

УДК 556.166

**М.Г. Сербов**, к.г.н.

*Одеський державний екологічний університет*

**С.Г. Кіріяк**

*Херсонський гідрометеорологічний технікум ОДЕКУ*

## **ПРО ОСОБЛИВОСТІ ЗАСТОСУВАННЯ СТАТИСТИЧНИХ МЕТОДІВ АНАЛІЗУ ЧАСОВИХ РЯДІВ МАКСИМАЛЬНОГО СТОКУ ВЕСНЯНОГО ВОДОПІЛЛЯ РІЧОК УКРАЇНИ**

*Наводяться результати статистичного аналізу, який включає використання методів багатовимірного аналізу (сумісного і факторного) для уточнення оцінок індивідуальних кривих розподілу.*

***Ключові слова:** статистичний аналіз, сумісний, факторний, максимальний стік водопілля річок України*

Обчислення максимальних витрат води весняного водопілля – одна з складних і відповідальних задач інженерної практики гідрологічних розрахунків. Відносна обмеженість вихідних часових рядів максимального стоку, а також багатофакторність розвитку гідрологічних процесів привели до того, що з різних областей математики переважний розвиток в гідрологічних розрахунках отримали теорія імовірності та математична статистика.

Статистичний аналіз, що включає вирішення низки питань (об'єктивний вибір оцінок кривих розподілу, аналіз та узагальнення індивідуальних статистичних характеристик в межах гідрологічно однорідних районів, оцінка точності обчислених розрахункових характеристик стоку заданої імовірності перевищення та багато іншого), є однією зі складових і більш важливих частин комплексу задач, що входять до розробки регіональних методик розрахунку максимального стоку.

Отримання надійних статистичних параметрів рядів спостережень дозволяє провести об'єктивний аналіз формування і розвитку досліджуваного процесу або явища, його факторну обумовленість в межах розглянутої моделі.

В основу статистичного аналізу максимальних витрат та шарів стоку весняного водопілля покладені матеріали спостережень по 230 пунктах опорної гідрологічної мережі України.

Довжина вихідних рядів спостережень складає від 23 (р.Ялпуг – с. Комрат) до 87 років (р. Дністер – с. Заліщики), причому, з цієї кількості 99% постів мають тривалість рядів максимального стоку повені більше 30 років, що забезпечує достатньо високу точність розрахунку імовірних характеристик максимального стоку.

Гідрологічна вивченість території України, а саме її південної частини, недостатня. Гідрологічна мережа розташована не рівномірно. Фізико-географічні особливості території України, а також антропогенний чинник сприяють тому, що річки різних регіонів відзначаються певними особливостями.

Аналіз структури часових рядів максимальних витрат водопілля показав відсутність внутрішньорядних зв'язків. Природа формування єдиної протягом року максимальної витрати водопілля настільки складна в факторній обумовленості, що відшукати аналогію розвитку процесу в інші роки практично неможливо.

Оцінка якості вихідних часових рядів максимальних витрат виконана за допомогою коефіцієнта  $K_\tau$ , що враховує співвідношення між строковими  $Q_{\max}$  та середньодобовими  $\tilde{Q}_{\max}$  максимальними витратами. Отримані аналітичні рівняння дозволяють визначити коефіцієнти  $K_\tau$  в межах площ водозборів, неосвітлених матеріалами спостережень.

Відповідно [4], в загальному вигляді рівняння можуть бути представлені як

$$K_\tau = \frac{Q_{\max}}{\tilde{Q}_{\max}} - \frac{1}{1 - \frac{1}{m+1} \left(\frac{1}{T_n}\right)^m}, \quad (1)$$

де  $T_n$  - тривалість повені ;  $m$  - параметр, що залежить від форми руслового гідрографа.

Значення параметрів  $T_n$  та  $m$  для описуваної території визначаються за допомогою рівнянь:

$$T_m = b(A+1)^{m^2} \quad (2)$$

та

$$m = \frac{1}{\frac{D}{(A+1)m_1} - 1}, \quad (3)$$

де  $A$  - площа водозбору, а  $D, b, m_2, m_1$  - емпіричні параметри, нормовані порайонно.

Використання формули (1) з її регіональними параметрами дозволило виконати суттєве уточнення  $K_\tau$  в порівнянні з існуючими рекомендаціями [7] і, перш за все, в області невеликих площ водозборів [4].

Методи теорії імовірності, що отримали широке використання в сучасній практиці розрахунків річкового стоку, описують коливання максимальних витрат води за допомогою кривих розподілу, які, як правило, асиметричні і опираються на три стандартні параметри: середнє  $\tilde{Q}_m$ , коефіцієнт варіації  $C_v$  і асиметрії  $C_s$  (або відношення  $C_s/C_v$ ). Тип кривої розподілу відображає тільки деякі загальні риси, які властиві коливанням розглянутих гідрологічних величин, а значення параметрів кривої розподілу характеризують індивідуальні особливості водозбору ( його розмір, географічне положення, характер рельєфу та ін.)

Статистичне описання вибірових кривих розподілу виконано за допомогою методів найбільшої правдоподібності та моментів. В якості основного закону прийнято три параметричний гама-розподіл. Статистичні оцінки рядів максимальних витрат води, розраховані різними методами, достатньо близькі між собою, але моментні оцінки, як правило, дещо знижені за рахунок від'ємного зсуення, яке при  $C_s/C_v < 2,0$  та  $C_v < 0,5$  не перевищує 5-6%.

При невеликих обсягах гідрологічних вибірок (менш 25 років) і великих коефіцієнтах варіації ( $C_v > 1,0$ ) від'ємне зсуення оцінок  $C_v$  значне і в деяких випадках становить до 20-25%.

Шляхом статистичного моделювання рядів [8] отримані поправки, що дозволяють ліквідувати зазначене зсуення моментних оцінок.

Досліджувані ряди максимальних витрат води по кожному конкретному пункту в силу відносної обмеженості періоду спостережень надають лише часткову інформацію для оцінки імовірних характеристик. Тому бажано збільшення обсягу спостережень, а під час розрахунків екстремумів рідкої повторності це стає необхідністю. Приведення гідрологічних характеристик до багаторічного періоду частіше виконується за

допомогою річок-аналогів із застосуванням парної множинної регресії [7]. Сумісний аналіз по ансамблях гідрологічних об'єктів, який використовує всю інформацію матриць вихідних даних, є одним з доступних засобів, що дозволяють знівелювати недоліки інформації по окремо взятому об'єкту і тим самим отримати більш надійні висновки щодо розрахункових значень характеристик стоку.

Використання гідрологічних об'єктів в груповому аналізі досить ускладнено завдяки існуючій значній неоднорідності в умовах формування стоку на кожному з водозборів. Характеристики стоку в порівнянні з метеорологічними величинами інтегрально враховують все якісне і кількісне різноманіття зв'язків на водозборах.

Дослідження однорідності вихідних рядів максимального стоку водопілля, виконані за допомогою критеріїв І.В.Дуніна-Барковського та М.В.Смірнова [10, 11], вказують на неправомірність використання в сумісному аналізі як модулів максимального стоку  $q_{\max}$ , так і параметра  $q'_{\max}$ , що визначається, як

$$q_{\max} = q'_{\max} / (A + 1)^{n_1}, \quad (4)$$

де  $n_1$  - тангенс кута нахилу лінії зв'язку  $q_{\max} = f(A)$ .

Враховуючи редукцію модуля максимального стоку по площі водозбору  $A$ , а також високу географічну складову мінливості, яку несе шар стоку  $h_{\max}$ , вихідна інформація наводилась у вигляді

$$\bar{K}_0 = \frac{\bar{q}_{\max}}{h_{\max}} (A + 1)^{n_2}. \quad (5)$$

Різниця в величинах  $\bar{K}_0$  в свою чергу, може бути обумовлена впливом комплексу фізико-географічних характеристик. Розрахунок природної факторної обумовленості виконано за допомогою лінійних регресійних рівнянь, що були побудовані за допомогою R- модифікації факторного аналізу [5]. В цьому плані факторний аналіз є одним з ефективних методів математичного опису взаємозв'язків природних комплексів, одночасно спрощуючи статистичне узагальнення рядів гідрологічних величин.

Сумісний аналіз, в основі якого лежить розділ повної дисперсії коливання оцінок вибірових параметрів  $\varepsilon_n^2$  на випадкову  $\varepsilon_c^2$  та на географічну  $\varepsilon_z^2$  складові [6], тобто  $\varepsilon_n^2 = \varepsilon_c^2 + \varepsilon_z^2$ , дозволяє суттєво уточнити параметри  $\bar{K}_0$  та  $C_v$  для цілого ряду об'єктів з періодом спостережень за максимальним стоком не більше 20 років. Випадкова складова дисперсії параметрів  $\bar{K}_0$  та  $C_v$  змінюється від 51 до 94%. В результаті цього оцінки даних параметрів розраховуються як середньозважені величини за [6]

$$t_{\text{сум}} = \frac{t_{\text{інд}} \varepsilon_{\text{сер}}^2 + t_{\text{сер}} \varepsilon_{\text{інд}}^2}{\varepsilon_{\text{інд}}^2 + \varepsilon_{\text{сер}}^2}, \quad (6)$$

де  $\varepsilon_{\text{інд}}$  - стандартна похибка індивідуальної оцінки  $t_{\text{інд}}$ ;  $\varepsilon_{\text{сер}}$  - стандартна похибка середньої по групі річок оцінки, яка розраховується за формулою

$$\varepsilon_{\text{сер}} = \varepsilon_e^2 / K + \varepsilon_z^2, \quad (7)$$

При визначенні випадкової складової розглянутих параметрів корелятивні зв'язки враховувались за формулою І.О. Сарманова [1, 9]

$$\varepsilon_{\theta}(t) = \varepsilon_c^H(t) \sqrt{1 - r_{cep}(t)}, \quad (8)$$

де  $\varepsilon_c^H$  - стандартна похибка параметра  $t$ ;  $r_{cep}(t)$  - середній по ансамблю розглянутих об'єктів коефіцієнт кореляції.

Значення коефіцієнтів кореляції між оцінками параметрів знайдено за співвідношеннями, що були отримані в роботі Є.Г.Блохінова [1]:  $r(\bar{K}_{0x}, \bar{K}_{0y}) = r_{xy}$ ;

$$r(C_{vx}, C_{vy}) = r_{xy}^2; \quad r[(C_s / C_v)_x, (C_s / C_v)_y] = r_{xy}^3.$$

При переході від сумісних оцінок  $\bar{E}_0$  до модулів максимального стоку  $\bar{q}_{max}$  виникають труднощі, обумовлені відповідністю відношень  $\bar{q}_{max} / \bar{h}_{max}$  з величинами  $\left(\frac{q_{max}}{h_{max}}\right)$ .

Виконані в межах розглянутої території дослідження показують, що різниця між  $\bar{q}_{max} / \bar{h}_{max}$  і  $\left(\frac{q_{max}}{h_{max}}\right)$  складає не більше 3-5%. Отримані результати дозволяють суттєво розширювати застосування сумісного аналізу під час дослідження ансамблів статистично не однорідних об'єктів.

Розсіювання оцінок відношення  $C_s/C_v$  практично повністю визначається випадковими варіаціями – величина географічної складової не перевищує 18,5%, що дає можливість при розрахунках максимальних витрат води приймати середні по районах значення  $C_s/C_v$ . Закріплені в межах району відношення  $C_s/C_v$  мають також важливе значення і для зменшення похибок розрахунку квантелів малої імовірності перевищення. Відповідно дослідженням Л.Ф. Сотникової [12], між оцінками  $C_v$  та  $C_s/C_v$ , що були розраховані за методом найбільшої правдоподібності, існує тісний позитивний корелятивний зв'язок. Наявність вказаного зв'язку приводить до того, що більшому вибірковому  $C_v$  відповідає більше  $C_s/C_v$  і, навпаки, в результаті чого похибки квантелів в зоні малої забезпеченості при індивідуальних оцінках  $C_v$  і  $C_s/C_v$  суттєво зростають у порівнянні з похибками квантелів при фіксованому  $C_s/C_v$ .

Аналіз рядів максимального стоку весняного водопілля, виконаний за допомогою критерію „забезпеченості забезпеченостей” [6], дозволяє встановити ступінь відповідності фактичних екстремумів теоретичної кривої розподілу. Теоретична крива „забезпеченості забезпеченостей” виражається залежністю [6]

$$P_{p_x} = 1 - (1 - p_x)^n, \quad (9)$$

де  $p_x$  - забезпеченість екстремального члена вибірки, яка визначається за кривими забезпеченості щорічних максимальних витрат повені;  $P_{p_x}$  - імовірність того, що щорічне перевищення екстремуму вибірки, яка складається з  $n$  членів, перебільшить величину  $p_x$ .

Ступінь відповідності характеризується співвідношенням між теоретичними та емпіричними значеннями середніх  $\bar{P}$  і стандартних відхилень  $\sigma_{(p)}$ , які визначаються як [ 6 ]

$$\bar{P} = \frac{1}{n+1}, \quad (10)$$

та

$$\sigma_{(p)} = \frac{1}{n+1} \sqrt{\frac{\bar{n}}{\bar{n}+2}}. \quad (11)$$

Результати розрахунків показують, що емпіричні криві імовірностей як для біноміального, так і для трипараметричного гама-розподілу узгоджуються з теоретичними. Але в зоні малої забезпеченості ( $P < 1 - 20\%$ ) емпіричні точки значно відхилені від теоретичних кривих забезпеченості.

Тим самим в цій області не обґрунтовано занижуються значення розрахункових забезпеченостей. В той час, як емпіричні криві розподілу, що були отримані за сумісними оцінками параметрів, цього недоліку не мають. Деяке заниження емпіричних  $\bar{P}$  та  $\sigma_{(p)}$  у порівнянні з їх теоретичними значеннями обумовлено неврахуванням синхронності коливань стоку.

В якості розрахункових далі рекомендується використовувати сумісні оцінки  $\bar{Q}_m$  і  $C_v$ , розраховані за методом найбільшої правдоподібності.

Оцінку точності вибірових квантилів 1%-ої імовірності перевищення виконано за даними статистичних випробувань, які дозволяють оцінити стандартні похибки за виразом

$$\sigma_{(Q1\%)} = \frac{E_\sigma}{\sqrt{n}}, \quad (12)$$

де  $E_\sigma$  - приведені похибки, що визначаються за номограмами [2, 13].

Середня величина відносних стандартних похибок  $\sigma_{(Q1\%)}$  складає 12-15% при максимальному її значенні до 27%.

Розподіл оцінок квантилів асиметричний, тому для повної характеристики випадкових похибок недостатньо тільки середнього значення оцінки квантиля і стандартного відхилення  $\sigma_{(Qp)}$ . Необхідно також знати і коефіцієнт його асиметрії. Для оцінки асиметрії вибірових квантилів, як і для визначення  $\sigma_{(Q1\%)}$ , використані результати статистичного моделювання рядів [8].

Шар стоку є одним з основних параметрів більшості розрахункових формул максимального шару, а оскільки він представляє собою площу гідрографа, то його визначення по матеріалах вимірювань не викликає суттєвих ускладнень.

Найчастіше розрахункові шари стоку і витрати заданої імовірності перевищення  $P\%$  отримують в результаті окремої статистичної обробки. За відсутності генетичного взаємозв'язку між величинами  $Q_{\max}$  та  $h_{\max}$  такий підхід може призвести до значних помилок. Більш зручним є підхід, запропонований в [3], де пропонується використовувати не рівнозабезпечені, а спряжені з максимальними витратами шари стоку, що відповідають більш дружним веснам.

## Список літератури

1. Блохинов Е.Г., Сотникова Л.Ф. Об оценке параметров распределения вероятностей годового стока рек СССР // Тр. ГГИ.-1970. – Вып. 180.-С.85-113.
2. Волкова З.Ф. О случайных погрешностях в оценке ординат кривой распределения вероятностей стока // Тр.ГГИ.-1970.-Вып.180.-С.124-137.
3. Гонченко Е.Д. Обоснование параметров редуцированной формулы для расчета максимального стока рек // Водные ресурсы .-1976.-№4-С.34-41.
4. Гонченко Е.Д. Сербов Н.Г. Соотношение между срочными и среднесуточными максимальными расходами весеннего половодья рек Западно-Сибирской равнины .- М.: 1986.-14с.- Деп в ИЦ ВНИИГМИ МЦД, 03.03.86г. № 496-гм;
5. Иберла К. Факторный анализ .-М.: Статистика, 1980.-397с.
6. Крицкий С.Н., Менкель М.Ф. Гидрологические основы управления речным стоком.- М.:Наука,1981.-255с.
7. Определение расчетных гидрологических характеристик // СНиП 2.01.14-83.-М., 1985.-36с.
8. Рождественский А.В. Оценка точности кривых распределения гидрологических характеристик.-Л.:Гидрометеиздат, 1977.-270с.
9. Сарманов И.О. О приближенном вычислении коэффициента корреляции между функциями от случайных зависимых величин // Матем.заметки. 1970- Т.7, вып.5. – С. 130-137;
10. Смирнов Н.В., Дунин – Барковский И.В. Курс теории вероятности и математической статистики для технических приложений . – М.: Наука, 1969.-511 с.
11. Сотникова Л.Ф. Исследования эффективности методов оценки параметров распределения вероятностей случайных величин применительно к анализу максимального стока рек // Водные ресурсы .- 1973.-№ 5.- С. 90-98;
12. Сотникова Л.Ф., Макарова Т.Ф. Оценка точности расчета гидрологических величин, моделируемых при посредстве трехпараметрического гамма-распределения// Проблемы изучения и комплексного использования водных ресурсов .-М.,1978.- С. 80-93;
13. Христофоров А.В. Оценка надежности карт гидрометеорологических элементов и пути повышения их точности// Экспресс- информ. ВНИИГМИ МЦД. – 1977. – Вып. 3(50) –С.5-11.

### **Про особенности применения статистических методов анализа временных рядов максимального стока весеннего половодья рек Украины. Сербов Н.Г., Кирияк С.Г.**

*На основе методов многомерного статистического анализа (совместного и факторного) выполнен статистический анализ рядов максимального стока весеннего половодья рек Украины. В основу статистического анализа максимальных расходов и слоев стока весеннего половодья положены материалы наблюдений по 230 пунктам гидрологической сети Украины.*

**Ключевые слова:** статистический анализ, совместный, факторный, максимальный сток весеннего половодья рек Украины

### **Statistical analysis of the series of maximum runoff of spring high water on the rivers of Ukraine.**

**M.G.Serbov, S.G.Kiriyak**

*On the basis of methods of multidimensional statistical analyses (integral and factor) a statistical analysis of the series of maximum runoff for spring high water on the rivers of Ukraine is conducted. Flow records from 230 stations of hydrological network of Ukraine formed the basis for statistical analysis of maximum discharges and runoff depths of spring high water.*

**Keywords:** statistical analysis, methods of multidimensional analysis, integral, factor, maximum runoff of spring high water on the rivers of Ukraine