

Метод восстановления среднемесячных значений прозрачности воды на примере северо-западной части Черного моря

А.С. Лубков¹, Е.Н. Воскресенская¹, А.С. Кукушкин^{2*}

¹Институт природно-технических систем РАН
299011, г. Севастополь, ул. Ленина, 28

²Морской гидрофизический институт РАН
299011, г. Севастополь, ул. Капитанская, 2

Поступила в редакцию 15.05.2015 г.

С целью восстановления данных наблюдений гидрофизических параметров предложена модель на основе нейронной сети с учителем. В качестве входных сигналов модели взяты индексы глобальных климатических колебаний в системе океан–атмосфера. Описан процесс обучения и адаптации модели, позволяющий найти наиболее точное решение задачи моделирования. Сопоставление среднемесячных значений объема стока Дуная, воспроизведенных с помощью модели, с данными наблюдений показало их хорошее согласие. Восстановлены пропущенные в различные годы данные наблюдений по прозрачности (глубине видимости белого диска). Отмечена близость абсолютных величин восстановленных и измеренных значений глубины видимости белого диска. Выявлены некоторые особенности межгодовых изменений восстановленных значений прозрачности, обусловленные климатическими факторами в период 1950–1962 гг. и влиянием изменений содержания хлорофилла *a* в период 1998–2010 гг.

Ключевые слова: прозрачность, нейронные сети, речной сток, гидрометеорологические условия, Черное море; transparency, neural network, runoff, hydrometeorological conditions, Black Sea.

Введение

В последние десятилетия экосистема Черного моря претерпела существенную трансформацию, связанную с климатическими изменениями и антропогенным воздействием. Северо-западная часть Черного моря (СЗЧМ), имеющая важное народно-хозяйственное значение, отличается от других районов моря высокой антропогенной нагрузкой. Большие объемы загрязненных береговых стоков существенно влияют на содержание в морской воде взвешенных и растворенных веществ, которые в основном определяют прозрачность воды [1]. В то же время распределение взвешенного вещества и прозрачности в зимне-весенний период зависит от региональных гидрометеорологических условий и объемов стока рек, аномалии которых, в свою очередь, обусловлены изменчивостью глобальных атмосферных процессов в системе океан–атмосфера.

В работах [1, 2] было показано влияние крупномасштабных атмосферных процессов в системе океан–атмосфера на межгодовые сезонные изменения распределения прозрачности в северо-западной части Черного моря. При этом немногочисленные данные

по измерению показателя ослабления направленного света, полученные в районе влияния стока Дуная в 1978–1995 гг., и неравномерное их распределение по годам и сезонам ограничивали возможности анализа сезонных межгодовых его изменений. Относительно регулярные наблюдения прозрачности (глубина видимости белого диска) в 1960–2000 гг. также имели пропуски данных, затрудняющие анализ долговременной ее изменчивости.

Для восстановления пропущенных данных по прозрачности может быть использован метод линейного регрессионного анализа. Однако его возможности восстановления этих данных, основанные на связи между прозрачностью и концентрацией хлорофилла *a* [3], были ограничены немногочисленными (в том числе по времени наблюдений) данными по измерению концентрации последнего. Более привлекательным является моделирование рядов данных по прозрачности с помощью метода нейронных сетей (НС) с использованием индексов крупномасштабных действующих атмосферных сигналов в системе океан–атмосфера [4]. Такой метод позволяет не только восстанавливать, но и прогнозировать значения прозрачности с заблаговременностью 1–3 мес.

Известно использование НС для моделирования среднемесячных и среднесуточных значений метеорологических [5–7] и гидрологических [5, 8, 9]

* Андрей Сергеевич Лубков (andrey-ls2015@yandex.ru);
Елена Николаевна Воскресенская (elena_voskr@mail.ru);
Александр Сергеевич Кукушкин (kukushkinas@mail.ru).

параметров. Исследование возможности статистического прогноза среднемесячного объема стока черноморских рек на основе НС было описано в работе [9]. В этой работе на основе пяти основных климатических мод (североатлантическая, восточноатлантическая, скандинавская, полярная и мода колебания между Восточной Атлантикой и Западной Евразией) было проведено моделирование среднемесячного стока впадающих в СЗЧМ крупных рек в паводковый период (май) и за два предыдущих месяца. Эти климатические моды оказывают заметное влияние на изменчивость гидрометеорологических условий в Европейском регионе в зимне-весенний период и поэтому существенно влияют на объем речного стока в мае, а также в марте и апреле. Так, величина корреляции среднемесячных значений стока Дуная в апреле и в мае была равна 0,8, а в марте и мае – 0,5 (использовались данные, зарегистрированные на метеостанции Измаил в течение 1921–2010 гг.). Полученный в работе [9] результат несущественно превышал автокоррелируемость (с задержкой в 1 мес) среднемесячных значений объемов стока Дуная. Возможно, это было связано с некорректно подобранной входной выборкой, содержащей взаимно-коррелированные входные параметры. При этом на фоне индексов дальнедействующих сигналов в системе океан–атмосфера существенно выделялась взаимосвязь прогнозируемого параметра со среднемесячным стоком в апреле и марте, что могло привести к заметным различиям изменения весов между нейронами скрытого и входного слоев [4]. Поэтому важной задачей настоящей работы было ответить на вопрос: «Возможно ли на основе информации о глобальных колебаниях в системе океан–атмосфера построить НС, позволяющую качественно восстанавливать выбранные параметры?». Предварительные результаты восстановления значений глубины видимости белого диска по данным наблюдений в СЗЧМ при помощи метода нейронных сетей были описаны в [10].

Целями настоящей работы являются описание метода моделирования (восстановления) значений прозрачности на основе нейронной сети, входными сигналами которой являются индексы глобальных климатических сигналов, и анализ результатов восстановления значений глубины видимости белого диска в СЗЧМ в районе влияния стока Дуная.

Материалы и методика

В работе использовались следующие данные:

– значения глубины видимости белого диска, полученные в период с 1963 по 2000 г. в западном районе СЗЧМ (28°30′–30°30′ в.д., 44°–45°40′ с.ш.), – 1552 наблюдения;

– среднемесячные данные об объеме расхода стока Дуная с 1950 по 2012 г.;

– среднемесячные значения индексов дальнедействующих атмосферных сигналов в системе океан–атмосфера (далее атмосферные индексы) за период 1950–2013 гг.: североатлантическое колеба-

ние (NAO), арктическое (AO), восточноатлантическое (EA), скандинавское (SCAND), полярное (POL), западнотихоокеанское (WP), колебания между Восточной Атлантикой и Западной Евразией (EATL/WRUS), между Тихим океаном и Северной Америкой (PNA), между восточной и северной частью Тихого океана (EP-NP) [11], Тихоокеанская декадная осцилляция (PDO) [12], южное колебание (SO) [13].

– среднемесячные спутниковые (цветовой сканер Sea WiFS) данные о концентрации хлорофилла *a* с 1998 по 2010 г.

Для восстановления пропущенных значений глубины видимости белого диска была применена модель нейронной сети, являющейся разделом искусственного интеллекта, в котором для обработки сигналов используются аналогичные происходящим в нейронах живых существ явления. Идея создания искусственной нейронной сети сформировалась в середине XX столетия, и одной из первых моделей нейрона была бинарная модель МакКаллока-Питтса, предложенная в 1943 г. [4]. До 1970-х гг. было представлено множество моделей нейронных сетей, имеющих линейные и рекуррентные связи между нейронами и использующих методику обучения с учителем или без него. В конце 70-х гг. метод НС получил дальнейшее развитие, связанное с разработкой алгоритма обучения НС с учителем – метода обратного распространения ошибки. Этот метод и в настоящее время считается одним из наиболее эффективных алгоритмов обучения многослойной НС.

Для моделирования значений прозрачности использовалась однонаправленная гетероассоциативная однослойная (один скрытый слой) НС с учителем. Нейроны входного, скрытого и выходного слоев представлены сигмоидальной биполярной функцией: $f(x) = \tanh(\beta x)$ [4, 14, 15]. Ее производная – деактивационная функция – используется для формирования целевой функции и отображения результата. Схема модели представлена на рис. 1. В качестве входных параметров был использован набор атмосферных индексов. На выходе модель имеет один нейрон. Единичная обучающая выборка представлена рядом значений глубины видимости белого диска.

Обучение модели проводилось на основе алгоритма обратного распространения ошибки. Ряд значений моделируемого параметра был поделен на обучающую выборку, используемую для обучения модели, и тестирующую (тестовую), по которой проверялся результат моделирования. Выбор временных промежутков для обучения и тестирования модели основывается на определении понятия климата, утвержденного ЮНЕСКО как статистический ансамбль состояний, наблюдаемых в климатической системе (атмосфера и деятельный подстилающий слой) в течение нескольких десятилетий (по принятым в 1957 и 1996 гг. международным соглашениям этот период равен 30 лет) [16]. Поэтому для обучения и тестирования модели выбирались 30-летние периоды. Если количество экспериментальных данных

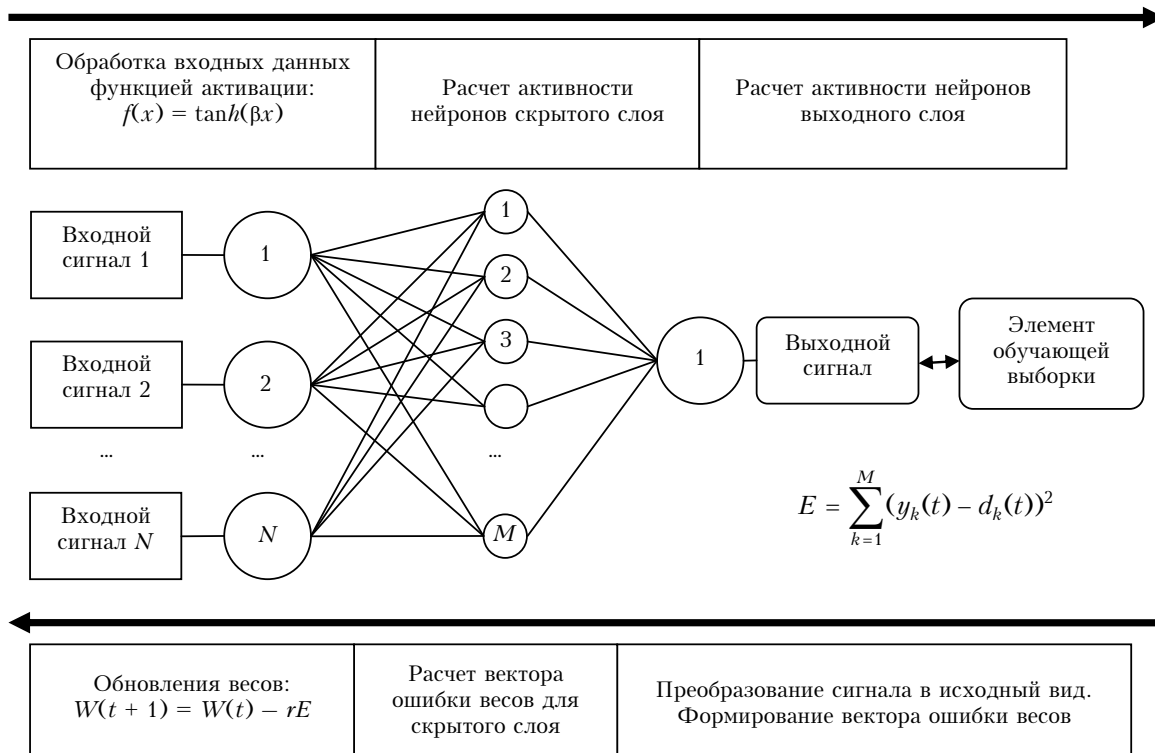


Рис. 1. Краткая схема обучения нейронной сети. W – веса активности нейронов; t – временной шаг; r – коэффициент скорости обучения; E – целевая функция; y_k – множество выходных сигналов; d_k – множество элементов обучающей выборки; N – общее количество нейронов во входном слое; M – общее количество нейронов в скрытом слое; k – номер нейрона скрытого слоя

было меньше 60 последовательных среднемесячных значений, то используемый ряд этих данных делился поровну.

Поскольку исходный ряд значений моделируемого параметра не всегда был достаточно статистически обеспечен, то применялись итерации обучения, на которых эмпирическое выражение моделируемого параметра изменялось. Целью обучения являлся поиск момента (точки) в процессе моделирования, называемого максимумом обучения, когда полученное выражение наиболее точно описывало бы моделируемый параметр. В основе классического алгоритма обратного распространения ошибки лежит целевая функция [4], выраженная в виде квадратичной суммы разностей между фактическими и ожидаемыми значениями преобразованных входных сигналов. В случае единичной выборки она определяется в виде

$$E = \sum_{k=1}^M (y_k - d_k)^2. \quad (1)$$

Практическое использование подобной целевой функции выявило некоторые проблемы. В сложных ситуациях, когда целевая функция находится в глобальном минимуме, разность значений моделируемого параметра и результата моделирования достаточно большая, а их значения взаимно-коррелированы, однозначно определить максимум обучения оказалось невозможно. Так же и в обратной ситуа-

ции, когда разность этих значений невелика, а значения не взаимно-коррелированы. По этой причине целевая функция была дополнена неким эмпирическим критерием, предложенным в процессе испытания модели.

Для испытания модели в качестве входных параметров были взяты четыре гармонические функции: $\sin(1,05x)$, $0,7\cos(9x)$, $0,3\sin(22x)$, $\cos(0,2x)$, подобранные в соответствии с известными климатическими периодами (60 лет, 11 лет, 3–4 года и теоретическое предположение – 320 лет; $\Delta x = 0,1$). Сумма этих сигналов моделировалась нейронной сетью. В испытаниях модели также использовались итерации обучающей выборки, когда на каждом их шаге вычислялась корреляция между элементами обучающей выборки и модельного расчета ($r_{обуч}$), корреляция между элементами тестовой выборки и смоделированными на основе полученного на обучающей выборке эмпирического выражения ($r_{тест}$), а также модуль разности фактических и рассчитанных на обучающей выборке значений параметра. В процессе тестирования были подтверждены известные свойства НС подобного типа: возможность моделирования зашумленной характеристики и снижение качества моделирования при зашумленных входных сигналах или при отсутствии одного из них. В результате испытания модели был получен важный вывод: минимум целевой функции выполняется при условии равенства коэффициентов корреляций тестовой и обучающей выборок. При этом максимум $r_{обуч}$ совпадает с минимумом целевой

функции, а максимум $r_{\text{тест}}$ находится в непосредственной близости.

Известно, что индексы дальнедействующих атмосферных сигналов в системе океан–атмосфера рассчитываются между географическими центрами действия атмосферных процессов, местоположение которых часто изменяется и не привязано к определенной точке пространства. Возможность регулярного мониторинга (временного и пространственного) за атмосферными явлениями появилась сравнительно недавно. Поэтому расчет индексов, как правило, производится по наблюдениям в стационарной точке пространства и, соответственно, требует корректировки с помощью различных методик. Отсюда следует, что входные сигналы НС, в качестве которых используются эти индексы, имеют ошибку, связанную с неточностью определенных центров действий атмосферных процессов. Кроме того, на любой климатический параметр, нелинейно связанный с глобальными атмосферными процессами, также действуют и другие не учтенные в модели факторы. В связи с этим моделирование значений различных параметров (например, прозрачности морской воды, объема стока рек и др.) с помощью метода НС, использующего в качестве входных сигналов атмосферные индексы, не позволяет абсолютно точно восстановить их значения и описать временные изменения этих величин. На практике эти особенности проявились в невозможности точно определить момент максимума обучения.

Поскольку определение динамики в климатических процессах имеет более значимый характер, чем евклидово расстояние между рассчитанными и фактическими значениями моделируемого параметра, то будем использовать $r_{\text{обуч}}$ как основной критерий поиска максимума обучения, при условии что $|r_{\text{тест}} - r_{\text{обуч}}| < 0,15$, где значение 0,15 выбрано эмпирически. В текущей работе не ставилась задача поиска оптимального критерия оценки нелинейных моделей. Подробное описание нескольких существующих критериев оценки моделирования представлено в работе [17].

Адаптация модели включала три стадии: предварительная обработка данных, моделирование и заключительная обработка результатов моделирования.

На этапе *предварительной обработки* данных выполняется поиск связей между значениями индексов дальнедействующих климатических сигналов и исследуемого параметра. Для этого рассчитывалась корреляция между его значениями и значениями атмосферных индексов, наблюдаемых в предшествующие месяцы текущего и предыдущего года. Далее значения исследуемого параметра и индексов нормировались и формировались выборки из их значений.

Моделирование. Практика моделирования показала, что поведение НС не всегда предсказуемо. На основе отобранных на предварительном этапе индексов модель не всегда показывала удовлетворительный результат. Это могло быть связано с использованием индексов, не имеющих связи с моделируемым параметром. Поэтому модель запуска-

лась многократно с различными комбинациями входных сигналов, выбранных на предварительном этапе. В то же время, используя один и тот же набор входных сигналов и изменяя количество нейронов скрытого слоя или стартовые веса связи между нейронами, можно получить различные результаты моделирования. Поэтому в процессе работы были перебраны все возможные комбинации входных сигналов с различным количеством нейронов скрытого слоя. НС, имеющая определенный набор входных сигналов и определенное количество нейронов скрытого слоя, называется *конструкцией модели*. В процессе моделирования для каждой такой конструкции определялась точка максимума обучения. Для упрощения эксперимента все стартовые веса связи между нейронами были приравнены к 0,6 [10].

Количество скрытых нейронов в скрытом слое изменялось следующим образом. Согласно теореме Колмогорова [18] о представлении непрерывных функций нескольких переменных суперпозицией непрерывных функций одной переменной, которая в 1987 г. была использована Хехт-Нильсеном для нейронных сетей, оптимальное количество нейронов в скрытом слое равнялось $2i + 1$ (i – количество входных сигналов) [10, 14, 19]. Однако на практике количество нейронов скрытого слоя, наиболее приемлемое для текущей входной выборки, определить было достаточно сложно (для различных выборок оно может варьироваться от трех до нескольких десятков нейронов). По этой причине для каждой выборки входных сигналов были рассчитаны конструкции модели с разным количеством нейронов скрытого слоя, которое изменялось от 3 до $4i + 2$. При этом найденный максимум обучения записывался в лог-файл. Общее количество различных конструкций НС можно определить как:

$$\text{Max} = \sum_{i=3}^N (4i + 2)C_N^i, \quad (2)$$

где C_N^i – сочетание i элементов по N , N – общее количество входных сигналов, равное количеству нейронов во входном слое НС (рис. 1), i – текущее количество входных сигналов, используемых НС на данном этапе ($i = 1, 2, \dots, N$).

На *ключительном* этапе экспертным путем отбиралась наилучшая из нескольких конструкций НС, выбранных программой на основе описанного ранее условия максимума обучения. Выбор основывался на характере описания моделируемого параметра на тестовой выборке. Поскольку в основе выбора наилучшей конструкции НС лежит корреляция, то есть вероятность, что евклидово расстояние между фактическими и рассчитанными значениями будет велико (к примеру, превышать 2–3 сигмы). Отметим, что все конструкции НС с низкими значениями $r_{\text{обуч}}$ и $r_{\text{тест}}$ отсеиваются в соответствии со свойствами линейных НС [4]. Это дает возможность отфильтровать ошибку, возникшую на этапе предварительной обработки и связанную с использованием ортогональных входных сигналов.

Результаты и обсуждение

Для проверки работоспособности модели использовался многолетний (1951–2010 гг.) массив данных о среднемесячных объемах стока Дуная. Для моделирования его временного хода этот ряд данных (60 значений) был разделен на обучающую (30 значений, 1951–1980 гг.) и тестовую (30 значений, 1981–2010 гг.) выборки. Эти 30-летние периоды, как отмечалось выше, соответствуют масштабу климатических изменений. Результат моделирования объема речного стока в июне для трех различных конструкций НС приведен на рис. 2.

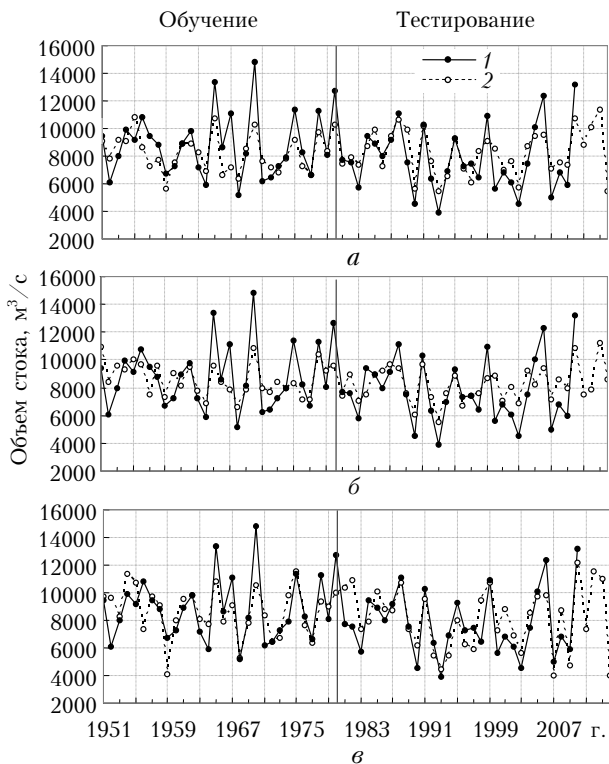


Рис. 2. Межгодовые изменения среднемесячного значения объема стока Дуная, измеренные в июне (1) и полученные в результате моделирования (2) в срок наблюдения (а) и с возможностью прогнозирования с заблаговременностью в 1 (б) и 2 мес (в)

На этапе тестирования все конструкции НС достаточно качественно описывают межгодовые изменения расхода стока Дуная. Первая конструкция (рис. 2, а) была сформирована на основе восьми индексов (АО-1-3, EA/WR-6, EP/NP-11, POL-5, SCAND-1-9-12; через дефис указан месяц используемого индекса) на входе. В скрытом слое было 32 нейрона, а количество итераций равнялось 385. Конструкция модели наиболее точно описывает наблюдаемые расходы, однако лишена возможности прогнозирования, так как в ее основе есть индекс месяца, который моделируется. Вторая конструкция (рис. 2, б) была сформирована на основе шести индексов (АО-2-3, POL-5, SCAND-1-9-12) на входе. В скрытом слое было 26 нейронов, а количе-

ство итераций равнялось 679. Конструкция модели сглаживает экстремальные значения объема речного стока. Третья конструкция (рис. 2, в) была сформирована на основе восьми индексов (АО-1-2-3-8, EP/NP-11, POL-7-10, SCAND-9) на входе. В скрытом слое находилось 24 нейрона, а количество итераций равнялось 810. Конструкция модели достаточно хорошо описывает ежегодные объемы речного стока, но не всегда отражает характер их временной изменчивости. Отметим, что в период с 2011 по 2014 г., когда отсутствовали данные по речному стоку, по всем трем конструкциям модели был получен подобный характер изменчивости восстановленных значений объема речного стока. Это также свидетельствует о корректной работе модели.

В таблице приведены результаты моделирования стока Дуная с различной заблаговременностью прогноза, определяемой количеством месяцев, за которое заранее можно составить модель. Хорошие результаты моделирования объема стока были получены в июне, июле, декабре и январе (коэффициент корреляции между измеренными и восстановленными значениями объема стока находился в пределах 0,81–0,85). Для февраля, марта, апреля и октября результаты также можно считать удовлетворительными (коэффициент корреляции был равен 0,76–0,77). В мае, августе, сентябре и ноябре модель описывала годовой ход объема стока несколько хуже (коэффициент корреляции был равен 0,71–0,75), что, видимо, было связано с ослаблением влияния некоторых климатических сигналов в системе океан–атмосфера на региональные метеосостояния, определяющие объем стока Дуная. Отметим, что качество моделирования падает с увеличением срока заблаговременности прогноза.

Для восстановления данных о прозрачности, содержащих пропуски, были использованы наиболее статистически обеспеченные наблюдения глубины видимости белого диска в мае и августе. Результаты восстановления ее значений приведены на рис. 3, где также выделены временные интервалы с использованными для обучения модели данными измерений. В отдельные годы наблюдаются различия в абсолютных величинах измеренных и полученных в результате моделирования среднемесячных значений глубины видимости белого диска. В целом отметим близость характера их межгодовых изменений. Для мая значение коэффициента корреляции с обучающей выборкой было равно 0,97, с тестовой выборкой – 0,94, а для августа значения этих коэффициентов равнялись 0,84 и 0,9 соответственно.

Для проверки качества результатов восстановления значений глубины видимости белого диска в годы отсутствия наблюдений (1951–1962, 1998–2013 гг.) и оценки характера межгодовых изменений этих значений может быть использована связь между прозрачностью и содержанием хлорофилла *a*. Известно, что концентрация хлорофилла *a* является важной количественной и качественной характеристикой фитопланктона, клетки которого поглощают и рассеивают свет и тем самым влияют

**Результаты моделирования среднемесячного объема стока Дуная
в срок наблюдения и с заблаговременностью 1–2 мес**

Месяц	Моделирование в срок наблюдения		Моделирование с заблаговременностью 1 мес		Моделирование с заблаговременностью 2 мес	
	$r_{\text{тест}}$	$r_{\text{обуч}}$	$r_{\text{тест}}$	$r_{\text{обуч}}$	$r_{\text{тест}}$	$r_{\text{обуч}}$
Январь	0,80	0,80	0,79	0,68	0,69	0,58
Февраль	0,75	0,72	0,75	0,72	0,75	0,72
Март	0,78	0,82	0,77	0,67	0,77	0,67
Апрель	0,75	0,71	0,75	0,71	0,64	0,49
Май	0,74	0,75	0,74	0,75	0,74	0,75
Июнь	0,82	0,70	0,77	0,65	0,76	0,63
Июль	0,78	0,82	0,80	0,71	0,80	0,71
Август	0,75	0,61	0,73	0,62	0,73	0,62
Сентябрь	0,71	0,63	0,71	0,63	0,71	0,63
Октябрь	0,73	0,74	0,70	0,72	0,64	0,60
Ноябрь	0,70	0,72	0,69	0,71	0,61	0,51
Декабрь	0,80	0,74	0,80	0,68	0,76	0,72

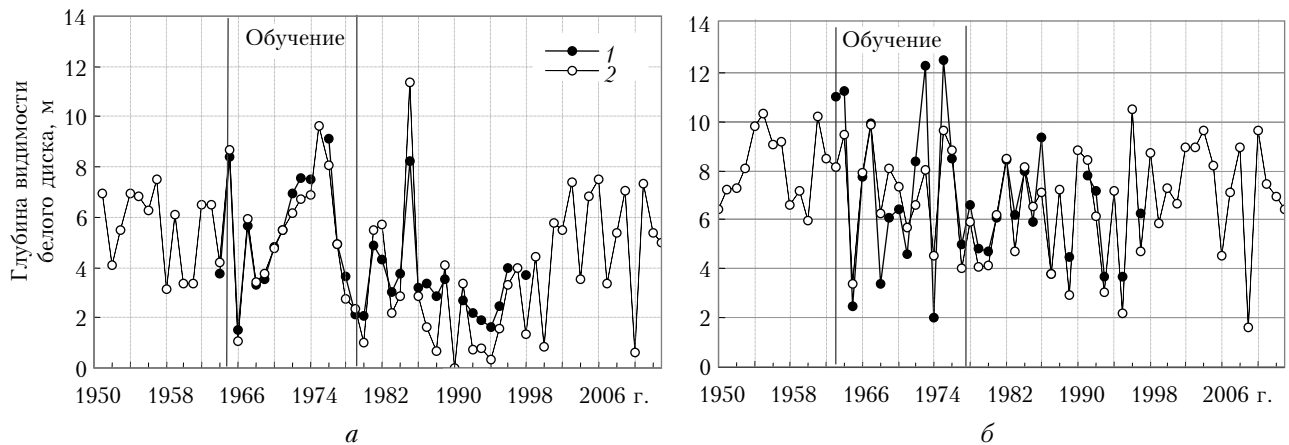


Рис. 3. Межгодовые изменения среднемесячных измеренных (1) и восстановленных (2) значений глубины видимости белого диска в мае (а) и августе (б)

на прозрачность морской воды. Таким образом, прозрачность и концентрация хлорофилла *a* связаны между собой обратной зависимостью. Проведенный в работе [3] регрессионный анализ показал наличие достаточно тесной линейной связи между показателем ослабления направленного света и концентрацией хлорофилла *a*. Поэтому сравнение характера межгодовых изменений восстановленных значений глубины видимости белого диска и концентрации хлорофилла *a* позволяет в первом приближении качественно оценить результаты моделирования. Однако в 1951–1962 гг. измерения концентрации хлорофилла *a* были крайне ограниченными, что не позволяет провести такое сравнение. В то же время известно о влиянии региональных метеоусловий на интенсивность развития фитопланктона в западном районе СЗЧМ, а следовательно, и на прозрачность [3]. Так, пониженные восстановленные значения глубины видимости белого диска были получены в мае 1958, 1960 и 1961 гг. (рис. 3, а). В эти годы зимы были теплыми, а в апреле и мае зарегистрирован выше среднего многолетнего для этих месяцев объем стока Дуная, что способствовало активному развитию фитопланктона. Повышенные значения глубины видимости белого диска были

получены в 1951, 1954, 1956, 1957 и 1962 гг., когда зимы были относительно холодными, что не способствовало развитию фитопланктона в весенний период.

Для оценки характера межгодовых изменений значений глубины видимости белого диска в исследуемом районе СЗЧМ в 1998–2010 гг. были использованы данные спутниковых наблюдений (цветовой сканер Sea WiFS) среднемесячной концентрации хлорофилла *a*. Приведенные на рис. 4 межгодовые изменения нормированных значений аномалий концентрации хлорофилла *a* и глубины видимости белого диска показали противофазный характер изменений этих показателей. В годы, когда концентрация хлорофилла *a* была повышенной (положительная аномалия), восстановленные значения прозрачности были пониженными (отрицательная аномалия), и наоборот. Коэффициент корреляции между аномалиями этих показателей в мае составил $-0,41$, в августе $-0,62$. В то же время коэффициенты корреляции, рассчитанные по одновременно измеренным в СЗЧМ в те же сезоны года значениям глубины видимости белого диска и концентрации хлорофилла *a*, составляли в среднем от $-0,7$ до $-0,85$.

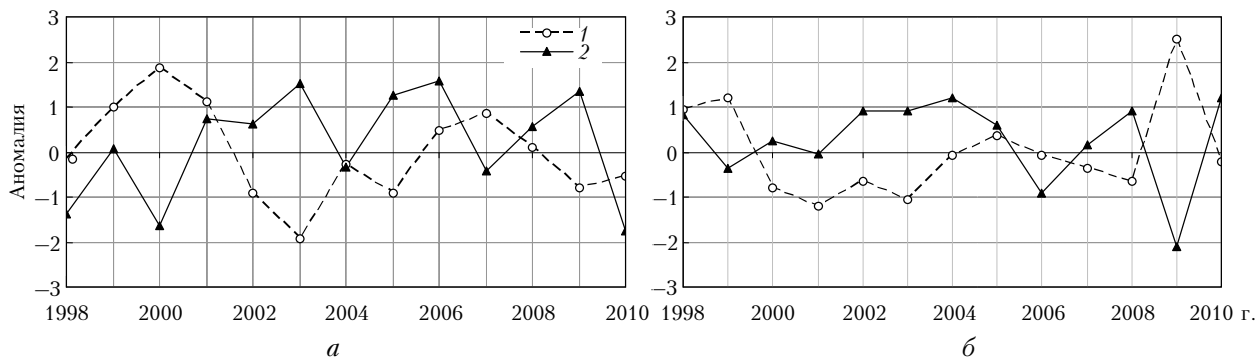


Рис. 4. Межгодовые изменения аномалии среднемесячных значений концентрации хлорофилла a (1) и глубины видимости белого диска (2) в мае (а) и в августе (б) в 1998–2010 гг.

Заклучение

Для восстановления пропущенных в различные годы данных измерений прозрачности (глубины видимости белого диска), необходимых для изучения ее межгодовых изменений, была успешно применена модель нейронной сети. Модель представляла собой однослойный перцептрон с учителем, в качестве которого использовалась часть ряда данных наблюдений, а ее обучение осуществлялось методом обратного распространения ошибки. В процессе многоступенчатого обучения модели на каждом шаге рассчитывался коэффициент корреляции между данными наблюдений и результатом модели. Этот процесс проводился до момента определения максимального значения коэффициента корреляции, что свидетельствовало о наиболее точном соответствии результатов модели с моделируемым параметром. В качестве входных сигналов модели использовались индексы глобальных климатических колебаний в системе океан–атмосфера, изменения которых определяли изменения гидрометеорологических, гидрологических и гидробиологических условий в западном районе СЗЧМ. Это, в свою очередь, влияло на изменчивость прозрачности вод в этом районе и позволяло не только восстанавливать пропущенные значения прозрачности, но и прогнозировать ее изменения.

В процессе практического использования модели нейронной сети была отработана процедура адаптации, состоящая из трех этапов. На этапе предварительной обработки выполнялся поиск сигналов, коррелированных с моделируемым параметром. Этап моделирования включал расчет эмпирических выражений всех конструкций НС, определяемых возможными комбинациями входных сигналов с различным количеством нейронов скрытого слоя. На заключительном этапе выбиралась конструкция НС, наиболее точно описывающая моделируемый параметр.

Работоспособность предложенной модели на основе нейронной сети была проверена с использованием массива многолетних данных регулярных среднемесячных наблюдений объема стока Дуная. Близость этих данных и полученных при моделировании среднемесячных значений объемов стока (коэффициент корреляции между ними в годовом

цикле находился в пределах 0,70–0,82) позволила оценить качество моделирования как удовлетворительное.

Предложенная модель была применена для восстановления среднемесячных значений глубины видимости белого диска в годы отсутствия измерений. Результат моделирования был сопоставлен с данными наблюдений глубины видимости белого диска с 1978 по 1998 г. (тестовая выборка), а также данными о концентрации хлорофилла a (по спутниковым наблюдениям) в период 1998–2010 гг., когда отсутствовали судовые наблюдения прозрачности. Коэффициент корреляции между восстановленными и измеренными значениями глубины видимости белого диска в одни и те же месяцы и годы в среднем был равен 0,9. Коэффициент корреляции между значениями межгодовых аномалий глубины видимости и аномалий концентрации хлорофилла a в мае был равен $-0,41$, в августе $-0,62$. Качество восстановленных данных можно считать удовлетворительным.

Использование в качестве входных сигналов модели индексов дальнедействующих атмосферных колебаний в системе океан–атмосфера дало возможность прогнозировать моделируемые величины с заблаговременностью от одного до нескольких месяцев. С увеличением заблаговременности прогноза качество моделирования существенно снижалось.

Моделирование с помощью НС имеет ограничение, связанное с количеством используемых данных. Для качественного моделирования климатических величин следует использовать не менее 60 значений (два климатических периода). При этом как минимум половина используемых для обучающей выборки значений должна быть последовательно расположена по времени. Минимальное количество данных составляет один климатический период, равный 30 годам.

Описанный метод может быть использован для восстановления различных гидрофизических и гидробиологических характеристик, изменчивость распределений которых зависит от изменчивости глобальных атмосферных процессов в системе океан–атмосфера.

Работа выполнена при финансовой поддержке гранта РФФИ (проект № 14-45-01549-юг-а).

1. Кукушкин А.С. Многолетняя и сезонная изменчивость распределения прозрачности в поверхностных водах северо-западной части Черного моря // Оптика атмосф. и океана. 2013. Т. 26, № 2. С. 113–123.
2. Кукушкин А.С., Воскресенская Е.Н., Маслова В.Н. Особенности формирования поля прозрачности в поверхностном слое северо-западной части Черного моря в зимне-весенний период // Оптика атмосф. и океана. 2010. Т. 23, № 8. С. 730–736; Kukushkin A.S., Voskresenskaya E.N., Maslova V.N. The Features of Transparency Field Generation in the Upper Sea Layer of the Northwestern Black Sea in the Winter-Spring Period // Atmos. Ocean. Opt. 2011. V. 24, N 1. P. 95–101.
3. Кукушкин А.С. Влияние распределений гидрологических и гидробиологических параметров на структуру поля прозрачности в верхнем слое пелагиали Черного моря // Метеорол. и гидрол. 2013. № 7. С. 74–86.
4. Осовский С. Нейронные сети для обработки информации. М.: Финансы и статистика, 2002. 344 с.
5. Ефимов В.В., Посошков В.Л. Применение метода искусственных нейронных сетей для регионализации прогноза осадков в Причерноморском регионе // Морской гидрофиз. ж. 2006. № 3. С. 23–35.
6. Посошков В.Л. Использование искусственных нейронных сетей для статистического оценивания ряда среднесуточных осадков // Системы контроля окружающей среды. Севастополь: МГИ НАН Украины, 2002. С. 146–152.
7. Посошков В.Л., Прусов А.В. Применение метода самоорганизующихся отображений признаков к анализу экстремальных осадков // Системы контроля окружающей среды. Севастополь: МГИ НАН Украины, 2003. С. 118–123.
8. Ефимов В.В., Шокуров М.В. Использование искусственной нейронной сети для статистического оценивания температуры поверхности Черного моря // Труды Междунар. конф., посвященной 75-летию академика Г.И. Марчука и 20-летию Института вычислительной математики. М.: ИВМ РАН, 2000. Т. 2. С. 267–274.
9. Полонский А.Б., Воскресенская Е.Н., Посошков В.Л. Статистический прогноз среднемесячного стока черноморских рек на основе циркуляционных атмосферных процессов // Доповіді НАН України. 2010. № 11, С. 95–101.
10. Лубков А.С., Воскресенская Е.Н., Кукушкин А.С. Восстановление значений прозрачности морской воды на основе моделирования методом нейронных сетей // XX Междунар. симп. «Оптика атмосферы и океана. Физика атмосферы». Сборник трудов [Электронный ресурс]. Томск: ИОА СО РАН, 2014. С. 238–241.
11. Northern Hemisphere Teleconnection Patterns // Climate Prediction Center. URL: <http://www.cpc.ncep.noaa.gov/data/teledoc/telecontents.shtml> (дата обращения 10.12.2014).
12. Pacific Decadal Oscillation (PDO) / Joint Institute for the Study of the Atmosphere and Ocean (JISAO). URL: <http://jisao.washington.edu/pdo/PDO.latest> (дата обращения 10.12.2014).
13. Southern Oscillation Index. Archives / Australia's official weather forecasts & weather radar Bureau of Meteorology. URL: <http://www.bom.gov.au/climate/current/soihtm1.shtml> (дата обращения 10.12.2014).
14. Haykin S. Neural networks, a comprehensive foundation. N.Y.: Macmillan College Publishing Company, 1994. 823 p.
15. Rosenblatt F. Principles of neurodynamics: Perceptrons and the theory of brain mechanisms. N.Y.: Spartan, 1962. 626 p.
16. Мочин А.С. Гидродинамика атмосферы, океана и земных недр. СПб.: Гидрометеоздат, 1999. 524 с.
17. Muleta M.K. Model performance sensitivity to objective function during automated calibrations // J. Hydrologic Engin. 2012. V. 17, N 6. С. 756–767.
18. Колмогоров А.Н. О представлении непрерывных функций нескольких переменных в виде суперпозиций непрерывных функций одного переменного и сложения // Докл. АН СССР. 1957. Т. 114, № 5. С. 953–956.
19. Hecht-Nielsen R. Neurocomputing. Amsterdam: Addison Wesley, 1991. 433 с.

A.S. Lubkov, E.N. Voskresenskaya, A.S. Kukushkin. Method for reconstruction of monthly average data on water transparency for the North-West Black Sea.

To reconstruct the data of hydrophysical parameter observations, a model based on a neural network with the teacher has been suggested. The indices of global climate modes of the ocean–atmosphere system were applied as the model input. The processes of the model teaching and adaptation, which allow one to find the most accurate solution of the problem of modeling, is described. The comparison of modeled monthly average Danube's runoff with data of observation demonstrated their good correspondence. Missed in some years observation data on sea water transparency (depth of white disc visibility) were reconstructed. The proximity of reconstructed and observed depths of white disc visibility was noted. Some features of interannual variability of reconstructed data on sea water transparency caused by both climatic factors during 1950–1962 and changes in the chlorophyll *a* concentration during 1998–2010 were found.