваннями в моменти $\tau, 2\tau, ..., (n+1)\tau$ у точках 1…3 концентрацій $U'_{k-1,n}, U'_{k,n}, U'_{k+1,n}$ (рис. 4.1) отримаємо матрицю спостережень X та вектор вихідної величини Y:

$$X = \begin{bmatrix} U^0_{k-1,n} & U^0_{k,n} & U^0_{k+1,n} \\ U^1_{k-1,n} & U^1_{k,n} & U^1_{k+1,n} \\ \dots\dots\dots & \dots\dots & \dots\dots\dots \\ U^{n_1}_{k-1,n} & U^{n_1}_{k+1,n} & U^{n_1}_{k+1,n} \end{bmatrix}; \qquad Y = \begin{bmatrix} U^1_{k,n} \\ U^2_{k,n} \\ \dots\dots \\ U^{n_1+1}_{k,n} \end{bmatrix} \tag{4.7.}$$

що дає можливість знайти вектор $a^T = (a_{-1}, a_0, a_1)$ з системи алгебраїчних рівнянь

$$\left(X^T X\right) a = X^T Y \ . \tag{4.8.}$$

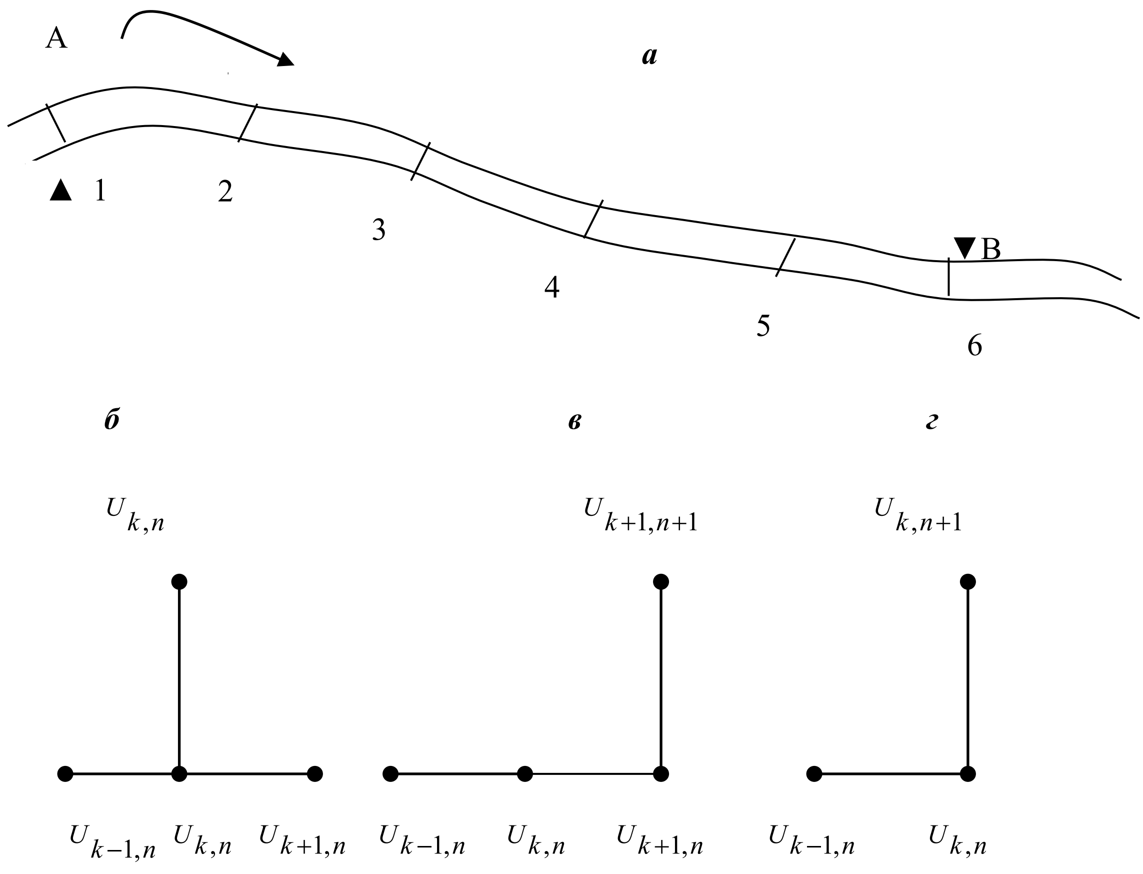

Рис. 4.1. Схема спостережень і розташування гідрохімічних постів на річці (*а*) та розрахункові комірки (шаблони) (*б-г*) для побудови моделей

Із рівнянь (4.4), (4.5) і (4.6.) за умови відомих значень a^T, визначаємо фізичні параметри a, V, λ, а також залежність концентрації забруднюючих речовин у створі повного змішування (точки 1):

$$U_{0,n} = b_0 Q + b_1 QM \qquad (4.9.)$$

як функцію викиду стічних вод M за час τ, м^3; витрати води в річці Q, м^3/с.

Оскільки перенесення речовини за рахунок дифузії мале порівняно з перенесенням течією річки, то :

$$\frac{dU}{dt} = -V\frac{dV}{dx} + f(x,t) + \lambda(x,t)U. \qquad (4.10.)$$

Тоді можна знайти невідомі коефіцієнти рівнянь :

$$U_{k+1,n+1} = a_0 U_{k-1,n} + a_1 U_{k,n} + a_2 U_{k+1,n}, \qquad (4.11.)$$

або $U_{k,n+1} = b_0 U_{k,n} + b_1 U_{k-1,n}$. (4.12.)

Для ідентифікації коефіцієнтів моделі (4.12) синхронні вимірювання можна здійснювати не на трьох створах, як у разі моделей (4.3.) або (4.11), а тільки на двох $U_{k,n}, U_{k-1,n}$.

Якщо швидкість течії або коефіцієнти самоочищення істотно відрізняються на різних ділянках річки, моделі типу (4.11) або (4.12) необхідно ідентифікувати для кожної точки (створу) окремо, тобто для розрахунку концентрації в наступний момент часу і побудови відповідних моделей динаміки.

Ідентифікація моделей динаміки для кожної точки створу проводиться наступним чином. Нехай маємо дані вимірювань гідрологічних характеристик і показника біхроматної окиснюваності на водпостах 1…6 (табл. 4.1) у три різні моменти часу з інтервалом Δt =1 год; Δx=1000 м при швидкостях течії 0,2…0,3 м/с.

Таблиця 4.1

Узгоджені вимірювання гідрологічних характеристик і показника біхроматної окиснюваності на водпостах річки

t	V, м/с	$U_{1,n}$, мг/л	$U_{2,n}$, мг/л	$U_{3,n}$,мг/л	U_{4n},мг/л	$U_{5,n}$,мг/л	U_{6n},мг/л
0	0,21	11,00	8,10	6,67	6,10	5,77	4,98
τ	0,25	12,30	9,00	7,41	6,78	6,41	5,50
2τ	0,23	9,10	7,90	6,49	5,94	5,67	4,86

В точці 1 покладено граничну умову $U_{1,n+1} = U_n$.

В інших точках (створах) річки

$$U_{2,n+1} = 0{,}81U_{2,n} + 0{,}01U_{1,n};$$
$$U_{3,n+1} = 0{,}85U_{3,n} + 0{,}053U_{2,n};$$
$$U_{4,n+1} = 0{,}89U_{4,n} + 0{,}051U_{3,n};$$
$$U_{5,n+1} = 0{,}81U_{5,n} + 0{,}045U_{4,n};$$
$$U_{6,n+1} = 0{,}84U_{6,n} + 0{,}041U_{5,n}$$

або в матричному вигляді

$$U_{k,n+1} = AU_{k,n} + BU_{k-1,n} \quad k = 1,\ldots,6$$

Задавши початкові умови $U_{k,0} = U_0, \quad k \in [2,6]$ та граничну умову $U_{k,n} = b_0 U_{k,n+1}$, розраховують концентрації забруднення в наступні моменти часу $n = 1,2,3,\ldots n_1$.

Контрольні запитання

1. Коли застосовується постійна схема спостережень, а у яких випадках періодична?
2. Які рівняння є базовим для моделювання і прогнозування забруднення водотоку?
3. Яким чином можна проводити ідентифікацію коефіцієнтів моделі?

ПРАКТИЧНА РОБОТА № 5

МОДЕЛЮВАННЯ І ПРОГНОЗУВАННЯ ЗАБРУДНЕННЯ ГРУНТІВ ХІМІЧНИМИ РЕЧОВИНАМ

МЕТА РОБОТИ: Ознайомлення студентів з моделюванням та прогнозуванням динаміки забруднення ґрунтів хімічними речовинами на прикладі пестицидів.

ЗАВДАННЯ ДО РОБОТИ: Дані до роботи надаються викладачем (це початкові точки граничні концентрації пестицидів для даної сільськогосподарської продукції, pH ґрунту), необхідно визначити: 1)час розпаду пестициду до безпечних концентрацій; 2)концентрації пестициду у заданих часових точках; 3) оцінку небезпечності пестициду за табл. 5.1.

ТЕОРЕТИЧНІ ВІДОМОСТІ

Моделювання і прогнозування поширення й розпаду шкідливих речовин дає змогу прогнозувати їх кількості і здійснювати контроль з метою обмеження їх впливу на людину.

Теоретичною моделлю розчинення, поглинання й розпаду пестицидів у ґрунтах у разі одновимірного руху розчину в пористому середовищі є рівняння дифузії:

$$m_0 \frac{\partial U}{\partial t} = D \frac{\partial^2 U}{\partial x^2} - V \frac{\partial U}{\partial x} - \frac{\partial b}{\partial t} - \alpha U + f(x), \qquad (5.1.)$$

де $D = D_M + \lambda |V|$ - дифузійна складова; α – коефіцієнт швидкості розпаду пестициду; U - концентрація пестициду в розчині; V – швидкість фільтрації; $\partial b / \partial t$ - швидкість розчинення пестициду у воді; $f(x)$ - функція поглинання пестициду кореневою системою. Для квазістаціонарного випадку, коли V не залежить від x і m_0=const, швидкість фільтрації розраховується з умови

$$\frac{\partial V}{\partial x} = \mu \frac{\partial^2 h}{\partial x^2}. \qquad (5.2.)$$

Кінетику процесу розчинення і розпаду можна описати рівнянням першого порядку:

$$\frac{\partial b}{\partial t} = \nu (U_m - U) - k_1 b \qquad (5.3.)$$

де k_1, ν – константи розпаду в твердій фазі (в сухих ґрунтах); U_m–- концентрація насичення ; b – концентрація пестициду в твердій фазі ґрунтів.

Виходячи з механізму явищ, основаного на тому, що процес розкладу речовини в ґрунтах здійснюється пропорційно поточній концентрації цієї речовини, а весь комплекс факторів, що діє на зміну концентрації пестицидів і радіонуклідів у часі, виражається через усереднений коефіцієнт *k*, кінетику розпаду пестицидів можна описати рівнянням

$$\frac{dU}{dt} = -kU(t), \qquad (5.4.)$$

розв’язок якого має вигляд

$$U(t) = U_0 e^{-kt}, \qquad (5.5.)$$

де $U(t)$ – кількість пестициду на момент часу t; U_0 - початкова концентрація пестициду; k – константа швидкості реакції розпаду пестициду; *t* – час.

Основним параметром хімічної кінетики є швидкість її реакцій, що обчислюється як

$$k=\frac{2{,}303}{t}\lg\frac{U_0}{U(t)}. \qquad (5.6.)$$

Час деструкції пестициду характеризується періодом напіврозпаду

$$T_{1/2}=0{,}693/k. \qquad (5.7.)$$

З рівняння (14) випливає, що час розпаду пестициду

$$t_p=\frac{(\lg U_0-\lg U(t))\cdot 2{,}303}{0{,}182}. \qquad (5.8.)$$

Таблиця 5.1

Шкала оціночних балів для рівнів небезпечності пестицидів

Екологотоксикологічні та гігієнічні показники x_i	Клас небезпечності $f_i(x_i)$	Характеристика класу	Оціночний бал $\mu(x_i)$
Персистентність у грунті	1	***До 1 місяця***	2
	2	1 ... 6 місяців	4
	3	0,5 ... 2 роки	6
	4	***Понад 2 роки***	8
Дія на грунтові ферментативні процеси й біоту	1	***Не впливає***	0
	2	Діє на одиничні процеси і популяції	1
	3	Діє на кілька процесів і популяцій	2
Транслокація в культурні рослини	1	Не надходить у рослини	0
	2	Надходить, але негативно не діє	1
	3	Надходить у продукти врожаю	2
	4	Проявляє фітотоксичну дію	3
Міграція за ґрунтовим профілем, см	1	***Не мігрує***	0
	2	Мігрує на відстань до 15	1
	3	до 50	2
	4	понад 50	3

Реакція на інсоляцію	1	Піддається фотохімічному розкладу	0
	2	***Не піддається***	1
ДЗК* для продуктів урожаю, мг/кг	1	1	0
	2	1 ... 0,1	1
	3	0,1 ... 0,01	2
	4	0,01	3
	5	0	4
ГДК для водоймищ, мг/л	1	1	0
	2	1 ... 0,1	1
	3	0,1 ... 0,01	2
	4	0,01	3
	5	0	4
Порогова концентрація у питній воді, мг/л	1	0,1	0
	2	0,1 ... 0,01	1
	3	0,01 ... 0,001	2
Дія на органолептичні якості продуктів урожаю	1	Не погіршує	0
	2	***Погіршує***	1
Леткість	1	Нелетка речовина	0
	2	Насичувальна концентрація нижче від порогової	1
	3	Насичувальна концентрація дорівнює пороговий	2
	4	Насичувальна концентрація дорівнює токсичній	3
Токсичність для теплокровних, мг/кг	1	1000	1
	2	201 ... 1000	2
	3	50 ... 200	3
	4	50	4
Коефіцієнт кумуляції в організмі теплокровних	1	5	0
	2	3 ... 5	1
	3	1 ... 3	2
	4	1	3

Згідно з даною класифікацією пестициди належать до однієї з трьох груп небезпечності за комплексом факторів, а саме :

$$F\left(\sum_{i=1}^{N}\mu(x_i)\right)=\begin{cases}3(\text{малонебезпечні}),\ \text{якщо}\sum_{i=1}^{N}\mu(x_i)\le 13;\\ 2(\text{середньонебезпечні}),\ \text{якщо}\ 13<\sum_{i=1}^{N}\mu(x_i)\le 21;\\ 1(\text{небезпечні}),\ \text{якщо}\sum_{i=1}^{N}\mu(x_i)>21.\end{cases} \qquad (5.9.)$$

Виходячи з даної моделі, проведено класифікацію і опис основних пестицидів, а також розроблено тактику їх застосування. Безпечність використання хімічних засобів досягаються за умови виконання двох вимог: необхідно, щоб продукти харчування і корм для тварин не містить залишків токсичних речовин у кількості понад гігієнічні нормативи ДЗК; не допускати нагромаджування залишків пестицидів у навколишньому середовищі. Стосовно тактики використання рекомендуються насамперед пестициди третьої екологотоксикологічної групи і тільки в разі, якщо за токсичністю або персистентністю вони непридатні для успішної боротьби зі шкідниками, слід використовувати препарати першої та другої груп.

Для того щоб не допускати вмісту пестицидів, вищого за ДЗК, треба враховувати час від початку обробки до запланованого строку збирання врожаю і застосовувати препарати, період детоксикації яких менший від гігієнічного нормативу часу розпаду пестициду до безпечних концентрацій. Нагромадження пестицидів в об’єктах навколишнього середовища не буде допущено, якщо застосовувати пестициди з таким розрахунком, аби нові надходження їх у ґрунт і рослини не перевищували темпів їхнього хімічного й біологічного розпаду.

Особливо жорстокі вимоги слід виконувати під час добору пестицидів на зрошувальних землях та в рисових сівозмінах, де забороняється скидання дренажних вод у водоймища, якщо вміст пестицидів перевищує ГДК.

Санітарними правилами заборонено застосування всіх пестицидів на відстані до 300 м від водоймищ санітарно-побутового використання; залежно від нахилу сільськогосподарських полів у бік водоймища захисна зона збільшується до 500м.

Контрольні запитання

1. В якому випадку необхідно виконувати особливо жорсткі вимоги під час вибору пестицидів?
2. Яке рівняння є базовим для моделювання і прогнозування забруднення ґрунтів хімічними речовинами?
3. Які припущення відносно механізму явищ розчинення і розпаду хімічних речовин у ґрунтах приймаються при моделюванні?
4. До якої з трьох груп небезпечності (за комплексом факторів відносяться пестициди?

ЧАСТИНА II. Моделювання стану довкілля у випадку недетермінованої базової інформації

ТЕОРІЯ КОРЕЛЯЦІЇ В АНАЛІЗІ ЗВ'ЯЗКІВ МІЖ ХАРАКТЕРИСТИКАМИ ЕКОЛОГІЧНИХ ЯВИЩ І ПРОЦЕСІВ

ТЕОРЕТИЧНІ ВІДОМОСТІ

При дослідженні екологічних явищ та процесів часто виникає необхідність аналізу зв'язків між різними кількісними та якісними ознаками.

Дві випадкові величини можуть бути пов'язані ***функціональною*** залежністю, або ***статистичною*** залежністю, або бути незалежними. При функціональному зв'язку значення однієї величини (аргументу) відповідає визначене значення другої величини (функції).

Статистичною називають залежність при якій:

1. кожному значенню однієї з випадкових величин (наприклад X) відповідає ряд розподілу іншої (наприклад Y);
2. із зміною X закономірно змінюються статистичні характеристики рядів розподілу Y: положення рядів, розподіл, розсіяння.

Сувора функціональна залежність реалізується рідко, оскільки обидві величини або одна з них піддаються ще дії випадкових факторів, причому серед них можуть бути і спільні для обох величин (під «спільними» тут розуміють такі фактори, які впливають і на Y і на X). В цьому випадку виникає ***статистична залежність***. Наприклад, якщо Y залежить від випадкових факторів Z_1, Z_2, V_1, V_2, а X залежить від випадкових факторів Z_1, Z_2, U_1, то між Y і X існує статистична залежність, оскільки серед випадкових факторів є спільні (Z_1 та Z_2). Статистичну залежність називають ***кореляційною***, якщо при зміні однієї з величин, змінюється середнє значення іншої. Будь яка кореляційна залежність є статистичною, але не всяка статистична залежність є кореляційною.

Нехай дані, що одержуються у спостереженні вивчаються з метою дослідження залежності між випадковими величинами X та Y і вони такі, що кожному значенню X відповідає декілька значень. Наприклад при $X = x$, величина Y приймає значення y_1, y_2,…, y_k. Знайдемо середня арифметичне цих значень та будемо позначати його через $\overline{y}_x$ тоді:

$$\bar{y}_x = \frac{\sum_{i-1}^{b} y_i}{k},$$

$\bar{y}_x$ називають умовною середньою.

Використовуючи поняття умовної середньої, кореляційну залежність Y від X можна визначити як функціональну залежність умовної середньої $\bar{y}_x$

$$\bar{y}_x = f(x). \tag{6.1}$$

Рівняння (2.1) називається ***рівнянням регресії Y на X (чи Y від X)***; функцію f(x) називають ***функцією регресії Y по X***, а її графік – ***лінією регресії Y на X (Y на X)***.

Аналогічно можна встановити кореляційну залежність умовної середньої $\bar{x}_y$ від y ($\bar{x}_y$ - середнє арифметичне значень x відповідно y

$$\bar{x}_y = \varphi(y) \tag{6.2}$$

Рівняння (6.2) називається ***рівнянням регресії X на Y (чи X від Y)***; функцію f(x) називають ***функцією регресії X по Y***, а її графік – ***лінією регресії X на Y (X на Y)***.

Існування (чи відсутність) зв'язку між випадковими величинами X та Y, а також силу (щільність) зв'язку між ними характеризує ***коефіцієнт кореляції.***

Коефіцієнтом кореляції називають відношення кореляційного моменту до добутку середніх квадратичних відхилень величин між якими розглядається зв'язок, тобто:

$$r_{x,y} = \frac{K_{x,y}}{\sigma_x \sigma_y}. \tag{6.3}$$

Кореляційним моментом $K_{x,y}$ («моментом зв'язку» чи коваріацією cov $\{X,Y\}$, ($\mathrm{cov}_{x,y}$) випадкових величин X та Y називають математичне сподівання добутку центрованих величин $\overset{\circ}{X}$ та $\overset{\circ}{Y}$:

$$K_{xy}=M\{\overset{\circ}{X},\overset{\circ}{Y}\}=M\{(X-m_x)(Y-m_y)\}\}.$$

Для незалежних випадкових величин кореляційний момент (а звідси і коефіцієнт кореляції) дорівнює нулю. Такі випадкові величини називають некорельованими.

Випадкові величини називаються ***корельованими***, якщо їх кореляційний момент (тобто і коефіцієнт кореляції) відмінний від нуля. Але якщо із незалежності випадкових величин завжди випливає їх некорельованість, то із некорельованості не завжди випливає їх незалежність. Інакше кажучи, якщо дві величини залежні, то вони можуть бути як корельованими так і некорельованими.

Вибірковий коефіцієнт кореляції безрозмірна величина, при цьому виконується:

$$|\,r_{xy}\,| \le 1.$$

Він служить для оцінки щільності тільки лінійного зв'язку між величинами X та Y: чим ближче абсолютна величина коефіцієнта кореляції до одиниці, тим зв'язок сильніший; чим ближче $|\,r_{xy}\,|$ до нуля, тим зв'язок слабший.

Якщо випадкові величини X та Y пов'язані точковою лінійною функціональною залежністю $Y = aX + b$, то $r_{xy} = \pm 1$(знак "+" чи "-" береться в залежності від того a>0 чи a <0).

Якщо $r_{xy} > 0$ кажуть про додатню кореляцію X та Y (при збільшенні однієї з величин друга має тенденцію у середньому збільшуватися). Якщо $r_{xy} < 0$ – кажуть про від'ємну кореляцію між X та Y (при збільшенні однієї з величин друга має тенденцію в середньому зменшуватися).

Про існування чи відсутність кореляції можна судити, у першому наближенні, за виглядом кореляційного зображення точок усіх одержаних вимірів (чи спостережень) пар значень випадкових величин.

Кореляційний аналіз вирішує дві основні задачі. Перша задача теорії кореляції – встановити форму кореляційного зв'язку, тобто вигляд функції регресії (лінійна, квадратична, показникова і т.п.). Найбільш часто функції регресії виявляються лінійними. Якщо обидві функції регресії лінійні, то кореляцію називають ***лінійною***, в протилежному випадку – ***нелінійною***. При лінійній кореляції обидві лінії регресії являються прямими лініями.

Друга задача теорії кореляції – оцінити щільність (силу) кореляційного зв'язку. Щільність кореляційної залежності Y від X оцінюється за величиною розсіювання значень величини Y навколо умовного середнього $\overline{y}_x$.

Велике розсіювання свідчить про слабку залежність Y та X, або про відсутність залежності; можливо навіть що Y та X пов'язані функціонально, але під дією другорядних випадкових факторів цей зв'язок виявився розмитим, в результаті чого при одному і тому ж значенні $X = x$ величина Y приймає різні значення. Аналогічно (за величиною розсіювання значень X навколо умовного середнього $\overline{x}_y$) оцінюється щільність кореляції зв'язку X від Y.

Нехай відомо, що випадкові величини X та Y зв'язані лінійною кореляційною залежністю (обидві лінії регресії прямі). Потрібно за даними спостережень (досліду) знайти рівняння прямих ліній регресії Y на X та X на Y та оцінити силу лінійного кореляційного зв'язку.

Нехай маємо найпростіший випадок, коли в результаті незалежних спостережень була одержана сукупність n пар чисел:
(x_1, y_1), (x_2, y_2), …, (x_n, y_n) (кожна пара чисел спостерігалась тільки по одному разу). Тоді шукане рівняння прямої лінії регресії Y на X буде мати вигляд

$$Y = \rho_{yx}x + b, \tag{6.4}$$

де ρ_{yx} - вибірковий коефіцієнт регресії Y по X.

Рівняння (6.4) називається ***ймовірним рівнянням прямої лінії регресії Y на X***.

Параметри ρ_{yx} та b в рівнянні (6.4) визначаються на основі методу найменших квадратів, тобто такими, щоб сума квадратів відхилень експериментальних значень y_i від значень Y_i (i=1, 2,…, n), обчислених за рівнянням (6.4), була мінімальна:

$$\sum_{i=1}^{n}(Y_i - y_i)^2 = \min.$$

Система нормальних рівнянь для визначення ρ_{yx} та b має вигляд:

$$\begin{cases} \left(\sum_{i=1}^{n} x_i^2\right)\rho_{yx} + \left(\sum_{i=1}^{n} x_i\right)b = \sum_{i=1}^{n} x_i y_i; \\ \left(\sum_{i=1}^{n} x_i\right)\rho_{yx} + nb = \sum_{i=1}^{n} y_i. \end{cases} \tag{6.5}$$

Розв’язавши цю систему, наприклад, за формулами Крамера, знайдемо шукані параметри:

$$\rho_{yx} = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}; \tag{6.6}$$

$$b = \frac{\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}. \tag{6.7}$$

Аналогічно можна знайти вибіркове рівняння прямої лінії регресії X на Y

$$X = \rho_{xy}\, y + c,$$

де ρ_{xy} - вибірковий коефіцієнт регресії X на Y.

Для характеристики сили лінійного кореляційного зв’язку між величинами X та Y за дослідними даними знаходимо ***вибірковий коефіцієнт кореляції***:

$$r_B = \frac{\sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y})^2}{n s_x s_y}, \tag{6.8}$$

де $\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$; $\bar{y}=\frac{1}{n}\sum_{i=1}^{n}y_i$;

$$s_x=\sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}}\,; \qquad s_y=\sqrt{\frac{\sum_{i=1}^{n}(y_i-\bar{y})^2}{n-1}}\,; \quad (i=\overline{1,n});$$

s_x, s_y - вибіркові середні квадратичні відхилення;

Для практичного використання більш зручними є формули:

$$\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})=\sum_{i=1}^{n}x_iy_i-\frac{1}{n}\sum_{i=1}^{n}x_i\sum_{i=1}^{n}y_i; \tag{6.9}$$

$$s_x=\sqrt{\frac{1}{n-1}\left[\sum_{i=1}^{n}x_i^2-\frac{1}{n}\left(\sum_{i=1}^{n}x_i\right)^2\right]}; \tag{6.10}$$

$$s_y=\sqrt{\frac{1}{n-1}\left[\sum_{i=1}^{n}y_i^2-\frac{1}{n}\left(\sum_{i=1}^{n}y_i\right)^2\right]}. \tag{6.11}$$

ПРАКТИЧНА РОБОТА №6

ОЦІНКА ЩІЛЬНОСТІ ЗВ’ЯЗКУ МІЖ ДВОМА ВИПАДКОВИМИ ВЕЛИЧИНАМИ У ВИПАДКУ ОДНАКОВОЇ КІЛЬКОСТІ ЇХ ЗНАЧЕНЬ

МЕТА РОБОТИ. кОпанування студентами навичок застосування елементів кореляційного аналізу для вивчення залежностей між випадковими величинами, які характеризують певні ознаки екологічних випадкових явищ та процесів і інформація одержана на основі спостережень.

ЗАВДАННЯ ДО РОБОТИ: за даними, наданими викладачем:

1) Знайти залежність між двома випадковими величинами *X* і *Y*;

2) Визначити коефіцієнт кореляції

Приклад 6.1

Деякі характеристики X та Y ряда з i компонентів параметрів довкілля, що досліджуються, пов'язані між собою, та мають значення, які подані у табл. 6.1. Знайти залежність між X та Y та обчислити вибірковий коефіцієнт кореляції.

Таблиця 6.1

x_i	5	8	10	7	5
y_i	4	5	9	6	1

Розв'язок. Побудуємо таблицю в якій крім вихідних даних розраховуються коефіцієнти для обчислення системи (табл. 6.2).

Таблиця 6.2

I	Компо-ненти	x_i	y_i	x_i^2	y_i^2	x_iy_i
1	1	5	4	25	16	20
2	2	8	5	64	25	40
3	3	10	9	100	81	90
4	4	7	6	49	36	42
5	5	5	1	25	1	5
$\sum$		35	25	263	159	197

Підставляючи значення з таблиці 6.2 в рівняння (6.5)-(6.7) знаходимо значення ρ_{yx}, b.

$$\rho_{yx} = \frac{5 \cdot 197 - 25 \cdot 35}{5 \cdot 263 - 35^2} = 1,2;$$

$$b = \frac{263 \cdot 25 - 35 \cdot 197}{5 \cdot 263 - 35^2} = -3,58.$$

Рівняння регресії, що визначається, буде мати такий вигляд:

$$Y = 1,2X - 3,58.$$

Вибірковий коефіцієнт кореляції знаходимо за формулою (6.8), попередньо зробивши обчислення за формулами (6.9)-(6.11).

$$\sum_{i=1}^{3}(x_i-\bar{x})(y_i-\bar{y})=\sum_{i=1}^{5}x_iy_i-\frac{1}{5}\sum_{i=1}^{5}x_i\sum_{i=1}^{5}y_i=197-\frac{1}{5}35\cdot 25=22;$$

$$s_x=\sqrt{\frac{1}{4}\left[\sum_{i=1}^{5}x_i^2-\frac{1}{5}\left(\sum_{i=1}^{5}x_i\right)^2\right]}=\sqrt{\frac{1}{4}\left[263-\frac{1}{5}\cdot 35^2\right]}\approx 2{,}12;$$

$$s_y=\sqrt{\frac{1}{4}\left[\sum_{i=1}^{5}y_i^2-\frac{1}{5}\left(\sum_{i=1}^{5}y_i\right)^2\right]}=\sqrt{\frac{1}{4}\left[159-125\right]}\approx 2{,}91;$$

$$r_B=\frac{22}{5\cdot 2.12\cdot 2.91}\approx 0{,}71.$$

Вибірковий коефіцієнт кореляції між *X* та *У* дорівнює 0,71.

ПРАКТИЧНА РОБОТА № 7

Оцінка щільності зв'язку між двома випадковими величинами при різній кількості вимірів їх значень

МЕТА РОБОТИ: Оцінити щільність зв'язку між двома недермінованими величинами при різній кількості вимірів їх значень

ЗАВДАННЯ ДО РОБОТИ. На основі даних спостережень, що надані для вивчення залежностей між випадковими величинами *X* та *Y*, і які подані у вигляді кореляційної таблиці:

1) знайти рівняння прямих ліній регресії *Y* на *X* та *X* на *Y*;
2) оцінити щільність зв'язку між *X* та *Y*

Розглянемо більш загальний випадок. Нехай в n незалежних спостереженнях одне і те саме значення x спостерігається n_x разів, а значення y – n_y разів. Такі дані групують, тобто підраховують частоти n_x, n_y, n_{xy} та подають їх у вигляді таблиці, яка називається кореляційною таблицею.

Приклад 7.1

В $n = 8$ незалежних вимірюваннях були одержані такі дані, що подають залежність густини водних розчинів від концентрації компонента, що досліджується, при t=10 °С. Подати ці дані у вигляді кореляційної таблиці.

Таблиця 7.1

Концентрація	15	15	16	16	17	17	17	18
Густина	1,11	1,12	1,12	1,13	1,12	1,13	1,13	1,14

Розв'язок. На перетині рядка та стовпця знаходять n_{xy}. Наприклад частота 1, що знаходиться на перетині першого рядка та першого стовпця, вказує, що пара чисел (15 та 1,11) спостерігається у досліді один раз.

Таблиця 7.2

Кореляційна таблиця

Значення y_i (густина) г/см^3	Значення x_i, % (концентрації компонента)				
	15	16	17	18	n_y
1,11	1				1
1,12	1	1	1		3
1,13		1	2		3
1,14				1	1
n_x	2	2	3	1	8

В останньому стовпці записані суми частот рядків, які вказують скільки раз спостерігалось значення y_i випадкової величини Y у комбінації (сполученні) з різними значеннями випадкової величини X. В останньому рядку записані суми частот стовпців, вони вказують скільки разів значення x_i величини X спостерігається з різними значеннями величини Y. У нижньому правому куті записана сума всіх частот (загальне число n проведених вимірів). Очевидно $\sum n_x = \sum n_y = n$, у нашому випадку $n = 8$.

Вибіркове рівняння прямої лінії регресії Y на X, що одержане за згрупованими даними, буде мати вигляд:

$$\bar{y}_x = \rho_{yx} x + b, \tag{7.1}$$

(тут використовується поняття умовної середньої $\bar{y}_x$). Враховуючи, що мають місце вирази:

$$\bar{x}=\frac{\sum_{i=1}^{n} x_i}{n}; \quad \bar{y}=\frac{\sum_{i=1}^{n} y_i}{n}; \quad \bar{x}^2=\frac{\sum_{i=1}^{n} x_i^{\,2}}{n},$$

маємо для згрупованих даних

$$\sum_{i=1}^{n} x_i=\bar{x}\cdot n; \sum_{i=1}^{n} y_i=\bar{y}\cdot n; \sum_{i=1}^{n} x_i^2=\bar{x}_i^2\cdot n,$$

а також те, що кожна пара чисел (x,y) спостерігалась n_{xy} разів ($\sum xy=\sum n_{xy}xy$), формули (6.6) і (6.7) можна подати у вигляді:

$$\rho_{yx}=\frac{\sum n_{xy}xy-n\cdot\bar{x}\bar{y}}{n(\overline{x^2}-\bar{x}^2)}; \quad b=\frac{n\overline{x^2}\bar{y}-\bar{x}\sum n_{xy}\bar{x}y}{n(\overline{x^2}-\bar{x}^2)}.$$

Так як $\overline{x^2}-\bar{x}^2=\sigma_x^2$, то

$$\rho_{xy}=\frac{\sum n_{xy}xy-n\cdot\bar{x}\cdot\bar{y}}{n\sigma_x^{\,2}}.$$

Помноживши обидві частини цієї рівності на σ_x/σ_y одержуємо:

$$\rho_{xy}\sigma_x/\sigma_y=\frac{\sum n_{xy}xy-n\bar{x}\bar{y}}{n\sigma_x\sigma_y}. \tag{7.2}$$

Таким чином, одержуємо вираз для вибіркового коефіцієнта кореляції:

$$r_B=\frac{\sum n_{xy}xy-n\bar{x}\bar{y}}{n\sigma_x\sigma_y}, \tag{7.3}$$

а формулу, яка зв'язує коефіцієнт регресії Y на X ρ_{yx} з коефіцієнтом вибіркової кореляції r_B:

$$\rho_{yx} = r_B \frac{\sigma_y}{\sigma_x}. \tag{7.4}$$

Вибіркове рівняння прямої лінії регресії Y на X:

$$\overline{y}_x - \overline{y} = r_B \frac{\sigma_y}{\sigma_x}(x - \overline{x}), \tag{7.5}$$

і аналогічно X на Y:

$$\overline{x}_y - \overline{x} = r_B \frac{\sigma_x}{\sigma_y}(y - \overline{y}). \tag{7.6}$$

Приклад 7.2

На основі спостережень, що були виконані протягом довгого періоду часу вимірювань компоненту, що досліджується, (X) в атмосферних викидах технологічного об'єкту, та даних проценту вилучення цього компонента (Y) при очищенні деяким методом складена кореляційна таблиця (табл. 7.3). Знайти вибіркове рівняння прямої лінії регресії Y на X та X на Y.

Таблиця 7.3

Y \ X	4-10	10-16	16-22	22-28	28-34	34-40
17-27	2	4				
27-37		6	2	1		
37-47			3	35	2	
47-57			1	10	6	
57-67				4	2	4

Розв'язок. Будемо шукати розв'язок у вигляді (7.5) та (7.6). Спочатку обчислимо вибірковий коефіцієнт кореляції. Результати розрахунку зручно записати у дві допоміжні таблиці, розташовані праворуч і знизу від вихідної кореляційної таблиці (будемо називати їх «правою» та «нижньою» таблицями).

Результати розрахунку наведені в табл.7.4.

Таблиця 7.4

Кореляційна та допоміжні («права» і «нижня») таблиці

$Y \backslash X$	4-10	10-16	16-22	22-28	28-34	34-40	n_y	y	v	$n_y v$	$n_y v^2$	q	qv
17-27	2	4					6	22	-2	-12	24	-14	28
27-37		6	2	1			9	32	-1	-9	9	-14	14
37-47			3	35	2		40	42	0	0	0	-1	0
47-57			1	10	6		17	52	1	17	17	5	5
57-67				4	2	4	10	62	2	20	40	10	20
n_x	2	10	6	50	10	4	82			16	90		
x	7	13	19	25	31	37							
u	-3	-2	-1	0	1	2							
$n_x u$	-6	-20	-6	0	10	8	-14						
$n_x u^2$	18	40	6	0	10	16	90						
t	-4	-14	-1	17	10	8							
tu	-12	28	1	0	10	16							

В якості значень y_i величини Y приймаємо середини інтервалів: $y_1 = 22$ (середина інтервалу 17-27), $y_2 = 32$ (середина інтервалу 27-37) і т. п. Запишемо ці значення в стовпець y «правої» таблиці. Аналогічно визначаємо значення x_i величини X: $x_1 = 7$ (середина інтервалу 4-10); $x_2 = 13$ (середина інтервалу 10-16) і т. п. Запишемо їх у рядок x «нижньої» таблиці. Для спрощення розрахунків, перейдемо від початкових варіант x_j та y_i (i=1,…,5; j=1,2,…,6) до умовних варіант:

$$u_j = \frac{x_j - c_1}{h_1} \quad \text{та} \quad v_i = \frac{y_i - c_2}{h_2},$$

де $h_1 = x_{j+1}-x_j$=10-4=16-10=…=6;

$h_2=y_{j+1}-y_j$=27-17=37-27=…=10.

У якості c_1 беремо варіанту x =25, яка має найбільшу частоту n_x=50, а у якості c_2 візьмемо варіанту y=42, що має найбільшу частоту n_y=40. Таким чином умовні варіанти

$$u_j=\frac{x_j-25}{6} \text{ та } v_i=\frac{y_i-40}{10}.$$

Запишемо їх відповідно в рядок u «нижньої» таблиці і в стовпчик v «правої» таблиці. Елементи рядка n_xu дорівнюють добуткам відповідних елементів рядків n_x і u. А елементи стовпця n_yv дорівнюють добуткам відповідних елементів стовпців n_y та v. Аналогічно заповнюємо рядок n_xu^2 і стовпець n_yv^2.

Кожний елемент стовпця q дорівнює сумі добутків частот n_{uv}, які знаходяться у відповідному рядку вихідної кореляційної таблиці, на відповідні елементи рядка u:

q_1=2·(-3)+4·(-2)= -14; q_2= 6·(-2)+2·(-1) = -14 і т. п.

У рядку t кожен елемент дорівнює сумі добутків частот n_{uv}, які знаходяться у відповідному стовпці початкової кореляційної таблиці, на відповідні елементи стовпця v:

t_1=2·(-2)=-4; t_2=4·(-2)+6·(-1)= -14 і т. п.

У стовпці qv кожний елемент дорівнює добутку відповідних елементів стовпців q та v, у рядку tu кожний елемент дорівнює добутку відповідних елементів рядків t та u. Рівність сум $\sum qv=\sum tu=67=\sum n_{uv}uv$ служить контролем вірності обчислень.

Після цього послідовно знаходимо:

$$\overline{u}=\frac{\sum n_x u}{n}=-0{,}171;\ \overline{v}=\frac{\sum n_y v}{n}=0{,}195;$$

$$\sigma_u = \sqrt{\frac{90}{82} - (-0,171)^2} = 1,034; \ \sigma_v = \sqrt{\frac{90}{82} - 0,195^2} = 1,029.$$

Враховуючи, що $\sum n_{uv}uv = 67$ знаходимо коефіцієнт кореляції:

$$r_B = \frac{67 - 82 \cdot (-0,171) \cdot 0,195}{82 \cdot 1,034 \cdot 1,025} \approx 0,799.$$

Знаходимо значення:

$$\bar{x} = \bar{u}h_1 + c_1 = 23,97;$$
$$\bar{y} = \bar{v}h_2 + c_2 = 43,95;$$
$$\sigma_x = \sigma_u h_1 = 6,204;$$
$$\sigma_y = \sigma_v h_2 = 10,29.$$

Підставивши всі знайдені величини у рівняння (7.5) одержуємо шукане вибіркове рівняння прямої лінії регресії *Y* на *X*:

$$\bar{y}_x = 1,325x + 12,184.$$

Аналогічно можна одержати вибіркове рівняння прямої лінії регресії *X* на *Y* використовуючи рівняння (7.6):

$$\bar{x}_y = 0,482y + 2,79.$$

Порівняємо умовні середні, що розраховані за цими рівняннями та задану кореляційну таблиці 7.4 ($\bar{y}_x^*, \bar{x}_y^*$ – умовні середні по даних таблиці 7.4.; $\bar{y}, \bar{x}$ – умовні середні обчислені за формулами).

Таблиця 7.5

x_i	$\bar{y}_x$	$\bar{y}_x^*$	$\bar{y}_x - \bar{y}_x^*$
7	21,459	22	0,541
13	29,409	28	-1,409
16	37,359	40,333	2,974
25	45,309	45,4	0,091
31	53,259	52	-1,259
37	61,209	62	0,791

$$\bar{y}_7 = \frac{2 \cdot 22}{2} = 22;$$

$$\bar{y}_{13} = \frac{4 \cdot 22 + 6 \cdot 32}{10} = 28;$$

$$\bar{y}_{19} = \frac{2 \cdot 32 + 3 \cdot 42 + 1 \cdot 52}{6} = 40{,}333;$$

$$\bar{y}_{25} = \frac{1 \cdot 31 + 35 \cdot 42 + 10 \cdot 52 + 4 \cdot 64}{50} = 45{,}4;$$

$$\bar{y}_{31} = \frac{2 \cdot 42 + 6 \cdot 52 + 2 \cdot 62}{10} = 52;$$

$$\bar{y}_{35} = \frac{4 \cdot 62}{4} = 62.$$

Таблиця 7.6

y_i	$\bar{x}_y$	$\bar{x}^*_y$	$\bar{x}^*_y - \bar{x}_y$
22	13,394	11	-2,394
32	18,214	15.667	-2,547
42	23,034	24.85	1,816
52	27,854	26.765	-1,089
62	32,674	31	1,674

$$\bar{x}_{22} = \frac{2 \cdot 7 + 4 \cdot 13}{6} = 11;$$

$$\bar{x}_{32} = \frac{6 \cdot 13 + 2 \cdot 19 + 1 \cdot 25}{9} = 15{,}667;$$

$$\bar{x}_{42} = \frac{3 \cdot 19 + 35 \cdot 25 + 2 \cdot 31}{40} = 24{,}85;$$

Контрольні запитання

1. Що таке статистична залежність?
2. Що визначає коефіцієнт кореляції?
3. Які випадкові величини називаються корельованими?
4. Які методи визначення коефіцієнтів регресійного рівняння ви знаєте?

СПИСОК ЛІТЕРАТУРИ

1. *Заграй Я.М.* Моделювання і прогнозування стану довкілля: Навчальний посібник / Я.М. Заграй, О.А. Котовенко – Київ : КНУБА, 2008. – 97 с.

2. *Лаврик В.И.,* Математическое моделирование в гидроэкологических исследованиях /В.И. Лаврик, Н.А. Никифорович – Київ : Фитосоциумцентр, 1998. – 287с.

3. *Лаврик В.І.* Методи математичного моделювання в екології / В.І. Лаврик Київ : «КМ Академія», 2002 . – 203 с.

4. *Заграй Я.М.* Статистичний аналіз в екології: Навчальний.посібник. /Я.М. Заграй, О.А. Котовенко, В.О. Карасьова – Київ : КНУБА, 2001. – 132 с.

Навчальне видання

МОДЕЛЮВАННЯ ТА ПРОГНОЗУВАННЯ СТАНУ ДОВКІЛЛЯ

Методичні вказівки
до виконання практичних робіт
для здобувачів першого (бакалаврського)
рівня вищої освіти за спеціальностями
101«Екологія», 183 «Технології захисту
навколишнього середовища»

Укладачі: КОТОВЕНКО Олена Андріївна
МІРОШНИЧЕНКО Олена Юріївна

Комп’ютерне верстання *А. П. Селівестрової*

Ум. друк. арк. 2,09. Обл.-вид. арк. 2,25
Електронний документ. Вид № 57/V-25.

Виконавець і виготовлювач

Київський національний університет будівництва і архітектури
Проспект Повітряних Сил, 31, Київ, Україна, 03680

Свідоцтво про внесення до Державного реєстру суб’єктів
видавничої справи ДК № 808 від 13.02.2002 р