

УДК 519.95

ПРОГНОЗИРОВАНИЕ ЧИСЛЕННОСТИ ПОПУЛЯЦИИ КРЕВЕТКИ В ОТКРЫТОЙ СИСТЕМЕ С ПОМОЩЬЮ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ

© А.А. Арзамасцев, А.С. Козадаев

Arzamastsev A.A., Kozadayev A.S. Forecasting the size of shrimp population in an open system with the aid of artificial neural networks. The article looks at the possibility to forecast open biological system with the aid of artificial neural networks on the basis of the experimental calculations undertaken by the authors.

ВВЕДЕНИЕ

Моделированию и прогнозированию развития биологических популяций в естественных условиях и биотехнологиях посвящено значительное количество публикаций [1–5]. Целью данной работы является оценка возможности использования аппарата искусственных нейронных сетей для прогнозирования плотности популяции креветок в открытой системе. Прогноз плотности популяций в открытых системах представляет значительную проблему ввиду того, что «территория» обитания такой популяции обычно не ограничена никакими естественными барьерами, она может свободно перемещаться как по самой рассматриваемой области, так и за ее пределы. Поэтому разработка методов прогноза в этих условиях представляет собой актуальную задачу.

МАТЕРИАЛЫ И МЕТОДЫ

Объект исследования – биоценоз, включающий смешанную популяцию креветок следующих видов: *Peneaeus indicus*, *Metapeneaeus monoceros*, *Peneaeus monodon*, *Peneaeus japonicus*. Место сбора экспериментальных данных – район промышленного лова Banco da Sofala, Республика Мозамбик, Индийский океан, 1500–2000 км от Мапуту. Экспериментальные данные любезно предоставлены компанией EFRIPPEL в виде электронных таблиц Microsoft Excel. Таблицы содержали данные по отлову креветки за 1996–1998 годы: даты начала и окончания циклов отлова (10 дней). Каждый год отлов велся в период с 1 марта по 31 декабря. К плотности популяции приравнено значение, равное массе выловленной в один цикл креветки разных видов (суммарная масса, табл. 1).

В настоящее время для решения задач прогнозирования временных рядов с успехом используется аппарат искусственных нейронных сетей (ИНС).

Для исследования и получения предположительных значений плотности популяции на 1999 год использовалась программа NNC v.3.02 (программа построения, настройки и эксплуатации нейронных сетей). Она написана на языке высокого уровня DELPHI III и работа-

ет под управлением операционной системы Microsoft Windows на IBM-совместимых компьютерах.

Таблица 1

Данные по плотности популяции креветки
за 1996–1998 годы

1996 год		1997 год		1998 год	
Дата начала интервала	Плотность популяции	Дата начала интервала	Плотность популяции	Дата начала интервала	Плотность популяции
01.03.96	196,25	01.03.97	238,83	01.03.98	196,66
11.03.96	135,95	11.03.97	181,59	11.03.98	238,99
21.03.96	141,44	21.03.97	182,53	21.03.98	155,79
01.04.96	172,25	01.04.97	130,84	01.04.98	133,72
11.04.96	180,80	11.04.97	224,45	11.04.98	156,31
21.04.96	173,70	21.04.97	223,44	21.04.98	165,27
01.05.96	136,72	01.05.97	173,77	01.05.98	155,65
11.05.96	120,55	11.05.97	186,96	11.05.98	136,73
21.05.96	133,11	21.05.97	171,37	21.05.98	128,73
01.06.96	98,65	01.06.97	179,78	01.06.98	114,19
11.06.96	165,35	21.06.97	129,84	11.06.98	92,56
21.06.96	160,15	01.07.97	115,41	21.06.98	104,10
01.07.96	126,32	11.07.97	102,66	01.07.98	90,21
11.07.96	121,63	21.07.97	127,06	11.07.98	72,87
21.07.96	109,52	01.08.97	138,63	21.07.98	81,95
01.08.96	83,94	11.08.97	119,12	01.08.98	63,66
11.08.96	73,24	21.08.97	89,62	11.08.98	53,12
21.08.96	70,43	01.09.97	70,53	21.08.98	59,49
01.09.96	86,14	11.09.97	49,18	01.09.98	43,85
11.09.96	68,50	21.09.97	47,30	11.09.98	57,57
21.09.96	60,37	01.10.97	42,91	21.09.98	53,26
01.10.96	74,18	11.10.97	39,05	01.10.98	34,00
11.10.96	50,00	21.10.97	55,06	11.10.98	41,16
21.10.96	45,92	01.11.97	49,06	21.10.98	49,74
01.11.96	73,71	11.11.97	49,42	01.11.98	66,28
11.11.96	46,74	21.11.97	39,53	11.11.98	67,63
21.11.96	54,05	01.12.97	50,04	21.11.98	76,06
01.12.96	52,18	11.12.97	57,07	01.12.98	51,34
11.12.96	56,71	21.12.97	133,00	11.12.98	81,28
21.12.96	94,77			21.12.98	127,84

Используя NNC, пользователь может создавать, обучать и использовать нейронные сети. NNC v.3.02 – свободное программное обеспечение, оно доступно в сети Интернет (адрес <http://vkrepets.chat.ru/NNC.exe>). Данная программа позволяет использовать до 60 входных нейронов, варьировать функцией нейрона, количеством слоев и связей сети. Она содержит в себе готовые алгоритмы обучения по следующим методам: по координатный спуск, случайный поиск, метод Ньютона, статистический градиентный метод, комбинации этих методов.

Вычислительный эксперимент проходил в два этапа. Первый включал в себя подготовительные операции, второй непосредственно эксперимент по прогнозированию численных значений плотности популяции на 1999 год.

Подготовительные операции включали в себя приведение исходных данных в удобный для исследования вид и их анализ, предварительный расчет. Предварительный расчет заключался в построении нейронных сетей с учетом данных только на 1996 год, с тем, чтобы получить значение плотности популяции на март 1997 года. Были построены 4 различные нейронные сети, с идентичными исходными данными и параметрами: параметр точности ($1 \cdot 10^{-5}$), параметр показателя степени метрического пространства (евклидово), одинаковым алгоритмом минимизации (метод покоординатного спуска). Различия заключались в использовании различных нейронов (и их количества), хотя во всех случаях не изменялось количество слоев.

В первом случае была построена нейронная сеть с шестью независимыми переменными и одной зависимой, одним нейроном типа «Constant». Была использована линейная функция состояния: $f(x) = x$ (ИНС первого типа). Вторая нейронная сеть содержала нейроны с функциями состояний: $f(x) = x$ и $f(x) = x^2$ (ИНС второго типа), третья: $f(x) = x$; $f(x) = x^2$ и $f(x) = x^3$ (ИНС третьего типа), и четвертая: $f(x) = x$; $f(x) = x^2$; $f(x) = x^3$; $f(x) = x/(1+|x|)$ (ИНС четвертого типа). Анализ результатов показал, что ИНС четвертого типа в большей степени подходит для решения данной задачи. В случае использования этой сети была получена наименьшая относительная погрешность, составляющая 8,89 %.

Затем эксперимент был повторен, но в качестве исходных данных были использованы значения за неполных три года (01.03.96 – 21.10.98), рассчитывалось значение плотности популяции на 21.12.98 для того, чтобы проверить, как повлияет изменение количества исходных данных на точность результата, как сильно расчетное значение будет отличаться от известного. Были использованы нейронные сети тех же типов. Анализ результатов подтвердил правильность выбора ИНС четвертого типа. Относительная погрешность в этом случае составила 5,74 %.

Структура сети в виде графа показана на рис. 1.

Коэффициенты передачи используемой ИНС показаны на рис. 2.

Таким образом, был выбран тип нейронной сети. Было показано также, что при увеличении количества исходных данных точность расчетных значений возрастает.

Для удобства проведения дальнейших экспериментов представим время в абсолютных значениях, т. е. в виде временного ряда (табл. 2).

В графе «Номер интервала» приведены номера 10 дневных интервалов отлова креветки. Нулевые значения в графе «Плотность популяции» означают, что в этот период времени отлов не осуществлялся.

Используя значения табл. 2, построены графики зависимости плотности популяции от времени (см. рис. 3–5). Для наблюдения периодичности явления был построен суммарный график за три года (рис. 6).

Анализ исходных данных позволил сделать вывод о периодическом характере изменения плотности популяции креветки с течением времени, поскольку использование автокорреляционной функции для исследования невозможно из-за имеющихся разрывов (рис. 4). На данном этапе исследования были рассчитаны коэффициенты корреляции для временных рядов (табл. 3).

Наибольшее значение коэффициента корреляции $\sim 0,87$. Это обстоятельство указывает на идентичность процессов, происходящих в объекте, в одни и те же календарные периоды, и относительное постоянство результирующей всех внешних воздействий в эти периоды.

Наблюдая периодический характер и большой «разброс» значений плотности популяции, была произведена попытка аппроксимировать данные. Результаты представлены на рис. 7.

Полученная кривая доказывает сезонность (периодичность) изменения плотности популяции, с периодом ~ 1 год. Для аппроксимации использовалась программа CurveExpert 1.34 © Microsoft Corporation.

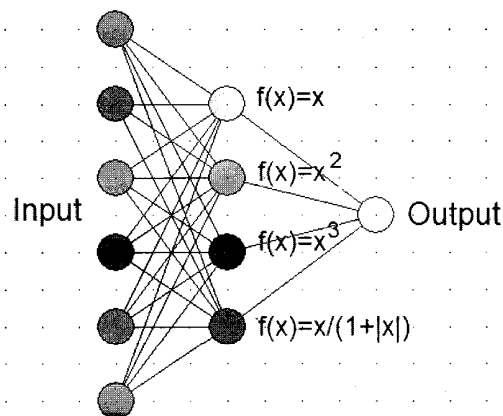


Рис. 1. Структура используемой сети

#	x	y	TYPE	DESCRIPTION	linkto	linkto	linkto	linkto	linkto
1	3	1	Input						
2	3	2	Input						
3	3	3	Input						
4	3	4	Input						
5	3	5	Input						
6	3	6	Input						
7	7	3	f=x			1			
8	11	6	OUT						
9	7	5	f=x^2			1,26687			
10	7	7	f=x^3			1			
11	7	9	Sigmoid			2,85252			

Рис. 2. Матрица инцидентности графа нейронной сети

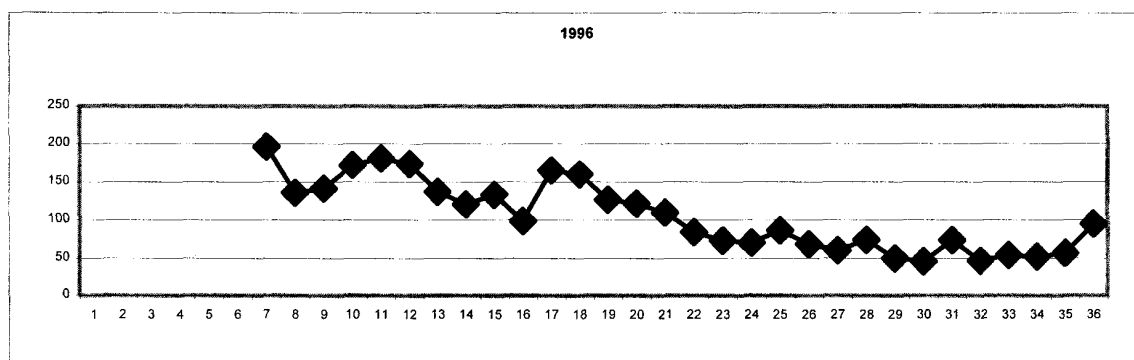


Рис. 3. Временной ряд плотности популяции (1996 год)

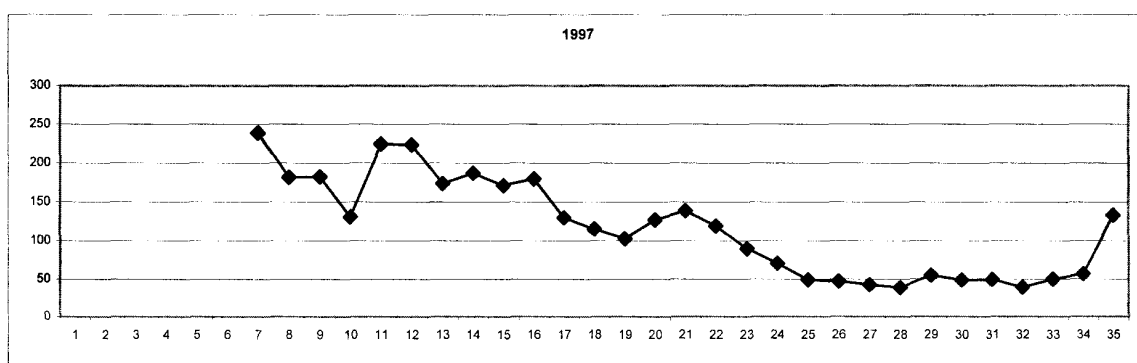


Рис. 4. 3 Временной ряд плотности популяции (1997 год)

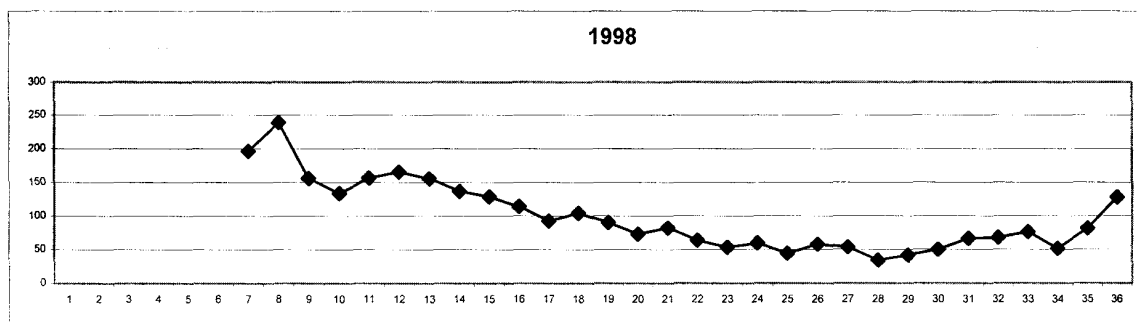


Рис. 5. Временной ряд плотности популяции (1998 год)

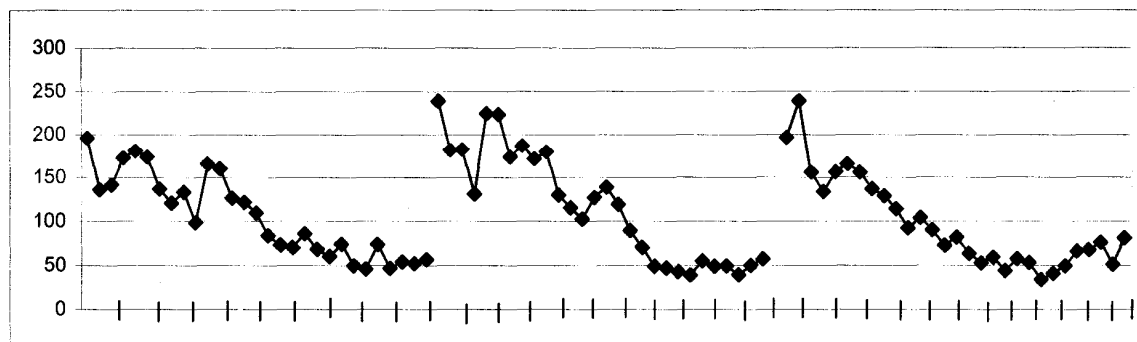


Рис. 6. Временной ряд плотности популяции (за три года)

Таблица 2

Временной ряд. Плотность популяции.
1996–1998 годы

1996 год		1997 год		1998 год	
Номер интервала	Плотность популяции	Номер интервала	Плотность популяции	Номер интервала	Плотность популяции
1	0	1	0	1	0
2	0	2	0	2	0
3	0	3	0	3	0
4	0	4	0	4	0
5	0	5	0	5	0
6	196,25	6	238,83	6	196,66
7	135,95	7	181,59	7	238,99
8	141,44	8	182,53	8	155,79
9	172,25	9	130,84	9	133,72
10	180,80	10	224,45	10	156,31
11	173,70	11	223,44	11	165,27
12	136,72	12	173,77	12	155,65
13	120,55	13	186,96	13	136,73
14	133,11	14	171,37	14	128,73
15	98,65	15	179,78	15	114,19
16	165,35	16	129,84	16	92,56
17	160,15	17	115,41	17	104,10
18	126,32	18	102,66	18	90,21
19	121,63	19	127,06	19	72,87
20	109,52	20	138,63	20	81,95
21	83,94	21	119,12	21	63,66
22	73,24	22	89,62	22	53,12
23	70,43	23	70,53	23	59,49
24	86,14	24	49,18	24	43,85
25	68,50	25	47,30	25	57,57
26	60,37	26	42,91	26	53,26
27	74,18	27	39,05	27	34,00
28	50,00	28	55,06	28	41,16
29	45,92	29	49,06	29	49,74
30	73,71	30	49,42	30	66,28
31	46,74	31	39,53	31	67,63
32	54,05	32	50,04	32	76,06
33	52,18	33	57,07	33	51,34
34	56,71	34	133,00	34	81,28
35	94,77	35		35	127,84

Таблица 3

Коэффициенты корреляции (исходные данные)

Годы	1996	1997	1998
1996	1		
1997	0,821	1	
1998	0,777	0,867	1

Последующий расчет производился только с помощью программы NNC v.3.02.

Для определения адекватности результатов вычислительного эксперимента был произведен расчет значений плотности популяции на 1998 год, с использованием данных за период времени 1996–1997 годы. Матрица значений обучающей выборки была построена, как показано в табл. 4. Все значения плотности популяции были представлены в следующем виде:

$x_1; x_2; x_3; \dots; x_m; \dots; x_{29}; \quad x_{30}; x_{31}; x_{32}; \dots; x_n; \dots; x_{58}$
(исключены нулевые значения)

данные за 1996 год

данные за 1996 год,

где $x_1; x_2 \dots; x_m; \dots; x_n; \dots; x_{58}$ – значения плотности популяции в десятидневные периоды времени (табл. 2).

Следующие значения получены путем смещения входов и выходов на 10 дней (табл. 5) и так далее.

Таблица 4

Матрица значений обучающей выборки

Входы						Выход	Примечание
X_1	X_2	X_3	X_4	X_5	X_6	Y	
x_1	x_2	x_3	x_4	x_5	x_6	x_{1+11}	
x_2	x_3	x_4	x_5	x_6	x_7	x_{1+12}	
x_3	x_4	x_5	x_6	x_7	x_8	x_{1+13}	
...	...	...	...	...	...	...	
x_m	x_{m+1}	x_{m+2}	x_{m+3}	x_{m+4}	x_{m+5}	x_{1+1+m}	
...	...	...	...	...	...	...	
x_n	x_{n+1}	x_{n+2}	x_{n+3}	x_{n+4}	x_{n+5}	x_{1+1+n}	
...	...	...	...	...	...	...	
x_{53}	x_{54}	x_{55}	x_{56}	x_{57}	x_{58}	x_{59}	x_{59} – прогнозируемое значение

Таблица 5

Матрица значений обучающей выборки

Входы						Выход	Примечание
X_1	X_2	X_3	X_4	X_5	X_6	Y	
x_2	x_3	x_4	x_5	x_6	x_7	x_{1+12}	
x_3	x_4	x_5	x_6	x_7	x_8	x_{1+13}	
x_4	x_5	x_6	x_7	x_8	x_9	x_{1+14}	
...	...	...	...	...	...	...	
x_{m+1}	x_{m+2}	x_{m+3}	x_{m+4}	x_{m+5}	x_{m+6}	x_{1+12+m}	
...	...	...	...	...	...	...	
x_{n+1}	x_{n+2}	x_{n+3}	x_{n+4}	x_{n+5}	x_{n+6}	x_{1+12+n}	
...	...	...	...	...	...	...	
x_{54}	x_{55}	x_{56}	x_{57}	x_{58}	x_{59}	x_{60}	x_{60} – прогнозируемое значение

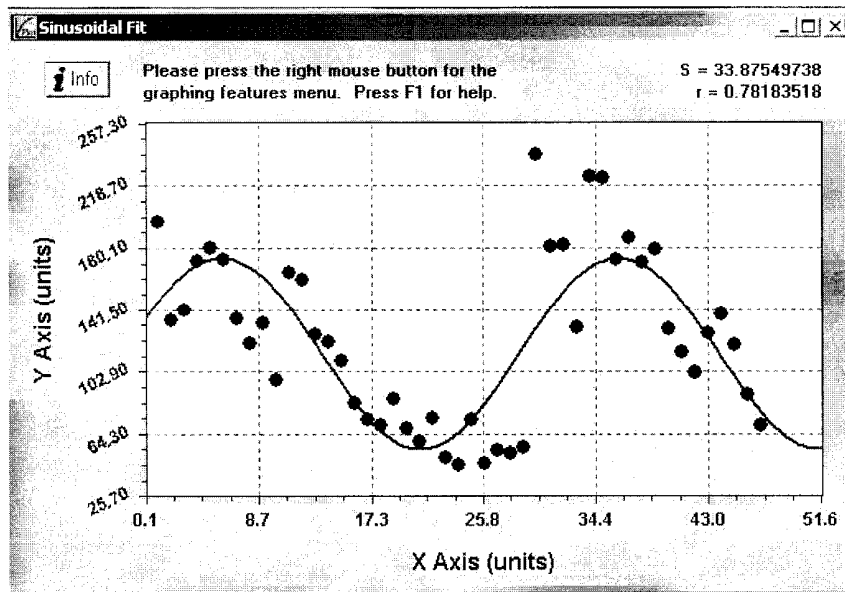


Рис. 7. Аппроксимация исходных данных

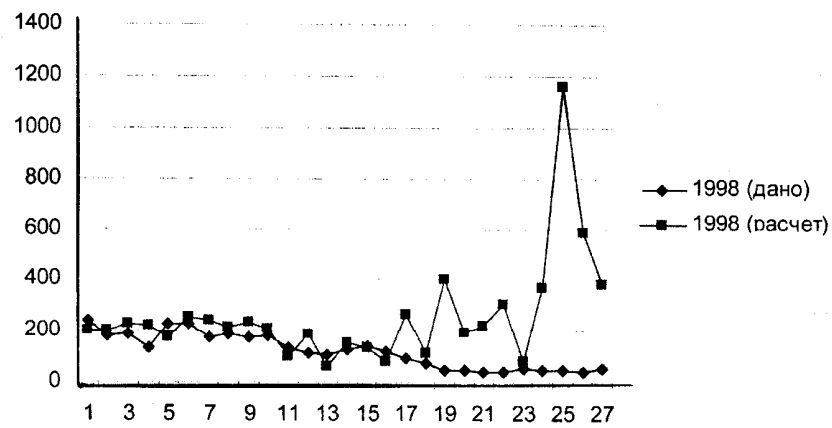


Рис. 8. Сравнительный график зависимости плотности популяции от времени

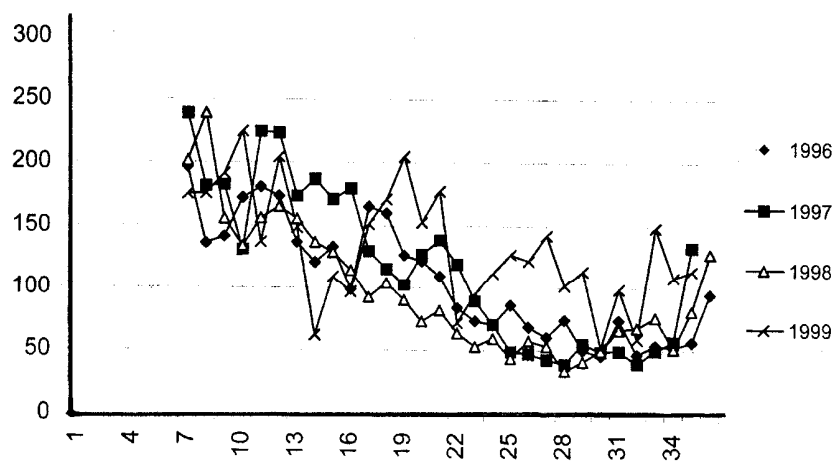


Рис. 9. Сравнительный график зависимости предполагаемой (1999 г.) и заданной (1996–1998 гг.) плотности популяции от времени

Таблица 6

Расчетная плотность популяции
(1999 года)

№ п/п	Дата начала интервала	Плотность популяции	№ п/п	Дата начала интервала	Плотность популяции
1	01.03.99	175,236	14	11.07.99	152,883
2	11.03.99	175,634	15	21.07.99	177,661
3	21.03.99	191,251	16	01.08.99	71,7396
4	01.04.99	224,664	17	11.08.99	95,2996
5	11.04.99	137,045	18	21.08.99	112,035
6	21.04.99	204,435	19	01.09.99	126,925
7	01.05.99	148,569	20	11.09.99	121,943
8	11.05.99	61,8347	21	21.09.99	142,417
9	21.05.99	109,131	22	01.10.99	101,813
10	01.06.99	96,8693	23	11.10.99	113,482
11	11.06.99	151,565	24	21.10.99	51,836
12	21.06.99	171,313	25	01.11.99	98,736
13	01.07.99	204,754	26	11.11.99	58,8108

Таблица 7

Коэффициенты корреляции (эксперимент)

	1996	1997	1998	1999
1999	0,6664	0,408436	0,504509	1

Таблица 8

Расчет ошибок

Годы	Ср. кв. ошибка	Примечания
1996	8,24	Год полностью
1997	11,62	-/-
1998	9,24	-/-
1999	9,49	Первые 210 дней (21 шаг)
1999	10,72	Год полностью (35 шагов)

По результатам расчета был построен сравнительный график зависимости плотности популяции от времени с использованием известных значений плотности популяции на 1998 год и полученных в ходе эксперимента (рис. 8).

На рис. 7 видно, что значения, близкие к реальным, были получены вплоть до 16 шага (коэффициент корреляции равен 0,65), максимальная разница между

соответствующими значениями плотности популяции ($x_{12(\text{дано})}$ и $x_{12(\text{расчет})}$) составляет 66,67 единиц.

На этом подготовительные операции закончены. В ходе их были подготовлены данные для вычислительного эксперимента, проведен их анализ, выбран тип ИНС.

Аналогичным образом были подготовлены значения выборки для прогнозирования значений плотности популяции на 1999 год. Для того чтобы количество переменных оставалось постоянным (учитывая смещение на 10 дней), в ходе всего вычислительного эксперимента в качестве входных значений использовались и расчетные значения плотности популяции.

В результате эксперимента были получены численные значения предполагаемой плотности популяции на 1999 г. (табл. 6), построен график зависимости (рис. 9).

Для экспериментально полученных значений плотности популяции были также рассчитаны коэффициенты корреляции (см. табл. 7).

Полученное значение коэффициента корреляции ~0,67 говорит о значительной схожести процессов, происходящих в объекте, относительно постоянство результирующей всех внешних воздействий.

Рост среднеквадратической ошибки значений плотности популяции с увеличением количества шагов вычислительного эксперимента иллюстрирует табл. 8.

ВЫВОД

Таким образом, в ходе вычислительного эксперимента показана возможность использования аппарата искусственных нейронных сетей для прогнозирования в открытых биологических системах.

ЛИТЕРАТУРА

1. Тютюнов Ю.В., Титова Л.И., Сенина И.Н. Модель динамики численности судака в Азовском море // Модели иерархического управления устойчивым развитием эколого-экономических систем / под. ред. Г.А. Угольного. М.: Вузовская книга, 2005 (в печати).
2. Сенина И.Н., Тютюнов Ю.В. Эвристическая модель миграционно-го цикла популяции тюльки Азовского моря // Компьютерное Моделирование. Экология: Вып. 2 / под. ред. Г.А. Угольного. М.: Вузовская книга, 2004. С. 92-100.
3. Говорухин В.Н., Моргулис А.Б., Сенина И.Н., Тютюнов Ю.В. Моделирование активных миграций пространственно распределенной популяции // Обозрение прикладной и промышленной математики. М.: Научное изд-во ТВП, 1999. Т. 6. Вып. 2. С. 271-295.
4. Домбровский Ю.А., Тютюнов Ю.В., Обущенко Н.И. Обзор методов моделирования рыбных популяций и сообществ. Приложение к ихтиофауне Азовского моря. Деп. в ВИНТИ. 1986. № 2841-В86. 83 с.
5. Абакумов А.И. Методология математического моделирования природных и экономических систем // Тр. Дальрыбвтуза (ТУ). Владивосток: Изд-во Дальрыбвтуза (ТУ), 1996. Вып. 7. С. 11-16.

Поступила в редакцию 15 июня 2005 г.