

Волчек А. А., Костюк Д. А., Петров Д. О., Шешко Н. Н.

МЕТОД ПРОГНОЗИРОВАНИЯ ПОЛОВОДИЙ НА ОСНОВЕ МНОГОФАКТОРНОГО НЕЙРОСЕТЕВОГО АНАЛИЗА

1. Введение. Мониторинг паводковой ситуации и прогноз рисков наводнения – это сложный комплекс задач, для решения которых требуется обработка больших объемов информации, а также важен учет гидрометеорологических особенностей конкретной местности. В Республике Беларусь влияние наводнений особенно ощутимо (в отдельные годы катастрофично) в пойме р. Припять и ее притоках. Их основную причину можно классифицировать как половодья или паводки редкой повторяемости [1], а характер типичен для равнинных рек с преимущественно снеговым питанием.

Задача своевременного прогнозирования наводнений в речных поймах, возникающих либо в результате продолжительных высоких половодий и кратковременных паводков, либо из-за возникновения аварийных ситуаций, является определяющей для своевременного проведения соответствующих мероприятий.

2. Нейросетевой подход к прогнозированию наводнений. Наряду с традиционными математическими моделями, для прогнозирования опасных гидрологических явлений и выявления сложных зависимостей между метеорологическими процессами и речным стоком активно используются искусственные нейронные сети (ИНС) [2, 3]. В первую очередь при помощи ИНС решаются задачи прогнозирования уровня подъема воды для участка реки в районе конкретного поста наблюдения в период половодья [4, 5].

В частности, в задачах прогнозирования паводков используются ИНС для прогнозирования и коррекции ошибок традиционных математических моделей при краткосрочном прогнозировании расхода воды на гидрологическом посту наблюдения [6]. На вход обученной ИНС поступает результат прогнозирования традиционной гидрологической моделью расхода воды на момент времени $t+1$ и фактический расход воды на момент времени t , на выходе же ИНС появляется прогноз расхода воды на момент времени $t+1$ с исправлением возможной ошибки гидрологической модели.

При этом для повышения качества прогнозирования при обучении ИНС можно так или иначе использовать не только многолетний временной ряд величин уровня подъема воды, зарегистрированных на посту наблюдения, но и временные ряды значений климатических факторов, влияющих на прохождение половодья – среднесуточной температуры воздуха, уровня осадков и снежного покрова в окрестностях гидрологического поста наблюдения, то есть параметров, оказывающих существенное влияние на картину наводнения. В частности, водный режим р. Припять (и ряда других рек, питание которых также относится к смешанному типу с преобладанием снегового) характеризуется длительным весенним половодьем, а потому учет динамики накопления снеготалпов позволяет существенно повысить точность прогноза. В настоящей работе предлагается метод подобного многофакторного анализа, учитывающего помимо гидрологических измерений расхода воды также такие климатические факторы, как содержание воды в снежном покрове на водосборной площади, температура воздуха и количество выпавших осадков.

Ряд источников сходится во мнении – одним из наиболее эффективных инструментов для анализа таких нелинейных и многофакторных зависимостей, как гидрологические временные ряды, является механизм многослойных ИНС с прямым распространением сигнала. В частности [7] и другие работы иллюстрируют преимущества такого подхода по сравнению с авторегрессионной интегриро-

ванной моделью скользящего среднего, известной также как метод Бокса-Дженкинса. Поэтому для прогнозирования значений расхода воды в створе реки в рамках настоящего исследования были выбраны нейросетевые механизмы.

Очевидно, что в представленной постановке решаемая задача ориентирована на равнинные реки с преимущественно снеговым питанием, примером которых в Республике Беларусь может служить р. Припять.

3. ИНС, используемая для оценки снеготалпов. Традиционный способ определения содержания воды в снеге предполагает взятие проб снега на снегомерных маршрутах. Его применение на водосборной площади является ограниченным в силу трудоемкости, и поэтому было принято решение воспользоваться результатами ежедневной оценки содержания воды в снежном покрове на основе пассивного радиотеплового сканирования поверхности водосборной площади с борта искусственного спутника Земли (ИСЗ).

Существующий подход определения водного эквивалента снежного покрова по данным радиотепловых спутниковых измерений основан на применении эмпирических регрессионных зависимостей, нередко дающих существенные расхождения по сравнению с непосредственными измерениями на снегомерных маршрутах [8].

С целью повышения точности вычисления содержания воды в снеге исследована на примере большой климатически неоднородной территории возможность замены регрессионных зависимостей математическим аппаратом ИНС. Для проведения исследований использованы данные, полученные при помощи микроволнового сканирующего радиометра-поляриметра SSM/I за промежуток времени с 1987 по 2014 гг. В качестве территории с требуемыми размерами и климатической неоднородностью была выбрана территория РФ [9]. В ходе экспериментов для обучения ИНС использовались метеорологические станции с дифференциацией снегомерных маршрутов по ландшафтному признаку: лес, поле и лес/поле. В качестве архитектуры ИНС использовался многослойный перцептрон с одним промежуточным слоем, а в качестве функции активации нейронов в скрытом слое используется гиперболический тангенс (рис. 1).

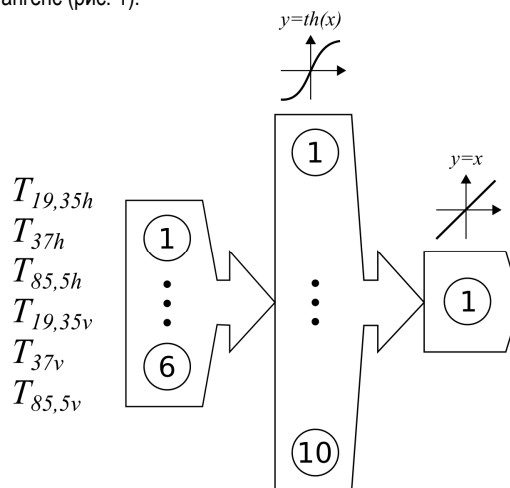


Рисунок 1 – Архитектура ИНС, используемая при вычислении водного эквивалента снежного покрова

Волчек Александр Александрович, д. т. н., декан факультета инженерных систем и экологии, профессор Брестского государственного технического университета.

Костюк Дмитрий Александрович, к. т. н., доцент кафедры ЭВМиС, доцент Брестского государственного технического университета.

Петров Дмитрий Олегович, ст. преподаватель кафедры ЭВМиС Брестского государственного технического университета.

Шешко Николай Николаевич, к. т. н., доцент кафедры СХГТМ, доцент Брестского государственного технического университета.

Беларусь, БрГТУ, 224017, г. Брест, ул. Московская, 267.

Физика, математика, информатика

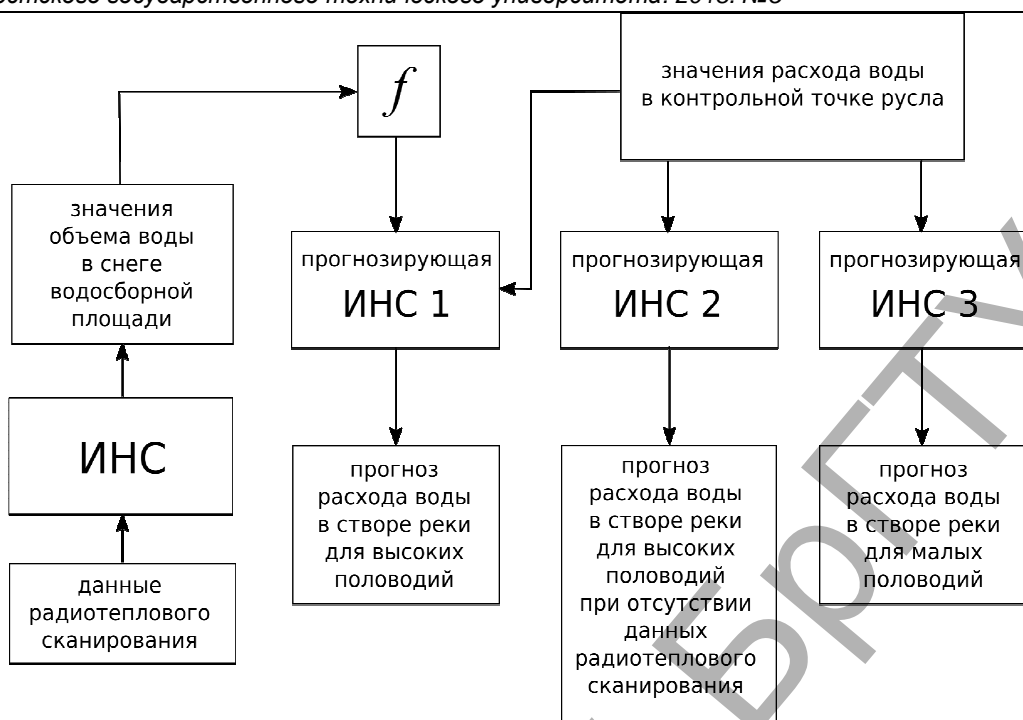


Рисунок 2 – Структурная схема метода прогнозирования расхода воды в створе реки при прохождении весенних половодий

В результате ИНС, индивидуально обученные для каждого снежного маршрута, для данных с расширенным набором используемых радиочастотных каналов (19,35; 37,0; 85,5 ГГц как горизонтальной, так и вертикальной поляризации) позволили достичь максимальной величины коэффициента корреляции Пирсона, r равной 0,79, что позволяет сделать вывод о предпочтительности использования ИНС для оценки величины водного эквивалента снежного покрова [9].

4. ИНС, используемые для выполнения прогноза. Структурная схема, иллюстрирующая представляемый метод нейросетевого прогноза наводнений, представлена на рис. 2. Предлагаемый метод позволяет для прогнозировать среднесуточный расход воды на выбранном створе реки во время прохождения как высоких весенних половодий, приводящих к возникновению наводнений, так и малых весенних половодий. В силу того, что высокие и малые половодья имеют сильно различающийся характер протекания, их прогнозирование выполняется отдельными ИНС, обученными на явлениях соответствующего класса.

Прогноз значений расхода воды вплоть до максимальных во время прохождения высоких половодий выполняется *прогнозирующей ИНС 1*, которая одновременно обрабатывает методом скользящего окна значения двух временных рядов: среднесуточных расходов воды, зарегистрированных на гидрологическом посту наблюдения, и суммарных величин накопления воды в снежном покрове на исследуемой водосборной площади. Расчет величины накопления воды в снежном покрове выполнялся описанной выше отдельной ИНС на основе радиояростных спутниковых измерений (рис. 1), с опциональной фильтрацией формируемого временного ряда.

На случай отсутствия доступа к результатам микроволнового сканирования поверхности водосбора (например, из-за неисправности оборудования на борту ИС3) предусмотрена *прогнозирующая ИНС 2*, которая способна выполнять прогноз высоких половодий по единственному временному ряду. Реальным примером ограниченной доступности таких данных является зарегистрированная 05.04.2016 и сохраняющаяся по сей день техническая неисправность канала 37 ГГц вертикальной поляризации сенсора SSMIS, размещенного на спутнике DMSP F-17 [10].

Наконец, малые половодья происходят при значительном впитывании почвой воды, выделяющейся в результате таяния снега, и для них не наблюдается существенной зависимости расхода воды в створе от накопленных снегозапасов. Для явлений данного класса

предусмотрена *прогнозирующая ИНС 3*, учитывающая только ряд значений расхода.

Признаком, позволяющим определить, что предстоит не низкое, а высокое половодье, является выпадение осадков осенью с последующим зимним промерзанием почвы (т. е. значительным снижением ее инфильтрационной способности) и наличием достаточного количества накопленных снегозапасов. Благодаря тому, что микроволновое излучение, регистрируемое орбитальным комплексом датчиков, полностью поглощается слоем воды в толще либо на поверхности снега (появившейся в результате таяния и/или выпадения осадков), оказывается возможным надежное и своевременное выявление момента начала интенсивного таяния снежного покрова.

Определение предполагаемой категории ожидаемого половодья основывается на результатах предварительного учета количества выпавших осенних осадков за период октябрь-ноябрь и доли дней с отрицательной среднесуточной температурой воздуха за период декабрь-январь. Необходимые характерные значения учитываемых величин для конкретной местности определяются на основе анализа метеорологических временных рядов, накопленных за соответствующие промежутки времени, непосредственно предшествующие зарегистрированным высоким половодьям.

В случае, когда данные об осенних осадках и зимних среднесуточных температурах свидетельствуют о предстоящем малом половодье, для прогнозирования используется третья ИНС, обученная на гидрологических данных, характерных для малых половодий.

Архитектура прогнозирующих ИНС аналогична показанной на рис. 1 (использован многослойный перцептрон с одним скрытым слоем). В качестве функции активации в скрытом и выходном слоях гиперболический тангенс. В соответствии с избранной функцией активации, элементы обучающих и тестирующих выборок масштабированы в диапазон $[-1, 1]$.

Размер скользящего окна для обработки входных данных прогнозирующими ИНС определяется эмпирическим путем с целью минимизации ошибки обученной нейронной сети на тестовой выборке. Наборы данных, используемых для обучения и тестирования ИНС, должны охватывать период, включающий момент пика среднесуточного расхода воды – самого позднего из зарегистрированных во время весенних половодий на выбранном створе реки.

Размер горизонта прогнозирования должен соответствовать заблаговременности среднесрочного метеорологического прогноза

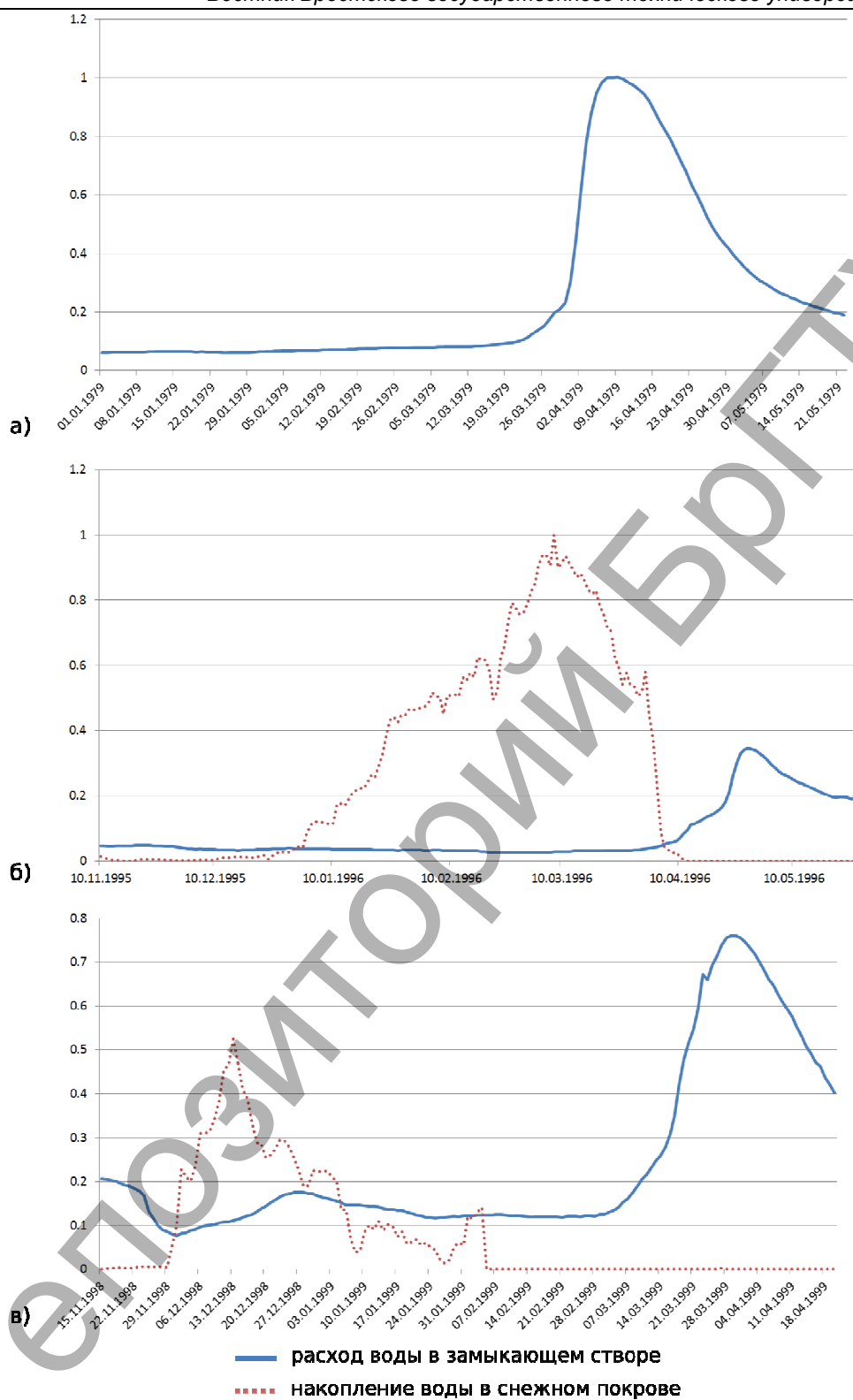


Рисунок 3 – Расход воды на гидрологическом посту г. Мозыря и накопление воды в снежном покрове 1979, 1996, 1999 гг.

(от 7 до 10 суток) [11]. Для оценки характера спада воды при достижении пика половодья целесообразно выполнять также прогноз нисходящей ветви гидрографа (не более чем на удвоенную величину горизонта прогнозирования).

5. Апробация метода на водосборной площади р. Припять.

Для апробации предложенной архитектуры и соответствующего метода прогнозирования был использован находящийся в открытом доступе архив зарегистрированных метеорологических станциями среднесуточных температур воздуха и величин выпавших осадков

на территории Европы с 1978 г. по 2005 г., предоставляемый международным проектом ECA&D (Europe Climate Assessment & Dataset) [12]. Временные ряды измерений расхода воды на гидрологическом посту наблюдения г. Мозыря (р. Припять) с 1978 г. по 2005 г. получены из архива Республиканского центра по гидрометеорологии, контролю радиоактивного загрязнения и мониторингу окружающей среды (РУП «Белгидромет»). Источником значений ежедневной оценки содержания воды в снежном покрове на основе пассивного

Таблица 6 – Результаты исследования вариантов архитектур ИНС для прогноза величины расхода воды в контрольном створе реки

ИНС	Структура слоев	Размер обучающей выборки	Размер проверочной выборки	Количество итераций обучения	MSE обучения	MSE прогноза
ИНС 1	92-92-7	18300	732	1000	0,0002	0,0007
	92-46-7	9150			0,0003	0,0012
	92-184-7	36478			0,0003	0,0011
	46-46-7	5040	1008		0,0002	0,0005
ИНС 2	46-46-7	5002	122		0,0004	0,0011
	46-23-7	2440			0,0004	0,001
	46-92-7	9760			0,0004	0,0011
	23-23-7	1428	168		0,0003	0,001
ИНС 3	46-46-7	5002	1968		0,0004	0,0008
	46-23-7	2460			0,0008	0,0011
	46-92-7	9758		0,0004	0,0012	
	23-23-7	1470	2520	0,001	0,0009	

радиотеплового сканирования земной поверхности за промежутки времени с 1979 г. по 2005 г. послужил архив проекта GlobSnow [13].

Анализ исходных данных позволил выявить три примечательных года: 1979, 1996 и 1999 (рис. 3). В 1996 году был зарегистрирован максимум накопления воды в снежном покрове на водосборной площади р. Припять, составивший 14,3 км³. В 1979 и 1999 году произошли катастрофические половодья с максимальными расходом воды по гидропосту г. Мозыря в 4300 м³/с и 3270 м³/с соответственно. Для 1999 года характерен продолжительный период интенсивного таяния снега сопровождаемый выпадением значительного количества осадков (110–255% нормы) в течение 51 дня [14].

Для обучения ИНС был выбран эвристический алгоритм RPROP [15], который позволяет сократить время, требуемое на обучение, по сравнению с стандартным методом обратного распространения ошибки. Параметры алгоритма RPROP были выбраны согласно рекомендациям его авторов [14]: $\Delta_{max} = 50$, $\Delta_{min} = 1,0 \cdot 10^{-6}$, $\eta^+ = 1,2$, $\eta^- = 0,5$. Начальную настройку весовых коэффициентов синаптических связей персептрона было решено производить согласно алгоритму, предложенному Нгуеном и Видроу [16]. В качестве меры величины ошибки обучения и достигнутой ошибки прогнозирования выбрана среднеквадратичная ошибка [17].

В целях эмпирического определения оптимального количества нейронов в слоях разрабатываемых ИНС проведен ряд вычислительных экспериментов.

Первоначально размер скользящего окна был выбран равным 46 дням (срок, предшествующий началу таяния снега) – поэтому размер входных слоев ИНС 2 и ИНС 3 составил 46 нейронов, а ИНС 1 – 92 (46 + 46). Размер скрытых слоев был установлен равным размеру соответствующих входных слоев и размер выходных слоев составил 7 нейронов согласно избранной величине горизонта прогнозирования. В результате структуры ИНС 1, ИНС 2 и ИНС 3 приобрели следующий вид: 92-92-7, 46-46-7 и 46-46-7. В таком случае общее число синаптических связей для указанных ИНС составило 9108, 2438 и 2438 соответственно. Но при этом размеры обучающих выборок, сформированных по доступным архивным гидрометеорологических данным, оказались меньше вычисленных для разрабатываемых ИНС количество синаптических связей. Для обеспечения как минимум двукратного превышения количества образцов в обучающих выборках над количеством межнейронных связей в соответствующих ИНС – наблюдаемые гидрометеорологические данные были дополнены промежуточными, вычисленными методом линейной интерполяции значений между соседними парами временных

рядов зарегистрированных в разные годы на гидрологическом посту наблюдения расходов воды. Для учета возможных расхождений в сроках наступления интенсивного таяния снега, набор соответствующих временных рядов, входящих в состав обучающей выборки ИНС прогнозирования высоких половодий, был сформирован путем смещения в обе стороны относительно момента наступления характерного перепада значений на случайное, но не превышающее по абсолютной величине 7, число дней, что соответствует размеру горизонта прогнозирования.

В целях выбора неповторяющихся величин параметра линейной интерполяции между смежными гидрографами и случайного смещения времени наступления момента таяния снега был применен генератор псевдослучайных чисел Mersenne twister [18].

Следует отметить, что обучающие выборки, в отличие от проверочных, сформированы на основе интерполированных гидрометеорологических данных. В проверочных выборках для ИНС 1 и ИНС 2 находятся значения расходов воды, зарегистрированные во время прохождения высоких половодий 1979 г. и 1999 г. Тестовый набор данных для ИНС 1 кроме гидрологических временных рядов содержит временные ряды изменения агрегатного состояния снежного покрова, в которых момент начала интенсивного таяния зафиксирован не только согласно сроку его обнаружения в 1999 г., но и смещен относительно него на 7 суток как вперед так и назад. Проверочные выборки для ИНС 3 составлены из 24 временных рядов расходов воды в контрольном створе реки, отобранных за период с 1980 по 2003 гг.

В проведенных вычислительных экспериментах исследованы архитектуры искусственных нейронных сетей со следующими изменениями размера скрытого слоя: число нейронов скрытого слоя соответствует числу нейронов во входном слое, размер скрытого слоя вдвое меньше размера входного слоя и размер скрытого слоя вдвое меньше размера входного слоя (см. таблицу 1). Также изучены последствия изменений в результатах работы ИНС при сокращении размера скользящего окна до 23 дней (см. таблицу 1). Пример прогнозирования с применением разработанной архитектуры приведен на рис. 4.

В результате проведенных исследований, основываясь на наименьших достигнутых величинах среднеквадратичной ошибки прогноза, для ИНС 1, ИНС 2 и ИНС 3 были признаны оптимальными следующие архитектуры: 46-46-7, 23-23-7 и 46-46-7 соответственно. При оценке качества прогнозирования расхода воды в контрольном створе реки во время прохождения весенних половодий разработанными ИНС получены высокие значения коэффициентов корреляции Пирсона, а именно: 0,99, 0,94 и 0,74.

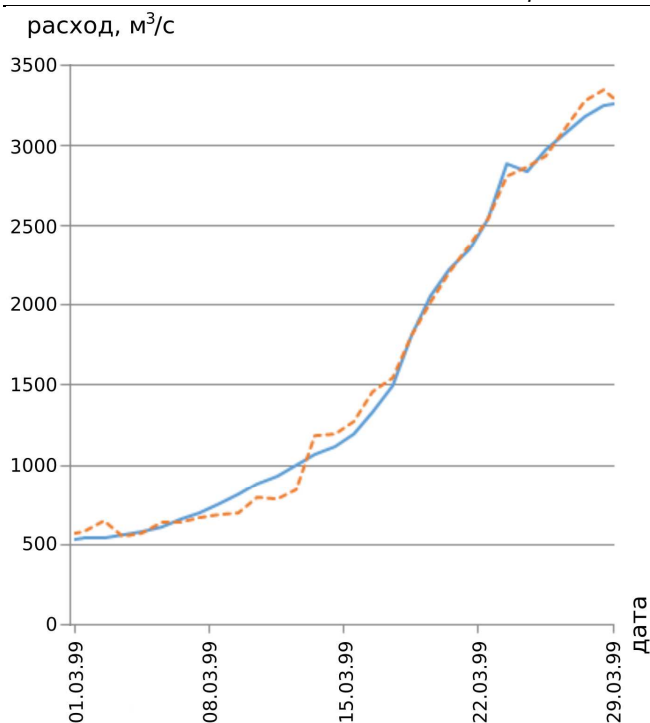


Рисунок 4 – Пример прогнозирования расхода воды (пунктир — значения, сгенерированные ИНС, сплошная линия — гидрограф, полученный с помощью натуральных измерений)

СПИСОК ЦИТИРОВАННЫХ ИСТОЧНИКОВ

1. Мухин, В. И. Типизация источников чрезвычайных ситуаций гидрологического характера / В.И. Мухин, В.Л. Шимитило // Научные и образовательные проблемы гражданской защиты. – 2010. – № 2. – С. 70–73.
2. Varoonchotikul, P. Flood Forecasting Using Artificial Neural Networks. – CRC Press, 2003. – 112 p.
3. Abraham, R. J. Neural Networks for Hydrological Modelling / R.J. Abraham, P.E. Kneale, L.M. See – A.A. BALKEMA PUBLISHERS, 2004. – 305 p.
4. Аракелян, С. М. Разработка методики прогноза весенних паводков на базе нейросистем / С. М. Аракелян [и др.] // Труды Международного семинара «Мягкие вычисления-96». – Казань, 1996. – С. 32–34.
5. Владимиров, В. А. Прогнозирование уровня весенних паводков и мониторинг зон затопления на основе ГИС-технологий и систем искусственного интеллекта / В. А. Владимиров [и др.] // Стратегия гражданской защиты: проблемы и исследования. – 2012. – № 2. – С. 519–540.
6. Shamseldin, A. Y. Hybrid Neural Network Modelling Solutions // Neural Networks for Hydrological Modelling. – A.A. BALKEMA PUBLISHERS, 2004. – Chapter 4. – P. 60–77.
7. Tang, Z. Feedforward neural nets as models for time series forecasting / Z. Tang, P. A. Fishwick // ORSA journal on computing. – 1993. – V. 5. – №. 4. – P. 374–385.
8. Volchak, A. Determination of the snow melting intensity in nowadays climate conditions by example of the Neman river basin / A. Volchak, D. Kostiuk, N. Sheshko, D. Petrov // 2nd International conference on Climate Change – The environmental and socio-economic response in the Southern Baltic region. Szczecin, Poland, 12-15 May 2014. – P. 37–38.
9. Волчек, А. А. Оценка водного эквивалента снега по данным пассивного микроволнового сканирования земной поверхности с использованием искусственных нейронных сетей для территории Российской Федерации / А. А. Волчек, Д. А. Костюк, Д. О. Петров // Лед и Снег. – 2016. – Т. 56. – № 1. – С. 43–51.
10. Brodzik, M. J. Near-Real-Time DMSR SSM/I-SSMIS Pathfinder Daily EASE-Grid Brightness Temperatures, Version 1 / M. J. Brodzik, R. Armstrong // USA NASA DAAC National Snow and Ice Data Center [Электронный ресурс]. – 2017. – Режим доступа: <https://nsidc.org/data/NSIDC-0342/versions/1#>. – Дата доступа: 18.11.2017.
11. Наставление по Глобальной системе обработки данных и прогнозирования // ВМО. – 2012. – № 485. – Т. 1: Глобальные аспекты. – 204 с.
12. Data Policy for ECA&D and E-OBS // European Climate Assessment and Dataset [Электронный ресурс]. – 2012. – Режим доступа: http://eca.knmi.nl/documents/ECAD_datapolicy.pdf. – Дата доступа: 08.06.2016.
13. Luojus K. Snow Water Equivalent (SWE) product guide / K. Luojus [et al.] // GlobSnow Consortium [Электронный ресурс]. – 2010. – Режим доступа: http://www.globsnow.info/swe/GlobSnow_SWE_product_readme_v1.0a.pdf. – Дата доступа: 08.06.2016.
14. Петров, Д. О. Использование климатических данных в задаче нейросетевого прогнозирования половодий [Электронный ресурс] / Д. О. Петров, А. А. Волчек, Н. Н. Шешко, Д. А. Костюк // Технологии информатизации и управления ТИМ-2016: материалы III Международной научно-практической конференции, Гродно, 14–15 апреля 2016 г. / ГУО «Институт повышения квалификации и переподготовки в области технологий информатизации и управления» Белорусского государственного университета. – Минск, 2016. – 1 электр. компакт диск (CD R). – 924 стр.: 306 ил., 97 табл. – Рус. – Деп. в ГУ «БелИСА» 06.10.2016 № Д201624.
15. Reidmiller, M. Direct Adaptive Method for Faster Backpropagation Learning: The RPROP Algorithm / M. Reidmiller, H. A. Braun // Neural Networks, 1993., IEEE International Conference on. – IEEE, 1993. – P. 586–591.
16. Nguyen, D. Improving the Learning Speed of 2-Layer Neural Networks by Choosing Initial Values of the Adaptive Weights / D. Nguyen, B. Widrow // Proceedings of IEEE International Joint Conference on Neural Networks. – 1992. – V. 5. – №. 4. – P. 595–603.
17. Волчек, А. А. Опыт прогнозирования весенних наводнений с использованием спутниковой информации о снеговых запасах на речном водосборе / А. А. Волчек, Д. А. Костюк, Д. О. Петров // Водные ресурсы и климат: материалы докладов V Международного Водного Форума: в 2 ч. – Минск: БГТУ, 2017. – Ч. 2. – С. 197–201.
18. Matsumoto M., Nishimura T. Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator // ACM Transactions on Modeling and Computer Simulation (TOMACS). – 1998. – V. 8. – №. 1. – P. 3–30.

Материал поступил в редакцию 06.02.2019

VOLCHEK A. A., KOSTIUK D. A., PETROV D. O., SHESHKO N. N. Flood forecasting method based on multivariate neural network analysis

A flood forecasting method is presented, based on artificial neural networks and multivariate analysis with use of the following parameters: hydrological measurements of water discharge at the river bed, precipitation and data on snow content in the river's catchment area. A cascade connection of a neural network, which determines the water equivalent of snow cover according to multi-band satellite radio thermal measurements, with a predictive neural network, is used. The results of testing the developed method are presented for the Pripyat' river basin.