

Метод відновлення пропущених даних в екологічних задачах на основі МГУА

В.С.Степашко*, Ю.В.Коппа*, Г.О.Іутинська**

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

України; 03680, Україна, Київ, проспект Академіка Глушкова, 40; E-mail: astrid@irtc.uran.net.ua

**Інститут мікробіології та вірусології НАН України; 03143, Україна, Київ, вул. Заболотного, 154

Для задач моніторингу екологічних процесів запропоновано метод відновлення пропусків у взаємозв'язаних рядах спостережень з використанням МГУА. Метод ґрунтується на побудові ієрархічного комплексу моделей за невеликим числом паралельних даних з використанням "опорних" рядів спостережень без пропусків і застосуванням цих моделей для відновлення пропущених даних. Для прикладу в роботі побудовано моделі залежності чисельності мікроорганізмів у ґрунті від основних факторів, що визначають їхню життєдіяльність в умовах забруднення важкими металами. При цьому для відновлення пропущених спостережень використано "опорні" ряди гідрометеорологічних даних про температуру та вологість повітря.

Однією зі значних екологічних проблем є забруднення земель важкими металами. Для виявлення і прогнозування критичних ситуацій у цій сфері проводиться мікробіологічний моніторинг ґрунтів. При цьому вивчають динаміку зміни чисельності мікроорганізмів у ґрунті під впливом різних екологічних факторів. Але такі дослідження проводяться нерегулярно через труднощі технології мікробіологічного аналізу. Це викликає необхідність відновлювати проміжні дані для отримання рівномірних рядів екологічних спостережень з метою подальшого математичного моделювання. Моделювання за даними спостережень є необхідною умовою екологічного моніторингу, оскільки дозволяє оперативно оцінювати поточні екологічні ситуації та прогнозувати їх розвиток.

Враховуючи, що природні процеси, крім техногенного впливу, обумовлюються вресітї-решт погодними чинниками, які вимірюються строго регулярно, можна запропонувати таку очевидну схему відновлення пропущених екологічних даних. Назвемо "опорними" рядами спостережень дані регулярних гідрометеорологічних вимірювань, які не мають пропусків. Тоді, відібравши з цих рядів дані, що відповідають моментам вимірювання показників контрольованих екологічних процесів, можна одержати таблиці даних без пропусків і за ними побудувати моделі залежності цих показників від усіх факторів впливу.

Якщо ж тепер одержати моделі залежності кожного з цих факторів від опорних погодних характеристик, то одержимо ієрархічний комплекс моделей, де на вході будуть гідрометеорологічні змінні, а на виході – екологічні показники. Очевидно, що такі моделі можна тепер використовувати як для відновлення пропущених даних, так і для прогнозування розвитку контрольованих екологічних процесів на основі прогнозів погодних умов. Для побудови всіх згаданих моделей слід використовувати МГУА як найефективніший метод за умов коротких вибірок даних.

Матеріали і методи досліджень

Для виявлення закономірностей впливу на мікрофлору ґрунту забруднень іонами важких металів Cu^{2+} , Cd^{2+} , Pb^{2+} , Zn^{2+} , Hg^{2+} були проведені польові досліді на моніторингових ділянках на темно-сірому опідзоленому ґрунті Київської області. Застосовувалась традиційна схема [1] проведення експерименту: виділяється кілька однотипних ділянок землі, одна з яких залишається в якості контрольної, тобто на неї не вносять забруднюючих речовин, а на інші вносять задану їх кількість.

Моделльне забруднення ґрунту здійснювалось шляхом щорічного одноразового внесення в ґрунт розчинів солей важких металів на початку вегетаційного сезону. Кількість внесених металів (з розрахунку вмісту їх іонів) відповідала дозам забруднення 2 і 4 гранично допустимих концентрацій (ГДК). Зразки ґрунту для аналізів відбирали упродовж вегетаційного періоду 1993-96 рр. на глибині орного шару (0-20 см) приблизно через 2, 30 та 90 діб після внесення солей металів. Таким чином, було одержано по три точки вимірювань за чотири роки – разом 12 точок.

Кількість органотрофних мікроорганізмів у контрольному й забрудненому ґрунтах визначали методом посіву ґрунтової суспензії на живильне середовище, до складу якого входили м'ясопептинний агар і водний агар у співвідношенні 1:10. Таке середовище відповідає оліготрофним умовам існування мікроорганізмів у ґрунті [1].

На моніторингових ділянках в ґрунті контрольного і дослідних (забруднених різними дозами важких металів) варіантів виявлена сезонна динаміка чисельності мікроорганізмів, яка, залежно від метеорологічних умов, з роками змінювалася. Загалом відзначали високу чисельність мікроорганізмів у червні-липні та спостерігали поступове зменшення її в серпні та вересні. На фоні сезонної динаміки чисельності мікроорганізмів виявлено вплив забруднення важкими металами, який виявляється у зменшенні чисельності мікроорганізмів у перші строки (2-3 доба) після внесення в ґрунт важких металів та в поступовому відновленні їхньої кількості в наступні строки спостережень [2].

Задача моделювання

Зважаючи на відсутність інформації щодо внутрішньої структури модельованих процесів, які зумовлюють динаміку чисельності мікроорганізмів в ґрунті, для побудови моделі була використана інтерактивна система АСТРІД [3]. Ця система базується на індуктивному методі групового урахування аргументів (МГУА) і дозволяє автоматично будувати оптимальні моделі за даними спостережень в умовах неповної апріорної інформації.

Вхідними (незалежними) змінними (факторами) для побудови моделей були такі: концентрація рухомих форм важких металів, середньодекадні значення температури, вологості ґрунту і повітря, чисельність мікроорганізмів в ґрунті контрольної незабрудненої ділянки. Вихідними (залежними) змінними були дані про чисельність мікроорганізмів еколого-трофічних груп на ділянках з експериментальним модельним забрудненням ґрунту важкими металами.

Для побудови моделі динаміки чисельності органотрофних мікроорганізмів був сформований такий список вхідних змінних:

X1 - чисельність мікроорганізмів на контрольній ділянці (млн. в 1 г сухого ґрунту)	X8 - середня температура повітря поточної декади (°C)
X2 - концентрація міді (мг/кг ґрунту)	X9 - середня температура повітря попередньої декади (°C)
X3 - концентрація стронцію (мг/кг ґрунту)	X10 - середня вологість повітря поточної декади(%)
X4 - концентрація кадмію (мг/кг ґрунту)	X11- середня вологість повітря попередньої декади (%)
X5 - концентрація ртуті (мг/кг ґрунту)	X12 - коефіцієнт вологості ґрунту
X6 - концентрація свинцю (мг/кг ґрунту)	
X7 - число діб від дати забруднення	

До переліку незалежних змінних були віднесені дані вологості та температури повітря попередніх декад, що було зроблено з огляду на певну інерційність біологічних процесів.

На базі даних про динаміку всіх перелічених змінних були побудовані лінійні моделі зміни чисельності органотрофних мікроорганізмів у ґрунті для кожного забруднювача окремо. При цьому будувалась спершу модель залежності чисельності мікроорганізмів на контрольній ділянці від усіх вимірюваних факторів, а далі - модель залежності цього ж показника на забруднених ділянках від вимірюваних значень чисельності на контрольній ділянці та всіх факторів. Після чого, підставивши отриману першу модель у другу, одержимо модель для забрудненої ділянки, де залишаться тільки зовнішні погодні фактори. Маючи таку модель, можна відновити значення досліджуваного параметра у пропущених точках.

Результати моделювання.

За даними дослідів були побудовані моделі чисельності органотрофних мікроорганізмів у контрольному і в забрудненому важкими металами ґрунтах.

Чисельність органотрофів у контрольному ґрунті описується рівнянням:

$$X_1 = -97.739 + 3.811X_6 - 0.00433X_7 + 0.369X_9 - 0.0852X_{10} + 0.0259X_{11} + 98.982X_{12},$$

Модель має такі статистичні характеристики: СКП=1.43 (середньоквадратична помилка моделі), R =3.5 (максимальна абсолютна помилка моделі), R_{відн}=24,7% (відношення максимальної помилки до величини найбільшої варіації модельованої величини). Ця модель не включає змінних від X₂ до X₅, тобто фонові значення вмісту важких металів (міді, стронцію, кадмію, ртуті) в контрольному ґрунті протягом періоду досліджень не впливали на чисельність органотрофів, за винятком концентрації свинцю (змінна X₆). Крім того, неінформативною була змінна X₈ - середня температура повітря поточної декади. Відповідні графіки експериментальних і модельних даних подано на рис.1.

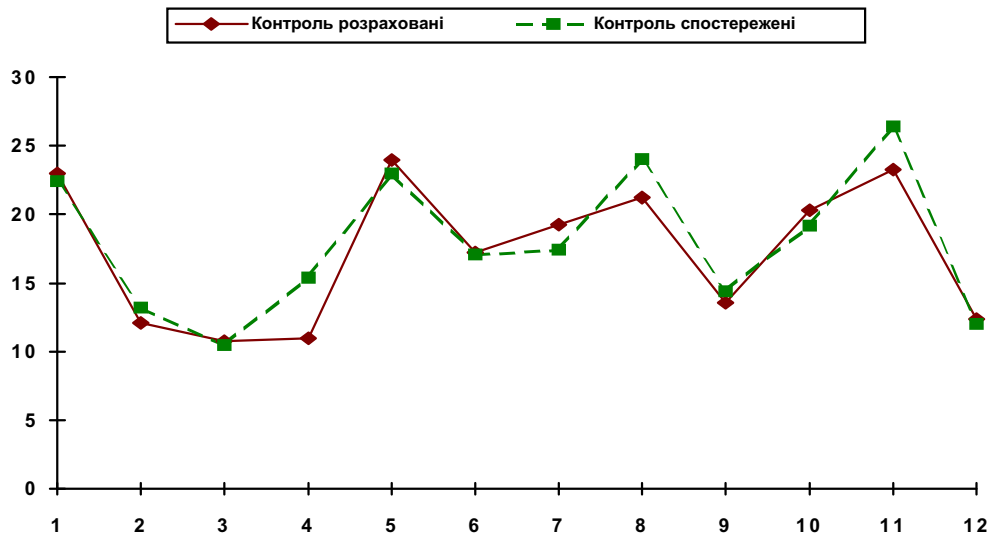


Рис.1 Графіки зміни чисельності органотрофних мікроорганізмів на контрольній ділянці в точках вимірювання (по три виміри за вегетаційний період 1993-96 рр.).

Моделі для чисельності мікроорганізмів у забрудненому ґрунті були побудовані з урахуванням чисельності органотрофних мікроорганізмів у ґрунті контрольної ділянки.

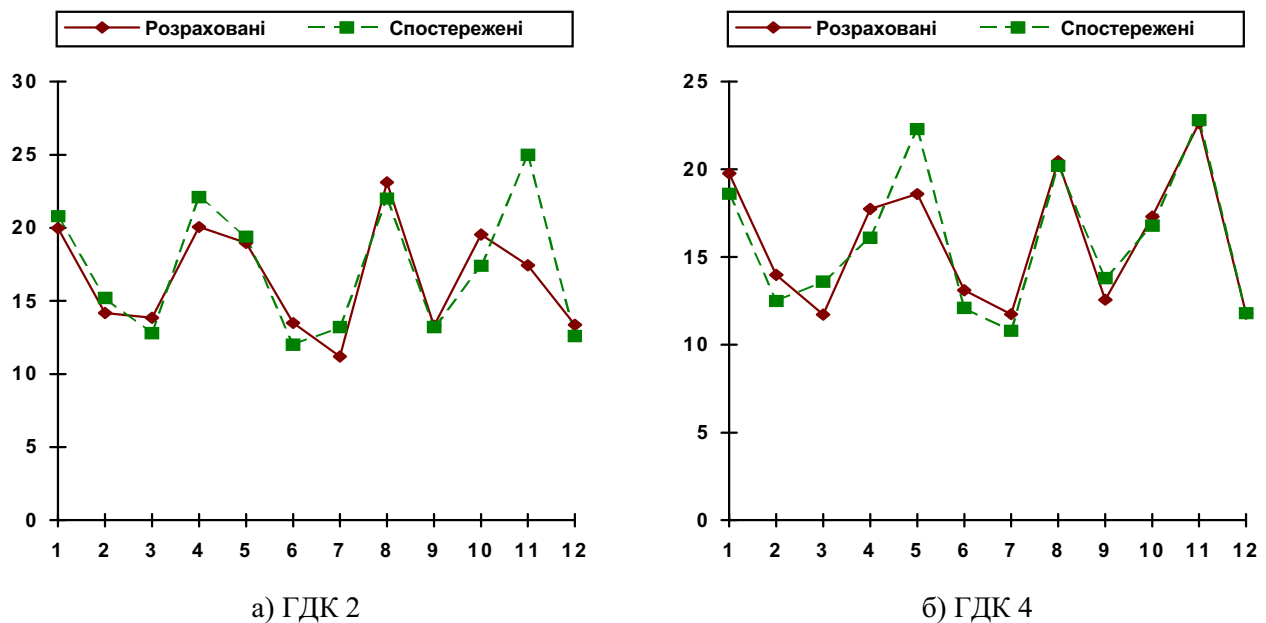
Модель динаміки чисельності органотрофів у ґрунті, забрудненому міддю, має вигляд:

$$Y = 105 + 1.138X_1 - 0.0539X_2 - 2.002X_3 - 1.019X_4 - 0.0545X_6 - 0.046X_7 + 0.287X_9 - 62.003X_{12}$$

з такими показниками її якості : СКО=1.431, R =3.505, R_{відн} =24,68%

Як бачимо, на чисельність органотрофів у забрудненому ґрунті істотно впливали всі показники вмісту важких металів, окрім X5 (концентрація ртуті), неінформативними були також змінні X8, X10 та X11, що характеризують температуру і вологість повітря.

Аналіз графіків на рис.2 показує, що в більшості точок експериментальні й розраховані за моделлю дані збігаються, тобто модель адекватно відображає динаміку чисельності мікроорганізмів. Про це свідчать три останні точки на графіку, які складала «екзаменаційну» вибірку для перевірки моделі в режимі прогнозування. Окремі розходження можна пояснити просторовою неоднорідністю ґрунту та іншими умовами, які могли викликати незакономірні коливання чисельності.



а) ГДК 2

б) ГДК 4

Рис. 2 Графіки зміни чисельності органотрофних мікроорганізмів на забрудненій міддю ділянці в точках вимірювань (по три вимірювання за вегетаційний період 1993-1996 рр.).

Чисельність органотрофів у ґрунті, забрудненому ртуттю, описується рівнянням:

$$Y = 17.852 + 0.872X_1 + 22.989X_2 + 0.154X_3 + 1.162X_4 - 34.492X_5 + 0.153X_8 - 0.096X_{11}$$

Ця модель має такі характеристики якості: СКП = 1.469, R = 3.369, R_{відн} = 24.96%.

Моделі чисельності органотрофів у ґрунті ділянок, забруднених міддю або ртуттю, мали як спільні, так і відмінні між собою інформативні змінні. Так, у ґрунті, забрудненому ртуттю, на чисельність органотрофних мікроорганізмів істотно впливали такі змінні, як концентрація міді, стронцію, кадмію, ртуті, середня температура та вологість повітря, а неінформативними були змінні, що характеризували концентрацію рухомих форм свинцю, а також температуру і вологість повітря.

Якість моделі ілюструє рис.3: видно, що за дози забруднення ртуттю 2 ГДК модель адекватно описує динаміку чисельності мікроорганізмів, про що свідчать три останніх контрольних точки.

За більшої дози забруднення (4 ГДК) спостерігається відхилення експериментальних даних від розрахованих за отриманою моделлю як на відрізку спостережень, так і в екзаменаційних точках. Таке відхилення можна пояснити тим, що за високої дози забруднення ртуттю (яка токсичніша для мікроорганізмів, ніж мідь) спостерігається втрата стабільності розвитку органотрофів, вони чутливіше реагують не тільки на ті фактори, що введені в модель, а й ті, що не були враховані при математичному моделюванні. В даному випадку потрібно розширювати кількість змінних або будувати нелінійну модель.

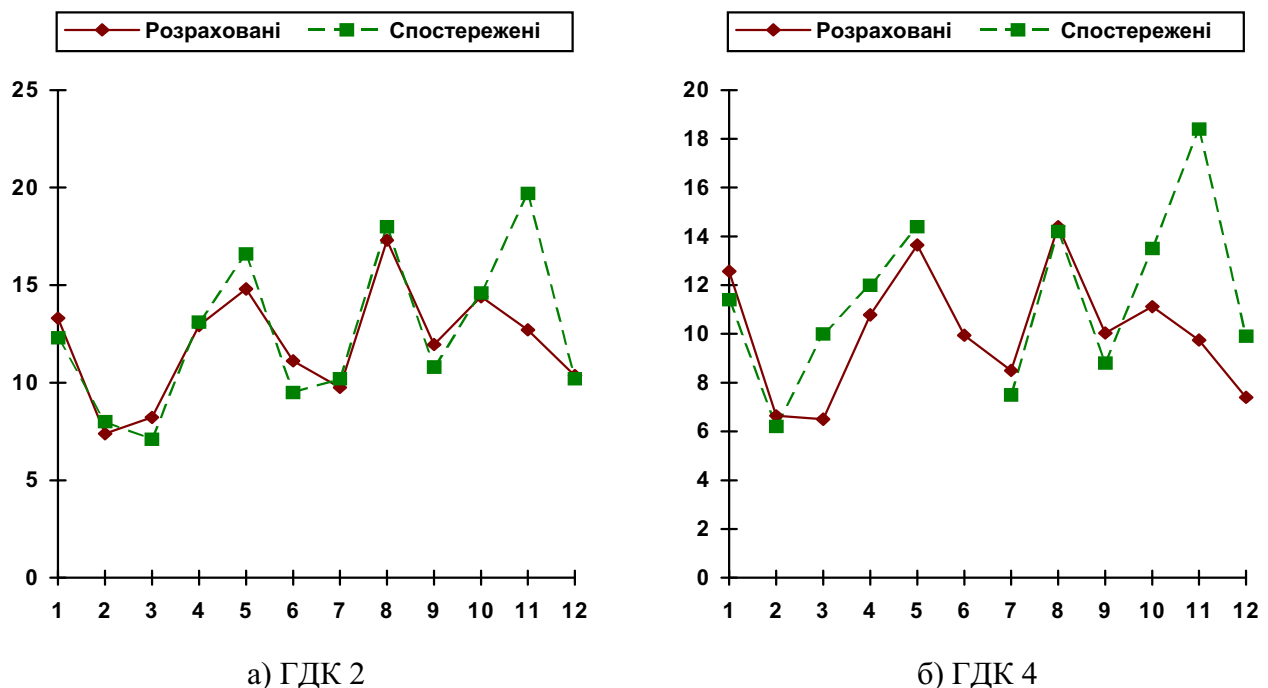


Рис.3 Графіки зміни чисельності органотрофних мікроорганізмів на забрудненій ртуттю ділянці в точках вимірювання (по три вимірювання за вегетаційний період 1993-1996 рр.).

Спільне використання моделей чисельності мікроорганізмів у контрольному і забрудненому ґрунтах дозволяє вирішити задачу відновлення відсутніх або проміжних даних. Наприклад, можна відновити подекадну динаміку чисельності мікроорганізмів протягом усього періоду проведення моніторингу, незважаючи на те, що спостереження проводили лише 3 рази за вегетаційний період. Можливість відновлення даних забезпечена тим, що моделі достатньо точно відповідають експерименту в точках спостережень, і, крім того, у проміжних точках наявні «опорні» дані про метеорологічні умови, які істотно впливають на динаміку чисельності мікрофлори. Для прикладу на рис.4 та 5 наведено графіки відновлених значень чисельності органотрофів на ділянках, забруднених міддю і ртуттю в дозі 2 і 4 ГДК. Для розрахунку цих результатів послідовно використовуються моделі для контрольної та забруднених ділянок.

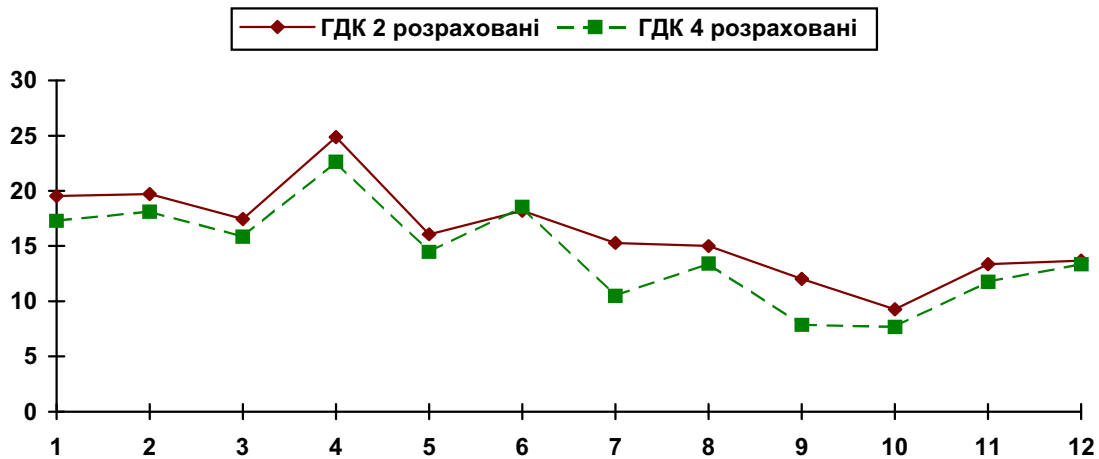


Рис.4 Відновлений графік зміни чисельності органотрофних мікроорганізмів по декадах за вегетаційний період у 1996 р. на ділянці, забрудненій міддю.

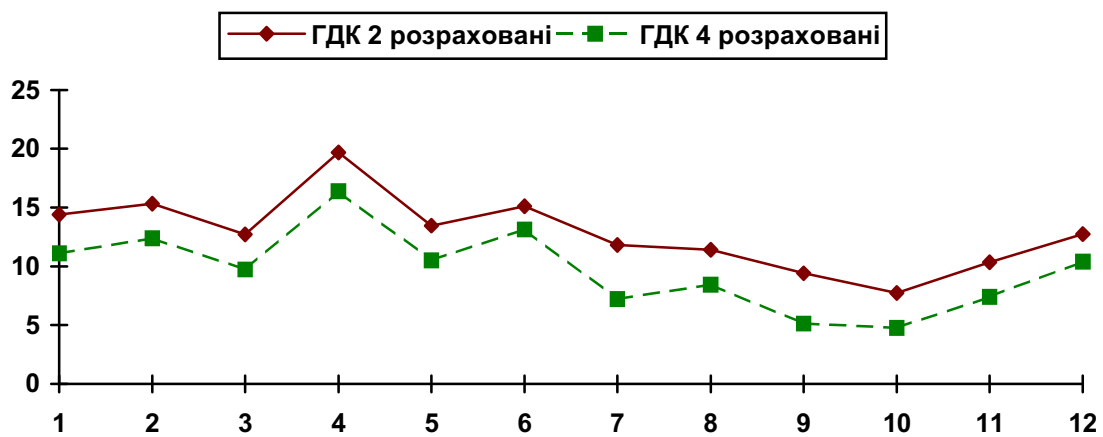


Рис.5 Відновлений графік зміни чисельності органотрофних мікроорганізмів по декадах за вегетаційний період у 1996 р. на ділянці забрудненому ртуттю.

Висновки. Проведені дослідження показали принципову можливість формалізації даних екологічних спостережень шляхом побудови математичних моделей. Для моделювання чисельності мікроорганізмів у ґрунті ефективним є застосування індуктивного підходу на основі МГУА. Отримані моделі достатньою мірою збігаються з експериментальними даними, що дає можливість використовувати їх у системі екологічного моніторингу для оцінки ступеня забруднення ґрунту, відновлення проміжних або відсутніх даних і оперативного прогнозування розвитку мікроорганізмів у певних екологічних умовах.

Список літератури

1. Андреюк Е.И., Иутинская Г.А., Петруша З.В. Гомеостаз микробных сообществ почв, загрязненных тяжелыми металлами // Микробиол. журн.-1999.- **61**, №6.-С.15-21.
2. Іутинська Г.О., Петруша З.В. Резистентність ґрунтових мікроорганізмів до забруднення ґрунтів важкими металами // Микробиол.журн.-1999.- **61**, №5.-С.72-77.
3. Степашко В.С., Коппа Ю.В. Опыт применения системы АСТРИД для моделирования экономических процессов по статистическим данным // Кибернетика и вычислительная техника.-1998.-вып.117.-С.24-31.