А.В. Сикан

МЕТОДИКА ПОСТРОЕНИЯ УСЕЧЕННЫХ КРИВЫХ ОБЕСПЕЧЕННОСТЕЙ ПРИ РАСЧЕТЕ МИНИМАЛЬНЫХ РАСХОДОВ И УРОВНЕЙ ВОДЫ

A.V. Sikan

METHOD OF CONSTRUCTING TRUNCATED EXCEEDANCE PROBABILITY CURVES, WHICH CAN BE USED FOR CALCULATING OF MINIMUM LEVELS AND WATER DISCHARGES

В работе представлена методика построения усеченных кривых обеспеченностей, которую можно использовать при расчетах минимального стока в случае генетической неоднородности рядов.

Ключевые слова: гидрологические расчеты, минимальные расходы и уровни воды, выбросы, усеченные кривые обеспеченностей, распределение Гумбеля.

The paper presents a method of constructing truncated exceedance probability curves, which can be used for calculating of minimum water discharge in the case of the heterogeneity of the series.

Key words: hydrological design, minimum levels and water discharges, outliers, truncated distribution curves, Gumbel distribution.

Если гидрологический ряд состоит из элементов (например, расходов или уровней воды) различного генетического происхождения, то это может быть причиной так называемых «выбросов».

Например, в рядах минимального 30-суточного летне-осеннего стока встречаются очень дождливые годы, когда один паводок сменят другой и межень практически отсутствует. В этом случае минимумы дождливых лет могут превышать минимумы засушливых лет в десятки раз.

Здесь мы сталкиваемся с так называемой генетической неоднородностью ряда, и эмпирическая кривая обеспеченностей резко уходит вверх в области максимальных значений. Такие кривые невозможно аппроксимировать традиционными кривыми Пирсона III типа или Крицкого-Менкеля, и в данной ситуации допустимо использовать усеченные кривые обеспеченностей.

При построении нижней части кривой обеспеченностей на первом этапе из ряда наблюдений исключаются расходы, которые резко отклоняются от эмпирической кривой обеспеченностей. Нормативные документы [2–4] рекомендуют выполнять эту процедуру с использованием критериев Диксона и Смирнова-Граббса обобщенных на случай асимметричных и автокоррелированных рядов. Однако наличие выбросов приводит к искажению статистических характеристик исследуемого ряда и оценки коэффициентов асимметрии и автокорреляции становятся крайне ненадежными.

В настоящей работе для исключения из ряда значений относящихся к другой генеральной совокупности предлагается новый критерий, который будем называть «медианный z-критерий для максимумов».

Предполагается, что имеется выборка подчиняющаяся гамма-распределению, в которую попало несколько значений из другой генеральной совокупности. В качестве анализируемой статистики рассматривается статистика:

$$z^* = \frac{x^*_{\max}}{x^*_{50\%}}, \tag{1}$$

где $x^*_{\max}$ — максимальное значение переменной x в выборке, состоящей из n членов; $x^*_{50\,\%}$ — выборочная медиана.

В качестве нормирующей величины в данном тесте выбрана медиана по следующей причине. Появление в выборке большого расхода (редкой повторяемости) может существенно повлиять на выборочное среднее. В тоже время значение медианы изменится мало. Поэтому вместо $\bar{x}$ используется $x_{50\,\%}$.

Гипотеза об однородности ряда не опровергается, если

$$z^* < z_\alpha, \tag{2}$$

где z_α — теоретическое значение статистики z при одностороннем уровне значимости α.

Расчет производится в следующем порядке:

1. По формуле (1) вычисляется z^*.
2. Рассчитывается параметр V_k, который представляет собой отношение интерквартильного размаха (*Interquartile range*) к медиане:

$$V_k = \frac{x_{75\%} - x_{25\%}}{x_{50\%}}\%. \tag{3}$$

Для гамма-распределения существует однозначная зависимость между параметром V_k и коэффициентом вариации (рис. 1), которую можно аппроксимировать кубическим сплайном:

$$C_v = 0{,}0079V_k^3 - 0{,}1266V_k^2 + 0{,}8123V_k. \tag{4}$$

3. По формуле (4) определяется условный коэффициент вариации C_v^* ненарушенного ряда.
4. Рассчитывается обеспеченность (в долях единицы) максимального модульного коэффициента, который с вероятностью α может присутствовать в выборке, состоящей из n членов (уровень значимости α рекомендуется принимать $\alpha = 0{,}1$ или $\alpha = 0{,}05$).

$$P = 1 - (1-\alpha)^{1/n}. \tag{5}$$

5. В зависимости от P и C_v^* для гамма-распределения определяется модульный коэффициент k_p и медиана $k_{50\,\%}$.

   Для определения этих параметров допустимо пользоваться таблицами Крицкого-Менкеля или Пирсона III типа при $C_s/C_v = 2$. При расчете в MS Excel можно воспользоваться функцией:

$$k_p = \text{ГАММАОБР}\left(F;a;b\right), \tag{6}$$

   где $F = 1 - P$ (вероятность непревышения в долях единицы); a, b — параметры гамма-распределения.

   Для модульных коэффициентов справедливы формулы:

$$a = \frac{1}{C_v^2};\quad b = C_v^2. \tag{7}$$

6. Рассчитывается теоретическое значение статистики z_α:

$$z_\alpha = \frac{k_p}{k_{50\%}}. \tag{8}$$

7. Проверяется выполнение неравенства (1).

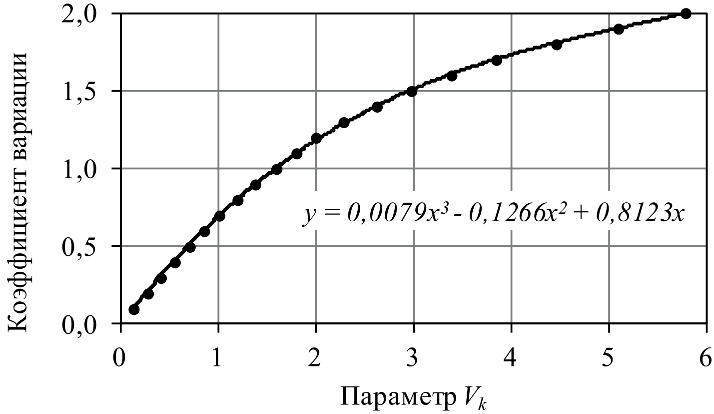

Рис. 1. График зависимости коэффициента вариации от параметра V_k для гамма-распределения

Если неравенство не выполняется, максимальный член удаляется из выборки. Затем вся процедура повторяется для выборки длиной ($n - 1$) и т. д., до тех пор, пока не будет выполняться неравенство (1).

Для удобства анализа строится график зависимости эмпирического и теоретического значения статистики z от числа последовательно удаленных максимумов (рис. 2).

В качестве примера на рис. 3 приводится хронологический график расходов воды, на котором отмечены исключенные значения.

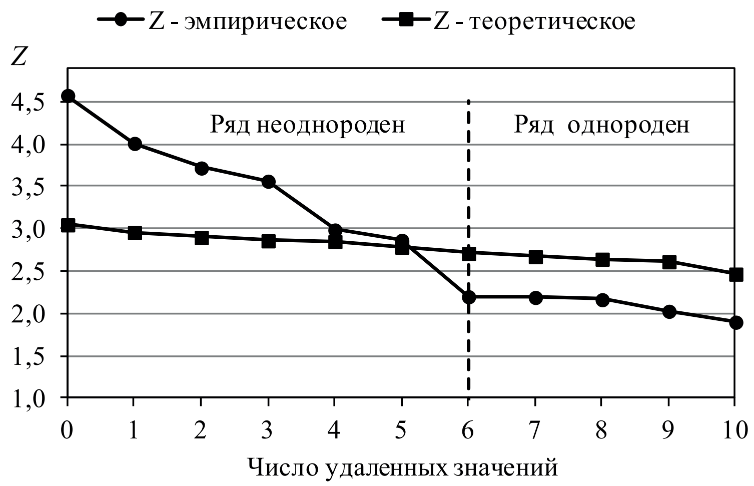

Рис. 2. Зависимость эмпирического и теоретического значения статистики z от числа последовательно удаленных максимумов; р. Устья — с. Бестужево. Ряд становится однородным после удаления первых 6 значений

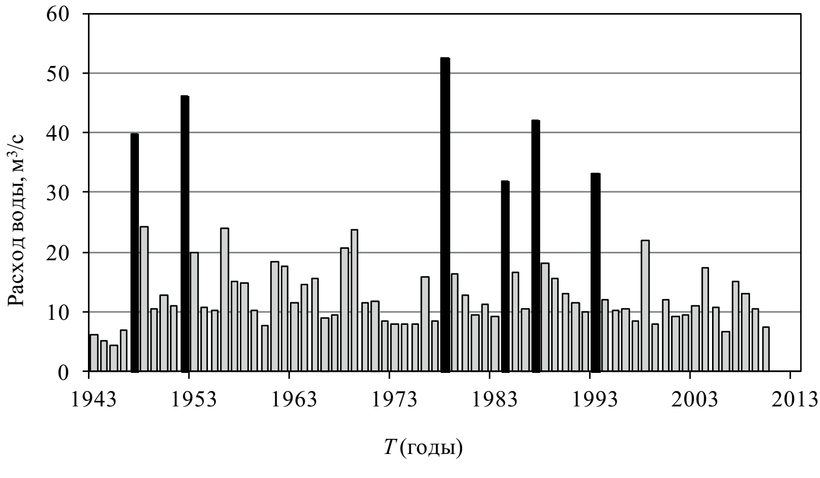

Рис. 3. Хронологический график минимальных 30-суточных летне-осенних расходов воды; р. Устья — с. Бестужево. Черным цветом выделены расходы, которые исключаются из ряда

После исключения из выборки k штук экстремальных значений, порождающих неоднородность ряда, получаем выборку объемом $n_1 = (n - k)$. Для этой выборки рассчитываются оценки параметров распределения аналитической кривой обеспеченностей (например, Крицкого-Менкеля).

С использованием полученных параметров на клетчатке вероятностей строится аналитическая кривая обеспеченностей (рис. 4).

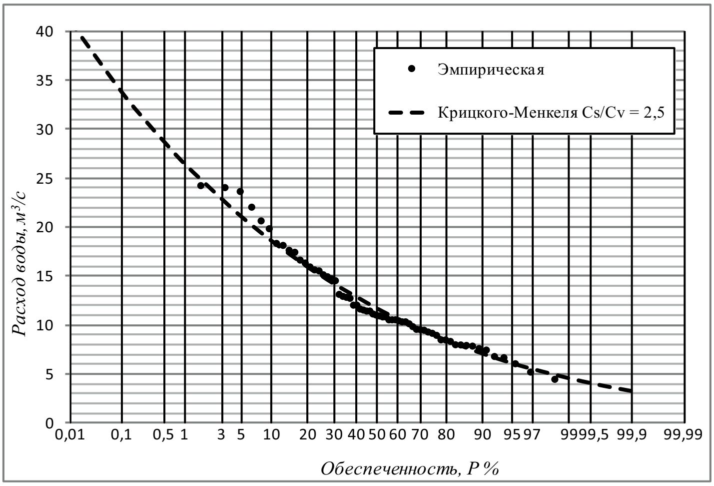

Рис. 4. Эмпирическая и аналитическая кривые обеспеченностей минимальных 30-суточных летне-осенних расходов воды после удаления 6 наибольших значений; р. Устья — с. Бестужево

Если кривая хорошо соответствует эмпирическим точкам, то на ее базе строится усеченная кривая обеспеченностей для исходной выборки объемом n.

Переход от обеспеченностей укороченной выборки (P_1) к обеспеченностям исходной выборки (P) производится по формуле:

$$P = \frac{n_1 P_1 / 100 + k}{n} 100\%. \tag{9}$$

Кривая обеспеченностей строится от точки усечения до $P = 99{,}9$ %. В качестве точки усечения принимается $P > P_k^*$, где P_k^* – эмпирическая обеспеченность k-ой (последней) удаленной точки (рис. 5).

К числу плюсов предложенного метода можно отнести то, что точка усечения не является строго фиксированной [1], а зависит от особенностей конкретного гидрологического ряда, и то, что для аппроксимации нижней части эмпирической кривой обеспеченностей можно использовать различные типы аналитических кривых.

Иногда точку усечения и число значений, относящихся к другой генеральной совокупности можно достаточно точно определить по перелому на эмпирической кривой обеспеченностей без использования статистических критериев. Тогда расчет выполняется по той же схеме — для однородной выборки рассчитываются оценки параметров распределения аналитической кривой обеспеченностей, а переход от обеспеченностей укороченной выборки к обеспеченностям исходной выборки производится по формуле (9).

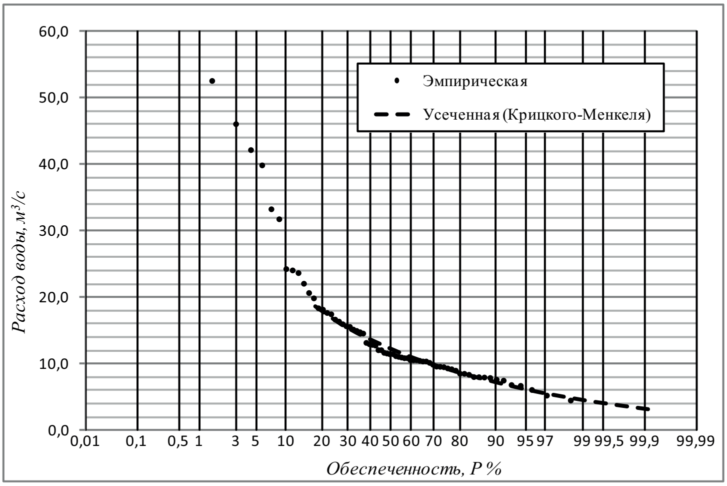

Рис. 5. Усеченная кривая обеспеченностей минимальных 30-суточных летне-осенних расходов воды; р. Устья — с. Бестужево

Особый случай, когда кривая обеспеченностей резко уходит вниз в области малых значений. Такой вид кривой может иметь место:

- для минимальных 30-суточных летне-осенних расходов воды, когда в ряду присутствуют несколько расходов воды, сформированных исключительно за счет грунтового притока в период его истощения;
- для минимальных 30-суточных летне-осенних расходов воды небольших озерных рек, если в очень засушливые годы уровень озер опускается до отметок близких к порогу стока;
- для некоторых рядов уровней минимального стока, где форма кривой обеспеченностей определяется особенностями поперечного профиля реки в расчетном створе.

В таких ситуациях в качестве усеченной кривой обеспеченностей можно рекомендовать распределение Гумбеля с отрицательной асимметрией (распределение Гумбеля для минимальных значений).

Функция плотности вероятности этого распределения имеет вид:

$$f(x) = \frac{1}{\lambda}\exp\left[y - \exp(y)\right], \tag{10}$$

где y — безразмерная величина, которая связана с x соотношением

$$y = \frac{x - \mu}{\lambda}, \tag{11}$$

где μ — мода случайной величины X; λ — параметр масштаба ($\lambda > 0$); $-\infty < x < +\infty$.

Интегральная функция распределения и функция обеспеченностей выражаются формулами:

$$F(x) = 1 - \exp\left[-\exp(y)\right], \tag{12}$$

$$P(x) = \exp\left[-\exp(y)\right]. \tag{13}$$

Из выражения (13) следует, что значения переменной y заданной обеспеченности можно рассчитать по формуле:

$$y_p = \ln\left[-\ln\left(\frac{P}{100}\right)\right], \tag{14}$$

где P — расчетная обеспеченность в процентах.

Ординаты кривой обеспеченностей с учетом выражения (11) определяются по формуле:

$$x_{p\%} = \mu + \lambda y_{p\%}. \tag{15}$$

Алгоритм построения усеченной кривой обеспеченностей в этом случае следующий:

1. Рассчитываются координаты эмпирической кривой обеспеченностей для всего ряда.
2. По формуле (14) рассчитываются эмпирические значения переменной y: $y^*_{P\%} = \ln[-\ln(P^*/100)]$.
3. Строится график зависимости $x = f(y^*_{P\%})$ и определяется точка перелома графика, которая принимается за точку усечения (рис.6.).
4. Повторно строится график зависимости $x = f(y^*_{P\%})$ для фрагмента ранжированного ряда, включающего нижнюю часть ряда до точки усечения (рис. 7).
5. Для построенного графика определяются коэффициенты уравнения линейной регрессии μ и λ, которые принимаются в качестве параметров распределения (см. формулу (15)).
6. С использованием формул (14, 15) рассчитываются ординаты усеченной кривой обеспеченностей от точки усечения до $x = 0$ (обеспеченность нуля в данном случае будет меньше 100 %).
7. На клетчатке вероятностей строятся эмпирическая кривая обеспеченностей и усеченная кривая для минимальных расходов (рис. 8).

При наличии в ряду нулевых значений следует использовать алгоритм, рекомендованный нормативными документами [2–4].

Применение усеченных кривых в большинстве случаев позволяет отказаться от использования сглаженных эмпирических кривых обеспеченностей и избежать субъективизма при расчетах минимального стока в случае генетической неоднородности рядов.

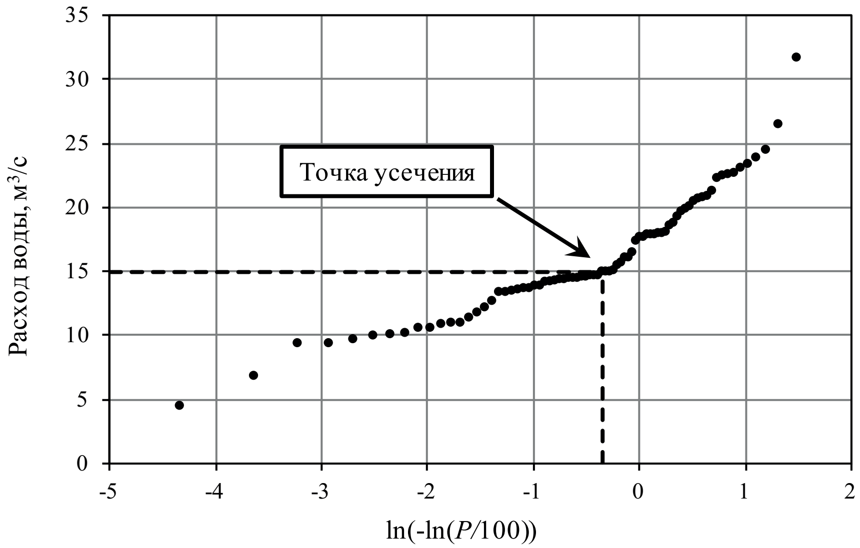

Рис. 6. Графики зависимости $x = f(y^*_{P\%})$ для всего ряда

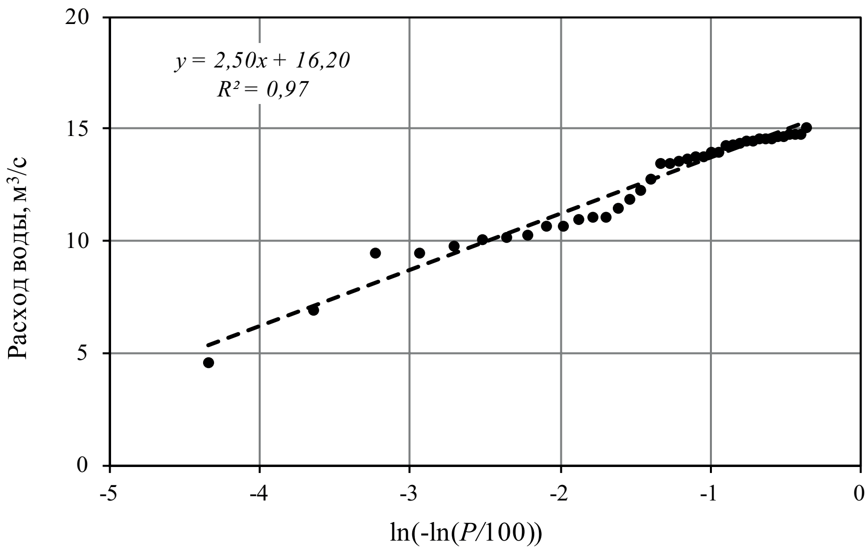

Рис. 7. Графики зависимости $x = f(y^*_{P\%})$ для нижней части ряда до точки усечения

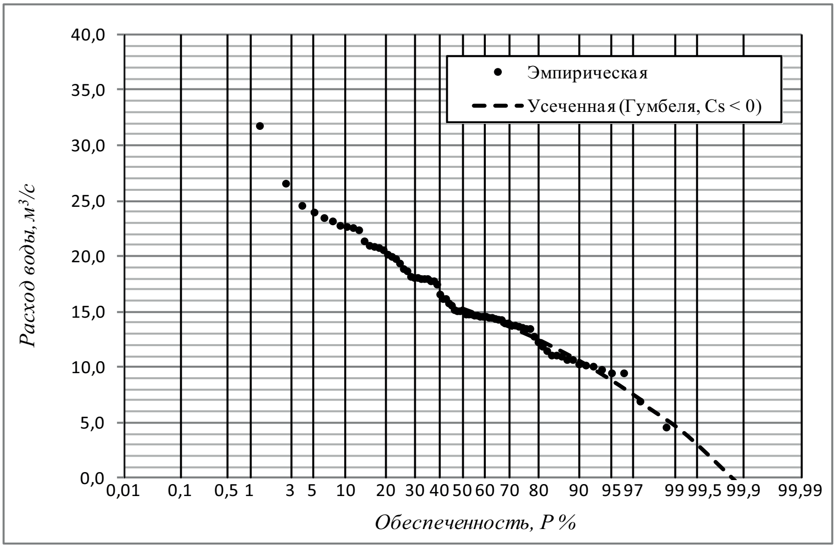

Рис. 8. Усеченная кривая обеспеченностей минимальных 30-суточных летне-осенних расходов воды; р. Суна — пгт Поросозеро

Литература

1. *Малышева Н.Г.* Способы аппроксимации эмпирических кривых обеспеченностей в условиях неоднородных выборок. // Ученые записки РГГМУ, 2011, № 21, с. 25–31.
   Malysheva N.G. Sposoby approksimatsii empiricheskikh krivykh obespechennostey v usloviyakh neodnorodnykh vyborok. // Uchenye zapiski RGGMU, 2011, № 21, s. 25–31.
2. Методические рекомендации по определению расчетных гидрологических характеристик при наличии данных гидрометрических наблюдений. — Нижний Новгород: Вектор-ТиС, 2007. — 134 с.
   Metodicheskie rekomendatsii po opredeleniyu raschetnykh gidrologicheskikh kharakteristik pri nalichii dannykh gidrometricheskikh nablyudeniy. — Nizhniy Novgorod: Vektor-TiS, 2007. — 134 s.
3. Методические рекомендации по оценке однородности гидрологических характеристик и определению их расчетных значений по неоднородным данным. — СПб.: Нестор-История, 2010. — 162 с.
   Metodicheskie rekomendatsii po otsenke odnorodnosti gidrologicheskikh kharakteristik i opredeleniyu ikh raschetnykh znacheniy po neodnorodnym dannym. — SPb.: Nestor-Istoriya, 2010. — 162 s.
4. СП 33-101-2003. Определение основных расчетных гидрологических характеристик. — М.: Стройиздат, 2004. — 72 с.
   SP 33-101-2003. Opredelenie osnovnykh raschetnykh gidrologicheskikh kharakteristik. — M.: Stroyizdat, 2004. — 72 s.