

УДК 556.166

**Н.Г. Сербов, к.г.н.**

*Одесский государственный экологический университет*

**С.Г. Кирияк, асп.**

*Херсонский гидрометеорологический техникум ОГЭКУ*

## **ПРИМЕНЕНИЕ СОВМЕСТНОГО АНАЛИЗА ДЛЯ УТОЧНЕНИЯ СТАТИСТИЧЕСКИХ ОЦЕНОК ВРЕМЕННЫХ РЯДОВ ВЕСЕННЕГО ПОЛОВОДЬЯ (НА ПРИМЕРЕ ЛЕВОБЕРЕЖНЫХ ПРИТОКОВ ДНЕПРА)**

*На основе применения методов многомерного совместного и R-модификации факторного анализа проведено уточнение параметров рядов весеннего половодья. Приводится анализ эффективности применения аппарата совместного анализа при уточнении статистических оценок характеристик весеннего половодья.*

**Ключевые слова:** многомерный статистический анализ, максимальный сток весеннего половодья, уточнение статистических параметров

Материалы наблюдений, имеющиеся по каждому гидрологическому объекту, представляют в силу относительной краткости периода наблюдений лишь ограниченную информацию для оценки вероятностных характеристик речного стока. Поэтому всегда желательно увеличение объема наблюдений, а при расчетах экстремумов редкой повторяемости это становится и необходимостью.

Приведение гидрологических рядов к многолетнему периоду при помощи рек-аналогов с применением парной и множественной регрессии не всегда возможно. Во-первых, из-за отсутствия надежных аналогов, имеющих достаточно продолжительные временные ряды, а, во-вторых, для некоторых стоковых величин (например, паводков ливневого происхождения) они вообще отсутствуют вследствие локальной ограниченности развития природных процессов в пространстве. Поскольку увеличение объема наблюдений на основе использования метода аналогий практически невозможно, остается единственный выход – использование принципа эргодичности, предположительно свойственного гидрологическим процессам и, как следствие этого, применение совместного анализа данных по нескольким более или менее однородным объектам.

Вопросы применения объединенного анализа в гидрологических расчетах рассмотрены во многих работах, но наиболее полно принципиальные положения объединения метеорологических и гидрологических данных и методика коллективного анализа обобщенных данных рядов изложены в монографии С.Н.Крицкого и М.Ф. Менкеля [ 5 ].

В основе метода совместного анализа лежит разделение дисперсии колебаний оценок выборочных параметров  $E^2_{полн.}$  на случайную  $E^2_{случ.}$  и географическую  $E^2_{геогр.}$  составляющие [ 5 ], причем

$$E^2_{полн.} = E^2_{случ.} + E^2_{геогр.} \quad (1)$$

Полная дисперсия оценок рассчитывается по формуле

$$E^2_{полн.} = \frac{\sum_{i=1}^k (t_i - \bar{t})^2}{k-1}, \quad (2)$$

где  $t_i$  - оценка параметра по  $i$ -му объекту;  $\bar{t}$  - средняя из оценок по коллективу совместно рассматриваемых объектов.

Случайная составляющая рассеяния исходных статистических параметров вычисляется по стандартным формулам математической статистики или по данным вероятностного моделирования рядов [ 7 ].

Как известно, одним из условий применимости совместного анализа является статистическая однородность коллектива исследуемых объектов. Рассеяние оценок метеорологических факторов практически полностью обусловлено изменчивостью макроатмосферных процессов и не зависит от морфометрических особенностей бассейнов и характера подстилающей поверхности [2]. Величина стока, в отличие от метеорологических факторов, интегрально учитывает все качественное и количественное многообразие связей на водосборах. Образно выражаясь, параметры стока – это результат наложения метеорологических явлений или процессов, измеряемых в отдельных пунктах, на пространственное многообразие характеристик водосборов. Все это обуславливает невозможность использования гидрологических величин в совместном анализе без их приводки к более или менее однородным условиям стока.

Простейший прием приводки применительно к стоковым рядам заключается обычно в переходе от расходов к модулям стока. Однако и модули максимального стока, являясь удельной характеристикой, также обнаруживают зависимость от площади водосбора вследствие расплывания паводочных волн в процессе добега, русло-пойменного водообмена и берегового регулирования.

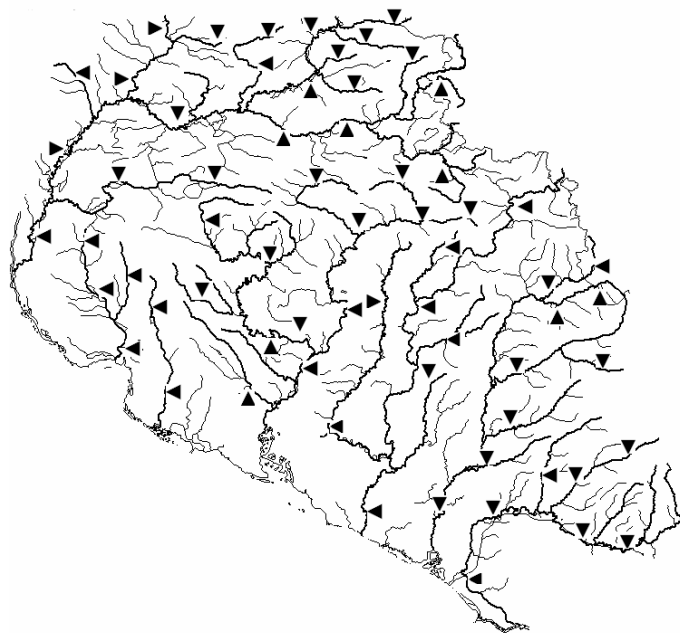


Рис.1 - Карта-схема расположения гидрологических пунктов наблюдений в пределах левобережных притоков Днепра.

Возможность применения совместного анализа для уточнения статистических параметров гидрологических рядов рассмотрена на примере левобережных притоков Днепра (рис.1). В исследовании использованы данные наблюдений за ежегодными максимальными срочными расходами воды и слоями стока за половодье по 70 постам опорной сети Государственной гидрометеорологической службы Украины.

Проверка на однородность исходных рядов модулей максимального половодного стока, выполненная при помощи критериев Дунина-Барковского [ 3 ], показала, что нулевая гипотеза об однородности как средних, так и дисперсий не только на 5%-ном, но и на 1%-ном уровне значимости отвергается.

Полученные результаты позволяют сделать вывод о неправомерности использования модулей стока в совместном анализе.

Собственно, ничего неожиданного в этом нет, т.к. известно, что модули максимального стока обнаруживают выраженную редукцию с ростом водосборной площади. Природа ее, как указывалось ранее, связана с расплыванием паводочных волн в процессе добега и русло-пойменным водообменом. Редукционные зависимости, в том числе и для средних многолетних модулей максимального стока  $q_m$ , описываются уравнением вида

$$\bar{q}_m = \frac{\bar{q}_m}{(F+1)^n}, \quad (3)$$

где  $F$  – площадь водосбора, км<sup>2</sup>;  $n$  – тангенс угла наклона линии связи к оси абсцисс.

Применение уравнение (3) позволяет исключить влияние площади водосбора на модули максимального стока. Процедура аналогичная той, которая представлена в [ 2 ], т.е. совместный анализ рядов гидрологических величин применен не к параметру  $\bar{q}_m$ , а к  $\bar{q}_m$ , который был рассчитан как

$$\bar{q}_m = \bar{q}_m (F+1)^n. \quad (4)$$

Данный прием, хотя и уменьшает неслучайную составляющую рассеяния  $E_{геогр.}^2$ , но на принятом исходном материале полностью не обеспечивает статистической однородности рядов максимального стока половодья рек рассматриваемого региона. Высокую остаточную дисперсию можно объяснить большим вкладом такой зонально изменяющейся характеристики весеннего половодья, как слой стока  $\bar{Y}_m$  [ 6 ], поскольку

$$\bar{q}_m = \tilde{K}_0 \bar{Y}_m, \quad (5)$$

где  $\tilde{K}_0$  – так называемый “коэффициент дружности весеннего половодья”.

Учитывая редукцию модуля максимального стока по площади  $F$ , а также высокую географическую составляющую, которую несет слой стока, исходная информация представляется как

$$\tilde{K}_0 = \frac{\bar{q}_m}{\bar{Y}_m} (F+1)^n. \quad (6)$$

Различия в величинах  $\tilde{K}_0$  также обусловлены не только случайными причинами, но и влиянием комплекса физико-географических характеристик, их пространственной неоднородностью. Приведение  $\tilde{K}_0$  к единым условиям формирования стока выполнено с помощью регрессионных уравнений, учитывающих факторную обусловленность  $\tilde{K}_0$ .

Одним из наиболее перспективных методов, позволяющих построить надежные эмпирические зависимости параметров речного стока с характеристиками условий его формирования, является факторный анализ, представляющий собой ту ветвь многомерного анализа, которая исследует внутреннюю структуру корреляционных матриц. В общем виде факторная модель может быть представлена в виде [ 4 ]

$$Z_{ij} = \sum_{L=1}^k a_{Lj} F_{iL} + e_j, \quad (7)$$

где  $Z_{ij}$  – нормированная величина  $j$ -ой переменной у  $i$ -го объекта;  $a_{Lj}$  – факторная нагрузка  $L$ -го обобщенного фактора у  $j$ -ой переменной;  $F_{iL}$  – значения  $L$ -го обобщенного фактора у  $i$ -го объекта;  $e_j$  – остатки, учитывающие как остаточную дисперсию, так и связанные с различными погрешностями расчетов.

Нормирование исходных данных, чтобы исключить влияние размерности и обеспечить сопоставимость переменных, производится по среднему арифметическому отклонению [ 2 ]

$$Z_{ij} = \frac{h_{ij} - \bar{h}_j}{\sigma_j}, \quad (8)$$

где  $\sigma_j$  - среднее квадратическое отклонение переменной  $h_{ij}$ .

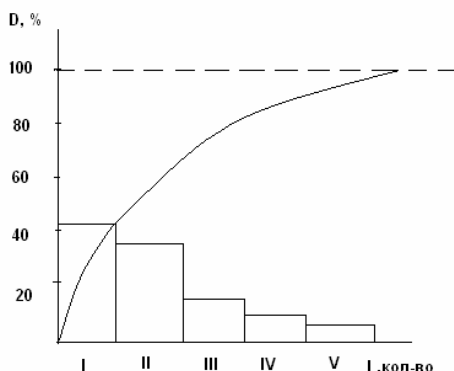


Рис. 2 Диаграмма распределения обобщенных факторов

Более подробно теория факторного анализа изложена в [4].

Регрессионная модель, с учетом (7) и (8), строилась на основе данных матрицы факторных нагрузок и уравнения обобщенных факторов с использованием R-модификации факторного анализа. Диаграмма распределения обобщенных факторов, из которых три первых являются эффективными и описывают 87,0% дисперсии исходных данных, представлена на рис. 2.

Расчетное уравнение имеет вид

$$\tilde{K}_0 = 26,2 + 0,110H_{cp} - 1,53f_b - 0,495f_l - 0,198f_p, \quad (9)$$

где  $f_b, f_l, f_p$  - соответственно, заболоченность, лесистость и распаханность водосбора, %;  $H_{cp}$  - средняя высота водосбора над уровнем моря, м.

Коэффициенты вариации  $Cv$  в значительно меньшей степени, чем параметр  $\tilde{K}_0$ , зависят от условий формирования стока. Тем не менее, в пределах исследуемой территории удалось обнаружить зависимость параметра  $Cv$  от площади водосбора  $F$  и слоя стока  $\bar{Y}_m$ , которая аналитически может быть аппроксимирована выражением вида:

$$Cv = 0,95 - 0,025 \lg(F+1) - 0,002 \bar{Y}_m. \quad (10)$$

Совместные оценки параметров  $\tilde{K}_0$ ,  $Cv$  и  $Cs/Cv$  по совокупности индивидуальных и объединенных наблюдений рассчитывались как средневзвешенные по каждой из оценок [ 5 ]

$$t_{сов} = \frac{t_{инд} \cdot E_{cp}^2 + t_{cp} \cdot E_{инд}^2}{E_{инд}^2 + E_{cp}^2}, \quad (11)$$

где  $E_{инд}$  - стандартная ошибка индивидуальной оценки  $t_{инд}$ ;  $E_{cp}$  - стандартная ошибка средней по группе рек оценки  $t_{cp}$ , определяемая по формуле

$$E_{cp}^2 = \frac{E_{случ}^2}{K} + E_{геогр}^2. \quad (12)$$

При количественной оценке случайной составляющей рассматриваемых статистических параметров корреляционные связи учитывались согласно И.О. Сарманову [ 1 ].

Значения коэффициентов корреляции между оценками параметров найдены по соотношениям, полученным Е.Г. Блохиновым [ 1 ]:

$$- r(\tilde{K}_{0x}, \tilde{K}_{0y}) = r_{xy}; \quad - r(C_{Vx}, C_{Vy}) = r_{xy}^2; \quad - r\left[\left(\frac{C_S}{C_V}\right)_x, \left(\frac{C_S}{C_V}\right)_y\right] = r_{xy}^3.$$

Результаты совместного анализа, соотношение между географической и случайной составляющими рассеяния параметров, представленные в табл.1, показывают хорошую согласованность индивидуальных и объединенных статистических оценок  $\tilde{K}_0$  и  $C_V$  - в большинстве случаев различия между оценками не превышают 12-15%. Тем не менее, применение совместного анализа позволило произвести существенное уточнение исследуемых характеристик по ряду объектов, имеющих в большинстве случаев период наблюдений за максимальным стоком половодья до 25 лет, что явно недостаточно для получения надежных статистических выводов.

Таблица 1 – Результаты совместного анализа по данным рек левобережья Днепра (при общем количестве исследуемых объектов n=70)

Параметр	$r_{cp}$	Составляющие дисперсии параметров		
		$E_{полн.}^2$	$E_{случ.}^2$	$E_{геогр.}^2$
$\tilde{K}_0$	0,34	67,8	42,6	25,2
$C_V$	0,12	0,046	0,027	0,019
$C_S/C_V$	0,04	0,471	0,465	0,006

Рассеяние оценок соотношения  $\frac{C_S}{C_V}$  практически полностью определяется случайными вариациями – величина географической составляющей не превышает 2% (табл.1). Полученные результаты дают основание при расчете максимального половодного стока принимать осредненные в пределах всего района значение соотношения  $\frac{C_S}{C_V} = 2,5$ .

При переходе от совместных оценок  $\tilde{K}_0$  к  $\bar{q}_m$  возникают трудности, обусловленные возможным несовпадением отношений  $\frac{\bar{q}_m}{\bar{Y}_m}$  с  $\overline{\left(\frac{q_m}{Y_m}\right)}$ . Выполненные исследования показывают, что существенной разницы между этими величинами нет, а отклонения данных от линии равных значений не превышают 5 %.

Значительный интерес представляет анализ эффективности метода коллективной оценки гидрологических характеристик. При статистически независимых колебаниях стока увеличение объема используемых для оценки наблюдений приводит к уменьшению стандартных ошибок пропорционально корню квадратному из числа наблюдений. Наличие же корреляции между колебаниями совместно рассматриваемых характеристик существенно меняет положение. Число лет статистически независимых наблюдений  $N_n$ , эквивалентное  $K_n$  годам, при среднем значении коэффициента корреляции  $r_{cp}$  может быть выражено из соотношения между соответствующими дисперсиями [ 5 ]

$$N_n = \frac{K_n}{1 - (K - 1)r_{cp}}. \quad (13)$$

На рис. 3 представлена графическая иллюстрация эффективности метода совместного анализа при различном количестве  $K$  совместно рассматриваемых объектов и  $r_{cp} = 0,34$ .

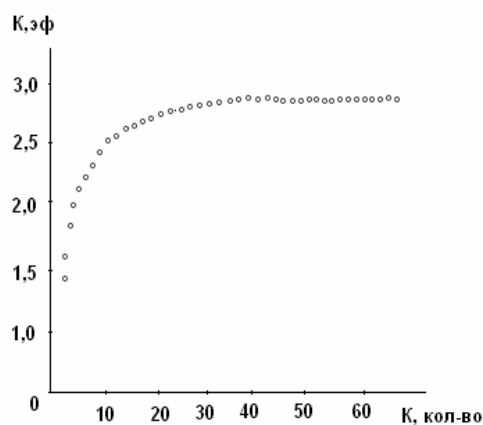


Рис. 3 Графічна ілюстрація ефективності спільного аналізу (при  $\gamma = 0,34$ )

В заключення ще раз необхідно підкреслити, що використання гідрологічних об'єктів в груповому аналізі дуже ускладнено через існуючу значущу неоднорідність в умовах формування стоку на кожному з водосборів. В цьому плані факторний аналіз є дуже ефективним методом опису взаємодії природних комплексів, в той же час спрощуючи статистичне узагальнення рядів стокових величин.

#### Список литературы

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. Гопченко Е.Д., Сербов Н.Г. Об особенностях применения совместного статистического анализа временных рядов весеннего половодья (на примере рек Западно-Сибирской равнины) // Водные ресурсы. - 1987. - № 2. - С. 10-17.
3. Дунин Барковский И.В., Смирнов Н.В. Курс теории вероятности и математической статистики для технических приложений. - М.: Физматгиз. -1959. - 511 с.
4. Иберла К. Факторный анализ (пер. с англ.). - М.: Статистика. - 1980. - 397 с.
5. Крицкий С.Н., Менкель М.Ф. Гидрологические основы управления речным стоком. - М.: Наука. - 1980. - 255 с.
6. Лобода Н.С., Гопченко Е.Д. Обоснование районирования статистических параметров стока, определяемых по наблюдаемым данным с малой степенью достоверности // Гідрологія, гідрохімія і гідроекологія: Наук. зб. - т.5. - Київ: Ніка-центр. - 2003. - С. 35-41
7. Рождественский А.В. Оценка точности кривых распределения гидрологических характеристик. - Л.: Гидрометеиздат. - 1977. - 270с.

#### Застосування сумісного аналізу для уточнення статистичних оцінок часових рядів весняного водопілля (на прикладі лівобережних притоків Дніпра).

Сербов М.Г., Кіряк С.Г.

На підставі застосування багатомірного сумісного та R-модифікації факторного аналізу проведено уточнення параметрів рядів весняного водопілля. Розрахунки виконані на прикладі 70 гідрологічних об'єктів, розташованих в лівобережній частині Дніпра. Наданий аналіз ефективності застосування апарату сумісного аналізу при уточненні статистичних оцінок характеристик весняного водопілля.

**Ключові слова:** багатомірний статистичний аналіз, максимальний стік весняного водопілля, уточнення статистичних параметрів

#### Application of integral analysis for improvement of statistical estimations for time series of spring high water (by the example of the left-bank tributaries of the Dnieper). M.G.Serbov, S.G.Kiriyak

Improvement of parameters for the series of spring high water is made on the basis of application of multidimensional integral analysis and the R-modification of factor analysis. Calculations are conducted with 70 hydrological objects, which are located on left bank of the Dnieper, taken as examples. Performance analysis for the application of the aid of integral analysis at improvement of statistical estimations for hydrological descriptions of spring high water is presented.

**Keywords:** multidimensional statistical analysis, maximum runoff of spring high water, improvement of statistical parameters