



УДК 556,531:556,18:311;

Член-корреспондент НАН України **В. И. Осадчий, Л. А. Ковальчук, Н. Н. Осадчая**

## Исследование структуры загрязнения водных объектов посредством процессов Ито

*Осуществлено достовірне розділення спостережених концентрацій хімічних компонентів поверхневих вод на природну і антропогенну складові шляхом розмежування двох гіпотез для процесів Іто. Експериментально підтверджено, що в основі зміни концентрацій інгредієнтів лежить процес Вінера. Відкрито шлях дослідженням напрямленості гідрохімічних процесів з допомогою апарату статистики випадкових процесів.*

**Состояние проблемы.** Оценка текущего состояния поверхностных вод и его прогнозирование неразрывно связаны с выделением в растворе веществ или элементов как природного (reference values), так и антропогенного происхождения (антропогенная составляющая). Использование внутри дисциплинарных и традиционных вероятностных методов для решения проблемы показало незначительную результативность классического подхода. В то же время статистика случайных процессов предоставляет необходимое теоретическое обоснование и аппарат оценивания [1].

В настоящем сообщении представлено выполненное нами исследование гидрохимических явлений посредством процессов Ито. В качестве примера выбрана задача о разделении наблюдаемых концентраций химических компонентов воды на природную и антропогенную составляющие, с помощью которой обосновано дополнение классических методов математической статистики методами статистики случайных процессов.

**Материал и методика.** В основу исследования положены материалы ежедневных гидрохимических наблюдений за содержанием нитратных ионов ( $\text{NO}_3^-$ , мгN/дм<sup>3</sup>) в замыкающем створе р. Десны с 1991 по 1994 гг. Выбор нитратных ионов обусловлен тем, что в наблюдаемых концентрациях содержится соизмеримое количество ионов как природного, так и антропогенного происхождения, хотя это существенно усложняет решение задачи. Для статистического анализа использованы аномалии ежедневных величин, вычисленные

---

© В. И. Осадчий, Л. А. Ковальчук, Н. Н. Осадчая, 2014

путем вычитания из фактических концентраций среднемноголетних значений (нормы) соответствующего дня.

Объективность выводов была достигнута использованием общеизвестного в химии метода добавок, реализованного путем числовых экспериментов. Для этого к значениям концентраций имеющегося ряда наблюдений экспертом были внесены числовые добавки исследуемых компонентов. При внесении добавок базировались на преимущественном поступлении  $N-NO_3^-$  от рассредоточенных источников, что установлено анализом многолетних данных мониторинга. Задача решалась в два этапа. В первую очередь в исходном временном ряду наблюдений была выявлена антропогенная составляющая. Затем устанавливались суммарные величины антропогенной составляющей и контрольных добавок в трансформированном ряду. В результате были вычислены величины добавок, известные только эксперту. Условие успешности экспериментов заключалось в том, чтобы коэффициенты корреляции между выявленными величинами добавок и их внесенными значениями не были ниже 0,85.

Известно, что наблюдаемые концентрации гидрохимических ингредиентов  $\xi$  являются суммой концентраций природного  $\omega$  и антропогенного происхождения  $\theta$ . Теоретические основы разделения наблюдаемых концентраций ингредиентов на природную и антропогенную составляющие заложены в аксиоматике теории вероятностей. Изначально полагалось, что величины концентраций антропогенной и природной составляющих пропорциональны их вероятностям. Сами вероятности вычислялись отдельно для природной и антропогенной составляющей на основе законов их статистических распределений, а законы идентифицировались по параметрам статистических распределений [2].

Закон статистического распределения наблюдаемых концентраций  $N_\xi (\mu_\xi, \sigma_\xi, \varepsilon_\xi)$  и закон статистического распределения концентраций антропогенного происхождения  $N_\theta (\mu_\theta, \sigma_\theta, \varepsilon_\theta)$  идентифицировались посредством трехпараметрической экспоненциальной модели [3]:

$$p(x) = \frac{\alpha}{2\lambda\sigma\Gamma(1/\alpha)} \exp\left(-\left|\frac{x-\mu}{\lambda\sigma}\right|^\alpha\right), \quad (1)$$

где  $\lambda = \sqrt{\Gamma(1/\alpha)/\Gamma(3/\alpha)}$ ;  $\Gamma(z)$  — гамма-функция;  $\sigma = \sigma(x)$  — стандартное отклонение;  $\mu = \mu(x)$  — центр распределения,  $\alpha = \alpha(x)$  — некоторая характеристика распределения, однозначно определяющая параметр формы экспоненциальных распределений — эксцесс ( $\varepsilon$ ):

$$\varepsilon = \frac{\Gamma(1/\alpha)\Gamma(5/\alpha)}{(\Gamma(3/\alpha))^2}. \quad (2)$$

Если  $\alpha < 1$ , то модель (1) описывает распределения, близкие по своим свойствам к распределению Коши;  $\alpha = 1$  — модель представляет распределение Лапласа, имеющее более островершинный вид, чем у нормального распределения;  $\alpha = 2$  — модель соответствует нормальному распределению;  $\alpha > 2$  — модель идентифицирует распределения, близкие по своим свойствам к трапецеидальным распределениям;  $\alpha \rightarrow \infty$  — модель отражает равномерное распределение.

Параметры  $\mu_\theta, \sigma_\theta$  закона статистического распределения антропогенной составляющей вычислялись по правилу развертки:

$$\begin{aligned} \mu_\theta &= \mu_\xi - \mu_\omega, \\ \sigma_\theta^2 &= \sigma_\xi^2 - \sigma_\omega^2, \end{aligned} \quad (3)$$

$\varepsilon_\theta$  принимался равным  $\varepsilon_\xi$ .

Статистическое распределение природных концентраций моделировалось посредством закона Гаусса  $N_\omega(\mu_\omega, \sigma_\omega)$ , а параметры  $\mu_\omega, \sigma_\omega$  вычислялись по той части ранжированного временного ряда  $\xi$ , которая представлена природными концентрациями со свойственным им распределением Гаусса. Для их выделения ранее было предложено интуитивное решение — определение выборочных значений эксцесса [2]. Когда эксцесс становился равным нулю, то очевидно, что подвергшаяся анализу часть временного ряда имела распределение Гаусса. Если же выборочные значения эксцесса отличались от нуля, то ряд считался смешанным, т. е. наблюдаемые концентрации ингредиентов были обусловлены действием как природных, так и антропогенных факторов. Однако возникала неопределенность, когда выборочные значения эксцесса совершали хаотичные колебания или приближались к нулевому значению и оставались такими на протяжении значительной части ранжированного ряда.

Для устранения неопределенности нами был использован аппарат статистики случайных процессов [1]: при выделении природной части ранжированного временного ряда  $\xi_t$  определяли соответствующий марковский момент  $\bar{\tau}(\xi)$ . В этой связи были сформулированы две гипотезы относительно наблюдаемого процесса  $\xi$ :

$$H_0: \quad d\xi_t = dW_t, \quad \xi_0 = 0, \quad (4)$$

$$H_1: \quad d\xi_t = \theta_t dt + dW_t, \quad \xi_0 = 0, \quad (5)$$

где  $W_t$  — процесс Винера.

Другими словами, задача была сведена к различению двух простых гипотез для процессов Ито относительно наличия в наблюдаемых концентрациях ингредиента антропогенной составляющей (гипотеза  $H_1$ ) или ее отсутствия (гипотеза  $H_0$ ). Затем была введена функция заключительного решения  $\delta$ : если  $\delta(\xi_t) = 0$ , принималось решение о справедливости гипотезы  $H_0$ , если же  $\delta(\xi_t) = 1$ , то верной считалась гипотеза  $H_1$ . Кроме того, учитывались вероятности ошибок первого рода  $\alpha = P_1(\delta(\xi) = 0)$  и второго рода  $\beta = P_0(\delta(\xi) = 1)$ . В формальном изложении алгоритм имеет следующий вид:

$$\bar{\tau}(\xi) = \inf\{t: \lambda_t(\xi)\}, \quad (6)$$

$$\bar{\delta}(\xi) = 0, \quad \lambda_{\bar{\tau}(\xi)} \geq A, \quad \bar{\delta}(\xi) = 1, \quad \lambda_{\bar{\tau}(\xi)} \leq B. \quad (7)$$

Здесь

$$A = \log \frac{\beta}{1 - \alpha}, \quad (8)$$

$$B = \log \frac{1 - \beta}{\alpha}, \quad \alpha = 0, 1, \quad \beta = 0, 2, \quad (9)$$

$$\lambda_t(\xi) = \int_0^t m_0(\xi) d\xi_s - \frac{1}{2} \int_0^t m_s^2(\xi) ds, \quad (10)$$

где  $m_s$  — математическое ожидание.

**Обсуждение результатов исследований.** Особое внимание следует уделить вопросу совместного представления статистических распределений химических ингредиентов водных объектов как законами статистических распределений временных рядов, так и законами случайных процессов. Приняв гипотезы  $H_0$  и  $H_1$ , мы тем самым признали, что основой

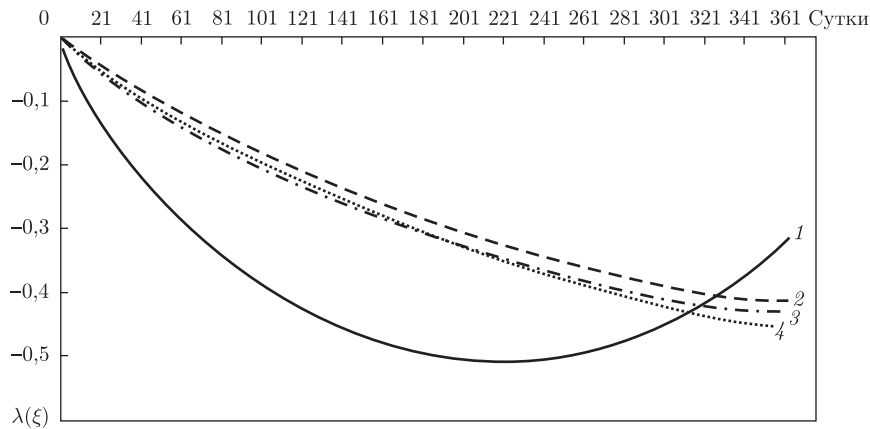


Рис. 1. Динамика статистики  $\lambda_t(\xi)$  за 1991 г. (1); 1992 г. (2); 1993 г. (3); 1994 г. (4)

гидрохимических процессов является процесс Винера. В таком случае за сколь угодно малый промежуток времени происходит бесконечное множество реализаций гидрохимического процесса. Временной же ряд наблюдаемых концентраций химической компоненты воды, хотя и является единственной реализацией из бесконечного множества реализаций, тем не менее позволяет определять параметры статистического распределения временного ряда. Вычисленные значения статистики  $\lambda_t(\xi)$  предоставили возможность однозначно проверить гипотезы  $H_0$  и  $H_1$ .

Для ранжированного ряда наблюдений за 1991 г. момент  $\bar{\tau}(\xi_1)$  соответствовал минимуму кривой  $\lambda_t(\xi)$  (рис. 1). Другими словами, момент  $\bar{\tau}(\xi_1)$  разделил ряд  $\xi_1$  на две части, где первая соответствовала распределению фоновых концентраций нитратных ионов, т. е. гипотезе  $H_0$ , а вторая — соответствовала смешанному распределению концентраций  $N-NO_3^-$  природного и антропогенного происхождения, т. е. гипотезе  $H_1$ . По значениям первой части ряда  $\xi_1$  были вычислены параметры распределения концентраций природной составляющей  $\mu_\omega, \sigma_\omega$ . Как и следовало ожидать, математическое ожидание  $\mu_\omega$  оказалось равным нулю, что является фундаментальной основой процесса Винера. На основании  $\mu_\omega, \sigma_\omega$  посредством модели Гаусса был идентифицирован закон статистического распределения концентраций нитратных ионов природного происхождения. На основании  $\mu_\theta, \sigma_\theta, \varepsilon_\theta$ , вычисленных согласно (3), был идентифицирован, посредством модели (1), закон статистического распределения концентраций нитратных ионов антропогенного происхождения (рис. 2). Хотя законы статистических распределений как природной, так и антропогенной составляющих сосредоточены в области низких концентраций, но у плотности вероятности антропогенной составляющей центр распределения имеет большее значение и выше значение дисперсии (см. рис. 2, а).

Для ранжированных рядов наблюдений за 1992, 1993, 1994 гг. момент  $\bar{\tau}(\xi_1)$ , как минимум кривой  $\lambda_t(\xi)$ , совпадал с концами рядов (см. рис. 1). Это свидетельствовало о том, что ряды наблюдаемых концентраций нитратных ионов не содержали в себе концентраций только природного происхождения, т. е. гипотеза  $H_0$  не подтвердилась. В этой связи центры распределений и дисперсии законов статистических распределений концентраций как природного, так и антропогенного происхождения совпадали (см. рис. 2, б).

Применение статистики  $\lambda_t(\xi)$  для различения гипотез  $H_0$  и  $H_1$  оказалось более эффективным, чем использованное нами ранее выборочное значение эксцесса  $\varepsilon$ . Хотя коэффициен-

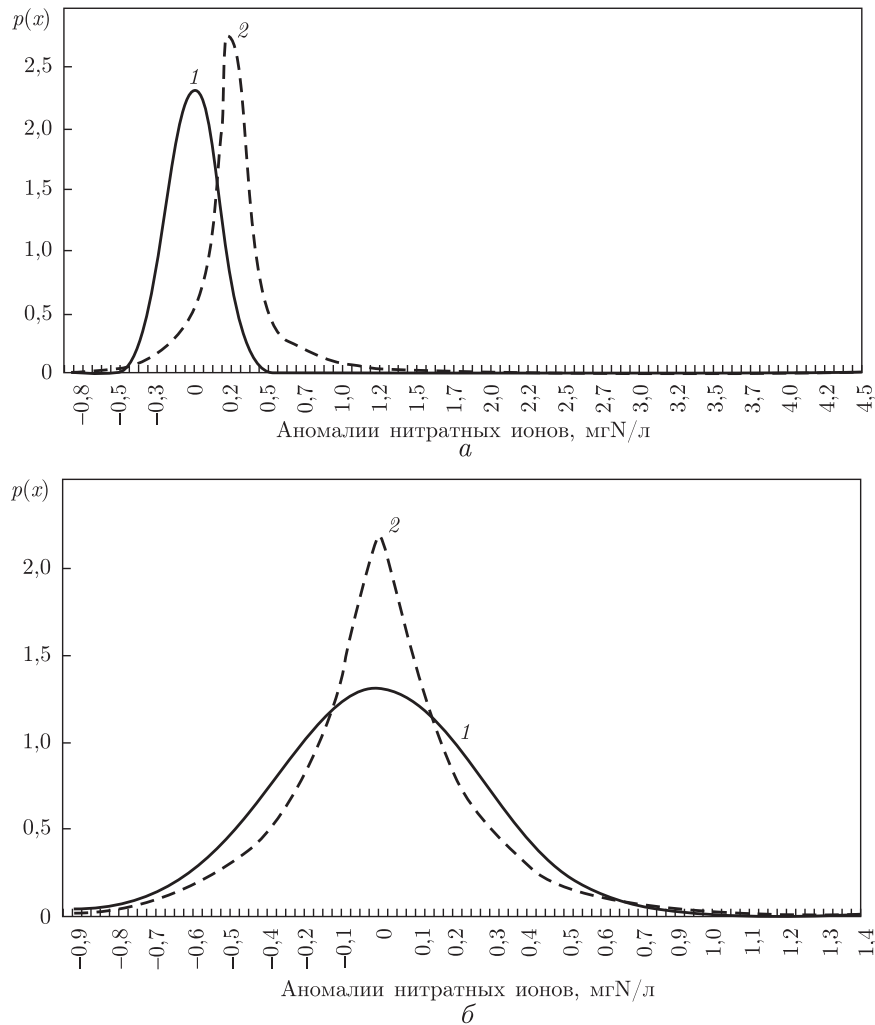


Рис. 2. Законы статистических распределений аномалий концентраций нитратных ионов: 1 — природная составляющая; 2 — антропогенная составляющая за 1991 г. (а) и 1994 г. (б)

ты корреляции между контрольными и выявленными добавками, в случаях применения  $\lambda_t(\xi)$  или  $\varepsilon$ , различались в пределах 10% (табл. 1), применение статистики  $\lambda_t(\xi)$  устранило неопределенности.

Более того, полученные результаты экспериментально подтвердили посылку о том, что в основе гидрохимических явлений лежит процесс Винера, а также открыли путь к исследованию направленности гидрохимических процессов с помощью аппарата статистики случайных процессов.

Таблица 1. Коэффициенты корреляции  $r$  между контрольными и выявленными добавками концентраций нитратных ионов в случаях применения статистики  $\lambda_t(\xi)$  или выборочных значений эксцесса  $\varepsilon$

Статистика	Коэффициенты корреляции по годам				Среднее значение $r$
	1991	1992	1993	1994	
$\varepsilon$	0,85	0,92	0,91	0,88	0,89
$\lambda_t(\xi)$	0,94	0,88	0,96	0,85	0,91

1. *Литцер Р. Ш., Ширяев А. Н.* Статистика случайных процессов. – Москва: Наука, 1974. – 696 с.
2. *Осадчий В. И., Ковальчук Л. А.* Теоретические основы вероятностно-статистического разделения величины показателей химического состава водных объектов на природную и антропогенную составляющие // Доп. НАН України. – 2013. – № 4. – С. 97–103.
3. *Булашев С. В.* Статистика для трейдеров. – Москва: Компания Спутник+, 2003. – 245 с.

Український гідрометеорологічний інститут  
ГСЧС України і НАН України, Київ

Поступило в редакцію 14.02.2014

Член-кореспондент НАН України **В. І. Осадчий, Л. А. Ковальчук,  
Н. Н. Осадча**

### **Дослідження структури забруднення водних об'єктів за допомогою процесів Іто**

*Виконано вірогідне розділення спостережених концентрацій хімічних компонентів поверхневих вод на природну та антропогенну складові шляхом розрізнення двох гіпотез для процесів Іто. Експериментально підтверджено, що в основі змін концентрацій інгредієнтів лежить процес Вінера. Відкрито шлях дослідженням спрямованості гідрохімічних процесів за допомогою апарата статистики випадкових процесів.*

Corresponding Member of the NAS of Ukraine **V. I. Osadchiy, L. A. Kovalchuk,  
N. N. Osadcha**

### **Study of the structure of a pollution of water bodies through the Itô process**

*The significant separation of the observed concentrations of chemical components of surface waters into the natural and anthropogenic components by testing two hypotheses for Itô processes is implemented. It is experimentally confirmed that the basis of changes in the concentrations of ingredients is the Wiener process, which opens the way to research the directedness of hydrochemical processes, by using the statistics of random processes.*