

УДК 556+550.423+519.25

**А.В. Клипа, М.Н. Жуков**

*Київський національний університет імені Тараса Шевченка,  
вул. Володимирська, 64/13, Київ 01601, Україна,  
e-mail: Klypaaa@bigmir.net, jukov42@ukr.net*

## **МОДЕЛЬ ЗРІЗАНОГО РОЗПОДІЛУ ВМІСТУ ГІДРОХІМІЧНИХ ДАНИХ**

Описано моделювання розподілів вмісту мікроелементів у природних водах на підставі зрізаного нормального та зрізаного логнормального законів. Застосування моделей для обробки гідрохімічних даних дало змогу розв'язати метрологічну проблему, пов'язану з недостатньою чутливістю вимірювального приладу. Доведено, що моделі зрізаного розподілу дають точніші оцінки числових характеристик розподілів і, як наслідок, інтегральних характеристик екологічного стану водних об'єктів. Розроблений метод може бути використаний для розв'язання подібних проблем під час дослідження інших геологічних об'єктів. Усі необхідні розрахунки можна проводити в автоматизованому режимі, застосовуючи створений програмний модуль.

**Ключові слова:** гідрохімічні дані, модель зрізаного розподілу, ймовірно-статистичний розподіл, інтегральні характеристики, мікроелементи, природні води.

**Постановка проблеми.** Метрологічна проблема недостатньої чутливості вимірювального приладу є суттєвою перешкодою для дослідження гідрогеологічних об'єктів, особливо, коли йдеться про дослідження закономірностей розподілів вмісту мікроелементів у природних водах. Як наслідок, частину некондиційних даних неможливо використати для проведення статистичної обробки. Ігнорування спостережень, які є нижчими за поріг чутливості, спотворює реальний розподіл, що призводить до хибних оцінок. Подібна проблема виникла під час проведення дослідження щодо з'ясування універсальної моделі розподілів вмісту хімічних елементів у водах Полтавської області.

Для пошуку універсальної моделі розподілів вмісту мікроелементів (Ba, Co, Cr, Cu, Mn, Mo, Ni, Pb, Sr, Ti, V, Zn) у природних водах було використано об'ємну базу даних (560 проб) з результатами гідролітохімічної зйомки масштабу 1 : 500 000, яка була проведена на території Полтавської області Казенним підприємством (КП) “Кіровгеологія” у період 1991–1993 рр. Вміст мікроелементів визначили за допомогою напівкількісного спектрального аналізу у спеціальній хімічній лабораторії КП “Кіровгеологія”.

Цей вид аналізу має значну похибку і недостатньо чутливий для визначення вмісту мікроелементів, особливо таких, які містяться у воді у малих концентраціях ( $n \cdot 10^{-4} \div 10^{-3}$  мг/дм<sup>3</sup>). Ця обставина була головною причиною того, що база даних містила значну кількість вимірювань, особливо вмісту Zn, Co, Pb, Mo, Ba, нижчих за поріг чутливості, та так званих “слідів” – елемент зафіксовано, але точний вміст відсутній.

**Аналіз останніх досліджень і публікацій.** Розв’язанню проблеми щодо встановлення закономірностей розподілів вмісту хімічних компонентів у природних водах і знаходження універсальної моделі для адекватного опису таких розподілів присвячені праці А.І. Гавришина та С.І. Смірнова, опубліковані ще у другій половині ХХ ст. Питання з того часу залишилося невирішеним. Здебільшого зазначені та інші автори для розв’язання проблеми використовували нормальну та логарифмічно-нормальну моделі, які виявилися частково дієздатними. Перевірка адекватності класичних моделей на такому обсязі даних [1, 2] показала, що в окремих випадках дещо кращою була нормальна модель, в інших, навпаки, – логнормальна, але результати перевірки на основі критеріїв адекватності та порівняння оцінок інтегральних характеристик, отриманих на підставі моделей, з прямою оцінкою за частотою свідчили про недостатню ефективність згаданих моделей. Причиною цього могли бути некондиційні дані.

Для перевірки цього припущення було застосовано зрізаний нормальний та логнормальний розподіли. Прикладів розв’язання задачі у практиці статистичного аналізу геологічних даних автори не знайшли, за винятком теоретичного обґрунтування такої можливості, поданого професором М. Жуковим [3].

**Виклад основного матеріалу дослідження з повним обґрунтуванням отриманих наукових результатів.** Першим етапом масштабного дослідження, яке дало змогу виявити проблему, була перевірка можливості застосування класичних законів розподілу. У середовищі MathCad було створено автоматизований програмний модуль, із використанням якого було проведено необхідні статистичні розрахунки та безпосереднє моделювання розподілів вмісту мікроелементів у природних водах. Значення вмісту мікроелементів, нижчі за поріг чутливості, було замінено на значення, що дорівнювали половині порогу чутливості приладу. Результати вимірювань показано на рис. 1.

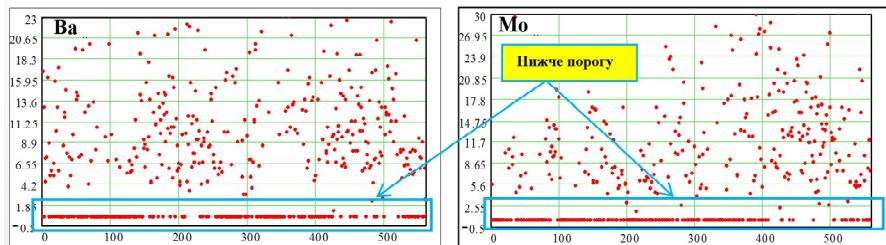


Рис. 1. Відображення даних вмісту Ва та Мо у поверхневих водах Полтавщини 1991–1993 рр. (виділено “некондиційні дані”, що були замінені значеннями нижнього порогу чутливості вимірювального приладу): вертикальна ось – концентрація елемента (Ва –  $n \cdot 10^{-2}$ , Мо –  $n \cdot 10^{-3}$ ); горизонтальна – порядкові номери спостережень

Доведено, що неврахування таких проб, їх відкидання або необгрунтована заміна призводять до спотворення реального розподілу вмісту мікроелементів у поверхневих водах області (рис. 2).

На рис. 2 чітко видно різницю між розподілами в разі відкидання проб, нижчих за поріг чутливості. Графік щільності, позначений пунктирною лінією 2, знаходиться значно правіше (нормальна модель, рис. 2, а) та значно вище (логнормальна модель, рис. 2, б) від графіка, позначеного пунктирною лінією 1. Наслідком таких розбіжностей є завищені оцінки інтегральних екологічних характеристик (за відкидання “некондиційних проб”), що не відповідає дійсності. Так, площа, на якій вміст молібдену перевищує  $21 \cdot 10^{-3}$  мг/дм<sup>3</sup> у разі заміни, дорівнює: за нормальною моделлю 11,63 %; за логнормальною – 9,4 %. У випадку відкидання таких проб вона дорівнює 28,98 і 19 % відповідно. Отже, оцінки відрізняються більш ніж удвічі.

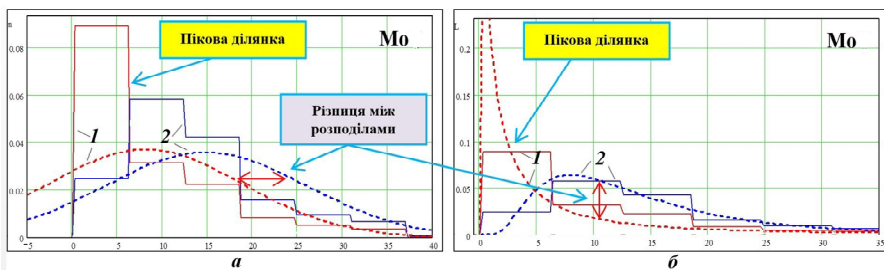


Рис. 2. Моделювання розподілу вмісту молібдену у природних водах за заміни даних нижче порогу чутливості (1) та без них (2) на основі нормальної (а) та логнормальної моделей (б): суцільна лінія – непараметрична щільність; пунктирна лінія – параметрична щільність

В результаті заміни “некондиційних вимірювань” на значення, що дорівнювали нижній межі чутливості приладу, отримали спотворення, що супроводжувалося “піковими ділянками” (рис. 2). Для усунення проблеми було з’ясовано можливості застосування моделей зрізаного нормального та логнормального розподілів. Такий розподіл може бути зрізаним зліва (за наявності значень, менших за визначений рівень) або справа (значення, більші за визначений рівень). Розглянемо модель розподілу, зрізаного зліва. Ступінь зрізу нормального закону розподілу дорівнює:

$$\alpha = \Phi\left(\frac{a - m}{\sigma}\right),$$

де  $\alpha$  – рівень зрізу,  $a$  – точка зрізу;  $\Phi(\dots)$  – функція  $(0, 1)$  – нормального розподілу;  $m$  – математичне сподівання;  $\sigma$  – середній квадратичний відхил. Щільність зрізаного зліва нормального розподілу має вигляд

$$p(x) = \begin{cases} \left[\sqrt{2\pi}\sigma(1-\alpha)\right]^{-1} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right), & \text{якщо } x \geq a. \\ 0, & \text{якщо } x < a \end{cases}$$

Щільність зрізаного зліва логнормального розподілу

$$p(x) = \begin{cases} \left[\sqrt{2\pi}\delta x(1-\alpha)\right]^{-1} \exp\left(-\frac{(\ln x - \mu)^2}{2\delta^2}\right), & \text{якщо } x \geq a, \\ 0, & \text{якщо } x < a. \end{cases}$$

Загальний вигляд графіків щільності зрізаного зліва нормального та логнормального розподілів показано на рис. 3.

Автори розробили метод оцінки параметрів обох моделей та реалізували його для обробки гідрогеохімічних даних. Суть методу полягає у тому, що його використання дає змогу екстраполювати щільність розподілу ліворуч від точки зрізування на основі оцінки параметрів за наявними спостереженнями у правій частині. Екстраполяцію ліворуч від точки зрізування нормального та логнормального розподілів здійснювали на основі методів симетризації; найменших квадратів (МНК); мінімуму критерію Колмогорова.

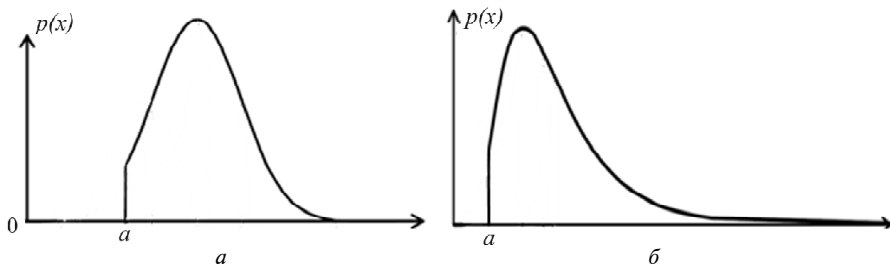


Рис. 3. Заміна щільності зрізаного зліва у точці  $a$  розподілу:  $a$  – нормальний закон розподілу;  $b$  – логарифмічнонормальний закон розподілу

Основною вимогою для побудови графіків функції розподілу за допомогою зрізаних моделей та отримання адекватних результатів було те, щоб кількість даних, нижчих за поріг чутливості, не перевищувала 50 %. У зв'язку з цим дослідження елементів Co, Cr, Pb та Zn не проводили.

В результаті застосування модуля отримано оцінки функції розподілу вмісту мікроелементів, де ліва частина кривої продовжена на підставі зазначених методів (рис. 4), на прикладі вмісту ванадію.

Для зручності сприйняття на рис. 4 показано оцінки функції розподілу на підставі найефективніших методів для ванадію – нормальна модель за оцінками МНК і нормальна модель на основі мінімуму критерію Колмогорова.

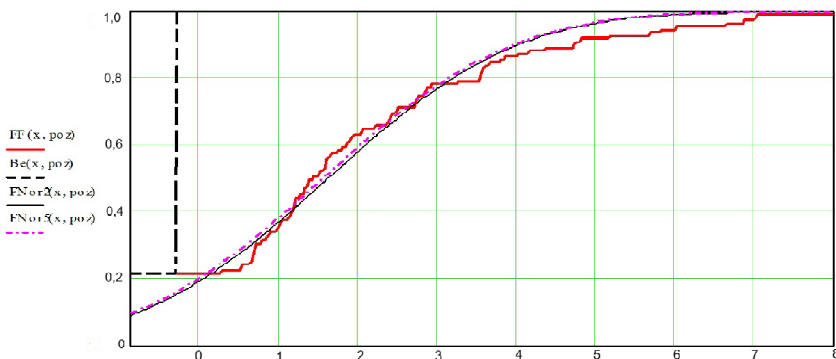


Рис. 4. Функції розподілу вмісту ванадію у природних водах Полтавщини: FF – вибірка; FNor2 – нормальна за оцінками МНК; FNor5 – мінімум критерію Колмогорова в інтервалі.

Примітка: Лінія графіка вибіркової функції ліворуч від першої сходинки означає, що там ця функція невизначена

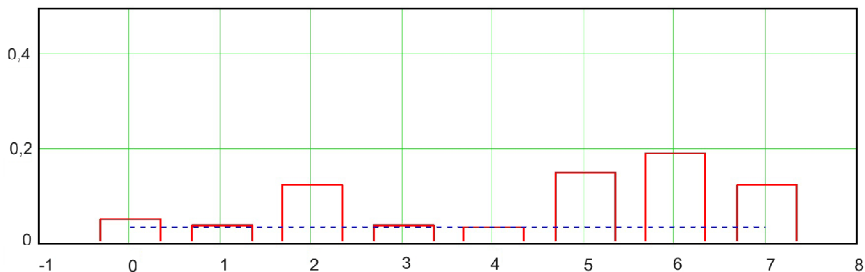


Рис. 5. Графічне відображення критерію Колмогорова. Для гіпотези нормального закону з оцінками: 0 – методом симетризації, 1 – МНК, ненормовані відхилення, 2 – МНК, нормовані відхилення, 3 – мінімуму критерію Колмогорова по точках спостережень, 4 – мінімуму критерію Колмогорова в інтервалі; для гіпотези логнормального закону з оцінками: 5 – методом симетризації, 6 – МНК, ненормовані відхилення, 7 – мінімуму критерію Колмогорова. Пунктирною лінією показано рівень критерію тієї моделі, у якої він виявився мінімальним

Для всіх мікроелементів, крім Co, Cr, Pb та Zn, було проведено ідентичну обробку. З метою полегшення сприйняття одержаних результатів для подальшої демонстрації вибрано результати, отримані цією процедурою за критерієм Колмогорова, ненормовані відхилення, оскільки оцінки, за іншими критеріями (МНК, нормовані та ненормовані відхилення, Колмогорова, нормовані відхилення) різняться несуттєво. Результати перевірки зазначеним критерієм зображено у графічному вигляді (рис. 5) на прикладі вмісту ванадію.

На рис. 5 видно, що оптимальною виявилася нормальна модель з оцінками, отриманими за методом мінімуму критерію Колмогорова (функцію розподілу, яку побудовано зазначеним методом, зображено на рис. 4). Дещо гіршу апроксимацію демонструють нормальні моделі з оцінками МНК (ненормовані відхилення) та оцінками мінімуму критерію Колмогорова по точках спостережень. Порівняння оцінок числових характеристик розподілів мікроелементів показало помітно вищу якість оцінок, одержаних на підставі моделей зрізаних розподілів.

**Висновки.** Розв'язано проблему адекватного оцінювання параметрів статистичних розподілів вмістів мікроелементів у природних водах за умов відсутності спостережень, нижчих за поріг чутливості. За створеним програмним модулем отримано стійкі оцінки функцій та щільностей розподілів без того спотворення, яке зазвичай виявляли у подібних ситуаціях. Запропонований метод дає змогу значно уточнювати оцінки інтег-

ральних характеристик, що мають суттєве значення для вирішення завдань екологічного моніторингу водних об'єктів.

Розроблений метод може бути використаний для розв'язання подібних проблем під час досліджень інших геологічних об'єктів і показників їх речовинного складу.

1. *Імовірнісна* оцінка забруднення природного середовища на основі моделювання розподілів вмісту хімічних елементів (на прикладі поверхневих вод Полтавської області) / М.Н. Жуков [та ін.] // Вісник Київського національного університету імені Тараса Шевченка. Геологія. – 2012. – № 58. – С. 52–57.
2. Жуков М.Н. Оцінка ефективності моделей розподілу вмісту важких металів у поверхневих водах (на прикладі Полтавської області) / М.Н. Жуков, А.В. Клипа, І.Р. Стахів // Вісник Київського національного університету імені Тараса Шевченка. Геологія. – 2012. – № 59. – С. 51–54.
3. Жуков Н.Н. Вероятностно-статистические методы анализа геолого-геофизической информации / Н.Н. Жуков. – К.: Вища шк., 1975. – 304 с.

### **Модель усеченного распределения содержания гидрохимических данных**

А.В. Клипа, Н.Н. Жуков

*Київський національний університет імені Тараса Шевченка,  
ул. Владимирская, 64/13, Киев 01601, Украина,  
e-mail: Klyraaa@bigmir.net, jukov42@ukr.net*

Описано моделювання розподілів вмісту мікроелементів в природних водах на основі усеченого нормального і логнормального розподілів. Застосування моделей для обробки гідрохімічних даних дозволило вирішити метрологічну проблему, пов'язану з недостатньою чутливістю вимірювального приладу. Доведено, що моделі усеченого розподілу дають більш точні оцінки числових характеристик розподілів і, як наслідок, інтегральних характеристик екологічного стану водних об'єктів. Розроблений метод може бути використаний для вирішення подібних проблем при дослідженні інших геологічних об'єктів. Застосування створеного програмного модуля дає можливість виробляти всі необхідні розрахунки в автоматизованому режимі.

**Ключевые слова:** гидрохимические данные, модель усеченного распределения, вероятностно-статистическое распределение, интегральные характеристики, микроэлементы, природные воды.

### **Model truncated distribution of content hydrochemical data**

A. Klypa, M. Zhukov

*Taras Shevchenko National University of Kyiv, 64/13 Volodymyrska Str., Kyiv 01601, Ukraine, e-mail: Klypaaa@bigmir.net, jukov42@ukr.net*

The object of the study was modeling distributions of trace elements in natural waters based on the truncated normal and lognormal distribution. To accomplish this goal was necessary to develop a method for estimating the parameters of both models and implement it for hydrochemical data processing. The method makes possible to extrapolate the density distribution curve in the left side of the cutoff point based on evaluation of the available parameters observations on the right side. Extrapolation in the left side of the cutoff point has been done based on symmetric method, method of the least squares and method of Kolmogorov. As a result, the method for estimating the parameters of both models was developed and implemented for hydrochemical data. The use of these models to handle hydrochemical data helped to solve metrological problem associated with insufficient sensitivity of measuring device. Proved that the truncated distribution models gives accurate estimates of numerical characteristics of distributions as a result integral characteristics of the ecological state of water bodies are more accurate. The software module provides a stable evaluation of function and density of distributions without distortion, which is usually practiced in similar situations. The method can be used to solve such problems in the study of other geological objects. The software module allows make all the necessary calculations in automatic mode.

**Keywords:** hydrochemical data, model truncated distribution, probability and statistical distribution, integral characteristics, trace elements, natural waters.

*Надійшла до редакції 10.06.2015 р.*