ООО «МЕТРОЛОГ»

ГЕОЭКОЛОГИЧЕСКИЕ ИССЛЕДОВАНИЯ

«Обработка и анализ данных наблюдений
река Ока – Костомарово за 1941-1980 гг.»

САНКТ-ПЕТЕРБУРГ
2004 г.

Сдержание:

1. Физико-географическая характеристика объекта исследования

Река Ока широко используется для промышленного и бытового водоснабжения, судоходства и лесосплава, гидроэнергетики и мелиорации. Ока протекает по району, который является исторически сложившимся промышленным центром России, с многочисленными крупными предприятиями черной и цветной металлургии, химической, текстильной и пищевой промышленности, потребляющими большое количество воды. Крупными водопотребителями являются промышленные центры с высокой концентрацией городского населения, такие как Москва, Кострома, Рязань, Калуга, Н. Новгород и другие. Вода после ее использования промышленными предприятиями, населением и в системах орошения возвращается в речную сеть в виде хозяйственных, промышленных стоков и возвратных вод с орошенных земель. Объем воды, отводимой обратно в речную сеть, составляет около 80% водопотребления. Около 60% возвратных вод условно чистые, а остальная часть в значительной степени загрязнена. Загрязнение поверхностных вод вызывает трудности в дальнейшем развитии водоснабжения и требует принятия серьезных мер по улучшению очистки сточных вод. Забор воды из реки для целей сельскохозяйственного водоснабжения и орошения составляет 4 % общего объема водопотребления. Потенциальный энергетический запас реки составляет около 30%.

Река Ока судоходна от г. Чекалина (1200 км). Выше Рязани Ока шлюзована: Белоотутовская и Кузьминская плотины. По Оке перевозят такие грузы как стройматериалы, лес, каменный уголь, нефтепродукты, хлеб, зерно, металлические изделия, текстиль, машины. Местные пассажирские перевозки – ниже г. Калуги. Транзитное судоходство: от устья р. Москвы до г. Н. Новгорода. По Оке существуют туристические рейсы. На Оке существуют предприятия рыболовецкой отрасли по разведению и лову различных видов рыб (стерлядь, язь, сом, щука, лещ, окунь и др.)

2.Характеристика данных наблюдений

2.1. Расположение пункта наблюдений, оборудование, период наблюдений

Пост «Ока – Костомарово» расположен в центре деревни Костомарово Орловской области, в 3 км ниже г. Орла. Прилегающая местность – волнистая равнина, пересеченная балками и оврагами, преимущественно занятая под посевы. Пойма правобережная, ровная, луговая, шириной 200 м, заполняется при уровне воды 350 см. Русло реки прямолинейное, песчано-галечное, деформирующееся, шириной 25 – 40 м. Левый берег сложен известняками, задернован, высотой 40 м; правый – суглинистый, высотой 2 – 2,5 м. В 0,8 км выше водпоста – плотина ГЭС, оказывающая влияние на уровненный режим реки. Вследствие попусков воды из вышерасположенного водохранилища сбора промышленных вод, а также выхода грунтовых вод на участке поста в течение зимы наблюдаются полыньи. В теплые зимы ледостав отсутствует. Водпост находится на левом берегу и состоит из свай, самописца уровня воды типа «Валдай» в деревенской будке, в 9 м ниже водпоста и реперов. Основной метеорологический репер № 1 - Курского УГСМ 1952 года в створе водпоста с высотой 155,000 м абс. Метеорологический репер № 2 – Курского УГМС 1955 года в 18 м выше водпоста с высотой 154,18 м абс. Высоты реперов получены нивелировкой IV Курского УГМС 1952 года от мет. репера б/н 1951 года в 0,8 км выше поста у плотины, с высотой 150,445 м абс. Сведений об исходном репере нет. Высота нуля графика – 145,26 м. абс. Гидростворы расположены: № 1 в створе, № 3 в 13 м выше водпоста. Температура воды измеряется в створе водпоста у берега, толщина льда – в створе водпоста на середине реки.

2.2. Оценка числовых характеристик

Оценка – приближенное значение искомой величины, полученное на основании результатов выборочного наблюдения, обеспечивающее возможность принятия обоснованных решений о неизвестных параметрах генеральной совокупности. Для полной характеристики оценки необходимо знать закон ее распределения. Однако некоторое представление о качестве оценки можно получить, исследуя ее общие свойства: состоятельность, несмещенность и эффективность. В тех случаях, когда закон распределения вероятностей (тип кривой) выбран, исходя из общих соображений, учитывающих, например, пределы изменения варьирующего признака, асимметричности распределения и т. д., в качестве основной возникает задача оценки параметров применительно к условиям данной конкретной совокупности. Очевидно, что решение этой задачи возможно лишь на основании той информации, которая содержится в материалах фактических наблюдений за рассматриваемым элементом гидрологического режима. Понятно так же, что, имея ряд такой же длительности, но за некоторый иной период времени, можно получить несколько иные значения параметров законов распределения рассматриваемого элемента режима. Значит, любое значение искомого параметра, вычисленное на основании ограниченного числа опытов, всегда будет содержать элемент случайности. Такое приближенное, случайное значение называют *оценкой* параметра.

Метод моментов.

Метод моментов заключается в приравнивании определённого количества выборочных моментов соответствующим моментам распределения, через которые могут быть выражены как функции неизвестные параметры распределения. Для оценки параметров распределения методом моментов используются те же формулы, что и для расчёта этих параметров по генеральной совокупности:

- эмпирическое математическое ожидание (среднее значение)

$$m_x^* = \bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}; \qquad (2.1)$$

(будем обозначать эмпирическое ожидание через $\bar{x}$);

дисперсия

$$D_x^* = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}; \qquad (2.2)$$

среднее квадратическое или стандартное отклонение

$$\sigma_x^* = \sqrt{D_x^*} = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}}; \qquad (2.3)$$

коэффициент вариации

$$\text{и} \quad C_v^* = \frac{\sigma_x}{\bar{x}} = \sqrt{\frac{\sum_{i=1}^{n}(k_i - 1)^2}{n}}; \qquad (2.4)$$

коэффициент асимметрии

$$C_x^* = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^3}{n \cdot \sigma_x^{*3}} = \frac{\sum_{i=1}^{n}(k_i - 1)^3}{n \cdot C_v^{*3}}. \qquad (2.5)$$

Достоинства метода моментов:

- моментные оценки параметров x, D_x, σ_x, C_v, C_s не зависят от закона распределения, тогда как другие методы оценки параметров от него зависят. Это свойство моментных оценок в немалой степени упрощает их практическое использование;
- эмпирическое математическое ожидание (среднее значение) является несмещённой и состоятельной оценкой математического ожидания.

Недостатки:

- Оценки дисперсии и коэффициента асимметрии смещены, и поправочные множители, вводимые в выражения для ликвидации погрешности в значениях соответствующих выборочных моментов μ_2 и μ_3 ,не всегда позволяют достигнуть соответствующего исправления оценок D_x^* и C_s^*;
- эффективность моментных оценок часто невысока, поэтому они не являются наилучшими.

Таким образом, применение метода моментов в задачах геоэкологии должно быть ограничено и иногда (когда есть необходимость) целесообразно заменять его другими, дающими оценки более высокой эффективности методами.

Метод наибольшего правдоподобия.

Сущность метода **наибольшего правдоподобия** состоит в том, что в качестве оценки неизвестного параметра a выбирается значение, максимизирующее функцию правдоподобия (совместная плотность вероятности, или вероятность получить данную выборку при данном a в результате n опытов). Это значение является функцией $x_1, x_2, \ldots x_n$ и называется оценкой наибольшего правдоподобия. Оценка эта зависит от закона распределения исследуемой величины и в каждом отдельном случае в зависимости от закона распределения должна рассматриваться заново.

Оценки параметров распределения нормального закона распределения методом наибольшего правдоподобия совпадают с моментными оценками.

При оценке параметров распределения закона Пирсона III-го типа вводят обозначение

$$\lambda = (1/n)\sum_i \ln k_i \ . \qquad (2.6)$$

Зависимость между λ и C_v обычно представляется в виде графика или таблицы $C_v=f(\lambda)$ в десятичных логарифмах, позволяющих с помощью простых вычислений находить оценку C_v при практических расчётах.

В явном виде коэффициент C_v выражается через λ соотношением

$$\widehat{C}_v \approx \sqrt{\frac{-6\lambda}{3-2\lambda}}\,, \qquad (2.7)$$

где λ задают в натуральных логарифмах.

Достоинства метода наибольшего правдоподобия:

- ✓ Если для параметра a существует эффективная оценка a^*, то она получается как единственное в этом случае решение уравнения правдоподобия;
- ✓ Статистики, найденные вышеупомянутым методом, основаны на использовании всей информации о данных параметрах, содержащейся в выборке;
- ✓ Всегда даёт состоятельные оценки;
- ✓ Оценки методом наибольшего правдоподобия могут быть смещёнными, но эта смещенность легко устраняется.

Недостатки метода наибольшего правдоподобия в следующем:

- ✖ Формулы оценки статистических параметров методом наибольшего правдоподобия зависят от закона распределения исследуемой величины. Так как закон распределения во многих случаях устанавливается по самим статистическим характеристикам, то это вносит некоторый элемент неопределённости в расчётную схему метода.
- ✖ Расчёты методом наибольшего правдоподобия часто требуют сложных вычислений.

Метод квантилей.

Метод квантилей основан на определении параметров кривых распределения непосредственно по сглаженной эмпирической кривой обеспеченности. При этом выборочные параметры $m_x^*, C_v^*,$ и C_s^* находятся как соответствующие функции от выборочных квантилей x_p^*, отвечающих заданному уровню вероятности превышения.

Порядок расчёта при применении метода квантилей по отношению к кривой обеспеченности Пирсона третьего порядка сводится к следующему:

1. В поле клетчатки вероятности строится эмпирическая кривая обеспеченности ряда X. В поле эмпирических точек проводится плавная осредняющая линия.

По формуле

$$S = \frac{x_5^* + x_{95}^* - 2x_{50}^*}{x_5^* - x_{95}^*}\,, \qquad (2.8)$$

где x_5^*, x_{95}^*, x_{50}^* – квантили эмпирической кривой обеспеченности обеспеченностью в 5, 50, 5%, вычисляется коэффициент скошенности.

2. По таблице C_s=f(S), разработанной для биномиального закона распределения, определяется коэффициент асимметрии.

3. По формуле $$\sigma = \frac{x_5^* - x_{95}^*}{t_5 - t_{95}} \qquad (2.9)$$

определяется среднее квадратическое отклонение. Здесь t_5 и t_{95} – нормированные отклонения от среднего значения, определяемые по C_s и данным вероятностям 5 и 95%. Формулу (39) получают из следующих выкладок:

$$t_5 = \frac{x_5^* - \bar{x}}{\sigma_x}; \quad t_{95} = \frac{x_{95}^* - \bar{x}}{\sigma_x}; \; t_5 - t_{95} = \frac{x_5 \;\; - x_{95}}{\sigma_x}. \qquad (2.10)$$

По формуле $$m_x^* = x_{50}^* - \sigma t_{50} \qquad (2.11)$$

Находится оценка математического ожидания.

Достоинство метода квантилей в том, что он прост в использовании, и во многих случаях позволяет получить результаты, не уступающие по точности расчётов другим методам.

Недостатки данного метода:

Он субъективен, т. к. результаты расчётов во многом зависят от проведения осредняющей линии в поле эмпирических точек. При коротком ряде наблюдений и значительном разбросе эмпирических точек в поле графика клетчатки вероятностей разные авторы могут провести сглаживающую линию по-разному.

На результаты расчётов оказывают влияние отдельные точки, расположенные в краевых частях графика.

Обоснование выбора метода и расчетных числовых характеристик.

Статистические характеристики, полученные методом моментов, отличаются от данных, указанных в условии задания (но незначительно). Коэффициент вариации максимального стока, полученный методом моментов, отличается от значения, полученного методом наибольшего правдоподобия. Если судить по оценке коэффициента вариации, то можно сказать, что метод моментов – точнее. Полагаю, можно порекомендовать применение метода моментов для числовых характеристик как наиболее простой в практическом использовании.

Статистические ряды для среднегодового и максимального стоков приведены в таблицах 1.1а и 1.1б в приложении 1. Расчёты числовых характеристик методом наибольшего прадоподобия приведены в таблицах 1.2а и 1.2б в приложении 1.

Оценка числовых характеристик приведены в таблицах 2.1 и 2.2.

***Доверительные интервалы числовых характеристик*.**

Доверительный интервал – доверительные интервалы для некоторой статистики –диапазон вокруг значения статистики, в котором находится истинное значение этой статистики.

Доверительный интервал для среднего – доверительные интервалы для среднего задают область вокруг среднего, в которой с заданным уровнем доверия содержится "истинное" среднее выборки. Ширина доверительного интервала зависит от размера выборки и дисперсии наблюдений. Вычисление доверительных интервалов основывается на предположении, что переменная в совокупности нормально распределена. Доверительный интервал для параметра *x* определяется по формуле

$$\bar{x} - t_{2\alpha}\frac{\sigma}{\sqrt{n}} < M < \bar{x} + t_{2\alpha}\frac{\sigma}{\sqrt{n}}. \qquad (2.12)$$

Для определения доверительного интервала дисперсии нормально распределенной случайной величины *X* используют следующее выражение:

$$p(\frac{nD*}{\chi_2^2} < D < \frac{nD*}{\chi_1^2}) = 1 - \alpha. \qquad (2.13)$$

где $D*$ – выборочное значение дисперсии, D – истинная оценка дисперсии; значения χ_2^2 и χ_1^2 выбираются так, чтобы выполнялось следующее условие:

$$p(\chi^2 < \chi_1^2) = p(\chi^2 > \chi_2^2) == \alpha/2.$$

С учетом этого условия доверительный интервал для дисперсии приближенно определяется формулой

$$\frac{nD*}{\chi_2^2} < D < \frac{nD*}{\chi_1^2}. \qquad (2.14)$$

Аналогичным образом находится истинная оценка для стандартного отклонения нормально распределенной случайной величины

$$\sqrt{n}\,\frac{\sigma^*}{\chi_2} < \sigma < \sqrt{n}\,\frac{\sigma^*}{\chi_1}. \qquad (2.15)$$

Числа χ_1^2 и χ_2^2 находятся по значениям $\nu = n - 1$ и из таблицы χ^2-распределения как квантили $\chi^2_{\frac{\alpha}{2},\nu}$ и $\chi^2_{(1-\alpha/2),\nu}$.

Доверительные интервалы для математического ожидания, дисперсии и среднего квадратического отклонения при 2α=10% вычислены для среднегодового и максимального стоков и приведены в таблице 2.3.

3. Оценка однородности исходных рядов наблюдений

Статистический анализ однородности рядов наблюдений включает следующие основные этапы: формулировку нулевой и альтернативных гипотез, определение уровня значимости, выбор критической области, браковку или признание нулевой гипотезы. Так как эти этапы являются, как правило, неотъемлемой частью любого статистического исследования однородности рядов наблюдений, остановимся кратко на них. Прежде всего условимся, что результаты наблюдений будем считать однородными тогда, когда они принадлежат к одной и той же генеральной совокупности. При этом все наблюдения будем считать независимыми как внутри рядов наблюдений (иными словами соблюдено условие случайности отбора), так и между исследуемыми рядами наблюдений.

Любое статистическое суждение об однородности тех или иных рядов наблюдений всегда имеет вероятностный характер. Статистический анализ однородности рядов наблюдений начинается с предположения отсутствия существенного различия между параметрами сравниваемых рядов (нулевая гипотеза). При этом обычно предполагается, что закон распределения сопоставляемых рядов наблюдений один и тот же, что следует из физических соображений или накопленного предыдущего опыта, но могут иметь место лишь различия в параметрах распределения, к числу которых могут быть отнесены: среднее значение, коэффициент вариации и коэффициент асимметрии. Во многих случаях подлежат проверке на нулевую гипотезу все параметры распределения. *Гипотезы, противоположные нулевой, называются альтернативными.*

Уровнем значимости будем считать такое достаточно малое значение вероятности, которое в том или ином конкретном случае может считаться характеризующим практически невозможное событие. Появление такого редкого события указывает на неправильность принятой нулевой гипотезы с вероятностью, не превышающей выбранный уровень значимости. В таком случае с вероятностью, равной выбранному уровню значимости, можно отвергнуть нулевую гипотезу, хотя она может оказаться правильной, или, как говорят, совершить ошибку первого рода. В другом случае, задаваясь некоторым достаточно малым уровнем значимости, можно принять неправильную альтернативную гипотезу или совершить ошибку второго рода. Очевидно, что полностью избежать ошибок первого и второго рода нельзя. При этом всегда имеет место некоторый риск. Можно лишь уменьшить риск совершения ошибки одного рода за счет увеличения ошибки другого рода. Обычно за уровень значимости принимают вероятность 5, 2 или 1%-ную. С уменьшением уровня значимости вероятность забраковать нулевую гипотезу уменьшается, когда она верна и, следовательно, уменьшается вероятность совершения ошибки первого рода. Но с уменьшением уровня значимости увеличивается область допустимых значений и, следовательно, увеличивается вероятность принятия нулевой гипотезы, когда она неверна, или увеличивается вероятность совершения ошибки второго рода. С другой стороны, увеличивая уровень значимости, мы увеличиваем вероятность совершения ошибок первого рода (т. е. отвергнуть исходную нулевую гипотезу, хотя она верна) и соответственно уменьшаем вероятность совершения ошибок второго рода.

Выбор **критической области** осуществляется таким образом, чтобы вероятность попадания в нее, когда гипотеза верна, в точности была равна уровню значимости. Область, которая дополняет критическую, обычно называют областью допустимых значений, или областью принятия. Выбор критической области при заданном уровне значимости необходимо осуществлять, исходя из тех или иных физических соображений и предполагаемых различий в параметрах их распределений. Иными словами, критическую область следует выбирать таким образом, чтобы вероятность попадания в нее критерия была наибольшей, когда справедлива альтернативная гипотеза, т. е. гипотеза, конкурирующая с нулевой гипотезой. Чем больше эта вероятность, которая часто называется мощностью критерия, тем меньше вероятность совершения ошибки второго рода.

3.1 Оценка однородности исходных рядов наблюдений по среднему значению

Оценка среднего значения

Оценка среднего значения производится при помощи статистики *t*, являющейся нормированным отклонением выборочного среднего $\bar{x} = m$, от математического ожидания m_x

$$\hat{t} = \frac{\bar{x} - m_x}{\sigma_{\bar{x}}} = \frac{\bar{x} - m_x}{\frac{\sigma_x}{\sqrt{n}}}. \tag{3.1}$$

Математическое ожидание статистики *t* и ее среднее квадратическое отклонение равны соответственно 0 и 1. Распределение статистики *t* подчиняется при некоторых условиях закону *t*-распределения Стьюдента.

Функция распределения статистики *t* описывается формулой

$$S_\nu(t) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu}\,\Gamma\left(\frac{\nu}{2}\right)} \int_{-\infty}^{t} \left(1 + \frac{\tau^2}{\nu}\right)^{-\frac{\nu+1}{2}} d\tau. \tag{3.2}$$

где ν – число степеней свободы $\nu = n - 1$.

Формула (3.1) выражает вероятность случайных значений τ меньших, чем заданные значения *t*, т.

$$S_\nu(t) = p(\tau < t).$$

Указанное соотношение может быть использовано для определения значимости или существенности расхождения между выборочным и предполагаемым генеральным значением параметра распределения *a,* или между действительными значениями параметра *a* по разным выборкам. Так, если вероятность данного расхождения выборочной характеристики с генеральной составляет менее 1–5 %, т. е. очень мало, то, по-видимому, расхождение существенно или значимо. Обычно вопрос о значимости решается с помощью таблицы значений *t,* соответствующих данному уровню значимости α при данном числе степеней свободы ν. При малых значениях степеней свободы ν *t*-распределение Стьюдента заметно отличается от нормального распределения.

t-распределение Стьюдента играет важную роль в гидрологических исследованиях. Наибольшее распространение оно получило для оценки доверительных границ математического ожидания, оценки значимости среднего, оценки расхождения средних значений по двум и более рядам значений исследуемых процессов.

Определение доверительных границ математического ожидания

Пусть имеется выборка значений случайной величины $X(x_1, x_n, \ldots, x_n)$ объемом в *n* членов. Среднее значение выборки $\bar{x}$, несмещенная оценка дисперсии $\overline{D}_x$. На основании этих данных требуется найти границы, внутри которых с определенной степенью надежности находится среднее значение генеральной совокупности m_x.

Выбрав двухсторонний 10 %-й уровень значимости (2α=10 %), при данном числе степеней свободы ν находим $t_{0,05}$ такое, что $p(|\hat{t}| \geq t_{0.05}) = 0{,}05$

т. е. вероятность того, что $|t|$ будет больше или равно табличному $t_{0,05}$ составляет 0,05. Тогда вероятность противоположного неравенства будет равна $p(|\hat{t}| < t_{0.05}) = 0{,}95$,

т. е. вероятность того, что $|t|$ будет меньше табличного или, иначе говоря, вероятность того, что *t* находится в пределах $\pm t_{0,05}$ $p(-t_{0.05} < \hat{t} < t_{0.05})$, составляет 90 %. Подставляя в это равенство значение $\hat{t}$, находим отсюда

$$\bar{x} - t_{0.05}\frac{\sigma}{\sqrt{n}} \le m_x < \bar{x} + t_{0.05}\frac{\sigma}{\sqrt{n}}. \qquad (3.3)$$

Таким образом, с вероятностью 0,90 можно утверждать, что среднее значение общей совокупности лежит между значениями $x_н$ и $x_в$, равными соответственно $x_н = \bar{x} - t_{0.05}\sigma/\sqrt{n}$, $x_в = \bar{x} + t_{0.05}\sigma/\sqrt{n}$. Эти значения называются доверительными границами среднего значения общей совокупности при двухстороннем 5%-м уровне значимости, а вероятность того, что промежуток с этими крайними значениями накроет среднее значение общей совокупности, называется доверительной вероятностью.

Оценка расхождения между средними значениями

Очень часто в исследованиях встает вопрос о сравнении математических ожиданий двух или более нормально распределенных случайных величин по их выборкам.

Пусть, например, имеются две независимые частичные совокупности объемом n_1 и n_2, взятые соответственно из генеральных совокупностей X и Y. Нулевая гипотеза состоит в предположении, что математические ожидания ряда X и Y равны, т. е.

H_0: $m_x = m_y$. В качестве критерия проверки H_o естественно взять нормированную разность средних выборочных значений:

$$\hat{t} = \frac{\bar{y} - \bar{x}}{\sigma_{\bar{y}-\bar{x}}}. \qquad (3.4)$$

где $\bar{y}$ и $\bar{x}$ – соответственно средние значения выборок ряда Y и X. Среднее квадратическое отклонение разности выборочных средних $\bar{x}$ и $\bar{y}$ при условии независимости рядов X и У может быть определено по формуле

$$\sigma_{(\bar{y}-\bar{x})} = \frac{\sigma_{\bar{y}}^2}{n_2} + \frac{\sigma_{\bar{x}}^2}{n_1} = \frac{n_1\sigma_{\bar{y}}^2 + n_2\sigma_{\bar{x}}^2}{n_1 \cdot n_2}. \qquad (3.5)$$

или

$$\sigma^2_{(\bar{y}-\bar{x})} = \frac{n_1 + n_2}{n_1 \cdot n_2} \cdot \hat{\sigma}^2. \qquad (3.6)$$

где

$$\hat{\sigma}^2 = \frac{n_2\sigma_{\bar{y}}^2 + n_1\sigma_{\bar{x}}^2}{n_1 + n_2} = \hat{D}. \qquad (3.7)$$

Отсюда

$$\hat{t} = \frac{\bar{y} - \bar{x}}{\hat{\sigma}} \cdot \sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}}. \qquad (3.8)$$

Разность $\bar{y} - \bar{x}$ двух выборочных средних следует, как известно, закону распределения Стьюдента–Госсета.

Отсюда по таблице распределения t может быть определено максимальное значение t при данном уровне значимости 2α и числе степеней свободы ν

$$\nu = n_1 + n_2 - 2. \qquad (3.9)$$

Обычно, как и в предыдущих случаях, в качестве уровня значимости принимается 2α равным 5 или 10%.

Если значение $\hat{t} > t$ при данном уровне значимости, то нулевая гипотеза о равенстве математических ожиданий рядов X и Y опровергается.

3.2. Оценка однородности исходных рядов наблюдений по дисперсии

Оценка производится при помощи статистики F, называемой дисперсионным отношением:

$$\hat{F} = \frac{\hat{D}_1}{\hat{D}_2}, \hat{D}_1 > \hat{D}_2. \qquad (3.10)$$

Функция обеспеченности статистики F, выражающая вероятность того, что значение $\hat{F}$ будет больше или равно некоторому значению F

$$P(\hat{F}) = P(\hat{F} > F) = \alpha, \qquad (3.11)$$

где α (в данных исследованиях уровень значимости) установлена Фишером.
Функция плотности вероятности этого распределения (F-распределение) описывается формулой

$$f(F) = \frac{\nu^{\frac{\nu_1}{2}} \nu^{\frac{\nu_2}{2}} \Gamma(\frac{\nu_1 + \nu_2}{2})}{\Gamma(\frac{\nu_1}{2}) \Gamma(\frac{\nu_2}{2})}, \qquad (3.12)$$

где ν_1 и ν_2 – числа степеней свободы первой и второй выборки, $\nu_1 = n_1 - 1$, $\nu_2 = n_2 - 1$.
Как следует из формулы (3.12), F-распределение не зависит от дисперсии исходных рядов, а зависит только от числа степеней свободы. Это обстоятельство является очень важным, так как именно дисперсию и требуется установить в результате тех или иных исследований.
Для F-распределения составлены таблицы значений в зависимости от числа степеней свободы и уровней значимости. Имеются таблицы значений *F,* которые могут быть превзойдены с вероятностью α 0,05; 0,025; 0,01; 0,005.
Для оценки доверительного интервала и значимости дисперсии используется отношение

$$\hat{\chi}^2 = \frac{(n-1)\hat{D}}{D}, \qquad (3.13)$$

которое удовлетворяет χ^2-распределению с $\nu = n-1$ степенями свободы, если гипотеза *Ho:* $\hat{D} = D$ верна.

Оценка равенства дисперсий

Оценки гипотезы *Ho:*$D_1 = D_2$ производятся по критерию Фишера (3.10). Если значение $\hat{D}$ меньше $\hat{D}_\alpha$, определенного по таблице при данном уровне значимости и числах степеней свободы $\nu_1 = n_1 - 1$ и $\nu_2 = n_2 - 1$, то расхождение между D_1 и D_2 можно считать случайным, а следовательно, и гипотезу о том, что выборки взяты из нормальных общих совокупностей с одинаковой дисперсией, в какой-то степени подтвердившейся. Если же F* больше *F,* то расхождение существенно и гипотеза о равенстве дисперсий должна быть опровергнута.
В некоторых случаях для оценки равенства дисперсий используется критерий Романовского. Для оценки равенства дисперсий с помощью критерия Романовского не требуется специальных таблиц. Расчет производится в следующей последовательности.
Сначала рассчитывается вспомогательная величина θ:

$$\theta = \frac{\nu_2 - 2}{\nu_2} F^*, \qquad (3.14)$$

где F^* – дисперсионное отношение, ν_2 – число степеней свободы больше 4.
Математическое ожидание θ для выборок из нормальной совокупности с одинаковой дисперсией равно единице, т. е.

$$M[\theta] = 1, \qquad (3.15)$$

а среднее квадратическое отклонение
$$\sigma(\theta) = \sqrt{\frac{2(\nu_1 + \nu_2 - 2)}{\nu_1(\nu_2 - 4)}}. \qquad (3.16)$$

С вероятностью $p \geq 0{,}99$ можно ожидать, что абсолютная величина отклонения θ от 1 по модулю не превышает 3σ. Отсюда следует, что, если выборки относятся к одной генеральной совокупности, то

$$R = \frac{|\theta - 1|}{\sigma(\theta)} < 3 \qquad (3.17)$$

т. е. расхождение выборочных оценок дисперсий несущественно. Если $R \geq 3$, то расхождение существенно.

3.3. Оценка однородности исходных рядов наблюдений по критерию Уилкоксона

Фактически этот критерий достаточно чувствителен по отношению к выборочным средним и почти не реагирует на изменение выборочных дисперсий. Поэтому точнее рассматриваемый критерий распространять на оценку однородности выборочных средних. Критерий Уилкоксона основан на подсчете числа так называемых инверсий, которые выявляются в результате следующей процедуры. Наблюдения, составляющие две выборки располагаются в общей последовательности в порядке убывания или возрастания их значений, например, в виде

$$y_1x_1x_2y_2y_3y_4x_3y_5y_6x_4,$$

где $x_1, ..., x_4$ – члены, принадлежащие первой выборке; $y_1, ..., y_6$ – члены второй выборки. Если какому-либо значению x предшествует некоторое значение y, то говорят, что эта пара образует инверсию. Так, в рассматриваемой последовательности x_1 и x_2 образуют по одной инверсии с y_1; x_3 образует четыре инверсии (с y_4, y_3, y_2 и y_1) и x_4 дает шесть инверсий (с y_6, y_5, y_4, y_3, y_2 и y_1). Всего инверсий в данном случае будет

u = 1+1+4 + 6=12. Теоретически доказано, что в однородных рядах, каждый из которых представлен выборкой объемом не менее 10 членов, число инверсий распределено приблизительно по нормальному закону с математическим ожиданием

$$Mu = \frac{m(m+n+1)}{2}, \qquad (3.18)$$

и дисперсией $$Du = \frac{mn}{12}(m+n+1), \qquad (3.19)$$

где n и m – число членов первой и второй выборки. В качестве нулевой гипотезы, учитывая сказанное в начале настоящего раздела относительно возможностей рассматриваемого критерия, примем гипотезу принадлежности выборочных средних к одной генеральной совокупности $(x = y)$. Теперь необходимо выбрать границу допустимых значений, выделяющую критическую область. Задавшись уровнем значимости q = 0,1; 1,0; 5% и т. д., выделяем область больших по абсолютной величине отклонений, вероятность попадания в которую в случае, когда гипотеза однородности верна, в точности равна уровню значимости. Тогда вероятность попадания в область допустимых значений при справедливости нашей гипотезы будет

$$\beta=(100-q)\%. \qquad (3.20)$$

Вероятность β называется уровнем доверительной вероятности.

Если значение критерия, вычисленное по данным наблюдений, окажется в критической области, то нулевая гипотеза однородности бракуется, и с вероятностью β принимается альтернативная гипотеза неоднородности. Если же значение критерия окажется в области допустимых отклонений от математического ожидания, то можно еще утверждать, что нулевая гипотеза подтверждается.

Критическая область для нулевой гипотезы однородности будет область больших по абсолютной величине отклонений:

$$u_н = Mu - t_p\sigma_u,$$
$$u_в = Mu + t_p\sigma_u,$$

где $\sigma_u = \sqrt{Du}$; t_p – нормированное отклонение при принятом уровне значимости q.

Отметим, что критерий однородности Уилкоксона применяется для попарного сравнения выборок в S пунктах наблюдений некоторого предполагаемого однородным региона. Известные обобщения этого критерия для случая более двух выборок отличаются большой громоздкостью и сложностью. Стремление к математической точности сильно усложняет расчет статистики критериев и ее критических значений. Это затрудняет применение таких критериев и делает их малоэффективными.

3.4. Анализ результатов оценки и влияния хозяйственной деятельности

Сравнение 20-летних выборок внутри среднегодового процесса показало, что расхождение между средними значениями и дисперсиями можно считать случайным, что доказывает их однородность и принадлежность к общим генеральным совокупностям.

Сравнение 20-летних выборок внутри максимального процесса показало, что расхождение между средними значениями велико, оценка равенства дисперсий по критериям Фишера и Романовского показала, что расхождение между ними случайно.

Для 40-летних выборок расхождения значений велики, и, значит, нулевые гипотезы опровергаются. Это значит, что изучаемые процессы неоднородны между собой по физическим свойствам и статистическим характеристикам. Сравнение представлено в таблицах 3.1, 3.2, 3.3, 3.4, 3.5.

Оценка однородности среднегодового стока по критерию Уилкоксона показала, что выборки принадлежат к одной генеральной совокупности, т. е. нулевая гипотеза не опровергается (Ho:

F_1=F_2).
Оценка однородности максимального стока по критерию Уилкоксона показала, что выборки не принадлежат к одной генеральной совокупности, т. е. нулевая гипотеза опровергается.
Расчёты приведены в таблицах 3.6, 3.7 и в приложении 1 (таблицы 1.3а и 1.3б).
При изучении процесса максимального стока отметим, что математическое ожидание второй 20-летней выборки почти вдвое меньше, чем первой, что может объясняться влиянием промышленной деятельности: забор воды для использования промышленными предприятиями, населением и в системах орошения. Объем воды, отводимой обратно в речную сеть, составляет около 80% водопотребления. Забор воды из реки для целей сельскохозяйственного водоснабжения и орошения составляет 4 % общего объема Отмечается увеличение значения дисперсий во вторых 20-летних выборках среднегодового и максимального стоков, что также говорит об искуственном вмешательстве. На уровненный режим реки оказывает влияние плотина ГЭС, расположенная в 0,8 км выше водпоста.

4. Оценка стационарности и случайности исходных рядов

Стационарность процессов и однородность имеющегося ряда наблюдений контролируются одинаковыми статистическими показателями, и различие их может констатироваться только на основе физического анализа.
Для оценки стационарности и однородности рядов наблюдений по среднему и дисперсии используются рассмотренные статистические критерии Фишера, при общей оценке однородности – критерии Уилкоксона. В последнее время для этого применяются также критерии Диксона и Смирнова–Граббса.
В некоторых случаях нестационарность процесса может определяться наличием тренда. В гидрологической практике *трендом* называется медленное, постепенное изменение случайной переменной X в течение анализируемого периода. Часто тренд является частью колебаний с периодами, длительность которых сравнима с продолжительностью периода наблюдений. Для оценки наличия тренда обычно используется коэффициент корреляции рассматриваемого ряда X со временем или критерий Спирмэна.

4.1. Оценка стационарности по связи интегральных кривых

Во многих случаях нам известна та точка отсчета (дата), начиная с которой произошло нарушение однородности. Такими точками отсчета могут быть: дата начала заполнения водохранилища, дата начала работы крупных водозаборных сооружений и т. д. Однако в некоторых случаях эта дата неизвестна. Для ее установления может использоваться график связи интегральной кривой
$u_j=\sum_{i=1}^{j} k_{xi}$ $(j= 1, 2, ..., n)$ по исследуемому ряду X с интегральной кривой $v_j=\sum_{i=1}^{j} k_{xi}$ $(j= 1, 2, ..., n)$ ряда Y, относительно которого однородность не вызывает сомнений.

4.2. Оценка наличия тренда по критерию Спирмэна

Одним из наиболее распространённых при исследовании зависимости процесса от времени является ранговый критерий Спирмэна

$$\hat{\rho} = 1 - 6\frac{\sum_{i=1}^{n} d_i^2}{n^3 - n} \tag{4.1}$$

где d_i – разность между порядковым номером и рангом каждого хронологического значения ряда длиной n. Если связь отсутствует, то ρ = 0; при прямой связи, т. е. при постепенном нарастании X, ρ = 1; при обратной связи, т. е. постепенном уменьшении X, ρ = –1.
Математическое ожидание ρ при отсутствии тренда M[ρ] = 0, дисперсия $D_\rho=1/(n-1)$.

4.3. Оценка тренда по коэффициенту корреляции

Для оценки тренда используется формула коэффициента корреляции

$$\widehat{r}_{xi} = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(i - \tau)}{(n-1)\widehat{\sigma}_x \sigma_i}, \tag{4.2}$$

где i–порядковый номер члена ряда; $\bar{i} = n(n+1)/2$; $\widehat{\sigma}_x$, σ_i,– средние квадратические отклонения ряда X и порядковых номеров i.

Если значение r_x, окажется значимым , то очевидно наличие тренда.

4.4. Оценка случайности по критериям

Для решения вопроса о наличии или отсутствии внутрирядных связей производится проверка рассматриваемых рядов наблюдении на случайность. С учетом выборочности рядов наблюдений такая проверка имеет статистический характер и основана на нулевой гипотезе об отсутствии внутрирядных связей. Если гипотеза верна, т. е. рассматриваемый ряд является чисто случайным, то он обладает некоторыми свойствами, связанными главным образом с характером группировок (серий) членов ряда по тем или иным признакам. В качестве признаков группировок могут выступать отклонения членов ряда от среднего значения, число подъемов и спадов, число экстремумов и т. д. Если выборочный ряд с большой степенью достоверности обладает указанными свойствами, то он является случайным, если нет, то нулевая гипотеза опровергается и с заданной степенью уверенности можно принять этот ряд не случайным, т. е. внутрирядно связанным.

Оценка случайности по критерию длин и числа серий

Серия – последовательность элементов (членов ряда) одинаковых по какому-либо признаку, непосредственно перед которой и после которой стоят элементы другого вида. Длиной серии называется число элементов в серии. Чаще всего в качестве признака серий берется знак отклонений элементов от медианы по ряду наблюдений.

Проверка случайности по данному критерию основана на нулевой гипотезе об отсутствии внутрирядных связей и проводится путем сопоставления длин и числа серий исследуемого ряда с длиной и числом серий выборок случайной величины той же продолжительности. Число серий различной длины в выборках случайной величины различной продолжительности при заданном уровне значимости устанавливается теоретически. Так, математическое ожидание числа серий с длиной, большей или равной k, состоящих из элементов больших и меньших медианы , определяется соответственно по формулам:

$$MR_{1,k} = (n_1 + 1)\frac{\prod_{i=1}^{k}(n_1 - i + 1)}{\prod_{i=1}^{k}(n - i + 1)}; \qquad MR_{2,k} = (n_2 + 1)\frac{\prod_{i=1}^{k}(n_2 - i + 1)}{\prod_{i=1}^{k}(n - i + 1)}, \tag{4.3}$$

где $R_{1,k}$ и $R_{2,k}$ – число серий, длина которого не меньше k, соответственно из членов ряда больших и меньших медианы; n_1 и n_2 – число членов ряда больших и меньших медианы; n – число членов ряда, П – знак произведения.

При ручном счете вместо формул (4.3) пользуются вытекающими из них рекуррентными формулами:

$$MR_{1,k} = MR_{1,k-1}\frac{n_1 - (k-1)}{n - (k-1)}; \qquad MR_{2,k} = MR_{2,k-1}\frac{n_2 - (k-1)}{n - (k-1)}, \tag{4.4}$$

Зная математическое ожидание числа серий, длина которых не меньше k, можно рассчитать математическое ожидание числа серий данной продолжительности k:

$$Mr_{1,k} = MR_{1,k} - MR_{1,k+1}; \qquad Mr_{2,k} = MR_{2,k} - MR_{2,k+1}, \tag{4.5}$$

Математическое ожидание общего числа серий с длиной, не меньшей k, можно рассчитать путем суммирования $MR_{1,k}$ и $MR_{2,k}$, т. е.

$$MR_k = MR_{1,k} + MR_{2,k} \tag{4.6}$$

Соответственно математическое ожидание общего числа серий данной продолжительности

$$Mr_k = Mr_{1,k} + Mr_{2,k} \tag{4.7}$$

Если по исследуемому ряду наблюдений общее число серий окажется меньше критических границ при 5 %-м или 2,5 %-м уровне значимости, то можно считать, что нулевая гипотеза о случайности ряда опровергается, и ряд является неслучайным.
При подсчете числа серий в исследуемом ряду при ручном счете члены исходного ряда больше медианы обычно обозначаются через «а», меньше медианы — через «в»; при машинном счете — через «0» и «1». Незамкнутые серии в начале и конце ряда наблюдений исключаются.

Оценка случайности по критерию числа повышений и понижений
Сущность проверки состоит в следующем. Пусть имеется выборка $x_1, x_2, \ldots, x_n$. Переход $x_{i-1} < x_i$ называется повышением (+), а $x_i > x_{i+1}$ – понижением (–). Общее число повышений (или понижений) ряда значений случайной величины распределено
асимптотически нормально с математическим ожиданием

$$m_+=m_-=n/2 \tag{4.8}$$

и дисперсией

$$D_+ =D_- = (n + 1)/12. \tag{4.9}$$

Зная m и D, можно рассчитать по данному ряду нормированную величину числа повышений или понижений

$$\widehat{t}_+ = \frac{\widehat{n}_+^* - m_+}{\sqrt{D_+}}, \ \widehat{t}_- = \frac{\widehat{n}_-^* - m_-}{\sqrt{D_-}}, \tag{4.10}$$

где n_+ и n_- – число повышений и понижений в исследуемом ряду соответственно. Дальше, учитывая асимптотически нормальный закон распределения числа повышений и понижений, сравниваем t_+ или t_- со значениями нормированных ординат таблицы нормального закона распределения. Если вероятность рассчитанных значений t_+ или t_- по таблице окажется меньше заданного уровня значимости, то гипотеза о случайности исследуемого ряда опровергается и считается, что ряд имеет устойчивую тенденцию к повышению или понижению. С помощью этого критерия хорошо обнаруживаются имеющиеся систематические изменения уровня ряда (тренда).

Оценка случайности по критерию числа экстремумов
Экстремумом называется любой элемент, принадлежащий последовательности $x_1, x_2, \ldots, x_n$, для которого выполняется одно из неравенств

$$x_{i-1}<x_i>x_{i+1} \text{ или } x_{i-1}>x_i<x_{i+1}.$$

В первом случае x_i – максимум, во втором – минимум. Общее число экстремумов случайного ряда распределено асимптотически нормально с математическим ожиданием

$$m_э = 2n/3 \tag{4.11}$$

и дисперсией

$$D_э = (16n - 29)/90. \tag{4.12}$$

Для проверки гипотезы случайности ряда X достаточно рассчитать по этому ряду фактическое нормированное число экстремумов

$$t_э^* = \frac{n_э^* - m_э}{\sqrt{D_э}} \tag{4.13}$$

и сравнить его с нормированным значением t_α нормального закона распределения при заданном уровне значимости α. При $t_э^* > t_\alpha$ гипотеза о случайности ряда, по-видимому, неверна. С помощью этого критерия хорошо обнаруживаются характерные циклы в ряду наблюдений. При подсчете числа повышений (понижений) и числа экстремумов может обнаружиться, что два соседних элемента x_j и x_{j+1} равны между собой. В этом случае один из них следует исключить, а общее число членов ряда уменьшить на 1.

4.5. Анализ и выводы о стационарности исследуемых процессов
Проведена проверка выборочных рядов среднегодового и максимального стоков на случайность по:
а) по критерию длин и числа серий.
Нулевая гипотеза: случайность процессов;
Для среднегодового и максимального стоков общее фактическое число серий входит в

допустимый интервал, следовательно, гипотеза о случайности процессов не опровергается. Расчёты – в таблице 4.1 и в приложении 1 (таблицы 1.4, 1.5а и 1.5б).
б) по числу повышений и понижений
нулевая гипотеза: процессы случайны;
вычисленные статистики $\widehat{t}_+$ и $\widehat{t}_-$ сравниваем с критическим значением t_α, результат – для максимального стока $\widehat{t}_+$ и $\widehat{t}_-$ $< t_\alpha$, следовательно, гипотеза о случайности процессов не опровергается.
Для среднегодового стока $\widehat{t}_+ > t_\alpha$, следовательно, гипотеза о случайности процесса опровергается и данный процесс не является случайным, а имеет тенденцию к повышению.
Расчёты – в таблице 4.2 и в приложении 1 (таблицы 1.6а и 1.6б).
в) по числу экстремумов
нулевая гипотеза: процессы случайны;
рассчитанные статистики t для среднегодового и максимального стоков, при сравнении их с t_α говорят о том, что гипотеза о случайности процессов не опровергается.
Расчёты – в таблице 4.3 и в приложении 1 (таблицы 1.6а и 1.6б).
При анализе исходных рядов среднегодового и максимального стоков на стационарность и однородность получены следующие результаты:
а) при определении даты нарушения однородности по интегральной кривой **(рис. 4.1)** можно сделать вывод о том, что в течение рассматриваемого периода наиболее существенное изменение характера связи наблюдалось в периоды: 1945 –1947 гг., 1950 – 1952 гг., 1962 – 1964 гг., 1970 – 1974 гг., 1975 – 1979 гг. Таким образом, во второй половине исследуемых рядов наблюдалось активное вмешательство человека.
Расчёты приведены в приложении 1 (таблица 1.7).
б) проведена оценка тренда по коэффициенту корреляции ($\underline{Ho: r_{i,Q}= 0}$, для среднегодового стока r_x= 0,108, для максимального стока r_y= -0,350) и по критерию Спирмэна (для среднегодового стока ρ_x= -0,092, для максимального стока ρ_y= 0,348).
В итоге, с определённой долей вероятности мы можем утверждать, что в рассматриваемых процессах *тренд отсутствует*, поскольку рассчитанные значения r_x и r_y оказались незначимыми. Кроме того, проведённая оценка с помощью критерия Спирмэна показала наличие слабой обратной связи.
Расчёты приведены в таблицах 4.4а, 4.4б и 4.5.
Таким образом, при исследовании процессов среднегодового и максимального стоков с помощью различных критериев наблюдались расхождения в оценке среднегодового стока, что, по моему мнению, является подтверждением искусственного вмешательства (плотина ГЭС).

5. Оценка парной связи между исследуемыми процессами.

Для исследования взаимосвязи между двумя природными процессами или явлениями чаще всего используется математическая модель в виде уравнения регрессии. При этом исследование состоит из двух этапов.

- Выявление на основе большого количества наблюдении того, как изменяется в среднем функция *Y* в зависимости от изменения её аргумента. Эта задача кратко формулируется как **задача определения формы и нахождения уравнения связи двух переменных величин.**

Общий вид линейной связи

$$\overline{y}\,(x)=ax+b \qquad (5.1)$$

где $\overline{y}\,(x)$–среднее из возможных значений *Y* при данном *x*.

Функция (5.1), выражающая связь между значениями аргумента и условным средним арифметическим исследуемой зависимой переменной, называется **уравнением линии регрессии**.

- Определение степени взаимосвязи между исследуемыми явлениями (если это связи сопряженности) или степени влияния факторов на исследуемое явление (если эти связи носят причинно-следственный характер).

Направление линий регрессии в поле графика определяется коэффициентами регрессии *a* и *a'*. Первый из них представляет собой тангенс угла наклона линии регрессии *y =f(x)* к оси *x*; второй – тангенс угла наклона линии регрессии *x = f(y)* к оси *y*. Обозначим эти углы через α и β.

Тогда коэффициенты регрессии

$$a = \mathrm{tg}\alpha, \quad a' = \mathrm{tg}\beta. \qquad (5.2)$$

Сумма углов α, β, φ составляет 90°. В случае, если связь между *Y* и *X* функциональная, то ∠φ = 0° и ∠α+∠β = 90°. Отсюда ∠α = 90°– ∠β и

$$\mathrm{tg}\alpha = \mathrm{tg}(90°-\beta) = \mathrm{ctg}\beta = 1/\mathrm{tg}\beta, \text{т. е.} \quad \mathrm{tg}\alpha\,\mathrm{tg}\beta = 1. \qquad (5.3)$$

Если связь между *Y* и *X* отсутствует, то ∠(β = 90°, а ∠α =∠β и tgαtgβ = 0.

Для стохастических связей с увеличением тесноты связи угол φ уменьшается от 90 до 0°; вместе с тем увеличиваются углы α и β, а, следовательно, тангенсы этих углов и их произведение. Таким образом, произведение тангенсов углов α и β может служить мерой тесноты связи между *X* и *Y*. Обычно в качестве критерия степени близости корреляционной связи к линейной функциональной зависимости используется корень квадратный из произведения tgαtgβ. Это произведение называется коэффициентом корреляции двух переменных величин и обозначается через *r* или *R*.

5.1. График связи значений исследуемых процессов и ранжирования, анализ зависимостей.

Графики связи неранжированных и ранжированных значений среднегодового и максимального стоков Y=*f*(X) (рис. 5.1 и рис. 5.2), представляют собой графики связи стохастически зависимых величин, и с большой долей вероятности мы можем утверждать, что связь присутствует, причём связь эта прямая. На графике связи неранжированных значений чётко прослеживается некоторая криволинейность, которая объясняется наличием поймы.

Проведя по точкам на рис. 5.1 и 5.2 линии связи, можем отметить, что угол наклона линии связи на рис. 5.1 меньше, чем на рис. 5.2. Это говорит о том, что на графике связи ранжированных значений представлена более тесная связь, чем на рис 5.1. Протяжённость поля точек по длине линии связи на обоих рисунках в два раза превышает наибольший разброс точек относительно линии связи, поэтому считаем, что данная теснота связи достаточна для наших дальнейших исследований. Логичным считаю отметить, что, несмотря на разность тесноты связи на рисунках 5.1 и 5.2, общим для них является «отскакивание» точек редкой повторяемости, в частности, точки (29,7;1510).

График связи неранжированных и ранжированных значений среднегодового и максимального стоков Y=*f*(X) изображён на рис.5.1 и 5.2.

5.2. Построение линии регрессии по рассчитанным параметрам.

Для исследования взаимосвязи между двумя природными процессами или явлениями чаще всего используется математическая модель в виде уравнения регрессии.

Общий вид линейной связи

$$\overline{y}\,(x) = ax + b \qquad (5.4)$$

где $\overline{y}\,(x)$–среднее из возможных значений *Y* при данном *x*.

Функция (5.4), выражающая связь между значениями аргумента и условным средним арифметическим исследуемой зависимой переменной, называется *уравнением линии регрессии*.

Для определения параметра *b* получаем выражение

$$\frac{\partial S_0^2}{\partial b} = \frac{1}{n}\frac{\left[\sum_{i=1}^{n}(y_i - ax_i - b)^2\right]}{\partial b} = 0 \qquad (5.5)$$

Отсюда

$$b = \overline{y} - a\overline{x} \qquad (5.6)$$

Аналогично,

$$\frac{\partial S_0^2}{\partial a} = \frac{1}{n}\frac{\partial}{\partial a}\left[\sum_{i=1}^{n}(y_i - ax_i - \overline{y} - a\overline{x})^2\right] = \frac{1}{n}\frac{\partial}{\partial a}\left[\sum_{i=1}^{n}(\delta y_i - a\delta x_i)^2\right] = 0 \qquad (5.7)$$

где $\delta y_i = y_i - \overline{y}$ □□□□□□□□□□□□□□, $\delta x_i = x_i - \overline{x}$ – отклонения y_i и x_i от среднего значения.

Отсюда

$$a = \frac{\sum_{i=1}^{n} \delta x_i \delta y_i}{\sum_{i=1}^{n} \delta x_i^2} \tag{5.8}$$

Подставив найденные значения a и b в уравнение регрессии (5.4), имеем

$$\overline{y}(x_i) - \overline{y} = \frac{\sum_{i=1}^{n} \delta x_i \delta y_i}{\sum_{i=1}^{n} \delta x_i^2} (x_i - \overline{x}) \tag{5.9}$$

В данном случае ряд X играл роль аргумента, ряд Y – роль функции. Однако возможны и такие случаи, когда X выступает в качестве функции, а Y–аргумента, т. е.

$$\overline{x}(y) = a'y + b \tag{5.10}$$

отсюда

$$a' = \frac{\sum_{i=1}^{n} \delta x_i \delta y_i}{\sum_{i=1}^{n} \delta y_i^2} \tag{5.11}$$

уравнение регрессии принимает вид

$$\overline{x}(y) - \overline{x} = \frac{\sum_{i=1}^{n} \delta x_i \delta y_i}{\sum_{i=1}^{n} \delta y_i^2} (y_i - \overline{y}) \tag{5.12}$$

Выражения (5.9) и (5.12) описывают две существенно отличные друг от друга прямые регрессии, пересекающиеся в точке *(x, y)* под углом φ.

Чем более тесная связь между Y и X и чем, соответственно меньше разброс точек в поле графика связи, тем меньше угол между линиями регрессии. Для функциональной связи угол $\varphi = 0°$; в случае отсутствия связи угол $\varphi = 90°$.

Направление линий регрессии в поле графика определяется коэффициентами регрессии a и a'. Первый из них представляет собой тангенс угла наклона линии регрессии $y = f(x)$ к оси x; второй – тангенс угла наклона линии регрессии $x = f(y)$ к оси y. Обозначим эти углы через α и β. Тогда коэффициенты регрессии

$$a = \text{tg}\alpha, \quad a' = \text{tg}\beta. \tag{5.13}$$

Сумма углов α, β, φ составляет 90°. В случае, если связь между Y и X функциональная, то $\angle\varphi = 0°$ и $\angle\alpha + \angle\beta = 90°$. Отсюда $\angle\alpha = 90° - \angle\beta$ и

$$\text{tg}\alpha = \text{tg}(90°-\beta) = \text{ctg}\beta = 1/\text{tg}\beta,$$

т. е.

$$\text{tg}\alpha\,\text{tg}\beta = 1. \tag{5.14}$$

Если связь между Y и X отсутствует, то $\angle(\beta = 90°$, а $\angle\alpha = \angle\beta$. Отсюда

$$\text{tg}\alpha\,\text{tg}\beta = 0. \tag{5.15}$$

Для стохастических связей с увеличением тесноты связи угол φ уменьшается от 90 до 0°; вместе с тем увеличиваются углы α и β, а, следовательно, тангенсы этих углов и их произведение. Таким образом, произведение тангенсов углов α и β может служить мерой тесноты связи между X и Y.

Обычно в качестве критерия степени близости корреляционной связи к линейной функциональной зависимости используется корень квадратный из произведения $\text{tg}\alpha\,\text{tg}\beta$. Это произведение называется коэффициентом корреляции двух переменных величин и обозначается через r или R. Как следует из определения,

$$r = \sqrt{tg\alpha \cdot tg\beta}\,. \tag{5.16}$$

Коэффициент корреляции принимается положительным, если Y возрастает с увеличением X, и отрицательным, если Y уменьшается с увеличением X.

Подставляя значения a и a', вычисленные по формулам (5.8) и (5.11), в формулу (5.16), получаем $$r = \frac{\sum_i \delta x_i \delta y_i}{\sqrt{\sum_i \delta x_i^2 \sum \delta y_i^2}}. \qquad (5.17)$$

Чаще используется другой вид записи формулы коэффициента корреляции: через σ_x и σ_y:

$$r = \frac{\sum_{i=1}^{N} \delta x_i \cdot \delta y_i}{N\sigma_x \sigma_y}. \qquad (5.18)$$

Расчёт параметров уравнения регрессии и коэффициента корреляции исходных рядов за 36 лет приведены в таблице5.1.
Расчёт параметров уравнения регрессии и коэффициента корреляции исходных рядов за первые и вторые 18 лет приведены в таблицах 5.2а и 5.2б.
По рассчитанным параметрам на рис.5.3 построены линии регрессии уравнений Y=f(X) и X=f(Y) для периода n = 36 лет и на рис. 5.4 и 5.5 построены линии регрессии уравнений Y=f(X) и X=f(Y) для периодов $n_1 = n_2 = 18$ лет.
Считаю нужным заметить, что график связи процессов среднегодового и максимального стоков за вторые 18 лет демонстрирует обратную связь, поскольку с увеличением X уменьшается Y, а так как угол между линиями регрессий очень близок к 0°, то можем утверждать, что связь очень слаба.

5.3. Построение доверительных границ линий регрессии с учетом и без учета погрешности определения линии связи.

Доверительные границы для некоторой статистики (например, среднего значения или линии регрессии) показывают диапазон вокруг значения статистики, в котором находится истинное значение этой статистики (с определенным уровнем надежности или доверия). Линия регрессии характеризует изменение условного математического ожидания результативного признака от вариации остальных признаков.

Если исходить из предположения о нормальном законе распределения *y* и *x,* то доверительный интервал значений $\bar{y}\,(x)$ может быть рассчитан по формуле

$$\bar{\bar{y}}(x) - t_\alpha \sigma_{\bar{y}(x)} \le \bar{y}(x) < \bar{\bar{y}}(x) + t_\alpha \sigma_{\bar{y}(x)}. \qquad (5.19)$$

Доверительный интервал значений *y(x)* может быть аналогично выражениям (5.19) и построен по формуле

$$\bar{\bar{y}}(x) - t_\alpha \sigma_\Delta \sqrt{1 + \frac{1}{n} + \frac{t^2}{n}} \le \bar{y}(x) < \bar{\bar{y}}(x) + t_\alpha \sigma_\Delta \sqrt{1 + \frac{1}{n} + \frac{t^2}{n}}. \qquad (5.20)$$

Из формул (5.19) и (5.20) следует, что доверительный интервал $\bar{y}\,(x)$ и *y(x)* растет вместе с увеличением отклонения *x* от $\bar{x}$, т. е. оценка этих величин с удалением *x* от *x* становится менее точной.

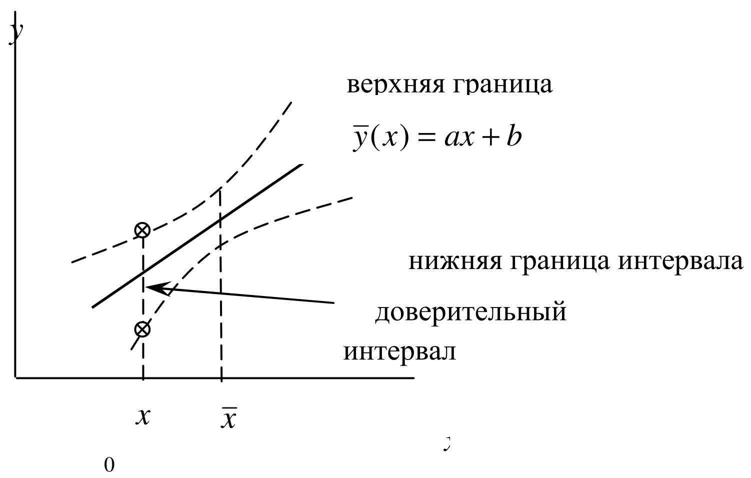

Рис. Доверительный интервал для уравнения регрессии вида $\bar{y}(x) = ax + b$

Из рисунка видно, что в точке $x_0 = \bar{x}$ границы интервала наиболее близки друг другу.

Расположение границ доверительного интервала показывает, что прогнозы по уравнению регрессии хороши только в случае, если значение *x* не выходит за пределы выборки, по которой вычислено уравнение регрессии; иными словами, экстраполяция по уравнению регрессии может привести к значительным погрешностям.
Расчет доверительных интервалов для уравнения регрессии Y=f(X) проведён по периоду в n=36 лет и по двум 18-летним периодам.
Расчеты параметров для построения уравнений регрессии Y=f(X) приведены в приложении 2 таб. 2.1, 2.2а и 2.2б.
Доверительные интервалы для уравнений регрессии Y=f(X) за 36 лет построены на рисунках 5.3, 5.4, 5.5.

5.4. Оценка стационарности связи во времени.

В результате расчётов, проведённых в главе 5, по 36-летнему периоду были найдены: коэффициент корреляции *r* – 0,382; параметры уравнения регрессии *a* – 19,91, *a*' – 0,007, *b* – 108, *b*' – 16. Также были определены среднеквадратические отклонения (ошибки): СКО(*r*) – 0,144, СКО(*a*) –8,6, СКО(Y(X)) – 306,8. Ошибка расчётов составляет 37,7 – 43,2%. Если допустить, что ошибка может достигать 20–25%, то применение корреляционного анализа не представляется возможным.
Расчёты приведены в таблице 5.1.
По первому 18-летнему периоду были найдены: коэффициент корреляции *r* – 0,841; параметры уравнения регрессии *a* – 47,5, *a*' – 0,015, *b* – -282, *b*' – 10. Также были определены среднеквадратические отклонения (ошибки): СКО(*r*) – 0,071, СКО(*a*) –7,48, СКО(Y(X)) – 185,03. Ошибка расчётов составляет 8,4 – 15,7%. Если допустить, что ошибка может достигать 20–25%, то применение корреляционного анализа возможно. Доверительные интервалы довольно узки, что говорит о малом разбросе значений, исключая точки редкой повторяемости.
Расчёты приведены в таблице 5.2а.
По второму 18-летнему периоду были найдены: коэффициент корреляции *r* – -0,022; параметры уравнения регрессии *a* – -0,8, *a*' – -0,001, *b* –359, *b*' – 20. Также были определены среднеквадратические отклонения (ошибки): СКО(*r*) – 0,243, СКО(*a*) –8,11, СКО(Y(X)) – 234,06. . Ошибка расчётов составляет 1013,7 – 1104,5%. Если допустить, что ошибка может достигать 20–25%, то применение корреляционного анализа не представляется возможным. Доверительные интервалы для данного периода очень широки, ошибки расчётов также очень велики – сказывается активное антропогенное вмешательство (ГЭС).
Расчёты приведены в таблице 5.2б.
Парметры уравнения регрессии по 36-летнему периоду приведены в таблице 5.3.
Поскольку коэффициент корреляции $r > 0{,}3$, то его доверительные границы определяем с помощью преобразования Фишера: $z = 0{,}403$ и доверительные границы $0{,}109 < z < 0{,}697$; $r = 0{,}382$, и доверительные границы $0{,}108 < r < 0{,}603$.
Расчёты приведены в таблице 5.4.
Произведена оценка значимости коэффициента корреляции через нулевую гипотезу (Ho: $r_{yx}=0$). В нашем случае СКО(*r*) – 0,144, величина $t_{2\alpha}=1{,}69$. Оценка гипотезы проводится с помощью соотношения $r_{yx} > \sigma_r \cdot t_{2\alpha}$ и она <u>опровергается</u>, т. е. коэффициент корреляции значим.
Расчёты приведены в таблице 5.5.
При оценке стационарности связи по 18-летним периодам применялась нулевая гипотеза $r_1=r_2$ и получены следующие результаты: $r_1=0{,}841$, $r_2= -0{,}022$. Вычисленная статистика t = 3,41 сравнивается с $t_{2\alpha}$ (критической), и порскольку $t > t_{2\alpha}$, то гипотеза о равенстве стационарности связи между 18-тилетними периодами опровергается
Расчёты приведены в таблице 5.6.
Считаю невозможным использование данной тесноты связи для определения одного процесса по другому, поскольку на уровенный режим реки существенное влияние оказывает ГЭС, находящаяся в 0,8 км выше по течению реки.
Точность определения параметров уравнений регрессии (СКО(*a*)) определяется точностью расчёта СКО(Y(X)). $\sigma_{y(x)}$, вычисленный по формуле

$$\sigma_{y(x)} = \sqrt{\frac{\sum_{i=1}^{n}(y_{i,н} - y_p)^2}{n}},$$

где $y_{iн}$– наблюдённая величина; $y_р$– величина, рассчитанная по уравнению регрессии, равен 303,06;

$\sigma_{y(x)}$, вычисленный по формуле $\sigma_{y(x)} = \sigma_{y_0}\sqrt{1-r^2}$,

где σ_{y0}– среднее квадратическое отклонение исходного ряда величин;
r – коэффициент корреляции уравнения регрессии, будет равен 306,8.
Стандартная ошибка свободного члена b уравнения регрессии определяется по формуле

$$\sigma_b = \frac{\sigma_{y(x)}}{\sqrt{n}} = \frac{303,06}{6} = 50,51$$

Или $$\sigma_b = \frac{\sigma_{y(x)}}{\sqrt{n}} = \frac{306,8}{6} = 51,1.$$

Стандартная ошибка коэффициента регрессии равна

$$\sigma_a = \frac{\sigma_{y(x)}}{\sqrt{\sum(x_i - \bar{x})^2}} = \frac{303,06}{39,6} = 7,67$$

или $$\sigma_a = \frac{\sigma_{y(x)}}{\sqrt{\sum(x_i - \bar{x})^2}} = \frac{306,8}{39,6} = 7,75.$$

В таком случае результаты расчётов параметров уравнения регрессии можно представить в следующем виде:

$$a \pm \sigma_a = 19,91 \pm 7,7,$$

$$b \pm \sigma_b = 108 \pm 50,0.$$

Стандартная ошибка ординаты уравнения регрессии по формуле

$$\bar{\sigma}^2_{\bar{y}(x_i)} = \sigma^2_{y(x)}\left[\frac{1}{n-2} + \frac{(x_i - \bar{x})^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right]$$

выразится так: $$\bar{\sigma}_{\bar{y}(x_i)} = \sigma_{y(x)}\sqrt{\frac{1}{n-2} + \frac{(x_i - \bar{x})^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}} = 306,8 \cdot \sqrt{\frac{1}{34} + \frac{(x_i - 20)^2}{1568,8}}.$$

Дисперсия $\bar{\sigma}_{\bar{y}(x_i)}$ характеризует рассеивание ординат выборочной линии регрессии относительно линии регрессии, соответствующей генеральной совокупности.
Погрешность расчётов зависит от продолжительности периодов наблюдений: чем длиннее период наблюдений, тем меньше погрешность.
Анализ остатков (погрешностей) является одним из методов проверки адекватности математической модели линейной регрессии. Поскольку остатки представляют собой расширяющуюся полосу, то модель можно рассматривать как гетероскедатичную, т. е., характеризующуюся отсутствием постоянства дисперсии σ^2_Δ. График остатков представлен на рис. 5.6 и рис. 5.7
Графики изменения процесса среднегодового и максимального стоков во времени представлены на рисунках 5.8 и 5.9.

6. Общие выводы

В ходе данной работы были проведены обработка и анализ данных наблюдений среднегодового и максимального стоков по реке Ока – Костомарово за период с 1941 по 1980 гг.

Речной сток – гидрологический процесс, он является интегральной характеристикой и результатом взаимодействия многих геофизических процессов на сравнительно больших

территориях: прямая и рассеяная радиация, осадки, температура воздуха и подстилающей поверхности, давление и влажность воздуха и т. д.; физико-географические условия бассейна: почвы, геологическое строение, растительность и т. д.; хозяйственная деятельность человека. Полное описание основных причин формирования речного стока приводится в *первой главе* данной курсовой работы.
Гидрологические процессы носят вероятностный характер, т. е. характеризуются наличием элементов случайностей, сопровождающих весь процесс формирования стока. Таким образом, обработка данных производится на основе теории случайных величин и её основных понятий. Важное значение в гидрологических расчётах имеют числовые характеристики и операции с ними. Во *второй главе* производится выбор методов расчёта числовых характеристик на основе сопоставлений их простоты в расчёте и в использовании. В результате проверок, из трёх методов: моментов, наибольшего правдоподобия и квантилей мы пришли к выводу, что **наиболее удобным для нас является метод моментов** из-за своей практичности и простоты в расчётах.
В *третьей главе* мы провели оценку однородности гидрологической информации. Максимальное использование информации при ограниченном объёме выборок достигается тем, что анализ основывается на гипотезах, которые должны быть доказаны или опровергнуты в ходе их проверки.
Сравнение 20-летних выборок внутри среднегодового процесса показало, что *расхождение между средними значениями* и *дисперсиями* можно считать случайным, **что доказывает их однородность и принадлежность к общим генеральным совокупностям.**
Сравнение 20-летних выборок внутри максимального процесса показало, что расхождение между средними значениями велико, оценка равенства дисперсий по критериям Фишера и Романовского показала, что рахождение между ними случайно.
Для 40-летних выборок расхождения значений велики, и, значит, *нулевые гипотезы опровергаются*. Это значит, что **изучаемые процессы неоднородны между собой по физическим свойствам и статистическим характеристикам.**
Оценка однородности среднегодового и максимального стоков *по критерию Уилкоксона* показала, что **выборки принадлежат к одной генеральной совокупности**, т. е. *нулевая гипотеза не опровергается.*
В *четвёртой главе* мы решаем вопрос о характере внутрирядных связей, является ли ряд стационарным, однородным или случайным. При проверке выборочных рядов среднегодового и максимального стоков на случайность:
а) по числу повышений и понижений –
гипотеза о случайности процесса для максимального стока не опровергается; гипотеза о случайности процесса для среднегодового стока опровергается (тенденция к повышению)
б) *по числу экстремумов* – **гипотеза о случайности процессов не опровергается**.
в) *по критерию длин и числа серий* – **гипотеза о случайности процессов не опровергается.**
При анализе исходных рядов среднегодового и максимального стоков на *стационарность и однородность* получены следующие результаты:
а) *при определении даты нарушения однородности по интегральной кривой* (рис. 4.1) можно сделать вывод о том, что в течение рассматриваемого периода наиболее **частое изменение характера связи** наблюдалось в период с **1962 по 1979** года
(во второй половине исследуемых рядов).
б) *при оценке тренда по коэффициенту корреляции и по критерию Спирмэна*, с определённой долей вероятности мы можем утверждать, что в рассматриваемых процессах **тренд отсутствует**. Рассчитанные значения ρ_{xi} и ρ_{yi} оказались незначимыми, что и говорит об отсутствии тренда. Кроме того, мы можем утверждать, что **присутствует обратная связь** (об этом говорит критерий Спирмэна), но она **незначительна (слаба)**.
В гидрологических исследованиях большое внимание уделяется анализу взаимосвязей между различными природными процессами, определяющими общий фон гидрологического режима исследуемых объектов. В пятой главе для исследования взаимосвязи природных процессов используется математическая модель в виде уравнения регрессии. Рассчитывается и анализируется критерий степени близости корреляционной связи к линейной функциональной зависимости, т. е. коэффициент корреляции. Строятся графики двух зависимых величин, по которым определяется теснота вероятностной зависимости.

Графики связи *неранжированных* и *ранжированных* значений среднегодового и максимального стоков Y=*f*(X) (рис. 5.1, рис. 5.2), представляют собой графики связи стохастически зависимых величин, и с большой долей вероятности мы можем утверждать, что **связь присутствует**, причём связь эта прямая. На графике связи неранжированных значений чётко прослеживается некоторая криволинейность, которая объясняется наличием поймы. Причём на графике связи ранжированных значений представлена более тесная связь, чем на графике неранжированных значений. Протяжённость поля точек по длине линии связи на обоих рисунках в два раза превышает наибольший разброс точек относительно линии связи, поэтому считаем, что данная теснота связи достаточна для наших исследований. Логичным считаю отметить, что, несмотря на разность тесноты связи на рисунках, общим для них является «отскакивание» точек редкой повторяемости.

В результате расчётов были найдены: коэффициент корреляции, параметры уравнения регрессии. Также были определены среднеквадратические отклонения (ошибки): СКО(r), СКО(a), СКО(Y(X)). Ошибка расчётов составляет 37,7 – 43,2%. Если допустить, что ошибка может достигать 20–25 то применение корреляционного анализа невозможно.

Поскольку коэффициент корреляции $r > 0{,}3$, то его доверительные границы мы определили с помощью преобразования Фишера.

Произведена *оценка значимости коэффициента корреляции* через нулевую гипотезу (Ho: $r_{yx}=0$). В нашем случае она опровергается, т. е. коэффициент корреляции значим. Несмотря на то, что коэффициент корреляции r между процессами среднегодового и максимального стоков > 0,3, то **считаю невозможным использование данной тесноты связи для определения одного процесса по другому.**

Точность определения параметров уравнений регрессии (СКО(a)) определяется точностью расчёта СКО(Y(X)).

Стандартная ошибка ординаты уравнения регрессии выразится так:

$$\bar{\sigma}_{\bar{y}(x_i)}=\sigma_{y(x)}\sqrt{\frac{1}{n-2}+\frac{(x_i-\bar{x})^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}}=306{,}8\cdot\sqrt{\frac{1}{34}+\frac{(x_i-20)^2}{1568{,}8}}\ .$$

Погрешность расчётов зависит от продолжительности периодов наблюдений: чем длиннее период наблюдений, тем меньше погрешность. Анализ остатков (погрешностей) является одним из методов проверки адекватности математической модели линейной регрессии. Поскольку остатки представляют собой расширяющуюся полосу с центром на оси абсцисс, то модель можно рассматривать как гетероскедатичную, т. е., характеризующуюся отсутствием постоянства дисперсии σ_{Δ}^2 .

Подводя общий итог проделанной работы, можно отметить, что получение конкретных результатов при обработке данных стало возможным при использовании теории вероятностей, математической статистики и их основных понятий и приёмов, а разногласия различных оценок и погрешности расчётов могут объясняться как небольшим объёмом выборки, так и активным вмешательством человека в природный процесс – забор воды для использования промышленными предприятиями, населением и в системах орошения. Объем воды, отводимой обратно в речную сеть, составляет около 80% водопотребления. Забор воды из реки для целей сельскохозяйственного водоснабжения и орошения составляет 4 % общего объема На уровненный режим реки, несомненно, оказывает большое влияние плотина ГЭС, расположенная в 0,8 км выше водпоста.

7. Список используемой литературы.

1. Гидрологический ежегодник. Бассейн Каспийского моря (без Кавказа и Ср. Азии), т. 4, вып. 5–7: Бассейн реки Камы; Свердловск, 1967 г.
2. Большая Советская Энциклопедия. М., изд. «Советская Энциклопедия», третье издание, 1976.
3. Шелутко В.А., Техника статистических вычислений в гидрологии. Изд. ЛПИ, 1977
4. Шелутко В.А., Численные методы в гидрологии. Л.: Гидрометеоиздат – 1991

5. ВайновскиЙ П.А., Малинин В.Н. Методы обработки и анализа океанологической информации. Одномерный анализ. Изд.РГГМИ, 1992.

Приложение 1

таб. 1.1а

статистический ряд
среднегодового стока
по р.Ока - Костомарово

№ п/п	Год	Q_i, м^3/с
1	1941	18,1
2	1942	29
3	1943	10,2
4	1944	13,6
5	1945	16,8
6	1946	23,4
7	1947	29,7
8	1948	17,9
9	1949	12,9
10	1950	13,6
11	1951	26,6
12	1952	29,7
13	1953	19,4
14	1954	15,2
15	1955	22,3
16	1956	17,8
17	1957	20,3
18	1958	24
19	1959	16,6
20	1960	29,1
21	1961	14,5
22	1962	19,9
23	1963	24,1
24	1964	18,8
25	1965	17,3
26	1966	26,6
27	1967	28,8
28	1968	19,4
29	1969	23
30	1970	35,9
31	1971	23,1
32	1972	12,3
33	1973	13,9
34	1974	18,4
35	1975	10,7
36	1976	11,4
37	1977	23,5
38	1978	25,5
39	1979	29,3
40	1980	24,1

таб. 1.1б

статистический ряд
максимального стока
по р.Ока - Костомарово

№ п/п	Год	Q_i, м^3/с
1	1941	404
2	1942	1190
3	1943	360
4	1944	450
5	1945	450
6	1946	910
7	1947	1510
8	1948	640
9	1949	404
10	1950	360
11	1951	1040
12	1952	1120
13	1953	616
14	1954	656
15	1955	632
16	1956	446
17	1957	421
18	1958	422
19	1959	522
20	1960	564
21	1961	284
22	1962	581
23	1963	155
24	1964	862
25	1965	111
26	1966	171
27	1967	185
28	1968	125
29	1969	148
30	1970	231
31	1971	726
32	1972	289
33	1973	562
34	1974	288
35	1975	68,2
36	1976	308
37	1977	686
38	1978	458
39	1979	1080
40	1980	352

таб. 1.2а

Оценка числовых характеристик методом наибольшего правдоподобия по р. Ока - Костомарово

№ п/п	$Q_{ср. г.}$	X_i, ранжир.	K_i, ранжир.	LnQ_i ранжир.
1	18,1	35,9	1,74	0,55
2	29	29,7	1,44	0,36
3	10,2	29,7	1,44	0,36
4	13,6	29,3	1,42	0,35
5	16,8	29,1	1,41	0,34
6	23,4	29	1,40	0,34
7	29,7	28,8	1,39	0,33
8	17,9	26,6	1,29	0,25
9	12,9	26,6	1,29	0,25
10	13,6	25,5	1,23	0,21
11	26,6	24,1	1,17	0,15
12	29,7	24,1	1,17	0,15
13	19,4	24	1,16	0,15
14	15,2	23,5	1,14	0,13
15	22,3	23,4	1,13	0,12
16	17,8	23,1	1,12	0,11
17	20,3	23	1,11	0,11
18	24	22,3	1,08	0,08
19	16,6	20,3	0,98	-0,02
20	29,1	19,9	0,96	-0,04
21	14,5	19,4	0,94	-0,06
22	19,9	19,4	0,94	-0,06
23	24,1	18,8	0,91	-0,09
24	18,8	18,4	0,89	-0,12
25	17,3	18,1	0,88	-0,13
26	26,6	17,9	0,87	-0,14
27	28,8	17,8	0,86	-0,15
28	19,4	17,3	0,84	-0,18
29	23	16,8	0,81	-0,21
30	35,9	16,6	0,80	-0,22
31	23,1	15,2	0,74	-0,31
32	12,3	14,5	0,70	-0,35
33	13,9	13,9	0,67	-0,40
34	18,4	13,6	0,66	-0,42
35	10,7	13,6	0,66	-0,42
36	11,4	12,9	0,62	-0,47
37	23,5	12,3	0,60	-0,52
38	25,5	11,4	0,55	-0,59
39	29,3	10,7	0,52	-0,66

40	24,1	10,2	0,49	-0,71
			S	-1,91
	1	-0,05	C*v	0,304

таб. 1.2б

Оценка числовых характеристик методом наибольшего правдоподобия по р. Ока - Костомарово

№ п/п	Q_{max}	Y_i, ранжир.	K_i, ранжир.	LnQ_i ранжир.
1	404	1510	2,91	1,07
2	1190	1190	2,29	0,83
3	360	1120	2,16	0,77
4	450	1080	2,08	0,73
5	450	1040	2,00	0,69
6	910	910	1,75	0,56
7	1510	862	1,66	0,51
8	640	726	1,40	0,34
9	404	686	1,32	0,28
10	360	656	1,26	0,23
11	1040	640	1,23	0,21
12	1120	632	1,22	0,20
13	616	616	1,19	0,17
14	656	581	1,12	0,11
15	632	564	1,09	0,09
16	446	562	1,08	0,08
17	421	522	1,00	0,00
18	422	458	0,88	-0,13
19	522	450	0,87	-0,14
20	564	450	0,87	-0,14
21	284	446	0,86	-0,15
22	581	422	0,81	-0,21
23	155	421	0,81	-0,21
24	862	404	0,78	-0,25
25	111	404	0,78	-0,25
26	171	360	0,69	-0,37
27	185	360	0,69	-0,37
28	125	352	0,68	-0,39
29	148	308	0,59	-0,53
30	231	289	0,56	-0,58
31	726	288	0,55	-0,60
32	289	284	0,55	-0,60
33	562	231	0,44	-0,82
34	288	185	0,36	-1,02
35	68,2	171	0,33	-1,11
36	308	155	0,30	-1,20
37	686	148	0,28	-1,27
38	458	125	0,24	-1,43
39	1080	111	0,21	-1,56
40	352	68,2	0,13	-2,04
			S	-8,50
	1	-0,213	C*v	0,610

таб.1.3а

Оценка однородности среднегодового стока по критерию Уилкоксона по р. Ока- Костомарово

№п/п	$Q_{ср.г}$	признак	Qранж	призн ранж	***u***
1	18,1	0	10,2	0	0
2	29	0	10,7	1	0
3	10,2	0	11,4	1	0
4	13,6	0	12,3	1	0
5	16,8	0	12,9	0	5
6	23,4	0	13,6	0	6
7	29,7	0	13,6	0	7
8	17,9	0	13,9	1	0
9	12,9	0	14,5	1	0
10	13,6	0	15,2	0	10
11	26,6	0	16,6	0	11
12	29,7	0	16,8	0	12
13	19,4	0	17,3	1	0
14	15,2	0	17,8	0	14
15	22,3	0	17,9	0	15
16	17,8	0	18,1	0	16
17	20,3	0	18,4	1	0
18	24	0	18,8	1	0
19	16,6	0	19,4	0	19
20	29,1	0	19,4	1	0
21	14,5	1	19,9	1	0
22	19,9	1	20,3	0	22
23	24,1	1	22,3	0	23
24	18,8	1	23	1	0
25	17,3	1	23,1	1	0
26	26,6	1	23,4	0	26
27	28,8	1	23,5	1	0
28	19,4	1	24	0	28
29	23	1	24,1	1	0
30	35,9	1	24,1	1	0
31	23,1	1	25,5	1	0
32	12,3	1	26,6	0	32
33	13,9	1	26,6	1	0
34	18,4	1	28,8	1	0
35	10,7	1	29	0	35
36	11,4	1	29,1	0	36
37	23,5	1	29,3	1	0
38	25,5	1	29,7	0	38
39	29,3	1	29,7	0	39
40	24,1	1	35,9	1	0
				S	394

таб.1.3б

Оценка однородности максимального стока по критерию Уилкоксона по р. Ока- Костомарово

№п/п	Q_{max}	признак	Qранж	призн ранж	***u***
1	404	0	68,2	1	0
2	1190	0	111	1	0
3	360	0	125	1	0
4	450	0	148	1	0
5	450	0	155	1	0
6	910	0	171	1	0
7	1510	0	185	1	0
8	640	0	231	1	0
9	404	0	284	1	0
10	360	0	288	1	0
11	1040	0	289	1	0
12	1120	0	308	1	0
13	616	0	352	1	0
14	656	0	360	0	14
15	632	0	360	0	15
16	446	0	404	0	16
17	421	0	404	0	17
18	422	0	421	0	18
19	522	0	422	0	19
20	564	0	446	0	20
21	284	1	450	0	21
22	581	1	450	0	22
23	155	1	458	1	0
24	862	1	522	0	24
25	111	1	562	1	0
26	171	1	564	0	26
27	185	1	581	1	0
28	125	1	616	0	28
29	148	1	632	0	29
30	231	1	640	0	30
31	726	1	656	0	31
32	289	1	686	1	0
33	562	1	726	1	0
34	288	1	862	1	0
35	68,2	1	910	0	35
36	308	1	1040	0	36
37	686	1	1080	1	0
38	458	1	1120	0	38
39	1080	1	1190	0	39
40	352	1	1510	0	40
				S	518

таб. 1.4

Таблица для проверки случайности по критерию длин и числа серий по р.Ока - Костомарово

№п/п	$Q_{ср.г}$	Q_{max}	преобразов. знач.	
			$Q_{ср.г}$	Q_{max}
1	18,1	404	1	1
2	29	1190	0	0
3	10,2	360	1	1
4	13,6	450	1	1
5	16,8	450	1	1
6	23,4	910	0	0
7	29,7	1510	0	0
8	17,9	640	1	0
9	12,9	404	1	1
10	13,6	360	1	1
11	26,6	1040	0	0
12	29,7	1120	0	0
13	19,4	616	1	0
14	15,2	656	1	0
15	22,3	632	0	0
16	17,8	446	1	1
17	20,3	421	1	1
18	24	422	0	1
19	16,6	522	1	0
20	29,1	564	0	0
21	14,5	284	1	1
22	19,9	581	1	0
23	24,1	155	0	1
24	18,8	862	1	0
25	17,3	111	1	1
26	26,6	171	0	1
27	28,8	185	0	1
28	19,4	125	1	1
29	23	148	0	1
30	35,9	231	0	1
31	23,1	726	0	0
32	12,3	289	1	1
33	13,9	562	1	0
34	18,4	288	1	1
35	10,7	68,2	1	1
36	11,4	308	1	1
37	23,5	686	0	0
38	25,5	458	0	1
39	29,3	1080	0	0
40	24,1	352	0	1
среднее	21	520		

таб.1.5а

Расчёт случайности по критерию длин и числа серий для среднегодового стока по р.Ока - Костомарово

длина серий	фактическое количество серий			ожидаемое количество серий				
k	0	1	S	$MR_{0,k}$	$MR_{1,k}$	$Mr_{0,k}$	$Mr_{1,k}$	Mr_k
1	5	2	7	9,00	8,80	5,56	3,62	9,18
2	3	4	7	3,44	5,18	2,19	2,19	4,39
3	1	2	3	1,25	2,98	0,82	1,30	2,12
4	0	0	0	0,43	1,68	0,29	0,76	1,05
5	0	1	1	0,14	0,92	0,09	0,43	0,53
6	0	0	0	0,04	0,49	0,04	0,49	0,53
S	9	9	18		S	9	9	18

n_1	n_2	кол-во лет (n)	предел 5%	предел 95%	общее фактическое количество серий	H_o:	результат
14	21	35	13	23	18	процесс случаен	гипотеза не опровергается

таб.1.5б

Расчёт случайности по критерию длин и числа серий для максимального стока по р.Ока - Костомарово

длина серий	фактическое количество серий			ожидаемое количество серий				
k	0	1	S	$MR_{0,k}$	$MR_{1,k}$	$Mr_{0,k}$	$Mr_{1,k}$	Mr_k
1	7	4	11	9,50	9,50	5,39	4,36	9,76
2	1	1	2	4,11	5,13	2,39	2,42	4,82
3	1	3	4	1,71	2,71	1,03	1,32	2,34
4	0	0	0	0,68	1,39	0,42	0,69	1,12
5	1	0	1	0,26	0,69	0,26	0,69	0,96
6	0	1	1	0,09	0,34	0,09	0,34	0,43
7	0	0	0	0,03	0,16	0,03	0,16	0,19
S	10	9	19		S	10	10	20
n_1	n_2	кол-во лет (n)	предел 5%	предел 95%	общее фактическое количество серий	H_o:	результат	
17	21	38	14	25	19	процесс случаен	гипотеза не опровергается	

таб. 1.6а

Расчёт числа повышений, понижений, экстремумов
для среднегодового стока по р.Ока - Костомарово

№п/п	год	$Q_{ср.г}$	повышение (+)/ понижение (-)	наличие экстремума
1	1941	18,1		
2	1942	29	+	+
3	1943	10,2	-	+
4	1944	13,6	+	
5	1945	16,8	+	
6	1946	23,4	+	
7	1947	29,7	+	+
8	1948	17,9	-	
9	1949	12,9	-	+
10	1950	13,6	+	
11	1951	26,6	+	
12	1952	29,7	+	+
13	1953	19,4	-	
14	1954	15,2	-	+
15	1955	22,3	+	+
16	1956	17,8	-	+
17	1957	20,3	+	
18	1958	24	+	+
19	1959	16,6	-	+
20	1960	29,1	+	+
21	1961	14,5	-	+
22	1962	19,9	+	
23	1963	24,1	+	+
24	1964	18,8	-	
25	1965	17,3	-	+
26	1966	26,6	+	
27	1967	28,8	+	+
28	1968	19,4	-	+
29	1969	23	+	
30	1970	35,9	+	+
31	1971	23,1	-	
32	1972	12,3	-	+
33	1973	13,9	+	
34	1974	18,4	+	+
35	1975	10,7	-	+
36	1976	11,4	+	
37	1977	23,5	+	
38	1978	25,5	+	
39	1979	29,3	+	+
40	1980	24,1	-	
среднее		21		
сумма (+)		24		

сумма (-)	15
число экстремумов	21

таб. 1.6б

Расчёт числа повышений, понижений, экстремумов
для максимального стока по р.Ока - Костомарово

№п/п	год	Qmax	повышение (+)/ понижение (-)	наличие экстремума
1	1941	404		
2	1942	1190	+	+
3	1943	360	-	+
4	1944	450	+	
5	1945	450		
6	1946	910	+	
7	1947	1510	+	+
8	1948	640	-	
9	1949	404	-	
10	1950	360	-	+
11	1951	1040	+	
12	1952	1120	+	+
13	1953	616	-	+
14	1954	656	+	+
15	1955	632	-	
16	1956	446	-	
17	1957	421	-	+
18	1958	422	+	
19	1959	522	+	
20	1960	564	+	+
21	1961	284	-	+
22	1962	581	+	+
23	1963	155	-	+
24	1964	862	+	+
25	1965	111	-	+
26	1966	171	+	
27	1967	185	+	+
28	1968	125	-	+
29	1969	148	+	
30	1970	231	+	
31	1971	726	+	+
32	1972	289	-	+
33	1973	562	+	+
34	1974	288	-	
35	1975	68,2	-	+
36	1976	308	+	
37	1977	686	+	+
38	1978	458	-	+
39	1979	1080	+	+
40	1980	352	-	
среднее		520		
сумма (+)		21		

сумма (-)	17
число экстремумов	23

таб 1.7

Определение даты нарушения однородности по р.Ока - Костомарово

№п/п	год	$Q_{ср.г}$ (X)	Qmax (Y)	K_x	K_y	□K_x □	□K_y □
1	1941	18,1	404	0,88	0,78	0,88	0,78
2	1942	29	1190	1,40	2,29	2,28	3,07
3	1943	10,2	360	0,49	0,69	2,77	3,76
4	1944	13,6	450	0,66	0,87	3,43	4,63
5	1945	16,8	450	0,81	0,87	4,24	5,49
6	1946	23,4	910	1,13	1,75	5,38	7,24
7	1947	29,7	1510	1,44	2,91	6,81	10,15
8	1948	17,9	640	0,87	1,23	7,68	11,38
9	1949	12,9	404	0,62	0,78	8,30	12,16
10	1950	13,6	360	0,66	0,69	8,96	12,85
11	1951	26,6	1040	1,29	2,00	10,25	14,85
12	1952	29,7	1120	1,44	2,16	0,00	17,01
13	1953	19,4	616	0,94	1,19	12,62	18,19
14	1954	15,2	656	0,74	1,26	13,36	19,45
15	1955	22,3	632	1,08	1,22	14,44	20,67
16	1956	17,8	446	0,86	0,86	15,30	21,53
17	1957	20,3	421	0,98	0,81	16,28	22,34
18	1958	24	422	1,16	0,81	17,44	23,15
19	1959	16,6	522	0,80	1,00	18,25	24,16
20	1960	29,1	564	1,41	1,09	19,65	25,24
21	1961	14,5	284	0,70	0,55	20,36	25,79
22	1962	19,9	581	0,96	1,12	21,32	26,91
23	1963	24,1	155	1,17	0,30	22,48	27,20
24	1964	18,8	862	0,91	1,66	23,39	28,86
25	1965	17,3	111	0,84	0,21	24,23	29,08
26	1966	26,6	171	1,29	0,33	25,52	29,40
27	1967	28,8	185	1,39	0,36	26,91	29,76
28	1968	19,4	125	0,94	0,24	27,85	30,00
29	1969	23	148	1,11	0,28	28,96	30,29
30	1970	35,9	231	1,74	0,44	30,70	30,73
31	1971	23,1	726	1,12	1,40	31,82	32,13
32	1972	12,3	289	0,60	0,56	32,41	32,68
33	1973	13,9	562	0,67	1,08	33,09	33,77
34	1974	18,4	288	0,89	0,55	33,98	34,32
35	1975	10,7	68,2	0,52	0,13	34,49	34,45
36	1976	11,4	308	0,55	0,59	35,05	35,04
37	1977	23,5	686	1,14	1,32	36,18	36,36
38	1978	25,5	458	1,23	0,88	37,42	37,24
39	1979	29,3	1080	1,42	2,08	38,83	39,32
40	1980	24,1	352	1,17	0,68	40,00	40,00
среднее		21	520				

Приложение 2

таб. 2.1

Расчет доверительного интервала для уравнения регрессии между процессами среднегодового и максимального стоков по р.Ока - Костомарово

№ п/п	Xранж	Xрасч	Yранж	Yрасч	t	$\square_{y(x)}$	$Y_{\text{ниж. гр.}}$	$Y_{\text{верх. гр.}}$
1	35,9	27	1510	794	3,02	347	207	1381
2	29,7	25	1190	675	2,06	328	120	1230
3	29,7	24	1120	675	1,85	325	126	1225
4	29,1	24	1040	664	1,61	322	120	1208
5	29	23	910	662	1,22	317	126	1198
6	28,8	22	862	658	1,07	316	124	1192
7	26,6	21	726	616	0,66	313	87	1145
8	26,6	21	656	616	0,45	312	89	1143
9	24,1	21	640	568	0,40	312	42	1095
10	24	21	632	566	0,38	312	40	1093
11	23,4	20	616	555	0,33	311	29	1081
12	23,1	20	581	549	0,23	311	23	1075
13	23	20	564	547	0,17	311	21	1073
14	22,3	20	562	534	0,17	311	8	1060
15	20,3	20	522	496	0,05	311	-30	1021
16	19,9	19	450	488	-0,17	311	-38	1014
17	19,4	19	450	479	-0,17	311	-47	1004
18	19,4	19	446	479	-0,18	311	-47	1004
19	18,8	19	422	467	-0,25	311	-59	993
20	18,4	19	421	459	-0,26	311	-67	986
21	18,1	19	404	454	-0,31	311	-73	980
22	17,9	19	404	450	-0,31	311	-76	976
23	17,8	19	360	448	-0,44	312	-79	975
24	17,3	19	360	438	-0,44	312	-89	965
25	16,8	18	308	429	-0,60	313	-99	957
26	16,6	18	289	425	-0,65	313	-104	954
27	15,2	18	288	398	-0,66	313	-130	927
28	14,5	18	284	385	-0,67	313	-144	914
29	13,9	18	231	373	-0,83	314	-157	904
30	13,6	17	185	368	-0,97	315	-164	900
31	13,6	17	171	368	-1,01	315	-165	901
32	12,9	17	155	354	-1,06	316	-179	888
33	12,3	17	148	343	-1,08	316	-191	877
34	11,4	17	125	326	-1,15	317	-209	861
35	10,7	17	111	312	-1,19	317	-223	848
36	10,2	16	68,2	303	-1,32	318	-235	841
□	724,3	709	18211,2	17722				

таб. 2.2а

Расчет доверительного интервала для уравнения регрессии между процессами среднегодового и максимального стоков по первым 18-ти годам по р.Ока - Костомарово

№ п/п	Хранж	Храсч	Yранж	Yрасч	t	$\square_{y(x)}$	$Y_{\text{ниж. гр.}}$	$Y_{\text{верх. гр.}}$
1	29,7	23,7	1510	1129	2,46	202,36	787	1471
2	29,7	22,9	1190	1129	1,53	193,40	802	1456
3	29	20,9	1120	1096	1,32	191,96	771	1420
4	26,6	19,8	1040	982	1,09	190,56	659	1304
5	24	19,6	910	858	0,71	188,85	539	1177
6	23,4	19,5	656	830	-0,04	187,59	512	1147
7	22,3	19,2	640	777	-0,08	187,60	460	1094
8	20,3	18,7	632	682	-0,11	187,61	365	999
9	19,4	18,5	616	640	-0,15	187,64	322	957
10	18,1	18,4	450	578	-0,64	188,61	259	896
11	17,9	17,8	450	568	-0,64	188,61	250	887
12	17,8	16,8	446	564	-0,65	188,65	245	882
13	16,8	16,8	422	516	-0,72	188,89	197	835
14	15,2	16,7	421	440	-0,72	188,90	121	759
15	13,6	16,3	404	364	-0,77	189,09	44	684
16	13,6	16,3	404	364	-0,77	189,09	44	684
17	12,9	16,1	360	331	-0,90	189,63	10	651
18	10,2	16,1	360	203	-0,90	189,63	-118	523
□	360,5	334,0	12031	12048				

таб. 2.2б

Расчет доверительного интервала для уравнения регрессии между процессами среднегодового и максимального стоков по вторым 18-ти годам по р.Ока - Костомарово

№ п/п	Хранж	Храсч	Yранж	Yрасч	t	$\square_{y(x)}$	$Y_{\text{ниж. гр.}}$	$Y_{\text{верх. гр.}}$
1	35,9	19,1	862	330	2,22	252,57	-97	757
2	29,1	19,3	726	336	1,64	245,73	-80	751
3	28,8	19,4	581	336	1,02	240,58	-71	743
4	26,6	19,4	564	338	0,94	240,13	-68	744
5	24,1	19,4	562	340	0,94	240,08	-66	745
6	23,1	19,5	522	341	0,76	239,16	-64	745
7	23	19,7	308	341	-0,15	237,36	-61	742
8	19,9	19,7	289	343	-0,23	237,46	-58	744
9	19,4	19,7	288	343	-0,24	237,47	-58	745
10	18,8	19,7	284	344	-0,25	237,49	-57	745
11	18,4	19,8	231	344	-0,48	238,02	-58	747
12	17,3	19,8	185	345	-0,68	238,75	-58	749
13	16,6	19,8	171	346	-0,74	239,01	-58	750
14	14,5	19,8	155	347	-0,80	239,35	-57	752
15	13,9	19,9	148	348	-0,83	239,51	-57	753

16	12,3	19,9	125	349	-0,93	240,06	-57	755
17	11,4	19,9	111	350	-0,99	240,42	-56	756
18	10,7	19,9	68	350	-1,17	241,67	-58	759
Σ	363,8	353,8	6180	6171				

Приложение 3

Средний годовой сток по р. Ока - Костомарово

№ п/п	Год	Q_i, м^3/с	K_i	Q_i, ранжир.	K_i, ранжир.	P_i, %	ΣKi	Σ(K_i-1)	LnQ_i
1	1941	18,1	0,88	35,9	1,74	2,44	0,88	-0,12	2,90
2	1942	29	1,40	29,7	1,44	4,88	2,28	0,28	3,37
3	1943	10,2	0,49	29,7	1,44	7,32	2,77	-0,23	2,32
4	1944	13,6	0,66	29,3	1,42	9,76	3,43	-0,57	2,61
5	1945	16,8	0,81	29,1	1,41	12,2	4,24	-0,76	2,82
6	1946	23,4	1,13	29	1,40	14,6	5,38	-0,62	3,15
7	1947	29,7	1,44	28,8	1,39	17,1	6,81	-0,19	3,39
8	1948	17,9	0,87	26,6	1,29	19,5	7,68	-0,32	2,88
9	1949	12,9	0,62	26,6	1,29	22,0	8,30	-0,70	2,56
10	1950	13,6	0,66	25,5	1,23	24,4	8,96	-1,04	2,61
11	1951	26,6	1,29	24,1	1,17	26,8	10,25	-0,75	3,28
12	1952	29,7	1,44	24,1	1,17	29,3	11,69	-0,31	3,39
13	1953	19,4	0,94	24	1,16	31,7	12,62	-0,38	2,97
14	1954	15,2	0,74	23,5	1,14	34,1	13,36	-0,64	2,72
15	1955	22,3	1,08	23,4	1,13	36,6	14,44	-0,56	3,10
16	1956	17,8	0,86	23,1	1,12	39,0	15,30	-0,70	2,88
17	1957	20,3	0,98	23	1,11	41,5	16,28	-0,72	3,01
18	1958	24	1,16	22,3	1,08	43,9	17,44	-0,56	3,18
19	1959	16,6	0,80	20,3	0,98	46,3	18,25	-0,75	2,81
20	1960	29,1	1,41	19,9	0,96	48,8	19,65	-0,35	3,37
21	1961	14,5	0,70	19,4	0,94	51,2	20,36	0,64	2,67
22	1962	19,9	0,96	19,4	0,94	53,7	21,32	-0,68	2,99
23	1963	24,1	1,17	18,8	0,91	56,1	22,48	-0,52	3,18
24	1964	18,8	0,91	18,4	0,89	58,5	23,39	-0,61	2,93
25	1965	17,3	0,84	18,1	0,88	61,0	24,23	-0,77	2,85
26	1966	26,6	1,29	17,9	0,87	63,4	25,52	-0,48	3,28
27	1967	28,8	1,39	17,8	0,86	65,9	26,91	-0,09	3,36
28	1968	19,4	0,94	17,3	0,84	68,3	27,85	-0,15	2,97
29	1969	23	1,11	16,8	0,81	70,7	28,96	-0,04	3,14
30	1970	35,9	1,74	16,6	0,80	73,2	30,70	0,70	3,58
31	1971	23,1	1,12	15,2	0,74	75,6	31,82	0,82	3,14
32	1972	12,3	0,60	14,5	0,70	78,0	32,41	0,41	2,51
33	1973	13,9	0,67	13,9	0,67	80,5	33,09	0,09	2,63
34	1974	18,4	0,89	13,6	0,66	82,9	33,98	-0,02	2,91
35	1975	10,7	0,52	13,6	0,66	85,4	34,49	-0,51	2,37
36	1976	11,4	0,55	12,9	0,62	87,8	35,05	-0,95	2,43
37	1977	23,5	1,14	12,3	0,60	90,2	36,18	-0,82	3,16
39	1979	29,3	1,42	10,7	0,52	95,1	38,83	-0,17	3,38
40	1980	24,1	1,17	10,2	0,49	97,6	40,00	0,00	3,18

Расчет статистических характеристик по выборкам различной продолжительности

n	Среднее	СКО	C_v	C_s	R_l	Сред. LnQ	СКО LnQ
20	20	6	0,302	0,221	0,000		

20	21	7	0,315	0,224	0,388

Максимальный сток по р. Ока - Костомарово

№ п/п	Год	Q_i, м^3/с	K_i	Q_i, ранжир.	K_i, ранжир.	P_i, %	ΣKi	Σ(K_i-1)	LnQ_i
1	1941	404	0,78	1510	2,91	2,44	0,78	-0,22	6,00
2	1942	1190	2,29	1190	2,29	4,88	3,07	1,07	7,08
3	1943	360	0,69	1120	2,16	7,32	3,76	0,76	5,89
4	1944	450	0,87	1080	2,08	9,76	4,63	0,63	6,11
5	1945	450	0,87	1040	2,00	12,20	5,49	0,49	6,11
6	1946	910	1,75	910	1,75	14,63	7,24	1,24	6,81
7	1947	1510	2,91	862	1,66	17,07	10,15	3,15	7,32
8	1948	640	1,23	726	1,40	19,51	11,38	3,38	6,46
9	1949	404	0,78	686	1,32	21,95	12,16	3,16	6,00
10	1950	360	0,69	656	1,26	24,39	12,85	2,85	5,89
11	1951	1040	2,00	640	1,23	26,83	14,85	3,85	6,95
12	1952	1120	2,16	632	1,22	29,27	17,01	5,01	7,02
13	1953	616	1,19	616	1,19	31,71	18,19	5,19	6,42
14	1954	656	1,26	581	1,12	34,15	19,45	5,45	6,49
15	1955	632	1,22	564	1,09	36,59	20,67	5,67	6,45
16	1956	446	0,86	562	1,08	39,02	21,53	5,53	6,10
17	1957	421	0,81	522	1,00	41,46	22,34	5,34	6,04
18	1958	422	0,81	458	0,88	43,90	23,15	5,15	6,05
19	1959	522	1,00	450	0,87	46,34	24,16	5,16	6,26
20	1960	564	1,09	450	0,87	48,78	25,24	5,24	6,34
21	1961	284	0,55	446	0,86	51,22	25,79	4,79	5,65
22	1962	581	1,12	422	0,81	53,66	26,91	4,91	6,36
23	1963	155	0,30	421	0,81	56,10	27,20	4,20	5,04
24	1964	862	1,66	404	0,78	58,54	28,86	4,86	6,76
25	1965	111	0,21	404	0,78	60,98	29,08	4,08	4,71
26	1966	171	0,33	360	0,69	63,41	29,40	3,40	5,14
27	1967	185	0,36	360	0,69	65,85	29,76	2,76	5,22
28	1968	125	0,24	352	0,68	68,29	30,00	2,00	4,83
29	1969	148	0,28	308	0,59	70,73	30,29	1,29	5,00
30	1970	231	0,44	289	0,56	73,17	30,73	0,73	5,44
31	1971	726	1,40	288	0,55	75,61	32,13	1,13	6,59
32	1972	289	0,56	284	0,55	78,05	32,68	0,68	5,67
33	1973	562	1,08	231	0,44	80,49	33,77	0,77	6,33
34	1974	288	0,55	185	0,36	82,93	34,32	0,32	5,66
35	1975	68,2	0,13	171	0,33	85,37	34,45	-0,55	4,22
36	1976	308	0,59	155	0,30	87,80	35,04	-0,96	5,73
37	1977	686	1,32	148	0,28	90,24	36,36	-0,64	6,53
38	1978	458	0,88	125	0,24	92,68	37,24	-0,76	6,13
39	1979	1080	2,08	111	0,21	95,12	39,32	0,32	6,98
40	1980	352	0,68	68,2	0,13	97,56	40,00	0,00	5,86

Расчет статистических характеристик по выборкам различной продолжительности

n	Среднее	СКО	C_v	C_s	R_l	Сред. LnQ	СКО LnQ
40	520	330	0,635	0,999	0,206	6,04	0,701
20	656	325	0,496	1,174	0,121		

20	384	280	0,730	0,936	-0,062

Расчет параметров уравнения регрессии $Q_{ср}=F(Q_{max})$ по выборкам различной продолжительности

n	среднее Y	среднее X	СКО Y	СКО X	R_{yx}
36	20	506	6	332	0,382
18	20	668	6	342	0,841
18	20	343	7	234	-0,022

Приложение 4

Расчёт коэффициента корреляции между процессами среднегодового и максимального стоков по р.Ока - Костомарово

№ п/п	год	Y	X	dY	dX	dY*dX	dY^2	dX^2
1	1941	404	18,1	-2,9	-116	336,4	8,4	13456
2	1942	1190	29	8	670	5360	64,0	448900
3	1943	360	10,2	-10,8	-160	1728	116,6	25600
4	1944	450	13,6	-7,4	-70	518	54,8	4900
5	1945	450	16,8	-4,2	-70	294	17,6	4900
6	1946	910	23,4	2,4	390	936	5,8	152100
7	1947	1510	29,7	8,7	990	8613	75,7	980100
8	1948	640	17,9	-3,1	120	-372	9,6	14400
9	1949	404	12,9	-8,1	-116	939,6	65,6	13456
10	1950	360	13,6	-7,4	-160	1184	54,8	25600
11	1951	1040	26,6	5,6	520	2912	31,4	270400
12	1952	1120	29,7	8,7	600	5220	75,7	360000
13	1953	616	19,4	-1,6	96	-153,6	2,6	9216
14	1954	656	15,2	-5,8	136	-788,8	33,6	18496
15	1955	632	22,3	1,3	112	145,6	1,7	12544
16	1956	446	17,8	-3,2	-74	236,8	10,2	5476
17	1957	421	20,3	-0,7	-99	69,3	0,5	9801
18	1958	422	24	3	-98	-294	9,0	9604
19	1959	522	16,6	-4,4	2	-8,8	19,4	4
20	1960	564	29,1	8,1	44	356,4	65,6	1936
21	1961	284	14,5	-6,5	-236	1534	42,3	55696
22	1962	581	19,9	-1,1	61	-67,1	1,2	3721
23	1963	155	24,1	3,1	-365	-1131,5	9,6	133225
24	1964	862	18,8	-2,2	342	-752,4	4,8	116964
25	1965	111	17,3	-3,7	-409	1513,3	13,7	167281
26	1966	171	26,6	5,6	-349	-1954,4	31,4	121801
27	1967	185	28,8	7,8	-335	-2613	60,8	112225
28	1968	125	19,4	-1,6	-395	632	2,6	156025
29	1969	148	23	2	-372	-744	4,0	138384
30	1970	231	35,9	14,9	-289	-4306,1	222,0	83521
31	1971	726	23,1	2,1	206	432,6	4,4	42436
32	1972	289	12,3	-8,7	-231	2009,7	75,7	53361
33	1973	562	13,9	-7,1	42	-298,2	50,4	1764
34	1974	288	18,4	-2,6	-232	603,2	6,8	53824
35	1975	68,2	10,7	-10,3	-451,8	4653,54	106,1	204123
36	1976	308	11,4	-9,6	-212	2035,2	92,2	44944
37	1977	686	23,5	2,5	166	415	6,3	27556
38	1978	458	25,5	4,5	-62	-279	20,3	3844
39	1979	1080	29,3	8,3	560	4648	68,9	313600
40	1980	352	24,1	3,1	-168	-520,8	9,6	28224
S		20787	826,7			33041,94	1555,4	4243408
СКО		330	6	корень квадратный из SdY^2SdX^2 =				81241,859
среднее		520	21				r =	0,41

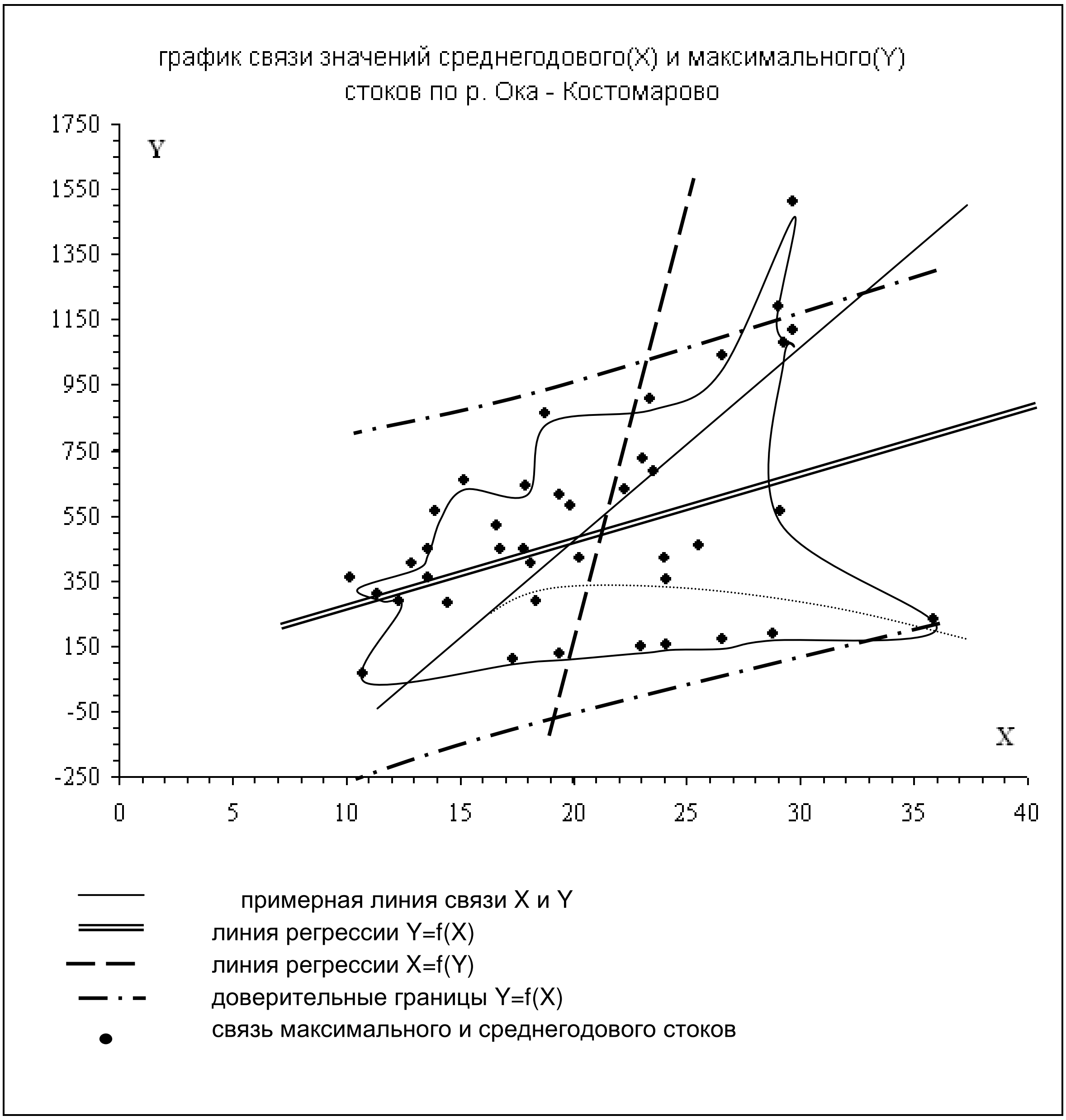
график связи значений среднегодового(X) и максимального(Y)
стоков по р. Ока - Костомарово
Y
1750
1550
1350
1150
950
750
550
350
150
-50
-250
X
0
5
10
15
20
25
30
35
40
примерная линия связи X и Y
линия регрессии Y=f(X)
линия регрессии X=f(Y)
доверительные границы Y=f(X)
связь максимального и среднегодового стоков

рис.1

Расчёт параметров уравнений регрессии Y=f(X) и X=f(Y) между процессами среднегодового и максимального стоков по р.Ока -Костомарово

№п/п	год	Y	X	dX	dY	dX^2	dY^2	dY*dX	Yрасч	Xрасч
1	1941	404	18,1	-2,9	-116	8,41	13456	336,4	462,1	20,0
2	1942	1190	29	8	670	64	448900	5360	678,4	25,8
3	1943	360	10,2	-10,8	-160	116,64	25600	1728	305,4	19,7
4	1944	450	13,6	-7,4	-70	54,76	4900	518	372,8	20,3
5	1945	450	16,8	-4,2	-70	17,64	4900	294	436,3	20,3
6	1946	910	23,4	2,4	390	5,76	152100	936	567,3	23,7
7	1947	1510	29,7	8,7	990	75,69	980100	8613	692,2	28,2
8	1948	640	17,9	-3,1	120	9,61	14400	-372	458,1	21,7
9	1949	404	12,9	-8,1	-116	65,61	13456	939,6	358,9	20,0
10	1950	360	13,6	-7,4	-160	54,76	25600	1184	372,8	19,7
11	1951	1040	26,6	5,6	520	31,36	270400	2912	630,7	24,7
12	1952	1120	29,7	8,7	600	75,69	360000	5220	692,2	25,3
13	1953	616	19,4	-1,6	96	2,56	9216	-153,6	487,9	21,6
14	1954	656	15,2	-5,8	136	33,64	18496	-788,8	404,6	21,9
15	1955	632	22,3	1,3	112	1,69	12544	145,6	545,4	21,7
16	1956	446	17,8	-3,2	-74	10,24	5476	236,8	456,2	20,3
17	1957	421	20,3	-0,7	-99	0,49	9801	69,3	505,8	20,1
18	1958	422	24	3	-98	9	9604	-294	579,2	20,1
19	1959	522	16,6	-4,4	2	19,36	4	-8,8	432,3	20,9
20	1960	564	29,1	8,1	44	65,61	1936	356,4	680,3	21,2
21	1961	284	14,5	-6,5	-236	42,25	55696	1534	390,7	19,1
22	1962	581	19,9	-1,1	61	1,21	3721	-67,1	497,8	21,3
23	1963	155	24,1	3,1	-365	9,61	133225	-1131,5	581,1	18,1
24	1964	862	18,8	-2,2	342	4,84	116964	-752,4	476,0	23,4
25	1965	111	17,3	-3,7	-409	13,69	167281	1513,3	446,2	17,8
26	1966	171	26,6	5,6	-349	31,36	121801	-1954,4	630,7	18,3
27	1967	185	28,8	7,8	-335	60,84	112225	-2613	674,4	18,4
28	1968	125	19,4	-1,6	-395	2,56	156025	632	487,9	17,9
29	1969	148	23	2	-372	4	138384	-744	559,3	18,1
30	1970	231	35,9	14,9	-289	222,01	83521	-4306,1	815,3	18,7
31	1971	726	23,1	2,1	206	4,41	42436	432,6	561,3	22,4
32	1972	289	12,3	-8,7	-231	75,69	53361	2009,7	347,0	19,1
33	1973	562	13,9	-7,1	42	50,41	1764	-298,2	378,8	21,2
34	1974	288	18,4	-2,6	-232	6,76	53824	603,2	468,1	19,1
35	1975	68,2	10,7	-10,3	-451,8	106,09	204123,2	4653,54	315,3	17,5
36	1976	308	11,4	-9,6	-212	92,16	44944	2035,2	329,2	19,3
37	1977	686	23,5	2,5						
38	1978	458	25,5	4,5						
39	1979	1080	29,3	8,3						
40	1980	352	24,1	3,1						
S		20787	826,7	-13,3	-508,8	1450,41	3870184	28778,74		
a	19,84	a'	0,0074	r	0,41					
b	103	b'	17							

Параметры уравнения регрессии по р. Ока - Костомарово

	коэффициент			СКО(r)			0,14
	корреляции ryx*		0,41	СКО(a)			7,63
a		19,84	a'	0,0074	СКО(уравнения регрессии y(x))		
b		103	b'	17			300,988

Расчет доверительного интервала для уравнения регрессии между процессами среднегодового и максимального стоков по р.Ока -Костомарово

№ п/п	Xранж	Xрасч	Yранж	Yрасч	t	□y(x)	Yниж. гр.	Yверх. гр.
1	35,9	28,2	1510	815,26	2,48	329,6	258	1372
2	29,7	25,8	1190	692,25	1,45	313,7	162	1222
3	29,7	25,3	1120	692,25	1,45	313,7	162	1222
4	29,1	24,7	1040	680,34	1,35	312,6	152	1209
5	29	23,7	910	678,36	1,33	312,4	150	1206
6	28,8	23,4	862	674,39	1,30	312,0	147	1202
7	26,6	22,4	726	630,74	0,93	308,7	109	1152
8	26,6	21,9	656	630,74	0,93	308,7	109	1152
9	24,1	21,7	640	581,14	0,52	306,3	64	1099
10	24	21,7	632	579,16	0,50	306,2	62	1097
11	23,4	21,6	616	567,26	0,40	305,8	50	1084
12	23,1	21,3	581	561,30	0,35	305,7	45	1078
13	23	21,2	564	559,32	0,33	305,6	43	1076
14	22,3	21,2	562	545,43	0,22	305,3	29	1061
15	20,3	20,9	522	505,75	-0,12	305,2	-10	1022
16	19,9	20,3	450	497,82	-0,18	305,3	-18	1014
17	19,4	20,3	450	487,90	-0,27	305,4	-28	1004
18	19,4	20,3	446	487,90	-0,27	305,4	-28	1004
19	18,8	20,1	422	475,99	-0,37	305,7	-41	993
20	18,4	20,1	421	468,06	-0,43	305,9	-49	985
21	18,1	20,0	404	462,10	-0,48	306,1	-55	979
22	17,9	20,0	404	458,14	-0,52	306,3	-59	976
23	17,8	19,7	360	456,15	-0,53	306,3	-62	974
24	17,3	19,7	360	446,23	-0,62	306,7	-72	965
25	16,8	19,3	308	436,31	-0,70	307,2	-83	955
26	16,6	19,1	289	432,34	-0,73	307,4	-87	952
27	15,2	19,1	288	404,57	-0,97	309,0	-118	927
28	14,5	19,1	284	390,68	-1,08	310,0	-133	915
29	13,9	18,7	231	378,78	-1,18	310,9	-147	904
30	13,6	18,4	185	372,82	-1,23	311,4	-153	899
31	13,6	18,3	171	372,82	-1,23	311,4	-153	899
32	12,9	18,1	155	358,94	-1,35	312,6	-169	887
33	12,3	18,1	148	347,03	-1,45	313,7	-183	877
34	11,4	17,9	125	329,18	-1,60	315,5	-204	862
35	10,7	17,8	111	315,29	-1,72	317,1	-221	851
36	10,2	17,5	68,2	305,37	-1,80	318,2	-232	843
□	724,3	746,8	18211,2	18078,11				

Оценка значимости коэффициента корреляции по р. Ока-Костомарово

H_0: r_{yx}*=0		r*	0,382
sr	0,141	ta	1,69
2a	0,1	ta*sr	0,2376

$|r_{yx}^*|$ > ta*sr = >Гипотеза опровергается

С вероятностью 90% можно считать, что корреляционная связь имеет место.

Оценка доверительных границ коэффициента корреляциии с помощью преобразования Фишера по р. Ока-Костомарово

z	0,436		sz	0,164
2a	0,1		ta	1,69
	0,158	< z <	0,713	
	0,156	< r_{yx} <	0,613	

Оценка стационарности связи по 18-летним периодам по р. Ока-Костомарово

период	H_0:	СКО		r	t	ta	s*
первые 18 лет	$r_1=r_2$	СКО (x)	20	0,841	4,823	1,69	0,258
		СКО (y)	668				
		z_1	1,223				
вторые 18 лет		СКО (x)	20	-0,022			
		СКО (y)	343				
		z_2	-0,022				

t > ta => Гипотеза опровергается.
При 10% уровне значимости стационарность связи не имеет места быть.

Приложение 5

таб. 2.1

Оценка числовых характеристик методом моментов
по р.Ока –Костомарово

название ряда	ОЦЕНКИ			
	m^*_x	s^*_x	C*v	C*s
$Q_{ср. г.}$	21	6	0,297	0,264
Qmax	520	326	0,626	1,047

таб. 2. 2

Оценка коэффициента вариации методом наибольшего правдоподобия
по р.Ока - Костомарово

ряд	лямбда	C*v
$Q_{ср. г.}$	-0,05	0,304
Qmax	-0,21	0,610

таб. 2.3

Доверительные интервалы числовых характеристик для среднегодового и максимального стоков за 40 лет по р. Ока – Костомарово

ряд	m_x		D		s	
	от	до	от	до	от	до
$Q_{ср. г.}$	19	23	27	57	5	8
Q_{max}	432	608	74146	156125	276	400

таб.3.1

Расчёт основных характеристик за различные периоды времени по р. Ока- Костомарово

период (лет)	среднегодовой сток			максимальный сток		
	m_x	s_x	C_v	m_y	s_y	C_v
40	21	6	0,305	520	330	0,635
20	20	6	0,302	656	325	0,496
20	21	7	0,315	384	280	0,730

таб.3.2

Оценка доверительных интервалов математического ожидания за различные периоды времени по р. Ока- Костомарово

период (лет)	среднегодовой сток			максимальный сток		
	min	max	C_v	min	max	C_v
40	19	23	0,305	432	608	0,635
20	18	22	0,302	533	779	0,496
20	18	24	0,315	278	490	0,730

таб.3.3

Оценка равенства математических ожиданий по выборкам различных продолжительностей по р. Ока- Костомарово

процесс	период	m_x	Ho:	статистика t	t_a	результат
$Q_{ср.г}$ (X)	20 лет	20	$M_{x1}=M_{x2}$	-0,49	1,68	гипотеза не опровергается
	20 лет	21				
	40 лет	21	$M_{x1}=M_{x2}$	9,56	1,67	гипотеза опровергается
Q_{max} (Y)	40 лет	520				
	20 лет	656	$M_{x1}=M_{x2}$	2,84	1,68	гипотеза опровергается
	20 лет	384				

таб.3.4

Оценка равенства дисперсий по критерию Фишера
за различные периоды времени по р. Ока- Костомарово

процесс	период	D_x	Ho:	статистика F	$F_{0,1}$	результат
$Q_{ср.г}$ (X)	20 лет	36	$D_{x1}=D_{x2}$	1,36	1,839	гипотеза не опровергается
	20 лет	49				
	40 лет	36	$D_{x1}=D_{x2}$	3025,00	1,545	гипотеза опровергается
Q_{max} (Y)	40 лет	108900				
	20 лет	105625	$D_{x1}=D_{x2}$	1,11	1,839	гипотеза не опровергается
	20 лет	117649				

таб.3.5

Оценка равенства дисперсий по критерию Романовского
по р. Ока- Костомарово

процесс	период	D_x	Ho:	статистика F	q	s(q)	R	результат
$Q_{ср.г}$ (X)	20 лет	36	$D_{x1}=D_{x2}$	1,36	1,224	0,49	0,457	гипотеза не опровергается
	20 лет	49						
	40 лет	36	$D_{x1}=D_{x2}$	3025,00	2873,750	0,33	8705,303	гипотеза опровергается
Q_{max} (Y)	40 лет	108900						
	20 лет	105625	$D_{x1}=D_{x2}$	1,11	0,999	0,49	0,224	гипотеза не опровергается
	20 лет	117649						

таб.3.6

Оценка однородности среднегодового стока по критерию
Уилкоксона по р. Ока- Костомарово

ряд	Ho:	$u_н$	u	$u_в$	s(u)	M(u)	D(u)	результат
20 лет 20 лет	$F_1=F_2$	337	394	483	37	410	1367	гипотеза не опровергается

таб.3.7

Оценка однородности максимального стока по критерию
Уилкоксона по р. Ока- Костомарово

ряд	Ho:	$u_н$	u	$u_в$	s(u)	M(u)	D(u)	результат
20 лет 20 лет	$F_1=F_2$	337	518	483	37	410	1367	гипотеза опровергается

таб.4.1

Проверка случайности по критерию длин и числа серий
пор.Ока - Костомарово

процесс	Ho:	$R_{0,k}$	$R_{1,k}$	Rниж	R_k	Rверх	результат
$Q_{ср.\ г.}$	процесс случайный	14	21	13	18	23	гипотеза не опровергается
Q_{max}		17	21	14	19	25	гипотеза не опровергается

таб.4.2

Проверка выборочных рядов среднегодового и максимального стоков
на случайность по числу повышений и понижений по р.Ока - Костомарово

процесс	Но:	n_+	n_-	m_+	m_-	D_+	D_-	t_+	t_-	$t_{0,1}$	результат
$Q_{ср.\ г.}$	процесс случайный	24	15	19,5	19,5	3,33	3,33	2,46	-2,46	1,68	гипотеза опровергается
Q_{max}		21	17	19	19	3,25	3,25	1,11	-1,11	1,68	гипотеза не опровергается

таб.4.3

Проверка выборочных рядов среднегодового и максимального стоков
на случайность по числу экстремумов по р.Ока- Костомарово

процесс	Но:	$n_э$	$m_э$	$D_э$	$t_э$	$t_{0,1}$	результат
$Q_{ср.\ г.}$	процесс случайный	21	26,00	6,61	-1,94	1,68	гипотеза не опровергается
Q_{max}		23	25,33	6,43	-0,92	1,68	гипотеза не опровергается

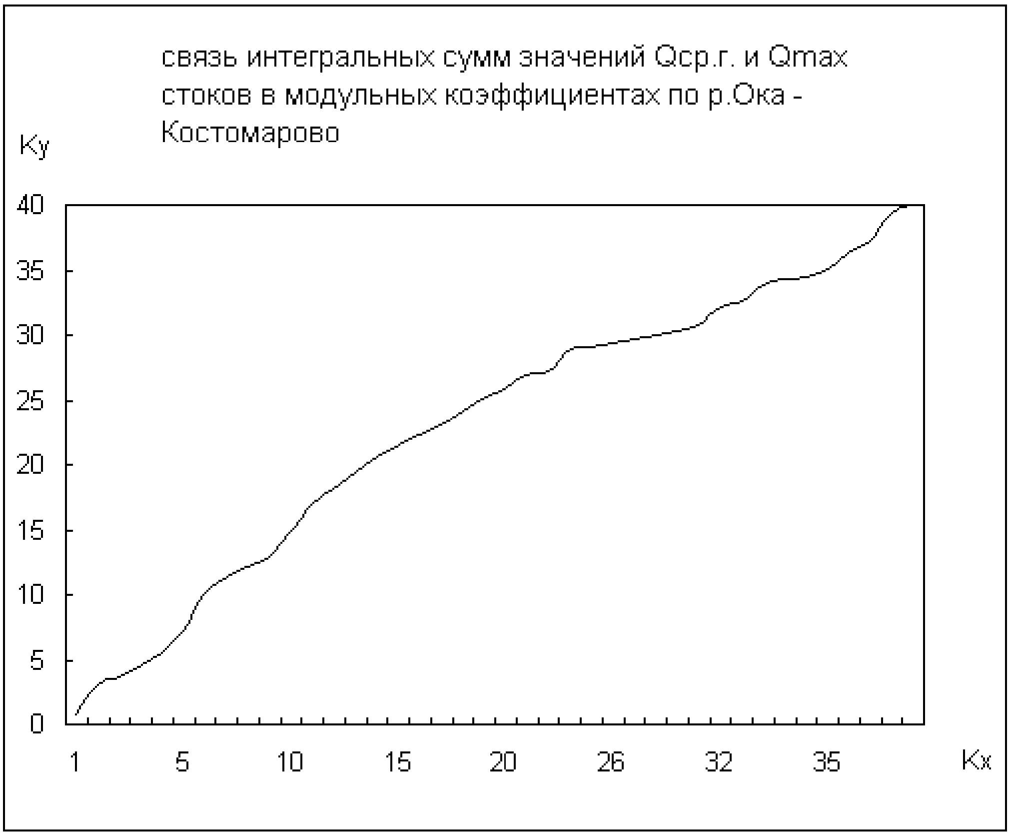

рис. 4.1

таб.5.1

Расчёт параметров уравнений регрессии Y=f(X) и X=f(Y)
по р.Ока - Костомарово

№п/п	год	Y(Qmax)	X(Qср.г.)	□X	□□Y	□X^2	□□Y^2	□X*□□Y	Yрасч	Xрасч
1	1947	404	18,1	-1,9	-102	3,61	10404	193,8	454	19
2	1974	1190	29	9	684	81	467856	6156	662	25
3	1965	360	10,2	-9,8	-146	96,04	21316	1430,8	303	19
4	1944	450	13,6	-6,4	-56	40,96	3136	358,4	368	19
5	1966	450	16,8	-3,2	-56	10,24	3136	179,2	429	19
6	1957	910	23,4	3,4	404	11,56	163216	1373,6	555	23
7	1942	1510	29,7	9,7	1004	94,09	1008016	9738,8	675	27
8	1968	640	17,9	-2,1	134	4,41	17956	-281,4	450	21
9	1950	404	12,9	-7,1	-102	50,41	10404	724,2	354	19
10	1943	360	13,6	-6,4	-146	40,96	21316	934,4	368	19
11	1971	1040	26,6	6,6	534	43,56	285156	3524,4	616	24
12	1958	1120	29,7	9,7	614	94,09	376996	5955,8	675	24
13	1948	616	19,4	-0,6	110	0,36	12100	-66	479	20
14	1972	656	15,2	-4,8	150	23,04	22500	-720	398	21
15	1970	632	22,3	2,3	126	5,29	15876	289,8	534	21
16	1962	446	17,8	-2,2	-60	4,84	3600	132	448	19
17	1946	421	20,3	0,3	-85	0,09	7225	-25,5	496	19
18	1969	422	24	4	-84	16	7056	-336	566	19
19	1956	522	16,6	-3,4	16	11,56	256	-54,4	425	20
20	1955	564	29,1	9,1	58	82,81	3364	527,8	664	20
21	1949	284	14,5	-5,5	-222	30,25	49284	1221	385	18
22	1959	581	19,9	-0,1	75	0,01	5625	-7,5	488	20
23	1941	155	24,1	4,1	-351	16,81	123201	-1439,1	568	17
24	1963	862	18,8	-1,2	356	1,44	126736	-427,2	467	22
25	1961	111	17,3	-2,7	-395	7,29	156025	1066,5	438	17
26	1945	171	26,6	6,6	-335	43,56	112225	-2211	616	17
27	1976	185	28,8	8,8	-321	77,44	103041	-2824,8	658	17
28	1960	125	19,4	-0,6	-381	0,36	145161	228,6	479	17
29	1964	148	23	3	-358	9	128164	-1074	547	17
30	1953	231	35,9	15,9	-275	252,81	75625	-4372,5	794	18
31	1954	726	23,1	3,1	220	9,61	48400	682	549	21
32	1951	289	12,3	-7,7	-217	59,29	47089	1670,9	343	18
33	1952	562	13,9	-6,1	56	37,21	3136	-341,6	373	20
34	1967	288	18,4	-1,6	-218	2,56	47524	348,8	459	18
35	1973	68,2	10,7	-9,3	-437,8	86,49	191668,8	4071,54	312	16
36	1975	308	11,4	-8,6	-198	73,96	39204	1702,8	326	18
□		18211	724,3	4,3	-4,8	1423	3862994	28330,14		
СКО		332	6	a =	19,91	a' =	0,007	*r* =	0,382	
среднее		506	20	b =	108	b' =	16		СКО(r) =	0,144
				СКО(a) =	8,6		СКО(y(x)) =	306,8		

таб. 5.2а

Расчёт коэффициента корреляции между процессами
среднегодового и максимального стоков и параметров уравнения регрессии
по первым 18-ти годам по р.Ока – Костомарово

№ п/п	год	Y	X	□Y	□X	□Y□□X	□Y□	□X^2
1	1941	404	18,1	-264	-1,9	501,6	69696	3,61
2	1942	1190	29	522	9	4698	272484	81
3	1943	360	10,2	-308	-9,8	3018,4	94864	96,04
4	1944	450	13,6	-218	-6,4	1395,2	47524	40,96
5	1945	450	16,8	-218	-3,2	697,6	47524	10,24
6	1946	910	23,4	242	3,4	822,8	58564	11,56
7	1947	1510	29,7	842	9,7	8167,4	708964	94,09
8	1948	640	17,9	-28	-2,1	58,8	784	4,41
9	1949	404	12,9	-264	-7,1	1874,4	69696	50,41
10	1950	360	13,6	-308	-6,4	1971,2	94864	40,96
11	1951	1040	26,6	372	6,6	2455,2	138384	43,56
12	1952	1120	29,7	452	9,7	4384,4	204304	94,09
13	1953	616	19,4	-52	-0,6	31,2	2704	0,36
14	1954	656	15,2	-12	-4,8	57,6	144	23,04
15	1955	632	22,3	-36	2,3	-82,8	1296	5,29
16	1956	446	17,8	-222	-2,2	488,4	49284	4,84
17	1957	421	20,3	-247	0,3	-74,1	61009	0,09
18	1958	422	24	-246	4	-984	60516	16
□		12031	360,5			29481,3	1982605	620,55
СКО		342	6	r =	0,841	СКО(r)	0,071	
среднее		668	20	a =	47,5	СКО(a)	7,48	
a' =	0,015	b' =	10	b =	-282	СКОy(x)	185,03	

таб. 5.2б

Расчёт коэффициента корреляции между процессами
среднегодового и максимального стоков и параметров уравнения регрессии
по вторым 18-ти годам по р.Ока – Костомарово

№ п/п	год	Y	X	□Y	□X	□Y□□X	□Y□	□X^2
1	1959	522	16,6	179	-3,4	-608,6	32041	11,56
2	1960	564	29,1	221	9,1	2011,1	48841	82,81
3	1961	284	14,5	-59	-5,5	324,5	3481	30,25
4	1962	581	19,9	238	-0,1	-23,8	56644	0,01
5	1963	155	24,1	-188	4,1	-770,8	35344	16,81
6	1964	862	18,8	519	-1,2	-622,8	269361	1,44
7	1965	111	17,3	-232	-2,7	626,4	53824	7,29
8	1966	171	26,6	-172	6,6	-1135,2	29584	43,56
9	1967	185	28,8	-158	8,8	-1390,4	24964	77,44
10	1968	125	19,4	-218	-0,6	130,8	47524	0,36
11	1969	148	23	-195	3	-585	38025	9
12	1970	231	35,9	-112	15,9	-1780,8	12544	252,81
13	1971	726	23,1	383	3,1	1187,3	146689	9,61
14	1972	289	12,3	-54	-7,7	415,8	2916	59,29
15	1973	562	13,9	219	-6,1	-1335,9	47961	37,21
16	1974	288	18,4	-55	-1,6	88	3025	2,56
17	1975	68,2	10,7	-274,8	-9,3	2555,64	75515,04	86,49
18	1976	308	11,4	-35	-8,6	301	1225	73,96
□		6180,2	363,8			-612,76	929508	802,46
СКО		234	7	r =	-0,022	СКО(r)	0,243	
среднее		343	20	a =	-0,8	СКО(a)	8,11	
a' =	-0,001	b' =	20	b =	359	СКО(y(x))	234,06	

таб. 5.3

Параметры уравнения регрессии по р. Ока – Костомарово

коэффициент корреляции *r*			0,382	СКО(r)	0,144
				СКО(a)	8,600
a	19,91	a'	0,007	СКО(уравнения регрессии y(x))	306,8
b	108	b'	16		

таб. 5.4

Оценка доверительных границ коэффициента корреляциии с помощью преобразования Фишера по р. Ока – Костомарово

коэффициент корреляции r			0,382	доверительные границы		
z	0,403	СКО(z)	0,174	нижняя		верхняя
2□	10%	$t_{□}$	1,69	для z	0,109	0,697
				для *r*	0,108	0,603

таб.5.5

Оценка значимости коэффициента корреляции по р. Ока – Костомарово

H_0:	r_{yx}*	□□r	t□	t□□□r	результат
r_{yx}* = 0	0,382	0,144	1,69	0,243	гипотеза опровергается

таб. 5.6

Оценка стационарности связи по 18-летним периодам по р. Ока – Костомарово

период	H_0:	СКО		r	t	t□	результат
первые 18 лет	r1=r2	СКО (x) СКО (y) СКО(r)	6 342 0,071	0,841	3,4	1,69	гипотеза опровергается
вторые 18 лет		СКО (x) СКО (y) СКО(r)	7 234 0,243	-0,022			

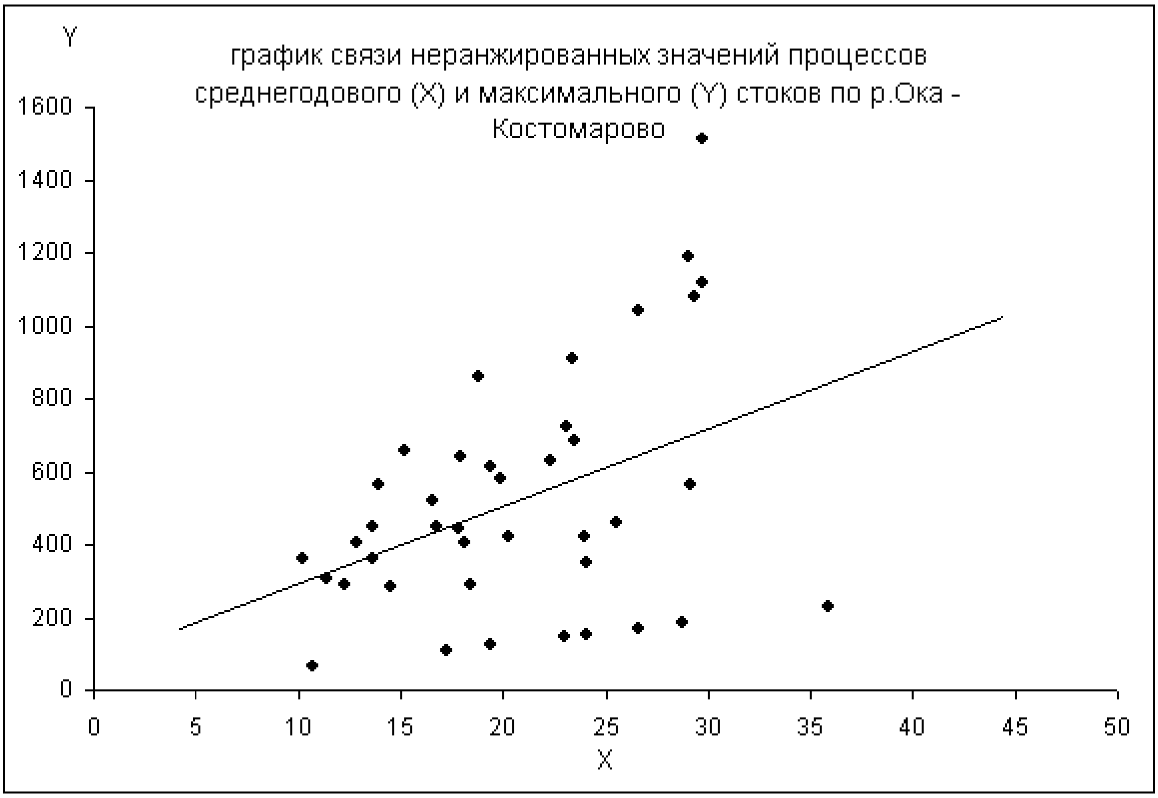
Y
график связи неранжированных значений процессов среднегодового (X) и максимального (Y) стоков по р.Ока - Костомарово
1600
1400
1200
1000
800
600
400
200
0
0
5
10
15
20
25
30
35
40
45
50
X

рис. 5.1

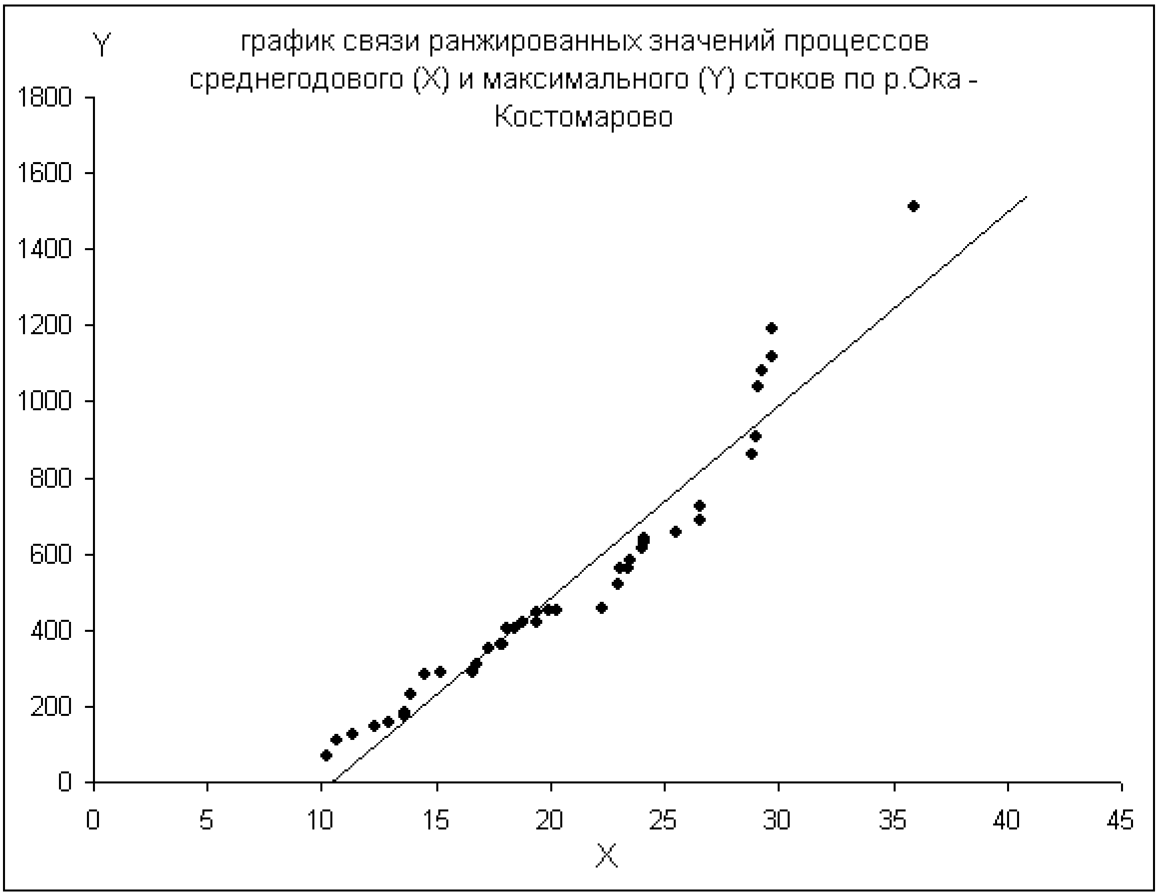
Y
график связи ранжированных значений процессов среднегодового (X) и максимального (Y) стоков по р.Ока - Костомарово
1800
1600
1400
1200
1000
800
600
400
200
0
0
5
10
15
20
25
30
35
40
45
X

рис. 5.2

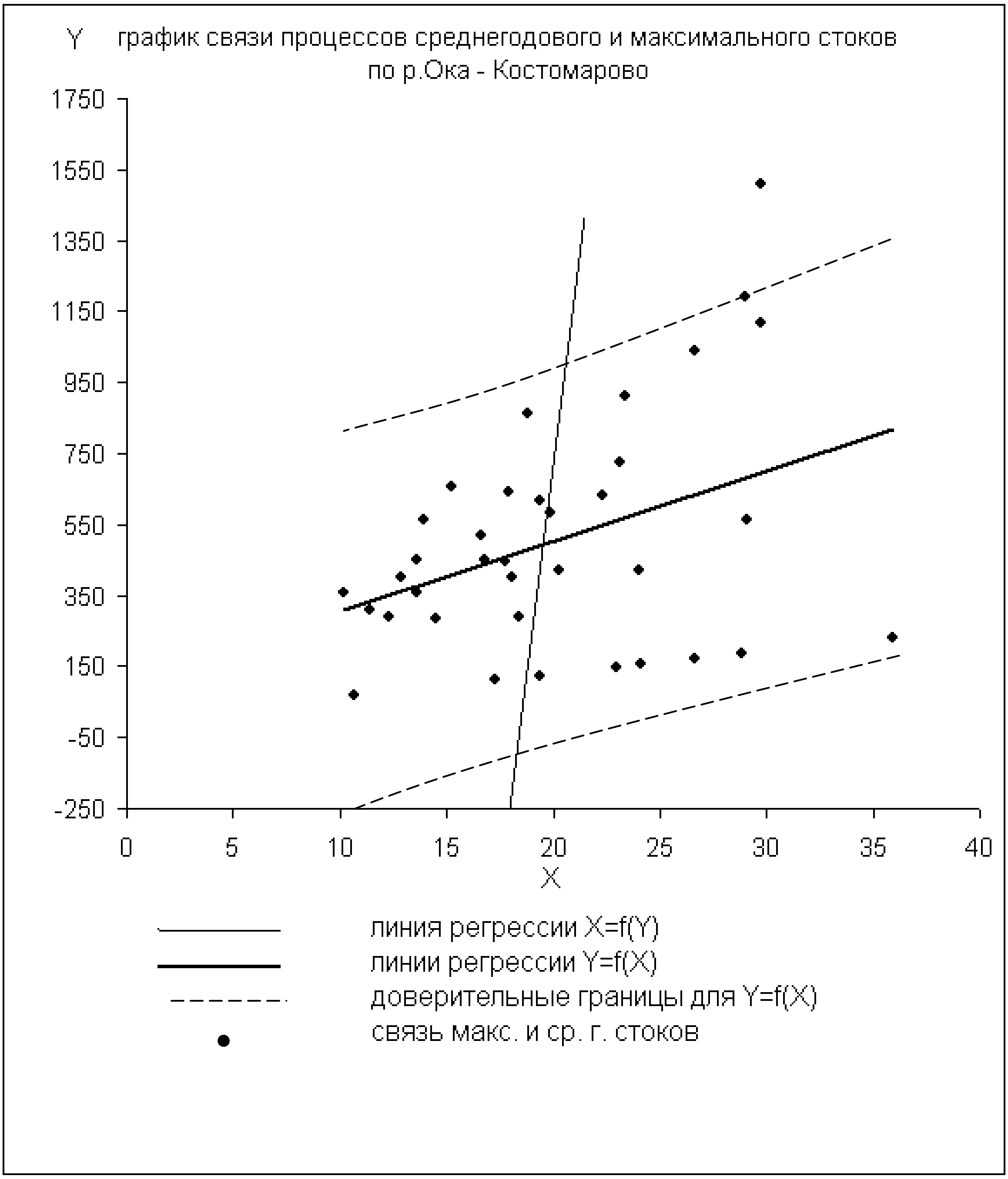
Y
график связи процессов среднегодового и максимального стоков
по р.Ока - Костомарово
1750
1550
1350
1150
950
750
550
350
150
-50
-250
0
5
10
15
20
25
30
35
40
X
линия регрессии X=f(Y)
линии регрессии Y=f(X)
доверительные границы для Y=f(X)
связь макс. и ср. г. стоков

рис. 5.3

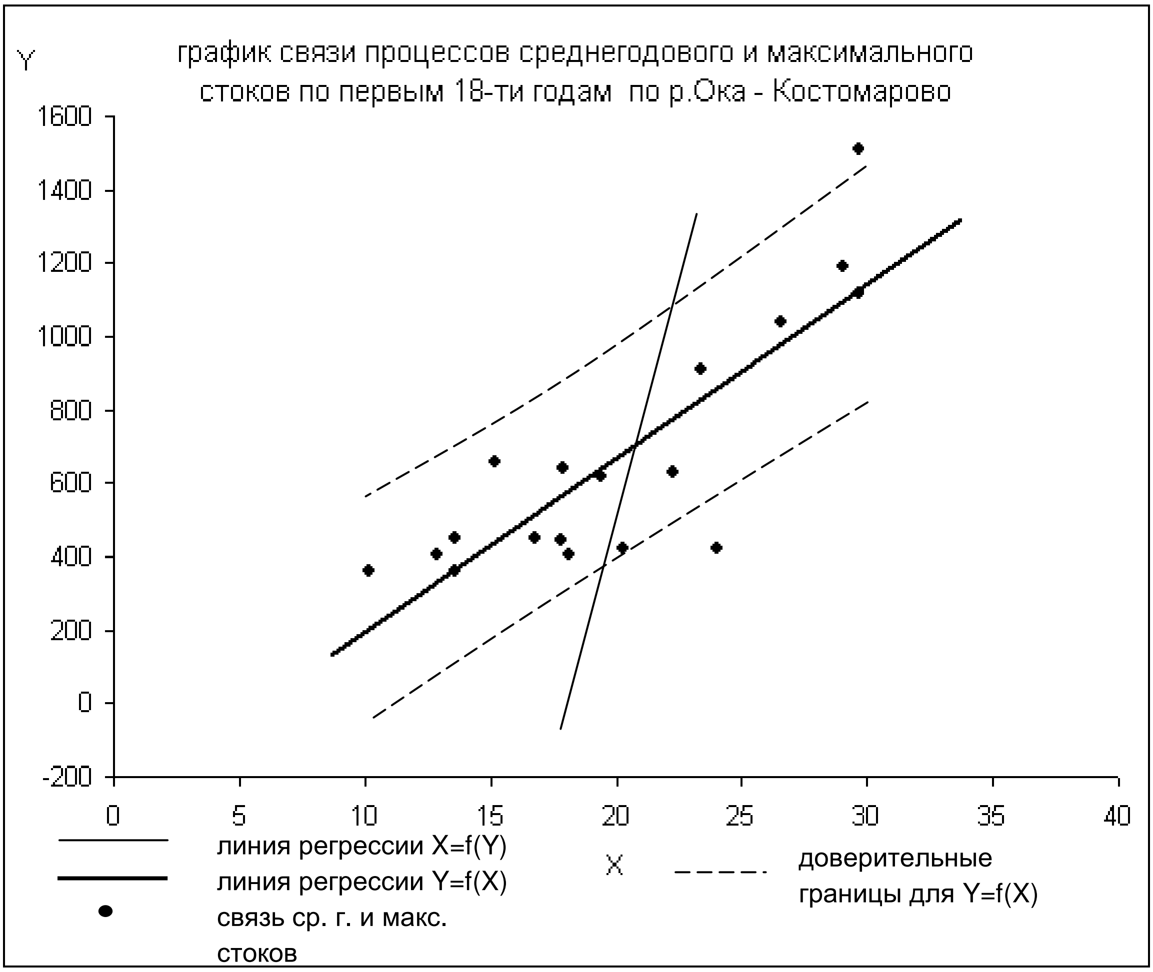
Y
график связи процессов среднегодового и максимального стоков по первым 18-ти годам по р.Ока - Костомарово
1600
1400
1200
1000
800
600
400
200
0
-200
0
5
10
15
20
25
30
35
40
X
линия регрессии X=f(Y)
линия регрессии Y=f(X)
связь ср. г. и макс. стоков
доверительные границы для Y=f(X)

рис. 5.4

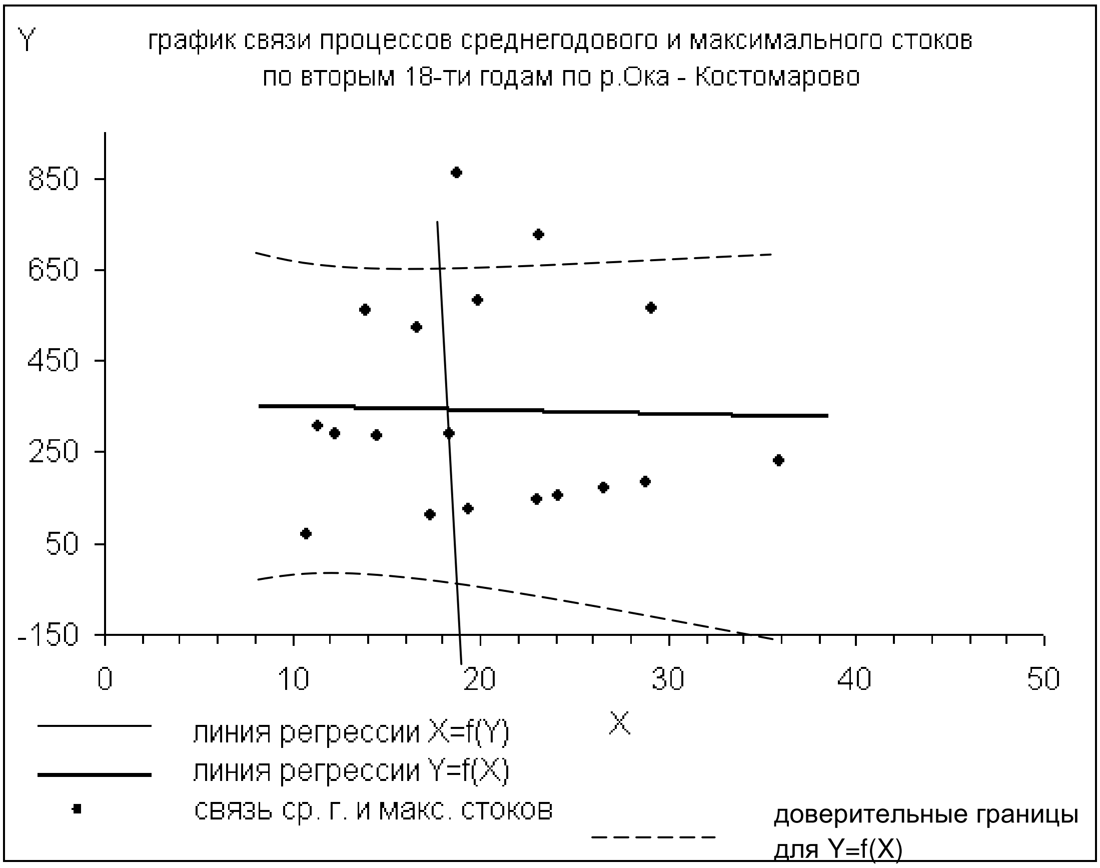
Y
график связи процессов среднегодового и максимального стоков по вторым 18-ти годам по р.Ока - Костомарово
850
650
450
250
50
-150
0
10
20
30
40
50
X
линия регрессии X=f(Y)
линия регрессии Y=f(X)
связь ср. г. и макс. стоков
доверительные границы для Y=f(X)

рис. 5.5

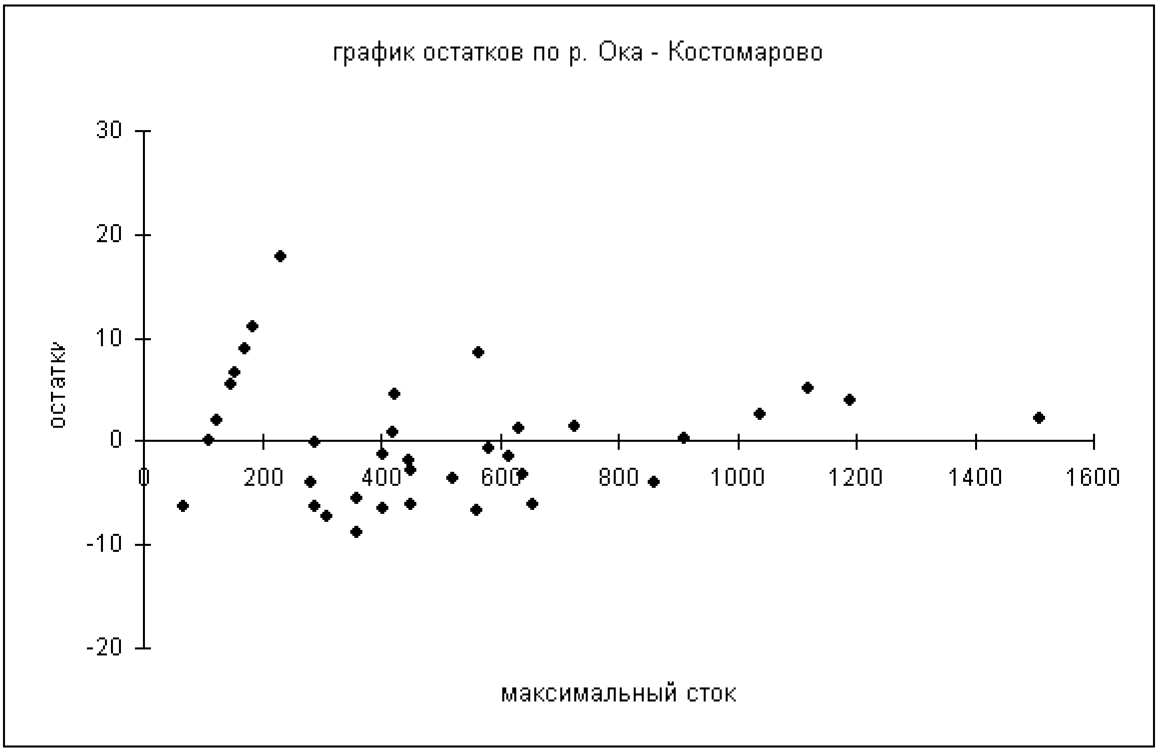
график остатков по р. Ока - Костомарово
30
20
10
0
-10
-20
остатки
0
200
400
600
800
1000
1200
1400
1600
максимальный сток

рис. 5.6

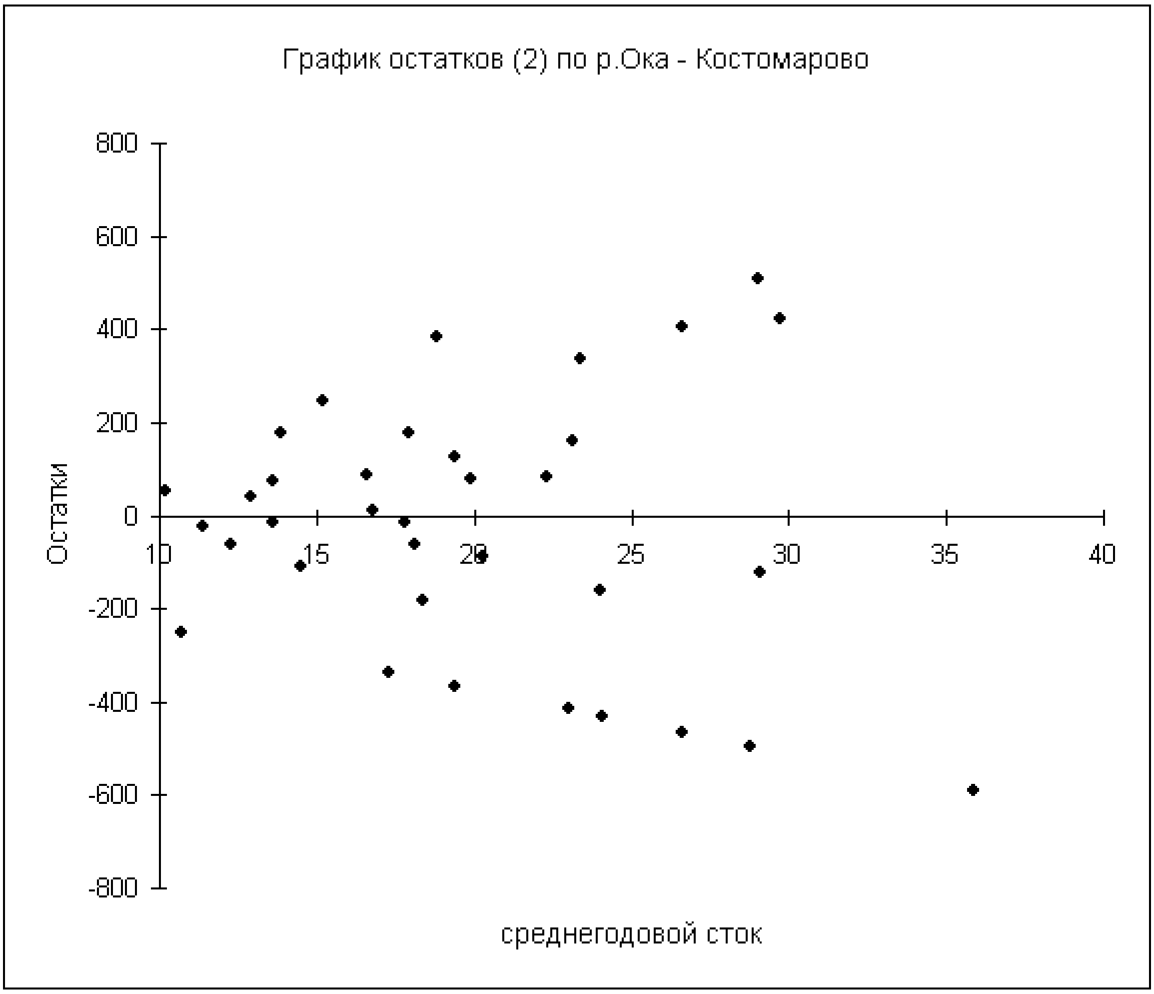
График остатков (2) по р.Ока - Костомарово
800
600
400
200
0
-200
-400
-600
-800
Остатки
10
15
20
25
30
35
40
среднегодовой сток

рис. 5.7

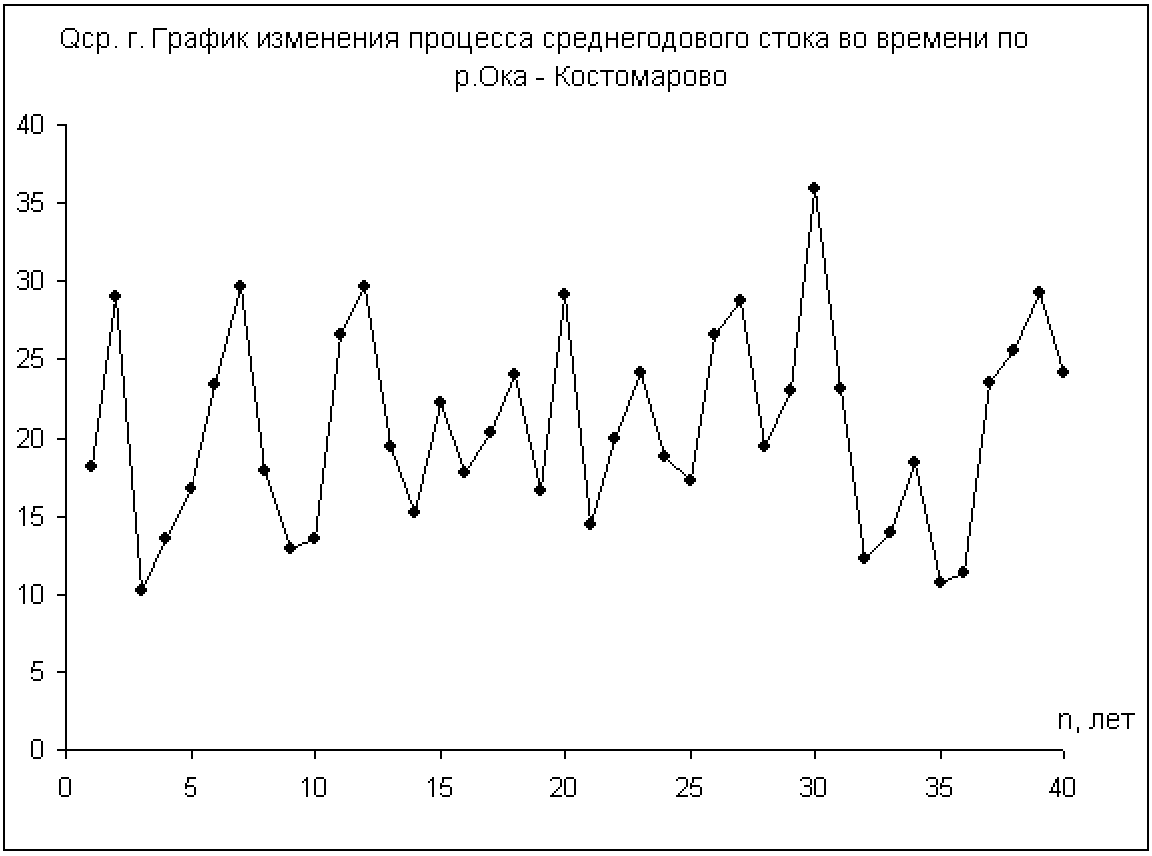
Qср. г. График изменения процесса среднегодового стока во времени по р.Ока - Костомарово
40
35
30
25
20
15
10
5
0
n, лет
0
5
10
15
20
25
30
35
40

рис. 5.8

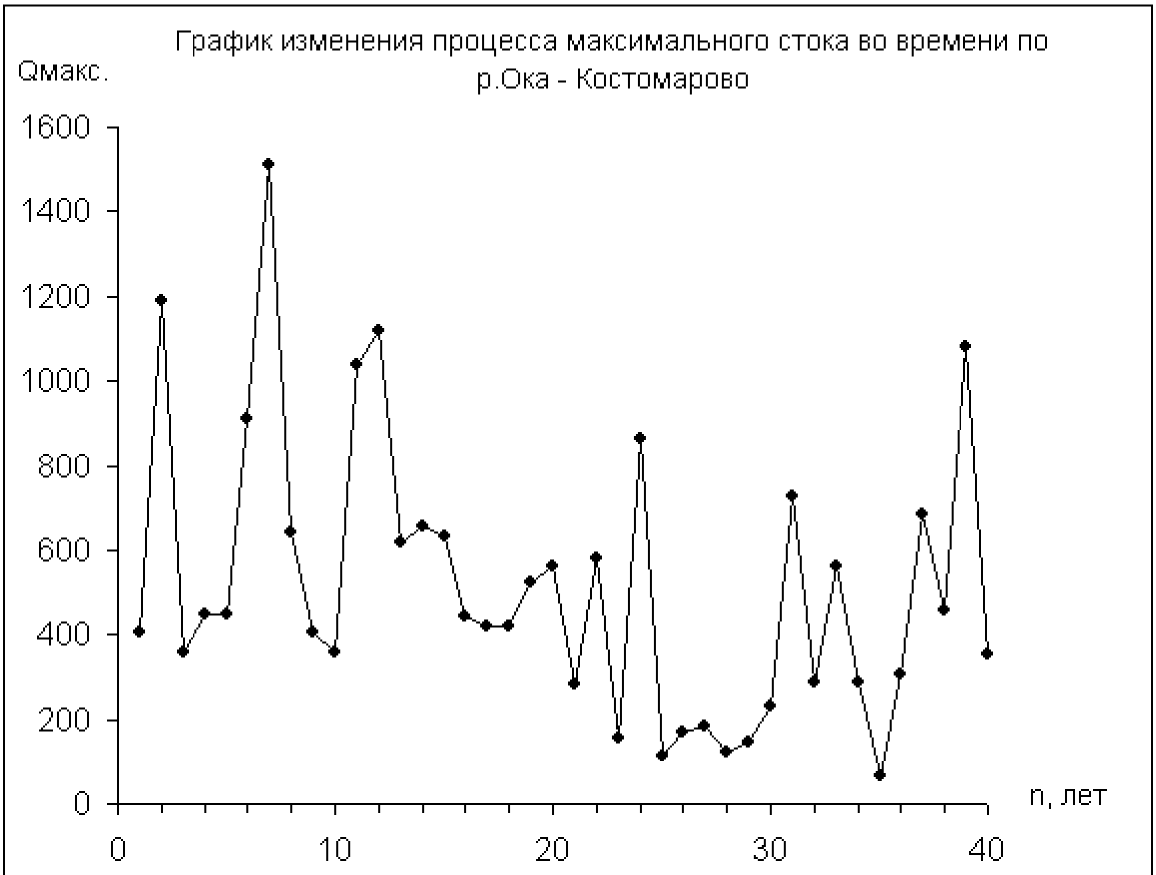
График изменения процесса максимального стока во времени по р.Ока - Костомарово
Qмакс.
1600
1400
1200
1000
800
600
400
200
0
n, лет
0
10
20
30
40

рис. 5.9