

## Применение методов обучения искусственных нейронных сетей при прогнозировании высших уровней воды на примере рек Двинско-Печорского бассейнового округа

А. Э. Сумачев\*, Л. С. Банщикова\*\*,  
С. А. Грига\*

*Рассмотрена реализация нейросетевых методик прогнозирования высших уровней воды за период весеннего ледохода на примере Сухоны, Северной Двины и Печоры. Все рассмотренные нейросетевые методики показали высокую эффективность по критериям, рекомендованным Гидрометцентром России, и по качеству выпускаемых прогнозов превзошли регрессионные зависимости. Стандартная ошибка прогнозирования при использовании метода обучения искусственных нейронных сетей снижается примерно на 10–20% по сравнению с регрессионными зависимостями.*

**Ключевые слова:** нейронные сети, прогноз, Двинско-Печорский бассейновый округ, высшие уровни воды.

**DOI:** 10.52002/0130-2906-2024-4-104-115

### Введение

Изменения характеристик ледового режима являются результатом сложного вариативного сочетания метеорологических и гидрологических условий и антропогенного воздействия. Существующие методики прогнозов этих характеристик не учитывают этих изменений и тем самым не всегда обеспечивают номинальное качество выпускаемых прогнозов. Причины понижения качества выпускаемых прогнозов сводятся к использованию нестационарных предикторов, общее число которых может быть значительно больше рекомендуемого при данной длительности рядов, и, как следствие, ведут к занижению стандартной ошибки прогнозирования. При этом ряд систем уравнений содержит предикторы, которые не всегда могут быть получены непосредственными измерениями на наблюдательной сети Росгидромета, что связано не только с сокращением числа пунктов наблюдений, но и с изменением программ наблюдений.

Все вышеупомянутое определяет необходимость разработки новых, более простых с точки зрения используемого фактического материала методик прогнозирования, но основанных на возможностях обучения искус-

\* Санкт-Петербургский государственный университет.

\*\*Государственный гидрологический институт; e-mail: BانشchikovaLS@yandex.ru (Банщикова Любовь Святославовна).

ственных нейронных сетей и обеспечивающих более высокую оправдываемость прогнозов при той же или большей заблаговременности.

Вопросу прогнозирования высших уровней воды, в том числе имеющих заторный генезис, посвящено большое количество работ, наиболее современной из которых является монография [2]. В этой работе приведено множество регрессионных зависимостей, в том числе внедренных на сети Росгидромета, однако к настоящему времени некоторые зависимости можно признать устаревшими в связи с сокращением программы наблюдений и обязательным требованием обновления регрессионных коэффициентов. Для дальнейшего развития идей прогнозирования, высказанных в работе [2], с учетом современного состояния гидрометеорологической сети наблюдений и развития вычислительной техники в данной работе авторы предложили подход к прогнозированию высших уровней воды, основанный на классификации участков рек [10] и возможностях обучения искусственных нейронных сетей.

Согласно классификации участков рек [10] по степени стохастичности условий формирования высшего уровня воды за период весеннего ледохода, основанной на многофакторном анализе ледового режима рек Двинско-Печорского бассейнового округа [1—5, 11, 12], и оценке степени влияния климатических изменений на ледовый и гидрологический режим рек района, были выделены три основных типа формирования высших за период весеннего ледохода уровней воды:

- первый тип — участки рек с динамическим типом формирования высших уровней воды за период весеннего ледохода;
- второй тип — участки рек со стохастико-динамическим типом формирования высших уровней воды за период весеннего ледохода;
- третий тип — участки рек со стохастическим типом формирования высших уровней за период весеннего ледохода.

Данная классификация была разработана на основе оценки коэффициентов корреляции между высшим предледоходным и высшим за период весеннего ледохода уровнем воды, дальнейшие исследования при этом показали сохранение данных связей для всех характерных уровней воды, в том числе высших за год. Классификация имеет физико-географическую интерпретацию: участки рек, отнесенные к первому типу, являются сильно зарегулированными. В пределах рассматриваемого района — это в первую очередь озерное регулирование, которое приводит к тому, что уровни формируются относительно плавно, а величина характерных уровней всецело зависит от предшествующих уровней. К данному типу в пределах изучаемого района относится р. Онега. Если рассматривать первый тип применительно к другим районам, то к данному типу могут быть отнесены реки с преимущественно ледниковым питанием, а также отдельные бесприточные участки крупных рек, на величину высших уровней воды которых уже не могут повлиять относительно маловодные притоки.

Второй тип является промежуточным. Строго говоря, данный тип должен быть наиболее распространен, однако в пределах изучаемого региона к нему можно отнести лишь устьевые участки крупных рек. Формирование характерных уровней воды на данных участках, с одной стороны, мало зависит от интенсивности снеготаяния, так как волна половодья на

них имеет уже достаточно распластанную форму, что придает условиям формирования стока сравнительно детерминированный вид. В то же время этой регулирующей способности недостаточно, чтобы с большой точностью предсказывать высшие уровни воды, что приводит к необходимости учета дополнительных предикторов.

Третий тип является полной противоположностью первого. Для участков рек, отнесенных к данному типу, характерны наличие существенной боковой приточности, асинхронность вскрытия, труднопредсказуемая ледовая ситуация, что приводит к необходимости использования при прогнозировании относительно большого количества предикторов, учитывающих разные факторы формирования характерных уровней воды. К данному типу в пределах исследуемого района относятся реки Северная Двина и Печора.

Для разных типов целесообразно применять те или иные наиболее подходящие методы прогнозирования и критерии оценки качества выпускаемых прогнозов. Таким образом, разработка методик прогнозирования сводится к выбору оптимального предиктивного состава. Разработка методик прогнозирования для рек второго и третьего типов имеет ряд сложностей, так как возникает необходимость учета различных факторов формирования высших уровней воды. При этом возможны варианты, когда добавление предикторов с вышележащих створов приводит к существенному уменьшению заблаговременности (вплоть до отрицательной).

Целью работы являются обобщение многолетнего авторского опыта прогнозирования характерных уровней воды на реках Двинско-Печорского округа с применением метода обучения искусственных нейронных сетей и выработка рекомендаций по использованию данного метода. Рассмотрены несколько регрессионных зависимостей для прогнозирования высших уровней воды за год и за период весеннего ледохода на участках рек, отнесенных к разным типам. В качестве исходных данных при разработке регрессионных моделей прогнозирования использованы режимные данные сети Росгидромета с 1950 по 2018 г. [7, 8] и оперативная информация за последующие годы. На их основе разработаны нейросетевые модели прогнозирования, которые были апробированы в течение последних лет на постах сети Росгидромета. В заключение приведены основные рекомендации по разработке методик прогнозирования характерных уровней воды в пределах исследуемого региона. В работе рассмотрены методика прогнозирования высшего уровня воды за год в г. Великий Устюг и на устьевом участке р. Печора, методика прогнозирования высших уровней воды за период весеннего ледохода на р. Онега и на Северной Двине в районе г. Котлас. При оценке качества выпускаемых прогнозов применены критерии, рекомендованные Гидрометцентром России [6]. Обучение искусственных нейронных сетей производилось в программном продукте Statistica 12, в котором реализован алгоритм подбора наилучшей архитектуры нейронной сети [14].

### **Особенности разработки методик прогнозирования на участках рек, отнесенных к первому типу**

Для первого типа рек характерно динамическое развитие волны половодья, вскрытия и ледохода. При этом можно отметить, что на всех участ-

Таблица 1

**Оценка качества выпускаемых прогнозов на независимом материале по сравнению с ориентированием на среднее значение высшего уровня за период ледохода**

Характеристика	Гидрологический пост		
	Поной — Краснощелье	Онега — Турчасово	Онега — Порог
$h$ , см	75	152	84
$M$ , см	407	643	66
$S$ , см	51	92	68
$S/h$	0,68	0,61	0,79

*Примечание.* — оценка среднеквадратического отклонения прогнозируемой величины;  $M$  — оценка математического ожидания прогнозируемой величины;  $S$  — среднеквадратическая ошибка прогнозирования при использовании формулы (1).

ках рек, относящихся к первому типу, волна половодья и вскрытия развивается подобно, что предполагает возможность использования универсальной прогностической формулы. Для разработки формулы данные о высших предледоходных и ледоходных уровнях воды необходимо привести к стандартному виду. Для этого от каждого значения уровня воды отнимается норма, после чего полученная разность делится на среднеквадратическое отклонение. Стандартизованный ряд при этом имеет среднее значение, равное 0, и среднеквадратическое отклонение, равное 1.

По стандартизованным данным для участков рек, относящихся к первому типу, была разработана единая формула для прогнозирования высшего уровня воды за период весеннего ледохода в зависимости от высшего предледоходного уровня:

$$H_x^{л\delta x} = 0,64H_x^{высш. пр\delta л\delta x} - \sigma_x^{л\delta x} m_x^{л\delta x}, \tag{1}$$

где  $H_x^{л\delta x}$  — высший уровень воды за период весеннего ледохода в пункте  $x$ ;  $H_x^{высш. пр\delta л\delta x}$  — стандартизованный высший предледоходный уровень воды в пункте  $x$ ;  $\sigma_x^{л\delta x}$  — среднеквадратическое отклонение высшего уровня воды за период весеннего ледохода в пункте  $x$ ;  $m_x^{л\delta x}$  — среднее значение высшего уровня воды за период весеннего ледохода в пункте  $x$ . Заблаговременность такого прогноза вариативна и в среднем составляет 4 дня. Оценка качества выпускаемых прогнозов на независимом материале дана в табл. 1.

Анализируя данные, приведенные в табл. 1, можно прийти к выводу о достаточно высокой эффективности предложенной методики, так как отношение  $S/h$  меньше 0,80 [6]. Приведенную методику прогнозирования высшего уровня воды за период весеннего ледохода можно применять и для прогнозирования других уровенных характеристик ледового режима. Например, она была успешно верифицирована для прогнозирования высших уровней воды в начале ледостава на гидрологических постах Каневка и Турчасово по данным об уровнях воды, приходящимся на день появления первичных ледовых явлений, без изменения весовых коэффициентов. В обоих случаях можно говорить о высоком качестве выпускаемых прогнозов, так как отношения  $S/h$  меньше 0,80 [6].

Таким образом, можно заключить, что оценка качества выпускаемых прогнозов показала эффективность предложенной методики. Улучшение качества выпускаемых прогнозов при использовании других методов в данном случае невозможно, поскольку берется всего один стандартизованный предиктор. Данная методика может быть использована для прогнозирования высшего уровня воды за период весеннего ледохода и других уровней характеристик на реках Кольского п-ова, на Онеге и на любых участках рек, характеризующихся высокими коэффициентами корреляции между высшими предледоходными и ледоходными уровнями воды.

**Особенности разработки методик прогнозирования на участках рек, отнесенных ко второму типу**

В работе [12] приведены зависимости для прогнозирования высших уровней за период весеннего ледохода на всем протяжении р. Печора до с. Оксина. В частности, отмечены следующие выявленные зависимости для прогнозирования высшего уровня воды за период весеннего ледохода: для с. Усть-Цильма — от высшего уровня ледохода на р. Ижма в с. Трусово и разницы в толщине льда в с. Усть-Цильма и р. п. Щельяюр и для с. Оксина — от высшего уровня ледохода в с. Ермицы и высшего предледоходного уровня на посту в с. Оксина. Также отмечается возможность улучшения качества выпускаемых прогнозов при использовании метода обучения искусственных нейронных сетей. В работе [13] представлены верифицированные нейросетевые и регрессионные методики прогнозирования для с. Оксина. Высший уровень воды вычисляется с помощью уравнения множественной линейной регрессии:

$$H_{\text{лдохО}} = 0,29H_{\text{предлдох}} + 0,59H_{\text{лдохЕ}} - 27, \tag{2}$$

где  $H_{\text{лдохО}}$  — высший уровень воды за период весеннего ледохода в районе с. Оксина;  $H_{\text{предлдох}}$  — высший предледоходный уровень воды в с. Оксина;  $H_{\text{лдохЕ}}$  — высший уровень воды за период весеннего ледохода в с. Ермицы.

Среднеквадратическая ошибка приведенной модели составляет 48 см, отношение  $S/\sigma = 0,53$ , средняя заблаговременность — 4 дня. На основе представленной модели возможно разработать нейросетевую методику, применение которой позволит значительно уменьшить ошибки. В частности, в работе [12] приведены сведения об опыте применения нейросетевой методики MLP 2-8-1 (сеть многослойного персептрона, содержащая два предиктора — входа, восемь скрытых нейронов и один выход), данные оценки качества выпускаемых прогнозов приведены в табл. 2.

Таблица 2

**Оценка качества выпускаемых прогнозов MLP 2-8-1 по сравнению с инерционным прогнозом**

Выборка	$\sigma, \text{ см}$	MLP 2-8-1	
		$S, \text{ см}$	$S/\sigma$
Полная	82	30	0,37
Обучающая	85	30	0,35
Тестовая	78	30	0,38

На основе приведенных сведений и анализа связей уровней можно заключить, что подобная зависимость характерна и для высших уровней воды за год на всех гидрологических постах на участке реки от с. Оксино до д. Осколково. Это предопределяет возможность разработки системы прогнозирования высшего уровня воды за год для данного участка реки, для чего целесообразно привести высший уровень воды за год в с. Оксино к стандартному виду, а предикторы перевести в метры над нулем поста. Для приведения предиктанта к стандартному виду необходимо из каждого значения прогнозируемой величины вычесть ее среднее значение и поделить на среднеквадратическое отклонение. Таким образом, при использовании фактических данных о двух предикторах прогнозируется отклонение прогностической величины от нормы в единицах среднеквадратического отклонения. Для перевода прогнозируемой величины из стандартизованного в фактический вид прогностическое значение необходимо домножить на среднеквадратическое отклонение конкретного поста и прибавить среднее многолетнее значение. Конечное прогностическое уравнение выглядит следующим образом:

$$H_{в.г} = (0,658H_{лдхЕ} + 0,141H_{прдлхО} - 7) + H_{ср.в.г}, \quad (3)$$

где  $H_{в.г}$  — высший уровень за год на конкретном посту (Оксино, Нарьян-Мар, Андег, Осколково);  $H_{прдлхО}$  — высший предледоходный уровень воды в с. Оксино;  $H_{лдхЕ}$  — высший уровень за период весеннего ледохода в с. Ермицы; — среднеквадратическое отклонение высшего уровня воды за год на конкретном посту (Оксино, Нарьян-Мар, Андег, Осколково);  $H_{ср.в.г}$  — средний за многолетний период высший уровень на конкретном посту (Оксино, Нарьян-Мар, Андег, Осколково). Значения и  $H_{ср.в.г}$  в этих пунктах следующие:

Пункт	с. Оксино	г. Нарьян-Мар	д. Андег	д. Осколково
$H_{ср.в.г}, см$	702	526	552	573
$\sigma, см$	80	77	60	59

Отметим стационарность данных величин по критериям Стьюдента и Фишера при заданном уровне значимости 5%.

Качество выпускаемых прогнозов для всех пунктов наблюдений является удовлетворительным (среднее значение  $S/\sigma = 0,72$ ), что позволяет составлять краткосрочные прогнозы со средней заблаговременностью от 5 до 7 дней. Качество может быть улучшено при использовании искусственных нейронных сетей, для чего в программном продукте Statistica 12 [14] по тем же предикторам произведено обучение искусственных нейронных сетей типа MLP [14]. Процедура обучения подразумевает минимизацию средней квадратической ошибки на обучающем подмножестве, к которому относится 70% лет наблюдений на посту Оксино. Процедура валидации подразумевает проверку прогнозов на независимом материале, к которому относится 30% лет наблюдений на посту Оксино и все фактические наблюдения на постах в г. Нарьян-Мар, деревнях Андег и Осколково. Таким образом, можно получить надежную модель для прогнозирования высшего за год уровня воды. Лучшая из обученных нейронных сетей имеет один скрытый слой с четырьмя нейронами. В качестве активационных функций

использованы экспонента и гиперболический тангенс. Можно заключить, что при использовании искусственных нейронных сетей качество выпускаемых прогнозов значительно улучшается, а среднее значение  $S/ = 0,62$ .

### **Особенности разработки методик прогнозирования на участках рек, отнесенных к третьему типу**

Разработана методика прогнозирования высшего за год уровня воды р. Сухона в районе г. Великий Устюг. При расположении Великого Устюга в верхней части водосборного бассейна Северной Двины доля высших уровней, сформированных на чистой воде, может оказаться значительной. Так, за периоды 1951—2013 и 2017—2019 гг. высший за год уровень р. Сухона в районе г. Великий Устюг формировался при ледовых явлениях (включая заторы) в 65% случаев, на свободном русле — в 35% случаев.

При прогнозировании высшего за год уровня воды в реке используют ряды большей однородности, увеличивается качество прогнозирования, задача упрощается. Модель становится более физически обоснованной, поскольку учитывает суть опасного гидрологического явления, так как при ледоходах не всегда формируется высший уровень — затопления не происходит.

Таким образом, в качестве прогнозируемой величины (предиктанта) был выбран высший за год уровень р. Сухона в районе г. Великий Устюг. В то же время значительная часть высших уровней воды за год формируется при заторах льда, основными факторами возникновения которых являются следующие [2]:

- мощность скопления шуги в русле, у верхней границы которого весной формируется затор;
- расход (уровень) воды у перемещающейся вниз по течению реки кромки ледяного покрова, который является характеристикой сил, развиваемых водным потоком в зоне торошения льда;
- максимальная за зиму толщина ледяного покрова на участке образования затора ниже по течению;
- температура воздуха в период весеннего таяния льда при образовании затора.

Набор предикторов определялся исходя из взаимосвязи физических процессов с учетом коэффициентов корреляции этих предикторов с высшими за год уровнями воды Сухоны в районе Великого Устюга, их статистической значимости и физической обоснованности. В связи с современными изменениями климата в формировании высшего заторного уровня все большую роль играет зашугованность речного русла. Именно от этого фактора могут зависеть высшие и низшие уровни в начале ледостава. Однако установлено, что низший среднесуточный уровень за декабрь образует более тесную связь с предиктантом. Фактор заторообразования представляет из себя сложную для мониторинга характеристику водного потока. Вместо него предлагается использовать среднесуточный уровень воды в первый день весеннего ледохода в Великом Устюге. Температура воздуха в период таяния характеризует увеличение уровня воды, но высший уровень может формироваться и на свободной воде, в условиях почти полного отсутствия снежного покрова. Так как данный участок реки является приточным, то третьим предиктором выбран среднесуточный уровень

в Кичменгском Городке на дату начала ледохода в Великом Устюге, характеризующий условия развития половодья наиболее крупного притока.

Таким образом, для решения регрессионной задачи наиболее целесообразно использование следующих предикторов:

- уровень воды в первый день весеннего ледохода в Великом Устюге (ВУ);
- низший уровень воды в декабре в Великом Устюге;
- уровень в Кичменгском Городке (КЧГ) на дату начала ледохода в Великом Устюге.

Подобная комбинация предикторов также позволяет определять заблаговременность прогноза только по первому из них.

После определения оптимального состава предикторов была построена регрессионная модель, при этом выборка 1950—2014 гг. разбивалась на зависимую и независимую части в соотношении 70 к 30. Получены следующие результаты регрессионного анализа: множественный коэффициент корреляции  $R$  — 0,70;  $R$ -квадрат — 0,49; нормированный  $R$ -квадрат — 0,46; средняя квадратическая ошибка — 105 см (число наблюдений — 45). Значение средней квадратической ошибки регрессионной модели составило 105 см. Дисперсия была объяснена на 46%, о чем говорит значение “нормированный  $R$ -квадрат”.

В результате по полученным данным было построено следующее регрессионное уравнение:

$$H_{\max \text{ ВУ}} = 0,34H_{\text{ВУ 1й лдхд}} + 1,07H_{\text{мин дек ВУ}} + 0,36H_{\text{КЧГ на лдхд ВУ}} + 245, \quad (4)$$

где  $H_{\max \text{ ВУ}}$  — высший за год уровень воды в Великом Устюге;  $H_{\text{ВУ 1й лдхд}}$  — уровень воды в первый день весеннего ледохода в Великом Устюге;  $H_{\text{мин дек ВУ}}$  — низший уровень воды в декабре в Великом Устюге;  $H_{\text{КЧГ на лдхд ВУ}}$  — уровень в Кичменгском Городке на дату начала ледохода в Великом Устюге.

Среднее арифметическое значение заблаговременности составило 6 сут. Первое модальное значение приходится на первые сутки. Оно связано с уровнями, формируемыми при ледовых явлениях. На значение среднего арифметического приходится еще одна мода, которая связана с уровнями, формируемыми на открытой воде. Заблаговременность прогнозов варьируется от краткосрочных (1 сут) до среднесрочных (42 сут).

Проверка качества данной методики, основанной на использовании регрессионного уравнения, произведена на независимом материале (тестовая выборка). На тестовом подмножестве уравнение показало среднюю квадратическую ошибку, равную 80 см. Результаты по обучающей выборке удовлетворяют обоим критериям качества, рекомендованными в работе [6], по тестовой — только первому критерию (для долгосрочных прогнозов).

Для улучшения качества прогнозирования использовали многослойные искусственные нейронные сети, обученные в программном продукте Statistica 12 [14]. В качестве исходных данных был взят тот же набор, что и для регрессионной модели с аналогичным разбиением на обучающее и тестовое подмножества. Были обучены 20 000 сетей с выводом пяти наилучших (наилучшими признавались нейронные сети с наименьшей средней квадратической ошибкой). Далее отбиралась одна искусственная ней-



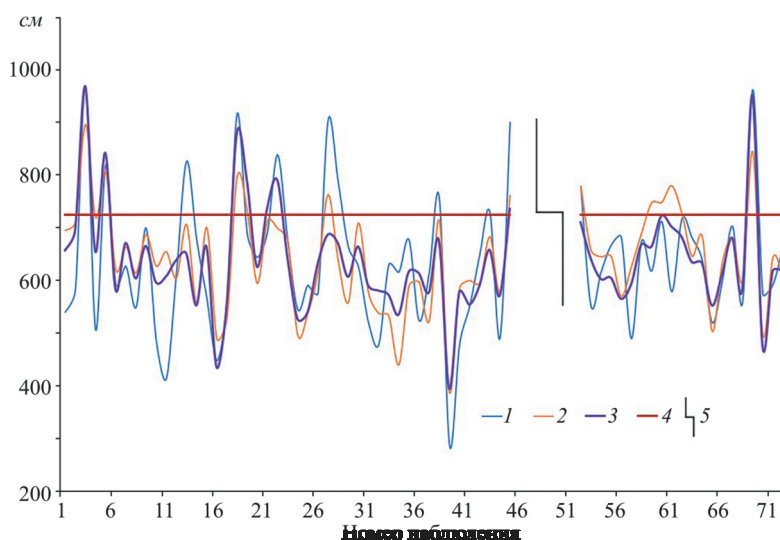


Рис. 1. Результаты прогнозирования высшего за год уровня воды р. Сухона у г. Великий Устюг с помощью нейронной сети и регрессионной модели, разделенные по выборкам: 1 — фактические значения; 2 — моделируемые регрессионным уравнением; 3 — моделируемые искусственной нейронной сетью; 4 — уровень начала опасного гидрологического явления (около 725 см [9]); 5 — знак-разделитель обучающей и тестовой выборки.

ронная сеть с наименьшей ошибкой, имеющая наименьшее число скрытых нейронов. Ей оказалась искусственная нейронная сеть с рабочим названием MLP 3-3-1. Архитектурные особенности сети состоят в следующем: число скрытых нейронов — 3; входящая функция активации — логистическая; выходящая функция активации — синусоидальная [9].

Результаты прогнозирования высшего за год уровня воды в Великом Устюге с помощью лучшей нейронной сети в сравнении с фактическими значениями и прогнозами по регрессионной модели для всего рассматриваемого периода (обучающая + тестовая выборки) представлены на рис. 1.

Модель соответствует обоим критериям качества Гидрометцентра России по обучающей, тестовой и полной выборкам, оправдываемость для них составила 67% (табл. 3).

Абсолютные ошибки искусственной нейронной сети имеют незначительную тенденцию к уменьшению. Из чего следует, что модель MLP 3-3-1 устойчива во времени, а качество выпускаемых прогнозов не падает; анализ ошибок показал отсутствие систематической ошибки прогнозирования. Таким образом, обученная искусственная нейронная сеть удовлетворяет всем критериям качества, рекомендуемым Гидрометцентром России. По этим показателям нейронная сеть значительно превзошла регрессионное уравнение. Также существенно увеличилась оправдываемость прогнозов: 67% у нейронной сети против 53—57% у регрессионного уравнения. Уменьшилась средняя квадратическая ошибка по сравнению с ошибкой ныне используемой на сети Росгидромета методики [2].

Для гидрологического поста в районе г. Котлас разработана методика прогнозирования высшего уровня воды за период весеннего ледохода в за-

Таблица 3

**Параметры и характеристики качества прогнозов высшего за год уровня воды в Великом Устюге с помощью регрессионного уравнения и модели MLP 3-3-1**

Параметр	Уравнение множественной регрессии			MLP 3-3-1		
	Выборка			Выборка		
	обучающая	тестовая	полная	обучающая	тестовая	полная
$S, см$	101	80	95	89	60	82
$S/$	0,71	0,75	0,71	0,63	0,57	0,61
$S/$	0,78	1,05	0,83	0,69	0,79	0,71
$N_{оп}/N$	24/45	12/21	36/66	30/45	14/21	44/66
$P, \%$	57	53	55	67	67	67

*Примечание.*  $S$  — средняя квадратическая ошибка;  $S/$  и  $S/$  — критерии Гидрометцентра России;  $N_{оп}$  — число оправдавшихся прогнозов;  $N$  — общее число прогнозов;  $P$  — оправдываемость.

висимости от уровня воды, приходящегося на первый день весеннего ледохода в г. Тотьма на р. Сухона, и низшего суточного уровня воды в г. Котлас за октябрь — март. Ошибка  $S$  прогностической зависимости при использовании уравнения множественной линейной регрессии составляет 65 см, что дает отношение  $S/ = 0,62$ , а отношение  $S/ = 0,72$ , что в целом удовлетворяет требованиям гидрометеорологической сети при известной простоте линейных моделей. Однако качество прогнозов может быть значительно улучшено при использовании нейросетевых зависимостей. Обученная в рамках данной работы нейронная сеть MLP 2-5-1, имеющая один скрытый слой с пятью нейронами и использующая гиперболический тангенс в качестве активационной функции, превосходит линейное уравнение по качеству выпускаемых прогнозов (табл. 4, рис. 2). В целом качество данной модели выше, чем модели, предложенной в работе [2], при этом можно отметить доступность и обоснованность использованных предикторов. Так, уровень воды, приходящийся на первый день ледохода в г. Тотьма, является интегральной характеристикой поступающих объемов воды, а низший уровень за октябрь — март характеризует дефицит воды в речной системе.

Таблица 4

**Оценка качества прогнозов высшего уровня воды за период весеннего ледохода для Северной Двины в районе г. Котлас, выпускаемых с помощью нейронной сети MLP 2-5-1**

Выборка	$S, см$	$S/$	$S/$	$N_{оп}$	$P, \%$
Полная	55	0,53	0,61	49	74
Обучающая	57	0,55	0,63	35	73
Тестовая	49	0,47	0,54	14	78

*Примечание.* Допустимая ошибка прогноза равна 61 см.

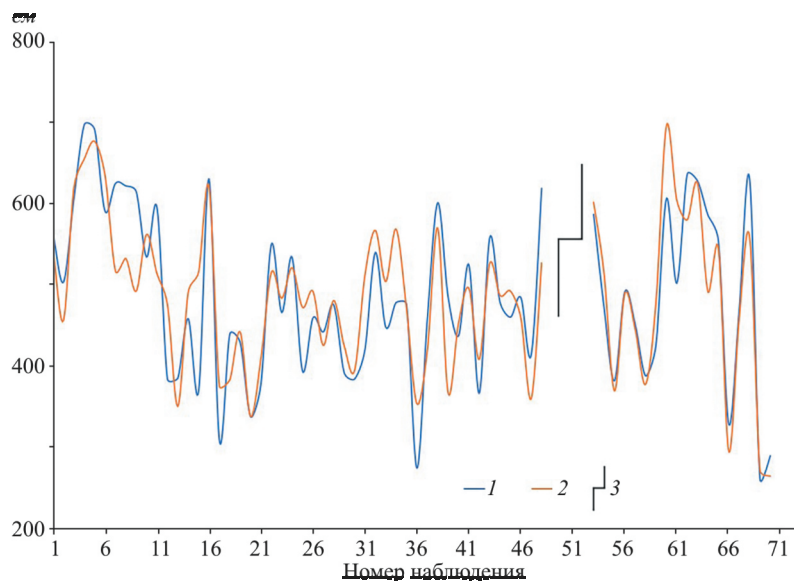


Рис. 2. Результаты прогнозирования высшего уровня воды за период весеннего ледохода в створе гидрологического поста г. Котлас: 1 — фактические значения; 2 — моделируемые с использованием нейросетевых моделей MLP 2-5-1; обучающая и тестовая выборка разделены условной линией 3.

Ошибки данной модели имеют нормальное распределение, статистически значимые тренды отсутствуют, что свидетельствует о возможности внедрения данной модели в практику прогнозирования высшего уровня воды за период весеннего ледохода на сети Росгидромета.

Таким образом, на участках рек, отнесенных ко второму и третьему типам, применение методов обучения искусственных нейронных сетей по сравнению с методами линейной регрессии для прогнозирования высшего уровня воды за период весеннего ледохода и за год дает уменьшение средней квадратической ошибки прогнозирования на 10—20% и увеличение оправдываемости прогнозов. Разработанные модели прогнозирования высшего уровня воды за период весеннего ледохода обеспечивают высокое качество выпускаемых прогнозов. В рамках сотрудничества с Вологодским ЦГМС полученные прогностические модели были апробированы на сети Росгидромета, были даны прогнозы высших за период весеннего ледохода уровней воды в районе г. Котлас и высшего за год уровня воды в районе г. Великий Устюг на 2022 и 2023 гг., ошибки которых не превысили допустимые.

В целом принципы, которым необходимо следовать при разработке методик прогнозирования, можно свести к следующим.

Необходимо создавать универсальные методики прогнозирования. Эта задача решается путем стандартизации прогностического уравнения. Данный пункт особенно актуален для участков рек, отнесенных к первому и третьему типам.

Необходимо использование предикторов, которые измеряются непосредственно на сети Росгидромета и обеспечивают приемлемое качество выпускаемых прогнозов при оптимальной заблаговременности.

Используемые методы прогнозирования должны быть минималистичны, но в то же время обеспечивать достаточно высокое качество выпускаемых прогнозов. В частности, при использовании метода обучения искусственных нейронных сетей следует отдавать предпочтение моделям с наиболее простой архитектурой и минимальным количеством скрытых нейронов, что обеспечивает высокую устойчивость нейросетевой модели. В случаях, когда нейронные сети не могут обеспечить значимого уменьшения средней квадратической ошибки прогнозирования по сравнению с более простыми методами, следует отдавать предпочтение наиболее простым моделям.

Заблаговременность и качество выпускаемых прогнозов при использовании наиболее строгих критериев качества должны быть не ниже, чем у существующих методик.

## Литература

- 1. Агафонова С. А., Фролова Н. Л.** Особенности ледового режима рек бассейна Северной Двины. — Водные ресурсы, 2007, т. 34, № 2, с. 141—149.
- 2. Бузин В. А.** Зажоры и заторы льда на реках России. — СПб, ГГИ, 2015, 240 с.
- 3. Вуглинский В. С.** Оценка изменений характеристик ледового режима водных объектов для различных регионов страны в современных климатических условиях. — Вестник Санкт-Петербургского университета. Науки о Земле, 2014, № 3, с. 32—45.
- 4. Двоеглазова К. С., Шелутко В. А., Горшкова Н. И.** Оценка термического и ледового режима рек бассейна Северной Двины. — Гидрометеорология и экология: достижения и перспективы развития, 2019, с. 274—277.
- 5. Козлов Д. В. и др.** Опасные ледовые явления на реках и водохранилищах России. — М., Изд-во РГАУ-МСХА, 2015, 348 с.
- 6. Попов Е. Г.** Основы гидрологических прогнозов. — Л., Гидрометеиздат, 1968, 294 с.
- 7. Свидетельство** о государственной регистрации базы данных № 2021621902 Российская Федерация. Характеристики ледового режима рек Двинско-Печорского бассейнового округа, № 2021621776, заявл. 30.08.2021, опублик. 07.09.2021. /А. Э. Сумачев, Л. С. Банщикова; заявитель Федеральное государственное бюджетное учреждение “Государственный гидрологический институт”.
- 8. Свидетельство** о государственной регистрации базы данных № 2021621912 Российская Федерация. Характеристики ледового режима рек Кольского полуострова, № 2021621808, заявл. 31.08.2021, опублик. 09.09.2021. /А. Э. Сумачев, Л. С. Банщикова; заявитель Федеральное государственное бюджетное учреждение “Государственный гидрологический институт”.
- 9. Свидетельство** о государственной регистрации базы данных № 2022622815 Российская Федерация. База данных элементов гидрологического режима для прогнозирования высшего за год уровня воды реки Сухоны в районе города Великий Устюг (VU MLP 3-3-1), № 2022622780, заявл. 31.10.2022, опублик. 09.11.2022.
- 10. Сумачев А. Э.** Совершенствование методов прогнозирования характеристик ледового режима рек бассейнов Баренцева и Белого морей. /Диссертация на соискание степени кандидата технических наук. — 2022, 173 с.
- 11. Сумачев А. Э.** Оценка возможности прогнозирования максимального ледоходного уровня воды с использованием максимальной толщины льда на примере рек бассейна Северной Двины. — Ученые записки Российского государственного гидрометеорологического университета, 2019, № 54, с. 48—57; doi: 10.33933/2074-2762-2019-54-48-57.
- 12. Сумачев А. Э., Банщикова Л. С.** Ледовый режим реки Печора и особенности прогнозирования высшего уровня ледохода. — Гидрометеорология и экология, 2020, № 61, с. 446—459; doi: 10.33933/2074-2762-2020-61-446-459.
- 13. Сумачев А. Э., Банщикова Л. С.** Ледовый режим реки Печоры в современных климатических условиях и принципы прогнозирования высшего уровня воды за период весеннего ледохода. — Успехи современного естествознания, 2021, № 10, с. 75—80; doi: 10.17513/use.37701.
- 14. Электронный учебник** по статистике; <http://statsoft.ru/home/textbook/default.htm> (дата обращения 27.07.2023).