

УДК 504.4.054

О. А. Проскурнин, канд. техн. наук, ст. научн. сотр., **С. А. Смирнова**
(УКРНИИЭП)

ИСПОЛЬЗОВАНИЕ МЕТОДА BOOTSTRAP ДЛЯ ОЦЕНКИ ВЛИЯНИЯ СБРОСА СТОЧНЫХ ВОД НА КАЧЕСТВО ВОДЫ ВОДНОГО ОБЪЕКТА

В статье рассматривается использование непараметрического статистического метода bootstrap для оценки влияния сброса сточных вод на качество воды водного объекта. Приводится пример расчета для предприятия-водопользователя ГКП «Бердянский водоканал».

Ключевые слова: сточные воды, контрольный створ, азот аммонийный, коэффициент корреляции, статистическая значимость.

Загрязнение водных объектов (ВО) сточными водами (СВ) является одной из основных экологических проблем экономически развитых стран. С целью оценки степени влияния отдельных точечных источников загрязнения на качество воды ВО в контрольных створах (КС) широко применяют различные методы детерминированного математического моделирования (Караушева, Лапшева и др.) [1]. Однако эффективность подобных методов существенно зависит от степени изученности ВО – его гидрологических и гидрохимических характеристик, что не всегда выполнимо. Кроме того, возникает неясность, по какому критерию следует делать вывод о значимости источника загрязнения? По этим причинам для решения данной задачи применяют также статистические методы, которые при наличии мониторинговых данных наблюдений не требуют специфической информации о ВО. В том случае, когда на качество природной воды в рассматриваемом КС по влиянию существенно выделяется какой-либо один фактор, либо факторов несколько, но только один из них отличается количественным непостоянством, может быть применен корреляционный анализ. В частности, на основании результатов исследования корреляционной зависимости между качественными характеристиками СВ и качеством воды ВО

можно сделать вывод о степени влияния СВ на состояние ВО по рассматриваемым показателям.

Целью данной работы является рассмотрение возможности использования корреляционного анализа для оценки влияния сброса СВ на качество воды ВО.

Традиционный подход к решению данной задачи заключается в расчете выборочного коэффициента корреляции и последующей проверке нулевой гипотезы, т. е. проверке на равенство нулю теоретического (по всей генеральной совокупности) коэффициента корреляции. В случае отклонения от нулевой гипотезы полагается, что корреляционная зависимость между качественными характеристиками СВ и качеством природной воды является статистически значимой, а коэффициент корреляции характеризует степень влияния источника загрязнения на состояние ВО.

Выборочный коэффициент корреляции рассчитывается по формуле:

$$R = \frac{\sum_{i=1}^N (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2 \cdot \sum_{i=1}^N (y_i - \bar{y})^2}}$$

где i , N – соответственно индекс и количество замеров; x_i – значение показателя качества СВ; y_i – значение показателя качества воды в КС с учетом времени перемещения водных масс; \bar{x} , \bar{y} – среднеарифметические значения.

Проверка статистической значимости выборочного коэффициента корреляции R основана на том факте, что величина

$$Z = \frac{R}{\sqrt{1 - R^2}} \sqrt{N - 2}$$

удовлетворяет t -распределению Стьюдента с $N - 2$ степенями свободы [2].

Таким образом, вывод о статистической значимости делался исходя из выполнения условия

$$|R| > |R_{кр}|,$$

где $R_{кр}$ – критическое значение, определяемое из уравнения

$$\frac{R_{кр}}{\sqrt{1 - R_{кр}^2}} \sqrt{N - 2} = t_{\alpha, N-2},$$

где $t_{\alpha, N-2}$ – квантиль распределения Стьюдента; α – принятый уровень значимости.

Однако приведенный метод определения статистической значимости является надежным лишь при условии нормального распределения случайных величин [2] x и y . В рассматриваемой же задаче это условие может не выполняться в силу специфики условий водоотведения.

При наличии сомнения в корректности применения традиционного подхода к проверке статистической значимости корреляция может быть исследована методом bootstrap, относящимся к группе непараметрических методов статистики (т. е. не требующих информации о законе и параметрах распределения случайных величин) [3, 4]. Суть метода bootstrap заключается в следующем. Для оценки некоторой величины T по выборке наблюдений $\{X_i\}$, $i = 1 \div N$, многократно составляется bootstrap-выборка $\{X_1^*, X_2^*, \dots, X_N^*\}$ из элементов начальной выборки. При этом некоторые элементы начальной выборки могут повторяться в каждой отдельной bootstrap-выборке, а некоторые – вовсе не включаться. Для каждой bootstrap-выборки рассчитываются оценки $T^*(\{X_1^*, X_2^*, \dots, X_N^*\})$, по которым составляется вариационный ряд $\{T_1^*, T_2^*, \dots, T_B^*\}$, где B – количество bootstrap-операций. Искомой оценкой является среднее арифметическое ряда

$$\hat{T} = \frac{1}{B} \sum_{b=1}^B T_b^*,$$

а границами доверительного интервала – значения $\{T_b^*\}$ с порядковыми номерами $[B \cdot \alpha / 2]$ и $[B \cdot (1 - \alpha / 2)]$, где α – выбранный уровень значимости. (Квадратные скобки означают целую часть числа.)

Если в качестве выборки $\{X_i\}$ взять множество пар случайных чисел $\{x_i, y_i\}$, то методом bootstrap можно оценить коэффициент корреляции между случайными числами x и y . В случае непопадания нуля в указанный доверительный интервал делается вывод о том, что влияние сброса СВ на ВО является статистически значимым при заданном уровне значимости. В противном случае полагается, что при анализе антропогенного загрязнения ВО влиянием рассматриваемого источника можно пренебречь.

Ниже приведен пример использования корреляционного анализа для оценки влияния сброса СВ предприятия-водопользователя ГКП «Бердянский водоканал» на качество воды Азовского моря по показателю азот аммонийный в КС, расположенных на расстоянии 500 м и 1000 м от места сброса (соответственно КС-1 и КС-2).

Исходные данные приведены в табл. 1.

Поскольку расход СВ предприятия приблизительно постоянен (около 800 тыс. м³ в месяц), а концентрация рассматриваемого вещества непостоянна, то установить факт влияния сброса на качество морской воды по данному показателю можно путем оценки коэффициента корреляции между концентрацией веществ в СВ и морской воде и последующей проверки статистической значимости рассчитанного коэффициента.

В силу малого объема выборки проверка закона распределения рассматриваемых случайных величин не может быть проведена с достаточным уровнем надежности. Это дает основание применить непараметрический метод для расчета. Для сравнения также был проведен расчет традиционным методом.

Результат расчета коэффициента корреляции традиционным методом приведен в табл. 2.

Как видно из результатов расчета, СВ предприятия оказывают влияние на качество морской воды в ближнем КС, а в КС, находящемся на расстоянии 1000 м от выпуска, влияние незначительное.

Результат расчета коэффициента корреляции методом bootstrap (при количестве bootstrap-операций $B = 500$ и уровне значимости $\alpha = 0,05$) приведен в табл. 3. Соответствующие гистограммы распределения величины T^* приведены на рис. 1 и 2.

Как видно из табл. 3 и рис. 1 и 2, коэффициент корреляции между концентрациями азота аммонийного в СВ и морской воде, оцененный методом bootstrap, является статистически значимым для обоих КС.

Из сравнения результатов расчета можно сделать следующий вывод. Для КС-1 результаты расчета обоими методами качественно совпали, а для КС-2 факт влияния СВ на состояние морской воды установлен только методом bootstrap. Поскольку, согласно вышеуказанным литературным источникам, достоверность результата расчета традиционным методом зависит от закона распределения рассматри-

1. Концентрация азота аммонийного в СВ ГКП «Бердянский водоканал» и в морской воде в районе водоотведения (мг/дм³)

СВ	КС-1	КС-2
0,27	0,16	0,25
0,16	0,11	0,07
0,22	0,13	0,08
0,22	0,13	0,07
0,22	0,08	0,08
0,98	0,6	0,3
0,73	0,37	0,17
0,23	0,09	0,09
0,28	0,19	0,28
0,79	0,25	0,21

2. Результат расчета коэффициента корреляции традиционным (параметрическим) методом

Контрольный створ	Выборочный коэффициент R	Критическое значение $R_{кр}$	Вывод о статистической значимости
КС-1	0,914	0,632	значимая
КС-2	0,613	0,632	незначимая

3. Результат расчета коэффициента корреляции методом bootstrap

Контрольный створ	Оценка коэффициента корреляции R	Доверительный интервал $[R_-, R_+]$	Вывод о статистической значимости
КС-1	0,919	[0,824; 0,992]	значимая
КС-2	0,610	[0,247; 0,980]	значимая

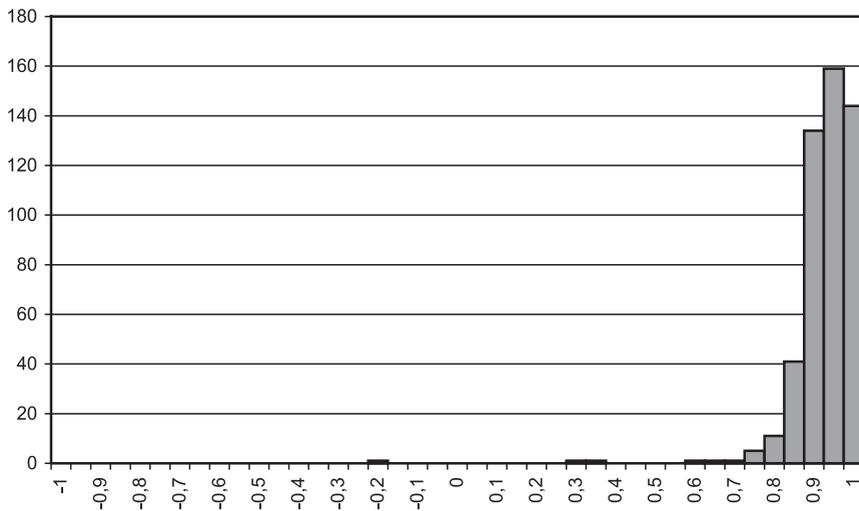


Рис. 1. Гистограммы распределения величины T^* для КС-1

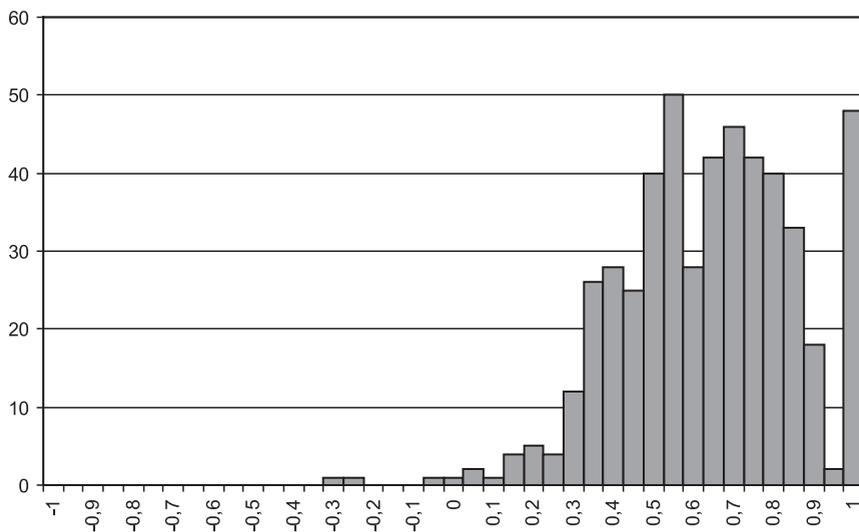


Рис. 2. Гистограммы распределения величины T^* для КС-2

ваемых случайных величин, а метод bootstrap является в этом плане универсальным, то вывод о наличии значимого влияния целесообразно делать по результатам расчета непараметрическим методом bootstrap.

Таким образом, в ряде случаев метод bootstrap может применяться для оценки влияния СВ на состояние ВО в районе водоотведения.

При этом следует заметить, что подобный расчет можно проводить при мало изменяющемся расходе СВ и при отсутствии иных существенных источников воздействия на состояние ВО. Если эти условия не выполняются, вывод о влиянии СВ на состояние ВО по статистическим данным возможен с помощью регрессионного анализа. Для этого также можно использовать непараметрический метод [5].

Кроме того, в данном примере использовался полностью имеющийся массив исходных данных без их проверки на репрезентативность.

Направлением дальнейших исследований является разработка предложений по выявлению и исключению из рассмотрения нехарактерных значений, т. е. подготовке репрезентативной выборки для оценки корреляционной зависимости между антропогенной нагрузкой на ВО и состоянием ВО в заданном КС.

1. Методические основы оценки и регламентирования антропогенного влияния на качество поверхностных вод / Под ред. А.В. Караушева. – Л.: Гидрометеиздат, 1987. – 285 с.
2. Орлов А. И. Прикладная статистика / А. И. Орлов. – М.: Экзамен, 2006. – 672 с.
3. Lo S.-H. The product-limit estimator and the bootstrap: some asymptotic representations / Lo S.-H., Singh K. // *Probab. Theory Rel/Fields.* – 1986. – Vol. 71. – № 3. – P. 455-466.
4. Смирнова С. А. Использование метода bootstrap для оценки влияния сброса сточных вод на качество воды водного объекта / С. А. Смирнова, О. А. Проскурнин // *Методы повышения ресурса городских инженерных инфраструктур* : VI Всеукр. научн. семинар, 15-16 октября 2014 г. – Х.: ЧФ «Михайлов», 2014. – С.127-128.
5. Проскурнин О. А. Прогнозирование влияния техногенного загрязнения на окружающую среду методом непараметрического регрессионного анализа: Дис... к-та техн. наук: 21.06.01 / УкрНИИЭП. – Харьков, 2007. – 171 с.

Проскурнін О. А., Смирнова С. А. ВИКОРИСТАННЯ МЕТОДУ BOOTSTRAP ДЛЯ ОЦІНКИ ВПЛИВУ СКИДАННЯ СТИЧНИХ ВОД НА ЯКІСТЬ ВОДИ ВОДНИХ ОБ'ЄКТІВ

У статті розглядається використання непараметричного статистичного методу bootstrap для оцінки впливу скидання стічних вод на якість води водного об'єкта. Наводиться приклад розрахунку для підприємства-водокористувача ДКП «Бердянський водоканал».

Ключові слова: *стічні води, контрольний створ, азот амонійний, коефіцієнт кореляції, статистична значимість.*

Proskurnin O. A., Smirnova S. A. USING BOOTSTRAP TO ASSESS THE IMPACT OF WASTEWATER DISCHARGES ON WATER QUALITY OF WATER BODIES

The article discusses the use of statistical nonparametric bootstrap method to assess the impact of wastewater discharge on water quality of the water body. An example of calculation for water users enterprise-DKP «Vodokanal Berdyansk».

Key words: *wastewater, control target, ammonia nitrogen, the correlation coefficient is statistically significant.*