

## К ПРОБЛЕМЕ ВЕРИФИКАЦИИ ГИПОТЕЗ В ГИДРОЛОГИИ

А.П. Лепихин, Горный институт УрО РАН;

Пермский государственный национальный исследовательский университет

Рассмотрена проблема оценки истинности научных знаний, намечены три подхода к определению научных гипотез: практичность, непротиворечивость, согласованность с положениями более высокого порядка. У каждого подхода, концепции выделены свои критерии оценки. Эти три подхода, три концепции раскрыты на примере верификации трех гидрологических проблем: использование статистических функций распределения для оценки экстремальных значений, параметризация коэффициентов Шези, эффективность использования средних арифметических значений для характеристики химического состава воды рек. Показано, что перенос технологий, хорошо отработанных на задачах интерполяции, на решение экстраполяционных задач оказывается некорректным. Разнообразие расчетных соотношений по оценке коэффициента Шези обусловлено сложностью и многопараметричностью процессов формирования гидравлического сопротивления в русловых потоках, а также стохастичностью морфометрических параметров естественных водотоков. Доказана несостоятельность средних арифметических оценок для описания процессов гидрохимического режима водных объектов. Параметрические оценки гидрохимических показателей качества воды характеризуются значительной неустойчивостью статистических характеристик. Наличие случайного параметра – расхода воды в знаменателе – приближает распределение гидрохимических показателей к распределению Коши, для которого не существует не только второго, но и первого статистических моментов.

**Ключевые слова:** истинность научных знаний, критерии истинности, гидрологические объекты, статистические функции распределения, параметризация коэффициентов Шези, гидрохимический режим водных объектов.

В 2018 г. была опубликована блестящая статья академика Е.Д. Свердлова, под весьма провокационным названием «Берегись! Высокий импакт-фактор» [13]. Хотя статья посвящена проблеме воспроизводимости научных результатов, опубликованных в статьях высокорейтинговых журналов, мне представляется проблема значительно глубже: как оценивать достоверность, корректность научных знаний.

Хотелось бы рассмотреть эти вопросы на некоторых примерах из такой, сугубо прикладной, науки, как гидрология.

Чтобы показать сложность рассматриваемой проблемы, приведем несколько высказываний по данному вопросу выдающихся ученых:

– истина понимается как знание, соответствующее действительности (Аристотель);

– свойством истинности обладают все знания, которые полезны и практически применимы (У. Джеймс);

– истина как результат условного соглашения (А. Пуанкаре);

– истинность как согласованность мышления с самим собой, непротиворечивость (И. Кант, Г. Гегель).

Любая наука, в том числе и гидрология, не может развиваться без выдвижения, принятия некоторых рабочих гипотез, предположений. Вопрос стоит в оценке их эффективной истинности применительно к задачам гидрологии: Фома Аквинский утверждал, что «аргументация от великого – наихудшая». Поэтому критерии истинности должны быть сформулированы средствами самой науки. Можно наметить несколько подходов к определению «истинности» научных знаний:

– истинно то, что обеспечивает решение практических задач. В марксистской философии – это утверждение выразилось в формулировке «практика – высший критерий истинности» или, в более общем виде, истинно то, что эффективно;

– истинно то, что внутренне не противоречиво;

– истинно то, что соответствует, согласуется с положениями более высокого порядка.

Для прикладных наук, какой является гидрология, вероятно, наиболее актуален первый, утилитарный, подход. Однако он предполагает определенный релятивизм истинности, связи с относительностью, изменчивостью характера и требований к решаемым практическим задачам. Так, гелиоцентрическая система Коперника на первых этапах была значительно более громоздкой и давала существенно худшие результаты, чем геоцентрическая система Птолемея. Для задач гидрологии, гидравлики это положение наиболее наглядно проявляется на примере древнеримских акведуков. Действующие на протяжении более 2000 лет, эти сооружения, даже по современным меркам, являются весьма совершенными,

однако их создатели не имели даже весьма смутных представлений о законах механики в их современном понимании.

В практической деятельности критерий первого типа, первого подхода выражается, в первую очередь, в требовании воспроизводимости результатов независимыми исследователями. Данный вопрос очень серьезный и сложный. В 2016 г. журнал «Nature» [2] провел опрос 1500 ученых относительно воспроизводимости и невозможности результатов. На вопрос «Существует ли кризис воспроизводимости?» 52% участников ответили «да, сильный», у 38% был ответ «да, незначительный». В качестве основных факторов, обуславливающих невозможность результатов, 60% корреспондентов указывали публикацию выборок данных, 55% – плохой статистический анализ [22].

Так как критерии первого типа не могут обеспечить в достаточно полной мере оценку «истинности», для повышения эффективности оценки целесообразно использовать комплексные оценки, включающие критерии всех трех типов, подходов.

При решении прикладных задач простейшим тестом на выполнение критерия второго типа является требование соответствия размерностей правой и левой частей расчетных соотношений.

Наиболее наглядным и характерным для критерия третьего типа является требование выполнения законов сохранения вещества, энергии, импульса.

В речной гидравлике данные законы представлены в виде уравнений сохранения количества движения и неразрывности.

Рассмотрим проблему верификации «истинности» гипотез на трех конкретных примерах:

– использование статистических функций распределения для оценки экстремальных значений;

– параметризация коэффициентов Шези;

– эффективность использования средних арифметических значений для характеристики химического состава воды рек.

### Статистические методы в гидрологии

В замечательном учебнике известного советского гидролога Б.А. Аполлова «Учение о реках» [1] в главе, посвященной статистическим методам, отмечается, что данные методы «позволяют, при наличии сравнительно коротких рядов в 20–50 лет, определять такие явления, которые возможны раз в 100, 1000 лет и т.д.». На этих взглядах, представлениях было воспитано не одно поколение гидрологов. В отечественной гидрологии такие взгляды стали популярны после работ Д.Л. Соколовского [14, 15]. Формально такие оценки, безусловно, возможны. Однако для их реализации требуется выполнение ряда весьма жестких условий:

- на протяжении последующих 1000 лет гидрологические ряды должны быть стационарны относительно, по крайней мере, первых трех статистических моментов, также как и за предшествующие 20–50 лет;

- должна быть априорно известна статистическая функция распределения рассматриваемых параметров на всем диапазоне их колебаний;

- задаваемый ряд должен позволять с абсолютной надежностью оценивать все требуемые для описания задаваемой функции распределения и параметры.

При решении практических задач такие условия практически не выполнимы, и поэтому отмечаемая теоретическая возможность оценки максимальных значений с повторяемостью 1 раз в 1000 лет очень сложно реализуема.

Однако и в отечественной, и в мировой практике мы имеем блестящие примеры таких оценок. Так, храм «Покрова на Нерли», воздвигнутый по распоряжению А. Боголюбского еще в середине XII века, является не только потрясающим шедевром древнерусского зодчества, но и блестящим примером знания древними строителями гидрологии рек. Будучи воздвигнутым на искусственном холме, в междуречье рр. Клязьма и Нерли, он ни разу не затапливался внешними водами. Как пишут летописи: при самых высоких паводках льдины только удалялись о камни его фундамента, но никогда

не поступали в сам храм [6]. В Андалузии уже на протяжении более 2000 лет успешно эксплуатируются несколько акведуков. Естественно, данные объекты создавались на принципиально других принципах, чем это указывалось в учебнике [1]. На наш взгляд, увлеченность статистическими методами в гидрологии, начало которой было положено работами У. Фуллера, А. Хазена, А. Фостера [18, 20, 21], характеризует некоторую романтическую эйфорию от могущества науки, возможностей математики и математической статистики, в частности.

Интересно отметить, что в конце XIX – начале XX века теорию вероятности часто, из-за очень большого вклада отечественных ученых в ее развитие, называли русской наукой. В то же время основоположниками статистических методов в гидрологии стали американские специалисты. В 30–50-е годы XX века эти методы получили, благодаря отечественным специалистам, очень быстрое, мощное развитие, заняв лидирующие позиции в мире. А дальше весьма злую шутку сыграло доминирование одного подхода, одних взглядов, характерное для социализма. Альтернативные качественные подходы, основанные на анализе максимальных, фиксированных уровней, стали восприниматься как весьма устаревшие, находящиеся вне майнстрима.

При проектировании ДнепроГЭСа, в 1926 г., в качестве максимально возможного, полученного на основе анализа экстремальных уровней воды, был принят расход 28 000 м<sup>3</sup>/с. Однако из-за того, что уже во время строительства, в 1931 г., был зафиксирован расход в р. Днепре 26 000 м<sup>3</sup>/с, данные оценки были принципиально пересмотрены. В качестве проектного был принят максимальный расчетный расход с обеспеченностью 1:10 000–41 000 м<sup>3</sup>/с. В то же время за последующий, более чем 90-летний, период ни разу не наблюдался расход, даже близкий к 26 000 м<sup>3</sup>/с.

При проектировании Камской ГЭС, исходя из статистического анализа удлиненных рядов наблюдений, было установле-

но, что расход с обеспеченностью 1:1 000 должен составлять 18 900 м<sup>3</sup>/с, а расход с обеспеченностью 1:10 000—21 500 м<sup>3</sup>/с [12]. В то же время в 1914 г. наблюдался в р. Каме в створе г. Пермь расход 18 600 м<sup>3</sup>/с, а в 1979 г. наблюдался приток к створу Камской ГЭС 19 100 м<sup>3</sup>/с. Возникает законный вопрос: как объяснить эти факты в рамках стандартных моделей статистической гидрологии, основанных на простых цепях Маркова и одно-модальных, асимметричных, неустойчивых функциях распределения, насколько объективно их использование при решении актуальных прикладных задач?

Значительные неопределенности возникают даже при оценке эмпирической повторяемости рассматриваемых значений. А. Хазен [18], автор первых работ по применению методов математической статистики, функций распределения к анализу колебаний стока, показал, что максимальная погрешность оценки эмпирической обеспеченности должна составлять

$$\max \Delta \bar{P} \sim \frac{1}{N} \quad (1),$$

соответственно, для эффективной оценки эмпирических обеспеченностей было предложено соотношение

$$P_i \sim \frac{m_i - K}{N}, \text{ при этом } K=0,5.$$

Совершенно очевидно, что величина  $K$  не может быть в общем случае постоянной величиной и она должна зависеть от характера распределения рассматриваемых параметров и изменяться в интервале  $0 \leq K \leq 1$ . Формально, без априорного задания функций распределения, эмпирическая погрешность оценивается как (1).

При характерной продолжительности рассматриваемого ряда  $N \sim 100$  лет, а при расчетах в середине прошлого века  $N \sim 10 - 30$  лет, совершенно проблематичными оказываются оценки с повторяемостью меньше  $P_i \leq 0,01$ .

В то же время многие технологические объекты требуют значительно большей надежности функционирования. Для корректных оценок с вероятностью  $< 1/N$  требуется привлечение дополнительной

информации, внешней по отношению к рассматриваемым рядам наблюдений.

В настоящее время данная задача в традиционной постановке решается достаточно просто: выбирается некоторое аналитическое распределение, допускающее изменение рассматриваемой переменной в требуемом диапазоне  $\min X \leq X \leq \max X$ , объявленное «теоретическим», оцениваются параметры этого распределения на основе данных наблюдений, затем на этой основе строится его экстраполяция в область экстремально малых повторяемостей.

При этом для упрощения выбора «теоретических» функций и последующей экстраполяции, как правило, принимается  $\max X \rightarrow \infty$ ,  $\min X \rightarrow 0$ . Такая схема весьма традиционна и была предложена еще А. Фостером [20]. Однако для обеспечения ее корректности:

- рассматриваемая выборка должна быть стационарной, по крайней мере относительно первых трех статистических моментов;

- обоснование «теоретичности» принимаемой кривой распределения должно быть получено вне рамок представленных статистических выборок.

Стационарность гидрологических рядов в условиях значительной динамики климатических эпох может носить только весьма локальный характер с характерной продолжительностью  $T_{Ni}$  лет. К сожалению, исходя из глобальных климатических моделей, нет корректных оценок стационарности конкретных гидрологических систем. Поэтому все оценки статистических параметров распределения носят априорно условный характер, полученные для условия «локальной» стационарности рассматриваемых рядов.

В настоящее время нет каких-либо серьезных исследований, направленных на обоснование статистических функций распределения конкретного водотока как элемента глобальной климатической системы. Поэтому выбор «теоретического» распределения определяется предпочтениями автора расчетов, удобством его проведения.

При этом совершенно открытым оста-

ется вопрос: формирование экстремальных значений малой повторяемости происходит под воздействием тех же самых механизмов, что и большой повторяемости. Если средняя часть распределения достаточно хорошо описывается большинством функций распределения, в том числе и нормальным, то вопрос о распределении в экстремальной области открыт, и вряд ли может быть решен в рамках традиционных подходов. При этом одним из эффективных критериев выбора «теоретического» распределения может быть требование его устойчивости. Это условие резко сокращает количество возможных претендентов на рассмотрение в качестве «теоретического» распределения.

В первую очередь, в качестве такого распределения может рассматриваться усеченное нормальное распределение. Однако в силу своей симметричности данное распределение не может использоваться для описания расходов воды в области экстремальных значений.

В этой области могут, по-видимому, в определенной мере использоваться степенные распределения, которые также устойчивы в асимптотике. Так как эти распределения характеризуются очень «тяжелыми хвостами», должен существовать некий критерий «отсечения хвостов», по аналогии с природными механизмами, ограничивающими максимально возможные расходы воды.

Таким образом, как уже упоминалось, вопрос выбора закона распределения для экстремальных значений остается открыт.

### Задачи речной гидравлики

Консул Ю. Фрондин, куратор водоснабжения г. Рима, писал: «Акведуки являются главным символом величия Рима, их нельзя даже сравнивать с бесполезными Египетскими пирамидами и прочими праздными постройками Греции» [19].

Возникают простейшие вопросы: как осуществлялось проектирование римских акведуков и, тем более, еще более совершенного акведука Кумбе–Майо в Перу? Как, исходя из требуемого количества по-

ставки воды, проводилась оценка параметров этих акведуков? Расчет даже простейшего канала требует весьма громоздких, многочисленных вычислений. А. Шези только в середине XVIII века при расчете параметров Иветтского канала, создаваемого для улучшения водоснабжения г. Парижа, предложил схему расчетов, получившую позднее известность как уравнение Шези. Проблема принципиально усложняется тем, что в древнем Риме использовалась непозиционная Римская система записи чисел, в которой выполнение даже простейших вычислений весьма затруднено. Вероятно, в основе их оценки лежали принципиально другие, не количественные, а качественные принципы.

Само уравнение Шези занимает центральное место в речной гидравлике, устанавливая связь между морфометрическими и динамическими параметрами потока. В общем случае оно отражает первый закон Ньютона, когда силы тяжести уравновешиваются силами гидравлического сопротивления.

По оценкам проф. П.Ф. Горбачева [4], уже в середине 30-х гг. XX века было известно и предлагалось для практического использования свыше 300 различных соотношений для оценки коэффициента Шези. В настоящее время таких расчетных соотношений значительно больше. Так как первые, чисто теоретические работы по оценке данного коэффициента были выполнены относительно недавно [16, 17], то возникает вопрос: как, почему, оперируя чисто эмпирическими данными, предлагается различными авторами такое большое количество расчетных соотношений для оценки одного и того же коэффициента. Причин, по-видимому, несколько:

- имеет место несколько различных механизмов формирования гидравлического сопротивления в водотоках;

- сложность и неоднозначность переноса экспериментальных данных, получаемых в лабораторных лотках и каналах, на естественные водотоки из-за инер-

ционности рассматриваемых гидравлических процессов;

– при определенном диапазоне значений различные аналитические зависимости могут давать близкие результаты, так, например, степенная и логарифмическая, что весьма существенно при построении эмпирических зависимостей.

Рассмотрим более подробно эти аспекты. Так как механизмы многих процессов, обуславливающих гидравлические сопротивления в русловых потоках, недостаточно изучены, на начальной стадии такой анализ следует выполнить на основе анализа размерности. При проведении анализа размерностей центральное место занимает объективность и корректность задания определяющих параметров.

При этом в качестве параметров, в достаточной мере описывающих гидравлическое сопротивление потока, можно принять:

$$\lambda \approx \Phi(V, R, g, \nu, d_{\Delta}, \rho_s, \rho, i, B, \omega_s, S), \quad (2)$$

где  $V$  – средняя скорость потока, м/с;

$R$  – гидравлический радиус, м;

$g$  – ускорение свободного падения, м/с<sup>2</sup>;

$\nu$  – кинематическая вязкость воды, м<sup>2</sup>/с;

$d_{\Delta}$  – характерный размер частиц донных отложений, м;

$\rho_s$  – плотность материала частиц, кг/м<sup>3</sup>;

$\rho$  – плотность воды, кг/м<sup>3</sup>;

$i$  – гидравлический уклон потока;

$B$  – характерная ширина потока, м;

$\omega_s$  – гидравлическая крупность взвешенных частиц, м/с;

$S$  – объемная концентрация взвешенных частиц.

Согласно П-теореме [2] при выборе 11 определяющих параметров и при наличии 3 независимых размерностей они должны быть сгруппированы в 9 безразмерных комплексов:

$$\begin{aligned} & \lambda(V, R, g, \nu, d_{\Delta}, \rho_s, \rho, i, B, \omega, S) \equiv \\ & \equiv \lambda \left( \text{Re}, \text{Re}_*, \Delta\rho, i, \text{Fr}, \text{Fr}_*, \frac{d_{\Delta}}{R}, \frac{B}{R}, K_0 \right), \quad (3) \end{aligned}$$

где  $V_* = (g \cdot R \cdot i)^{1/2}$  – динамическая скорость потока;

$\text{Re} = (R \cdot V) / \nu$  – глобальное число Рейнольдса, представляющее собой отноше-

ние сил инерции к силам вязкости, построенное по параметрам, отнесенным ко всему сечению потока;

$\text{Re}_* = (d_{\Delta} \cdot V_*) / \nu$  – локальное число Рейнольдса, характеризующее отношение сил инерции к силам вязкости, построенное по параметрам, характеризующим придонный слой;

$\text{Fr} = V^2 / (g \cdot R)$  – глобальное число Фруда, характеризующее отношение силы инерции к силе тяжести, построенное по параметрам относительно всего сечения водотока;

$\text{Fr}_* = V_*^2 / (g \cdot d_{\Delta})$  – локальное число Фруда, характеризующее отношение силы инерции и силы тяжести, построенное по параметрам, характеризующим придонный слой;

$d_{\Delta} / R$  – относительная шероховатость русла, в ряде литературных источников рассматривается как число Штриклера;

$\Delta\rho = (\rho_s - \rho) / \rho$  – относительная плотность частиц;

$B / R$  – параметр формы поперечного сечения русла;

$K_0 = g \cdot H \cdot \omega \cdot S / V_*^3$  – число Колмогорова [2, 5], характеризующее долю турбулентной энергии, расходуемой на взвешивание частиц.

Данные показатели тесно связаны с критериями, используемыми в гидродинамике и динамике речных потоков, такими как  $\text{Re} \cdot \text{Re}_*^2 / \text{Fr}_* = \text{Ga}$  – число Галилея, характеризует отношение сил тяжести и вязкости, при этом  $\text{Ga} \cdot \Delta\rho = \text{Ar}$  – число Архимеда.

В то же время  $\text{Ar}^{1/3} = D_*$  – число Боннефилле, широко используется в динамике русловых потоков для характеристики устойчивости частиц на дне потока,  $\text{Fr}_* / \Delta\rho = 1 / \theta$ ,  $\theta$  – число Шильдса, характеризует отношение сдвигающих и удерживающих сил частиц на дне потока, при

$$\text{этом } D_* = \left( \frac{\text{Re}_*}{\sqrt{\theta}} \right)^{2/3}.$$

При этом может реализоваться несколько механизмов гидравлического сопротивления. В первую очередь можно выделить следующие механизмы формирования гидравлического сопротивления:

- на гладкой стенке (вязкий подслои),
- на неподвижной шероховатости,
- на грядовых структурах,
- на когерентных структурах,
- на транспортировке взвешенных наносов.

В качестве первого приближения, как правило, принимается аддитивность их взаимодействия  $\lambda_{нач} = \sum_{i=1}^m \lambda_i$ . [5].

Однако такая конструкция для решения практических задач малоэффективна. Из-за большого количества входящих в нее параметров ее калибровка весьма затруднена. В то же время степенная аппроксимация, как следствие неполной автомодельности, приемлема только для описания асимптотик.

В связи с этим особое значение имеет анализ отдельных частных случаев.

В достаточно общем виде модели расчета гидравлического сопротивления можно для отдельных частных случаев представить следующим образом (таблица).

Анализ коэффициентов гидравлического сопротивления для гидравлически гладких русел дается в работе [9].

При решении практических задач наибольший интерес представляют потоки с зернистой шероховатостью. Большинство эмпирических расчетных соотношений для оценки коэффициента Шези

было получено для данного режима. В последние годы значительный интерес проявляется к оценке роли и параметризации грядового сопротивления [5, 10].

В зависимости от особенностей конкретного гидравлического режима водотока доминирующими могут быть различные механизмы и, соответственно, различные расчетные соотношения. В целом значительное количество их возможных комбинаторных сочетаний стало одним из основных факторов, обусловивших наличие очень большого количества расчетных соотношений для оценки коэффициента Шези.

Из-за сложности проведения натуральных экспериментов на естественных водотоках: реках, ручьях – как правило, большинство экспериментов по оценке параметров гидравлического сопротивления, транспортирующей способности потоков проводят на специальных потоках, каналах. Однако установленные в этих экспериментах зависимости очень сложно, с большими погрешностями работают в реальных потоках. Возникает естественный вопрос, в чем дело? Так как принципы гидродинамики должны работать вне зависимости от того, естественный или искусственный исследуется объект.

Принципиальными особенностями ес-

**Классификация моделей расчета коэффициента гидравлического сопротивления в русловых потоках**

№ п/п	Режим течения и характер сопротивления	Область реализации	Определяющие параметры, расчетные соотношения
1	Ламинарные течения	$Re/Re_{кр} \ll 1$	$\lambda = f(Re), \lambda = 64/Re$
2	Турбулентные течения	$Re/Re_{кр} \gg 1$	$\lambda = f(Re, \Delta/R, (R \cdot i)/\Delta)$
2.1	Гидравлически гладкие русла	$Re_* / Re_{*кр} \ll 1$	$\lambda = f_2(Re), \lambda_{тр} = \frac{2 * e}{(\ln(Re))^{2 * e}}$
2.2	Зернистая шероховатость	$Re_{с*} / Re_{*кр} \gg 1$ $F_{г*} / F_{г*кр} \ll 1$	$\lambda = f_3(\Delta/R, Re)$ , при этом, как правило, используются степенная или логарифмическая аппроксимации. $\frac{1}{\sqrt{\lambda_{зер}}} = A_2 (R / \Delta)^{1/\gamma}$ , $C \sim K^* (R/\Delta)^{1/\gamma}$ – Маннинга-Штриклера $\gamma = 6$ – Маннинг, $\gamma = 5$ – Форхгеймер, $4 \leq \gamma \leq 8$ – Павловский. $K \sim 20 \div 23$ .
2.3	Грядовое сопротивление	$Re_* / Re_{*кр} \gg 1$ $F_{г*} / F_{г*кр} \geq 1$	$\lambda_{тр} = f(\psi), \psi \approx (i \cdot R)/d, \lambda_{тр} = \psi \exp(-k \cdot \psi)$

тественных водотоков являются сложность, стохастичность их геометрии. Достаточно широко обсуждается фрактальность как береговых полос, так и характерной протяженности рек [8].

В условиях существенной геометрической неоднородности русел естественных водотоков принципиальное значение приобретает корректность оценок динамических характеристик потока на основе их связи с морфометрическими параметрами.

Традиционно такие показатели, как скорость потока, его транспортирующая способность рассматриваются как локальные характеристики, относящиеся к некоторому расчетному створу и определяемые морфометрическими параметрами, задаваемыми в этом створе. Такой подход, в строгой постановке, не корректен, так как он не учитывает инерционность потока. Нетрудно построить простейший масштаб инерционного потока по скорости  $L_V \sim H/\lambda$ , где  $H$  – характерная глубина потока,  $\lambda$  – коэффициент гидравлического сопротивления.

Если характерные линейные масштабы изменения морфометрических характеристик потока  $L_{\mu d} = \left( \frac{1}{\mu} \frac{\partial \mu}{\partial X} \right)^{-1}$ ,

где  $\mu$  – морфометрическая характеристика потока, то нетрудно видеть, что при  $L_{\mu} \gg L_V$ , что характерно для искусственных водотоков: лотков, каналов – наличие инерционности потока никак не сказывается на механизме параметризации расчетных параметров. Однако если  $L_{\mu} \sim L_V$ , то скорость потока в рассматриваемом створе определяется морфометрическими характеристиками потока не в данном створе, а осредненными за весь предшествующий инерционный интервал, а сами динамические характеристики потока не могут рассматриваться как локальные.

При этом морфометрические неоднородности с характерным размером  $L_{\mu} \ll L_V$  будут потоком восприниматься как дополнительный «шум», вызывающий дополнительное гидравлическое сопротивление.

Естественные водотоки характеризуются, как правило, очень широким спектром изменения морфометрических показателей. Поэтому весьма серьезной причиной, обуславливающей сложность переноса зависимостей, полученных на экспериментальных лотках, каналах, на естественные водотоки, является некорректность оценки интегральных динамических показателей на основе локальных морфометрических характеристик.

ются, как правило, очень широким спектром изменения морфометрических показателей. Поэтому весьма серьезной причиной, обуславливающей сложность переноса зависимостей, полученных на экспериментальных лотках, каналах, на естественные водотоки, является некорректность оценки интегральных динамических показателей на основе локальных морфометрических характеристик.

### Особенности применения статистических методов в гидрохимии поверхностных водных объектов

Использование статистических методов при решении практических задач априорно предполагает, что погрешность оцениваемого параметра  $\sigma$  уменьшается с увеличением объема выборки, т.е. оценка является состоятельной. На этом постулате строится требование к увеличению объема выборки с условием повышения надежности оценок. Однако это условие справедливо для эффективных оценок. Нормальное распределение и близкие к нему статистические распределения обеспечивают состоятельность оценок статистических моментов.

Вывод из центральной предельной теоремы (ЦПТ) что, когда распределение показателей формируется в результате сложения множества однородных факторов, оно асимптотически приближается к нормальному, породил уверенность, что оценка всех расчетных параметров состоятельна. Однако это далеко не так.

При этом существуют и другие распределения, в частности распределение Коши [7], для которого не существует устойчивых не только второго и третьего статистических моментов, но и первого. При этом необходимо подчеркнуть, что данные распределения не экзотика, и это необходимо учитывать при решении практических задач. Так, при решении прикладных задач гидрохимии представляется приемлемым отказ от нормальности распределения [11], однако совершенно недопустимой считается статистическая неустойчивость оцениваемых пара-



метров, при этом данные распределения очень тесно связаны между собой. Если имеется два независимых процесса  $\xi_1(t)$  и  $\xi_2(t)$ , описываемых нормальным распределением, то распределение  $Z(t)=\xi_1(t)/\xi_2(t)$  будет асимптотически приближаться к распределению Коши при  $Cv_1 \rightarrow \infty$  и  $Cv_2 \rightarrow \infty$  [3]. Это положение имеет принципиальное значение для обработки и анализа информации.

Рассмотрим данную ситуацию на примере обработки гидрохимической информации. Исходя из простейших балансовых соотношений имеем

$$C_i(t) = \frac{q_i(t)}{Q(t)}, \quad (4)$$

где  $C_i(t)$  – осреднённая по сечению потока концентрация  $i$ -го загрязняющего вещества;  $q_i(t)$  – интенсивность поступления  $i$ -го поллютанта в водоток.

Если  $q_i(t)$  также зависит от  $Q(t)$ , что характерно для эндогенных поллютантов, т.е.  $q_i(Q(t)) \approx q_{i0} \cdot \psi(Q(t))$ , соответственно,

$$C_i(t) = \frac{q_{i0} \cdot \psi(Q(t))}{Q(t)} \quad \text{или, принимая}$$

$$f(Q) = \frac{\psi(Q)}{Q},$$

$$C(Q) = q \cdot f(Q). \quad (5)$$

Уравнения типа (4)-(5) весьма широко используются в инженерной гидрохимии. Для построения эффективной технологии их статистической обработки принципиальное значение имеет характер их функций распределения. При этом статистические распределения химических показателей качества воды, как правило, не имеют нормального распределения [11]. Так как содержание гидрохимических показателей в воде очень тесно связано с динамикой расходов воды, то представляет непосредственный интерес анализ функций распределения, исходя из характера распределений  $Q$ , а в более общем случае и распределения  $q_i(t)$ .

Очень часто рассматривается простейшая ситуация, когда содержание рассматриваемого поллютанта определяется только расходом водотока, в этом случае не-

трудно получить функцию распределения рассматриваемого химического показателя, исходя из функций распределений расходов воды  $Q$  в самом водотоке –  $P_Q(Q)$ . Так, в этом случае функция распределения  $P(C)$  для концентрации поллютанта в воде  $C$  будет иметь следующий вид [7]:

$$P_c(C) = P_Q(f^{-1}(C)) \cdot \left| \frac{df^{-1}(C)}{dC} \right|.$$

Нетрудно видеть, что плотность распределения химических показателей качества воды будет иметь в этом случае нормальное распределение, только если нормальным будет распределение  $Q(t)$ , а связь между концентрацией  $C$  и расходом воды  $Q$  – линейной.

Если расход сброса  $q$  независим от  $Q$  и является константой, то получается  $C(t) \sim q/Q(t)$ . Такого типа зависимость достаточно часто встречается и используется в прикладных гидрологических исследованиях. В этом случае, соответственно, будем иметь

$$P_c(C) = P_Q(q/C) \cdot \frac{1}{C^2}. \quad (6)$$

Более сложная ситуация, когда  $q(t)$  также является случайной величиной с интегральной функцией распределения  $F_q$ . В этом случае функция распределения должна описываться соотношением

$$F_{q/Q}(C) = \int_0^{\infty} P_Q(X_1) F_q(C \cdot X_1) dX_1. \quad (7)$$

При этом весьма актуальна задача, какими свойствами должно обладать распределение  $P(C)$ , описываемое соотношениями (6)–(7).

К сожалению, даже в простейшем случае, когда распределения процессов  $q(t)$  и  $Q(t)$  нормальны, функция распределения  $F_{q/Q}(C)$  имеет весьма громоздкий вид и затруднительна для анализа. В то же время при  $Cv_q \gg 0$  и  $Cv_Q \gg 0$  она асимптотически приближается, как уже отмечалось, к распределению Коши.

На рисунке представлена зависимость относительной среднеквадратичной погрешности  $\delta_x$  оценки среднеарифметических значений от коэффициента вариации  $Cv$  процесса  $Q(t)$ , полученная на базе чис-

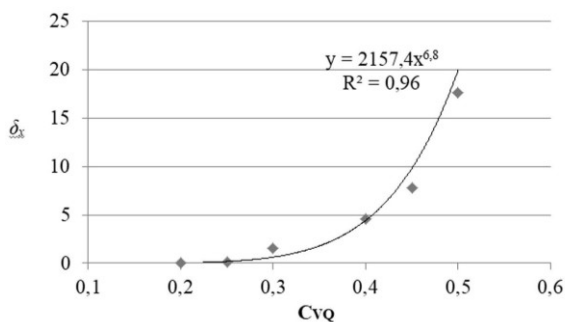


Рис. Зависимость относительной среднеквадратичной погрешности  $\delta_x$  оценки среднеарифметических значений в зависимости от коэффициента вариации  $C_v$  расхода воды водотока-приемника  $Q(t)$

ленных экспериментов на основе метода Монте-Карло. Как видно из рисунка, наблюдается очень резкий рост статистической неустойчивости при  $C_v Q \geq 0,25$ .

Таким образом, в зонах активного техногенеза, где содержание поллютантов двойного генезиса описывается соотношением (4), параметрические оценки в виде средних арифметических, среднеквадратичных значений неэффективны и даже могут быть несостоятельными.

Определяющим критерием эффективности моделей в прикладных областях знаний, в том числе гидрологии, является «близость» расчетных и фактических значений.

Наиболее отработана процедура оценки близости расчетных и фактических значений в задачах прогнозирования на основе использования критерия Нэша – Сатклифа  $NSE$ , систематической ошибки  $BIAS$ .

При этом критерий  $NSE$  оценивается

$$\text{как } NSE = 1 - \frac{\sum_{i=1}^N (T_{fi} - T_{si})^2}{\sum_{i=1}^N (T_{fi} - T_f)^2},$$

где  $T_{fi}$ ,  $T_{si}$  –соответственно фактические и расчетные значения расчетного параметра  $C$ , а  $BIAS$  определяется как

$$BIAS = \frac{\bar{T}_f - \bar{T}_s}{\bar{T}_f} \cdot 100\%.$$

Рассматриваемые критерии строятся на основе стьюдентовской схемы оценки значимости отклонений средних значений.

Основными недостатками данной схе-

мы являются:

- несостоятельность среднеарифметических оценок для рассматриваемых выборок. Как было показано выше, средние арифметические оценки далеко не всегда состоятельны. Если эти оценки несостоятельны, то средняя квадратичная погрешность оценки  $\sigma_{\bar{x}}$  не уменьшается с увеличением объема выборки  $N$  и операции осреднения не эффективны.

- равноценность и равнозначность для всех рассматриваемых значений решаемой задачи. В то же время при решении конкретных прикладных задач требуется, в первую очередь, оценка значимости лимитирующих величин.

Если речь идет о прогнозе уровня воды в водном объекте, то определенное значение имеют эффективность прогноза и надежность оценок рассматриваемых лимитирующих значений не за весь период наблюдений, а за какой-то временной интервал. Аналогично в задачах расчета распределения по акватории водного объекта наибольший интерес представляет распределение рассматриваемых показателей не по акватории водного объекта, а только в зонах лимитирующего водопользования, например, в некоторых контрольных створах или даже точках у водозаборов.

Если значимость расчетных величин неоднозначна для решения прикладных задач, то мера эффективности расчетных моделей должна строиться, исходя из их надежности и эффективности при решении прикладных задач.

В качестве такой оценки могут рассматриваться абсолютные или относительные отклонения:

$$\Delta c \sim \max_{i \in N} |C_{Ri} - C_{\Phi i}|,$$

$$\sigma_c \sim \max_{i \in N} \frac{|C_{Ri} - C_{\Phi i}|}{C_{\Phi i}},$$

где  $C_{Ri}$ ,  $C_{\Phi i}$  – соответственно расчетные фактические значения лимитирующих показателей состояния водного объекта.

Так как  $\Delta C_i$ ,  $\sigma_{C_i}$  в общем случае представляют собой случайные величины, то при достаточно больших значе-

ниях  $N$  в качестве эффективного критерия целесообразно рассматривать соответственно квантили  $\Delta_{Cr}$ ,  $\sigma_{Cr}$  порядка  $p$ , принимая их за нормативные.

### Выводы

1. Перенос технологий, хорошо отрабатанных на задачах интерполяции, на решение экстраполяционных задач оказывается, на примере оценки экстремальных значений гидрологических параметров, некорректным. Для обеспечения их эффективного решения необходимо привлечение дополнительной информации, введение критериев, обеспечивающих внутреннюю непротиворечивость используемых методов расчетов, исходя из особенностей формирования стока, необходимо обеспечение устойчивости их функций распределения.

2. Многопараметричность процессов формирования гидравлического сопротивления в русловых потоках, сложность,

стохастичность морфометрических параметров естественных водотоков являются основными факторами, обуславливающими разнообразие расчетных соотношений по оценке коэффициента Шези.

3. Неустойчивость статистических характеристик, достаточно характерная для параметрических оценок гидрохимических показателей качества воды, в значительной мере, обуславливается несостоятельностью данных оценок для описания подобных процессов. Наличие случайного параметра – расхода воды в знаменателе – приближает распределение гидрохимических показателей к распределению Коши, для которого не существует не только второго, но и первого статистического момента.

4. В целом критерии «близости» расчетных и наблюдаемых показателей должны строиться с учетом особенностей рассматриваемой задачи с выделением лимитирующих показателей.

### Библиографический список

1. Аполлов Б.А. Учение о реках. – Изд-во. МГУ, 1952. – 512 с.
2. Баренблатт Г.И. Автомодельные явления – анализ размерностей и скейлинг / пер. с англ.: учеб. пособие. – Долгопрудный МО: Изд. Дом Интеллект. – 2009. – 216 с.
3. Вадзинский Р.Н. Справочник по вероятностным распределениям. – СПб.: Наука, 2001. – 294 с.
4. Горбачев П.Ф. Формулы скорости течения жидкости. – М. – Л., ОНТИ. Гл. ред. строительной лит, 1936. – 171 с.
5. Гришанин К.В. Гидравлическое сопротивление естественных русел. – СПб.: Гидрометеоздат, 1992. – 184 с.
6. Заграевский С.В. К вопросу о реконструкции и датировке церкви Покрова на Нерли // Новые исследования памятников архитектуры Владимиро-Суздальского музея-заповедника. – М.: Алев-В, 2008. – С. 129–146.
7. Кобзарь А.И. Прикладная математическая статистика. – М.: Физматлит, 2006. – 813 с.
8. Лепихин А.П. К анализу структуры гидрографических сетей // Географический вестник Geographical bulletin. – 2017. – №3 (42. doi 10.17072/2079-7877-2017-3-53-60).
9. Лепихин А.П. К оценке коэффициента гидравлического сопротивления в гладких трубах // Вычислительная механика сплошных сред. – 2015. – Т. 8. – № 4. С. 369–375.
10. Лепихин А.П., Богомолов А.В. История установления и современные представления об основной закономерности равномерного установившегося течения в водотоках (к 240-летию формулы А. Шези) // Водное хозяйство России: проблемы, технологии, управление. – 2015. – № 6. – С. 76–92.
11. Лепихин А.П., Возняк А.А. Статистические функции распределения гидрохимических показателей качества воды поверхностных водных объектов // Водное хозяйство России. – 2012. – № 4. – С. 21–32.
12. Машкович Л.А. План эксплуатации водного хозяйства Камской ГЭС // Материалы совещания по вопросам эксплуатации Камского водохранилища. – Пермь, 1959. – С. 1–14.
13. Свердлов Е.Д. Берегись! Высокий импакт-фактор // Вестник Российской академии наук. – 2018. – Т. 88. – № 6. – С. 531–538.
14. Соколовский Д.Л. Нормы максимального стока весенних паводков рек СССР. – Л.–М.: Гидрометеоздат, 1937.
15. Соколовский Д.Л. Применение кривых распределения к установлению вероятностных колебаний годового стока рек в Европейской части СССР. – М.: Гостехиздат, 1930. – 78с.

16. Gioia G., Bombardelli F.A. Scaling und Similarity in Rough Channel Flows // Physical Review Letters. – Vol. 88. – № 1. – P. 014501-1-014501-4.
17. Gioia G., Chakraborty P. Turbulent Friction in Rough Pipes and the Energy Spectrum of the Phenomenological Theory Physical Review Letters // PRL 96, 044502 (2006). – P. 044502-1-044502-4.
18. Hazen A. Discussion of ‘The probable variations in yearly runoff as determined from a study of California streams’ by L. Standish Hall. – ASCE Trans., 84, 1921. – P. 214–222.
19. Herschel C. Frontinus and the water supply system of the city of Rome. – Boston, 1899. – P. 106.
20. Foster H.A. Theoretical frequency curves. – ASCE Trans., 87. 1924. – P. 142–203.
21. Fuller W.E. Flood flows. – ASCE Trans., 77, 1914. – P. 567–617.
22. Challenges in irreproducible research URL: <http://www.nature.com/news/reproducibility-1.17552>.

**TO THE PROBLEM OF VERIFICATION  
OF HYPOTHESES IN HYDROLOGY**

A.P. Lepikhin

*Mining Institute UB RAS  
Perm State National Research University*

The article considers the problem of assessing the truth of scientific knowledge and outlines three approaches to determining the truth of scientific hypotheses: practicality, consistency, and coherence with the provisions of a higher order. Each approach or concept has its own evaluation criteria. These three approaches or concepts are revealed on the example of verification of three hydrological problems: the use of statistical distribution functions for estimating extreme values, the parameterization of the Chézy coefficients and the efficiency of using arithmetic averages for characterizing the chemical composition of rivers. The study shows that the transfer of technologies well-tested on interpolation tasks to solving extrapolation problems is incorrect. The variety of calculated ratios for estimating the Chézy coefficient resulted from the complexity and multiparameter nature of hydraulic resistance formation processes in channel flows, as well as the stochasticity of the morphometric parameters of natural watercourses. The article also proves the inconsistency of arithmetic average evaluations for describing the processes of hydrochemical regime of water bodies. Parametric estimates of hydrochemical indicators of water quality are characterized by significant instability of statistical characteristics. The presence of a random parameter, that is the water discharge in the denominator, makes the distribution of hydrochemical parameters closer to the Cauchy distribution for which not only the second, but also the first statistical moment do not exist.

*Keywords:* the truth of scientific knowledge, truth criteria, hydrological objects, statistical distribution functions, parameterization of the Chezy coefficients, hydrochemical regime of water bodies.

**Сведения об авторе**

*Лепихин Анатолий Павлович*, доктор географических наук, профессор, заведующий лабораторией проблем гидрологии суши, Горный институт УрО РАН – филиал Пермского федерального исследовательского центра УрО РАН (ГИ УрО РАН), 614007, г. Пермь, ул. Сибирская, 78А; профессор кафедры гидрологии и охраны водных ресурсов, Пермский государственный национальный исследовательский университет (ПГНИУ), 614990, г. Пермь, ул. Букирева, 15; e-mail: [lepikhin49@mail.ru](mailto:lepikhin49@mail.ru)

*Материал поступил в редакцию 15.04.2019 г.*