

ВОССТАНОВЛЕНИЕ КЛИМАТИЧЕСКИХ ДАННЫХ ГИДРООПТИЧЕСКИХ ХАРАКТЕРИСТИК НА ОСНОВЕ МОДЕЛИРОВАНИЯ

*А.С. Лубков, Е.Н. Воскресенская,
А.С. Кукушкин*

Морской гидрофизический институт
НАН Украины
г. Севастополь, ул. Капитанская, 2
E-mail: andrey-ls@bigmir.net

В статье приводятся результаты восстановления массива глубины видимости диска Секки по данным наблюдений с 1963 по 2000 годы при помощи метода нейронных сетей. Дан анализ полученных результатов, как на статистическом основании, так и на основании сравнения с концентрацией хлорофилла.

Введение. Прозрачность воды, помимо традиционно используемых гидрологических и гидрохимических характеристик вод, может служить хорошим индикатором распространения загрязненных береговых стоков в море из-за больших различий в концентрациях взвешенных и растворенных веществ в этих стоках и морских водах, сохраняющихся на значительных удалениях от берега [1]. Поэтому изучение оптического режима вод в прибрежной и шельфовой зонах Черного моря в условиях их прогрессирующего загрязнения и эвтрофикации представляет несомненный интерес.

Относительно немногочисленные данные по измерению прозрачности вод в северо-западной части Черного моря, неравномерно распределенные по годам и сезонам, не позволили подробно и с достаточной степенью достоверности рассмотреть ее межгодовую изменчивость [2]. Компенсировать их отсутствие возможно расчетным путем с использованием модельных расчетов.

Данные. В работе использовались среднемесячные данные измерений глубины видимости диска Секки за период с 1963 по 2000 годы в Придунайском районе Черного моря. При этом в результате нерегулярности выполнения

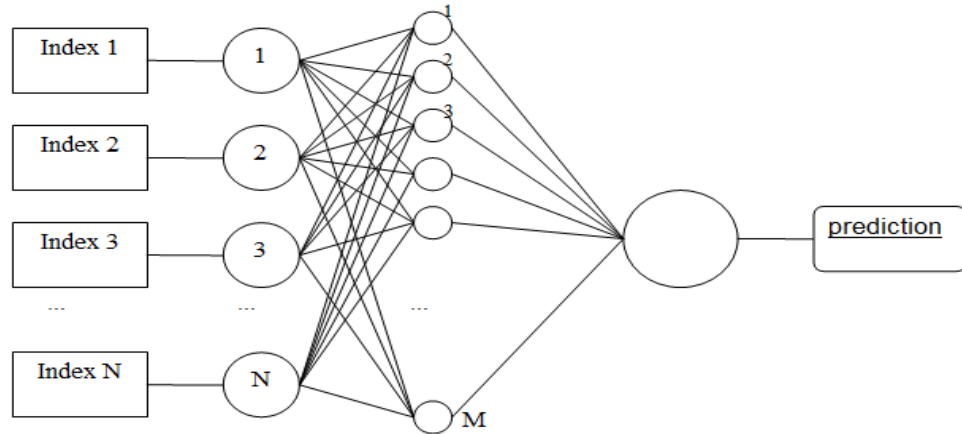
экспедиционных работ в указанном регионе в рядах данных встречаются пропуски. Тем не менее, для установления закономерностей формирования аномалий оптических характеристик важно иметь непрерывные ряды данных. Для этого была предпринята попытка восстановить ряд на основе имеющихся данных. В качестве предикторов использовались климатические сигналы системы океан-атмосфера, которые включали индексы следующих осцилляций, представляющих собой нормированные разности атмосферного давления/температуры поверхности океана в соответствующих специфических регионах: NAO (северо-Атлантическое колебание), AO (Арктическое колебание), EA (Восточно-Атлантическое колебание), EA/WR (осцилляция между восточной Атлантикой и западной Россией), SCAND (Скандинавское колебание), P/E (Полярно-Евразийское колебание), PNA (Тихоокеанско-Североамериканское колебание), WP (колебание на западе Тихого океана), SOI (Южное колебание) и PDO (Тихоокеанская декадная колебание), взятые с сайтов [4, 5, 6].

Методика. Для восстановления массивов данных использовалась нейронная сеть (НС). Ее тип представляет собой однослойный перцептрон с учителем. Обучение НС осуществлялось методом обратного распространения ошибки. В качестве входных сигналов взяты индексы перечисленных выше глобальных осцилляций с опережением на 1 и более месяцев проявления их в анализируемом гидрофизическом параметре Черноморского региона. На выходе НС имеет один нейрон (рис. 1). В качестве учителя использовалась часть ряда данных наблюдений глубины видимости диска Секки. На основе имеющихся входных данных, НС может не только *восстанавливать* ряд, но и *прогнозировать* его изменения. Функция активации, используемая в модели, – сигмоидальная биполярная: $f(x) = \tanh(\beta x)$. Подробно НС данной архитектуры описана в [3].

В настоящей работе на предварительном этапе выполнялся поиск связей между индексами и глубиной видимости диска. В ходе обучения НС на каждом шаге итера-

ций оценивались и сопоставлялись коэффициенты корреляции между данными наблюдений и продуктом модели. После завершения цикла определялся максимальный коэффициент

корреляции и соответствующей ей номер итерации, на основе которого далее строили НС. Величина коэффициента корреляции используется для процесса отбора лучшего результата.



Входные данные	Функция Активации	Скрытый слой	Выходной сигнал	Преобразованный выходной сигнал
Атмосферные индексы и осадки	Биполярная $f(x) = \tanh(\beta x)$	Суммирование входного слоя, формирование активности нейронов	Суммирование скрытого слоя	Результирующая $F(x) = \sqrt{\ln\left(\frac{-x}{x-1} - \frac{1}{x-1}\right)}$

Рис. 1. Структура однослойного персептрона с одним нейроном на выходе, Index - используемый предиктор, 1..N – преобразованные сигналы предикторов, 1..M – скрытый слой, prediction – обратно преобразованный сигнал выходного нейрона

Практическое решение задачи показало, что, поведение НС не всегда предсказуемо. Дело в том, что, используя одни и те же предикторы, часто можно получить существенное различие прогнозов, меняя количество нейронов скрытого слоя или стартовые веса. В то же время, не все предикторы, показавшие возможную связь с глубиной видимости, следует использовать одновременно из-за возможной неортогональности.

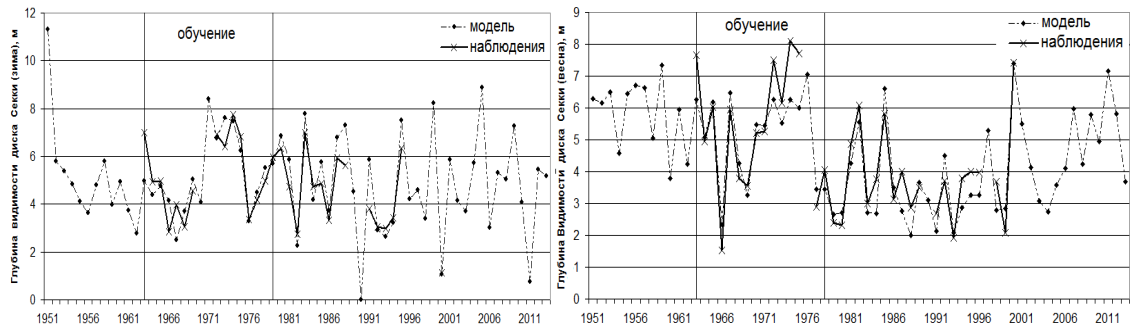
Поэтому в ходе работы были перебраны все возможные комбинации предикторов и для каждой из них находились максимальные величины корреляции и соответствующие им итерации. Все стартовые веса приравнены к константе, величина которой должна лежать в пределах от $-\sqrt{M}$ до \sqrt{M} где M – количество нейронов скрытого слоя, n – количество входных сигналов в нейрон скрытого слоя [3], однако, поскольку количество нейронов в скрытом слое и входных сигна-

лов в нейрон этого слоя в нашем случае не постоянно, были взяты средние значения, лежащие в диапазоне от -1,3 до 1,3. В работе стартовая постоянная весов связей составляла 0,6. По итогам анализа отображены самые удачные конструкции НС, на основе которых восстанавливались массивы данных. При этом получено, что в случае ограниченного количества осредненных за месяц исходных данных не всегда возможно восстановить среднемесячные ряды. Минимальное количество, необходимое для обучения, должно составлять 13-15 значений, а для достаточно надежного теста – не менее 15. Однако, зачастую ряды среднемесячных данных содержат немалое количество пропусков в наблюдениях. Поэтому, в случаях недостатка исходных данных в некоторых среднемесячных рядах, приходилось восстанавливать ряды по сезонам. В нашем случае рассматривались гидрологические сезоны: зима представлена месяцами январь, февраль, март, весна – апрель, май, июнь и так далее. Поскольку средний для осени

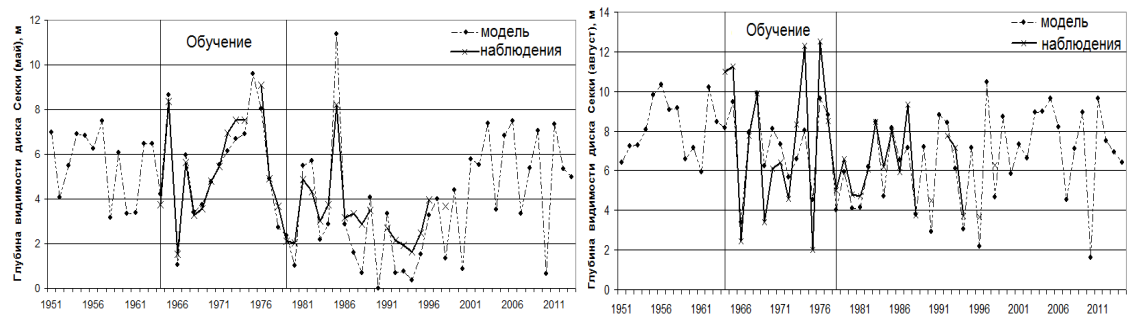
ряд составил всего 25 значений, его восстановление не проводилось. Аналогичная ситуация складывается и со среднемесячными рядами в осенний период.

Результаты и их анализ. Восстановление данных прозрачности морской воды проводилось для сезонов и отдельно для нескольких месяцев, в достаточной степени обеспеченных данными. На рис. 2 и 3 приведены не-

которые результаты восстановления. Оценка качества результатов восстановления выполнялась сопоставлением с данными наблюдений, приведена в таблице 1. Необходимо отметить, что для повышения качества восстановления следует брать промежутки времени не менее двух климатических периодов: один период для обучения, второй период для тестирования модели.



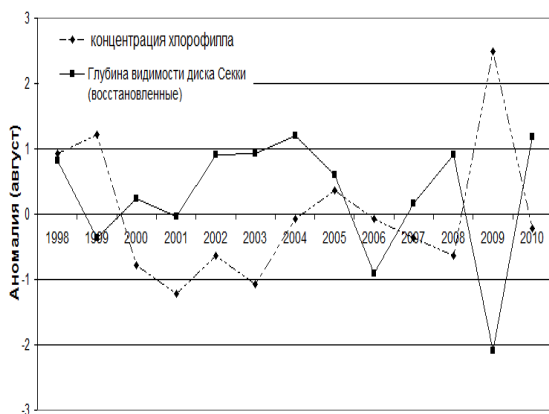
Р и с. 2. Годовая изменчивость усредненной за зимний сезон (слева) и за весенний сезон (справа) наблюдаемой (сплошная линия) и восстановленной (штрихпунктирная линия) глубины видимости диска Секки



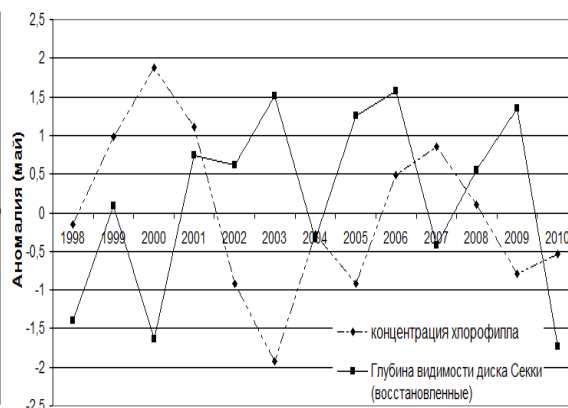
Р и с. 3. Годовая изменчивость среднемесячной наблюдаемой (сплошная линия) и восстановленной (штрихпунктирная линия) глубины видимости диска Секки в мае (слева) и в августе (справа)

Поскольку восстановление проводилось по данным в диапазоне с 1963 по 2000 годы, то промежутки с 1950 по 1962 и с 2001 по 2013 годы, для которых данные наблюдений отсутствовали, следует проверить. Для проверки использовались спутниковые данные (сканер Sea WiFS) среднемесячных значений концентрации хлорофилла, наблюдаемые в исследуемом районе, которые в основном определяют про-

зрачность воды [2]. Приведенные на рис. 4(а) и 4(б) межгодовые изменения нормированных аномалий концентрации хлорофилла и глубины видимости диска Секки показали, что ход этих изменений находится в противофазе. Увеличение концентрации хлорофилла (положительная аномалия) приводит к понижению прозрачности (отрицательная аномалия) и наоборот. Коэффициент корреляции между ними в мае составил $-0,41$, в августе $-0,62$.



Р и с. 4 (а). Среднемесячные аномалии концентрации хлорофилла и глубины видимости диска Секки за период с 1998 по 2010 годы в августе



Р и с. 4 (б). Среднемесячные аномалии концентрации хлорофилла и глубины видимости диска Секки за период с 1998 по 2010 годы в мае

Т а б л и ц а 1

Оценка результатов моделирования

Сезон/месяц	тест	Длина ряда тестовой выборки	Обучение	Длина ряда обучающей выборки	Среднеквадратическое отклонение модели, м
Зима	0,95	15	0,83	15	1,48
Весна	0,90	20	0,92	15	0,88
Лето	0,91	22	0,81	15	0,89
Май	0,94	17	0,97	15	1,44
Август	0,90	16	0,84	15	1,06
Сентябрь	0,94	15	0,91	13	1,21

Заключение. В работе восстановлены ряды глубины видимости диска Секки методом нейронных сетей для некоторых сезонов и месяцев с достаточной, для корректной работы модели, обеспеченностью. Сравнение восстановленных рядов прозрачности и измеренной концентрации хлорофилла, определяющей прозрачность в морской среде, показало приемлемое для практики качество такого восстановления.

Более качественное восстановление возможно при условии использования для анализа промежутка времени, не менее двух климатических периодов.

СПИСОК ЛИТЕРАТУРЫ

1. Агафонов Е.А., Кукушкин А.С., Прохоренко Ю.А. Исследование распространения речных вод в северо-

западной части Черного моря в весенний гидрологический сезон // Метеорология. и гидрология, 2000, № 8. – С. 93 – 101.

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 атмосферы и океана, 2010, т. 23, № 8. – С. 730 – 736.
3. Осовский С. Нейронные сети для обработки информации – М.: Финансы и статистика, 2002. – 344 с.
4. <http://www.cpc.ncep.noaa.gov/data/teledoc/telecontents.shtml>
5. <http://www.cgd.ucar.edu/cas/catalog/climind/SOI.signal.ascii>
6. <http://jisao.washington.edu/pdo/PDO.latest>