

УДК 574

**ПРОГНОЗИРОВАНИЕ ЧИСЛЕННОСТИ СУСЛОРОИДА В ПОЙМЕННОМ ОЗЕРЕ
Р. САМАРА В 2006 Г. С ПОМОЩЬЮ МОДЕЛЕЙ АВТОРЕГРЕССИИ**© 2013 В.Б. Дмитриев¹, Ю.Л. Герасимов²¹Самарский государственный колледж сервисных технологий и дизайна, г. Самара²Самарский государственный университет, г. Самара

Поступила 12.04.2013

Изучена сезонная динамика численности популяций планктонных веслоногих ракообразных в пойменном озере р. Самара в 2006 году. Были построены стохастические модели для моделирования и прогнозирования численности циклопов, при этом для прогнозирования наилучшей из выбранных оказалась модель AR(1).

Ключевые слова: модель авторегрессии, пойменное озеро, циклопы, численность.

Математическое моделирование изменений численности животных с целью прогнозирования её изменений применяется давно. К настоящему времени разработаны разнообразные принципы и подходы к созданию моделей. Чаще всего исследователи пытались оценивать изменения численности как результат соотношения рождаемости и смертности, разрабатывали методы расчёта этих показателей, выясняли, под влиянием каких внешних условий они меняются [1]. Однако, хотя в ходе сезонной сукцессии зоопланктона в большинстве водоемов наблюдается определенная, более или менее повторяющаяся из года в год, последовательность появления отдельных видов и достижения ими максимумов, кривая хода численности для каждого года обследования оказывается уникальной.

Колебания численности планктонных ракообразных наблюдаются не только в водоемах умеренной зоны с их ярко выраженной сезонностью, но и в тропиках, где колебания температуры воды не превышают 2-3 °С, а численность ракообразных, несмотря на это, демонстрирует перепады иногда на 1-2 порядка [2]. Дело в том, что размножение и гибель планктонных ракообразных относятся к процессам сложным, вероятностным образом изменяющимся во времени. Однако, описание и предсказание численности планктонных ракообразных возможно при помощи стохастической модели – подобные модели построены для описания динамики популяций ряда видов, социально-экономических процессов и явлений, речных потоков, солнечных пятен и

др.

Все эти процессы носят случайный характер, что и определяло особенности описывающих их систем. Традиционно системы рассматриваются как детерминированные или стохастические.

Принципиальным недостатком детерминированных моделей систем является отсутствие эффективного метода сравнения различных возможных моделей, создаваемых на основе одних и тех же эмпирических данных. Кроме того, детерминированная модель крайне неэффективна при наличии сильно нерегулярных флуктуаций или в тех случаях, когда она основывается на неправильной или неполной форме: игнорируется вероятностный характер данных.

В случае численности циклопов, случайных факторов, влияющих на неё, очень много (в том числе погодных), при этом факторы, контролирующие одну популяцию, меняются в течение сезона или год от года, поэтому нецелесообразно рассматривать детерминированные модели. Исследователю приходится считаться с тем, что разные популяции одного вида в один и тот же период могут лимитироваться совершенно разными факторами. Стохастические же модели были применены для описания изменений численности наземных позвоночных [3, 4], и мы попытались применить их к водным беспозвоночным. При этом, однако, понадобились некоторые изменения моделей. Поясним, кроме того, в чём проявляется непредсказуемость всех факторов.

Именно сильно нерегулярными флуктуациями вызвана трудность предсказания численности ракообразных путём соотношения рождаемости и смертности: слишком много факторов нужно учитывать. В нашем случае известна рождаемость циклопов, однако смертность зависит от очень многих трудно учитываемых факторов, поэтому представляется целесообразным сделать прогноз,

Дмитриев Виктор Борисович, кандидат математических наук, преподаватель, dmitriev.v.b@mail.ru; Герасимов Юрий Леонидович, кандидат биологических наук, заведующий кафедрой зоологии, генетики и общей экологии, yuger55@list.ru

основываясь только на прошлых значениях общей численности половозрелых циклопов и копепоидитов всех возрастов – а также на некоторых общих особенностях изменения численности циклопов в течение сезона.

Следует подчеркнуть это преимущество стохастических моделей: не нужно часто измерять многие факторы, учитывается их совместное влияние путём измерения большой выборки прошлых значений численности.

В предпринятой нами работе мы попытались применить для моделирования сезонных изменений численности копепоид программу R – среду анализа наборов статистических данных. Разумеется, подобные модели возможны и для других ракообразных.

РАЙОН ИССЛЕДОВАНИЯ

В черте города Самары в районе Южного моста существует сложный комплекс пойменных озёр, расчлённый дамбой, по которой проходит автомобильная дорога с оживлённым движением. В период паводка эти озёра соединяются с основным руслом, с середины июня часть их изолируются и существуют далее как самостоятельные водоёмы. Изучение их зоопланктона начато в 2005 г. [5, 6]. В 2006 г. нами проведены сборы зоопланктона в данных озёрах, в том числе в озере, расположенном в 200 м выше по течению реки относительно дамбы. Участок между дамбой и озером занят деревьями и кустарниками, по нему проходят высоковольтная ЛЭП и грунтовая дорога. Озеро овальной формы площадью более 2 га, глубиной до 4 м. На озере происходит активный любительский лов рыбы, вследствие чего на берегах постоянно стоят автомобили, много кострищ, накапливается бытовой мусор. В 2006 г. в озере выявлено более 20-ти видов ракообразных, относящихся к подклассам Copepoda (семейства Cyclopoidea и Eudiaptomidae) и Cladocera (семейства Bosminidae, Chydoridae, Daphniidae, Sididae, Leptodoridae и Polyphemidae). Нами были проанализированы данные по численности ракообразных, относящихся к сем. Cyclopoidea, поскольку они весь сезон присутствовали в пробах и доминировали по численности среди остальных семейств ракообразных. Циклопы – важная пища рыб, в том числе промысловых, а также промежуточные хозяева личинок плоских

червей – паразитов рыб и других позвоночных. Численность циклопов является одним из важных критериев оценки состояния экосистем водоёмов [7].

МАТЕРИАЛ И МЕТОДЫ

Пробы отбирали еженедельно 17 раз с конца апреля до конца октября на 5-ти станциях стандартными гидробиологическими методами [8, 9]. Для анализа данные по станциям усредняли. В озере обнаружено 9 видов циклопов: *Acanthocyclops* (*Diacyclops*) *bicuspidatus* Claus, 1857, *A. vernalis* (Fisher, 1851), *Eucyclops serrulatus* (Fisher, 1851), *Macrocyclops albidus* (Jurine, 1820), *Cyclops abissorum* Sars, 1862, *C. kolensis* Lilljeborg, 1901, *C. strenuus* (Fisher, 1851), *Mesocyclops leuckarti* (Claus, 1857) и *Thermocyclops* (*Mesocyclops*) *oithonoides* Sars, 1863 [10]. Большинство этих видов эвритопны. Наибольшей средней численности достигали популяции *C. kolensis* (до 2,58 тыс.экз/м³), *E. serrulatus* (до 2,24 тыс.экз/м³) и *M. leuckarti* (до 1,13 тыс.экз/м³), представители остальных видов встречались в гораздо меньших количествах. Сезонная динамика численности разных видов циклопов различалась достаточно сильно, происходила смена доминирующих видов, поэтому для моделирования мы взяли всё семейство Cyclopoidea. Для расчетов использовали численность половозрелых циклопов и копепоидитов всех возрастов. Численность науплиев для моделирования не учитывали, в связи с наличием в пробах науплиев Calanoida.

Цель нашего исследования состояла в разработке модели, позволяющей прогнозировать численность циклопов.

Нами решались следующие задачи: 1) примерное предсказание численности циклопов на момент следующего наблюдения; 2) построение стохастических моделей и выбор наилучшей из них для моделирования численности циклопов.

Полученные результаты были обработаны с помощью программы R – среды анализа наборов статистических данных.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Средняя численность циклопов, взрослых и копепоидитов (экз/л) приведена в таблице 1.

Таблица 1. Средняя численность циклопов, взрослых и копепоидитов (экз/л)

Даты	24.04	5.05	15.05	26.05	6.06	17.06	29.06	10.07	21.07
Cyclopoida	0,06	0,43	1,99	3,08	3,57	3,36	1,91	2,11	0,75
Даты	1.08	12.08	25.08	4.09	17.09	26.09	11.10	25.10	
Cyclopoida	0,07	0,15	0,86	0,45	0,26	0,13	0,14	0,04	

Следует отметить модели авторегрессии (AR), которые оказались удачными в случае моделирования численности циклопов. Это важный класс стохастических динамических моделей, а в вычислительном отношении это еще и простейший класс.

Авторегрессионная (AR-) модель — модель временных рядов, в которой значения временного ряда в данный момент линейно зависят от предыдущих значений этого же ряда. Авторегрессионный процесс порядка p (AR(p)-процесс) – определяется следующим образом

$$y(t) = c + \sum_{i=1}^p a_i y(t-i) + w(t),$$

где a_1, \dots, a_p — *параметры* модели (коэффициенты авторегрессии), c – постоянная (часто для упрощения предполагается равной нулю), $\{w(t)\}$ — белый шум, то есть последовательность независимых и одинаково распределённых случайных величин (как правило, нормальных), с нулевым средним.

Мы провели анализ численности популяции циклопов подобно тому, как это проводилось для американского журавля [3,4]. Но эксперименты проводились примерно раз в десять дней, что вызвано спецификой вида – их жизненные циклы существенно меньше, чем у журавля. Тем не менее, стохастические модели в целом напоминают такие же для журавля, но для учёта затухания численности циклопов в октябре (и возрождения в апреле) вводится множитель $e^{D(t-9)^{2.5}}$, где $D=1/5$. Выбор множителя с именно таким значением $D=1/5$ служит улучшению авторегрессионной модели и обусловлен общими особенностями графиков численности циклопов, в частности, данными за прошлые годы. Таким образом, исследовалась величина

$$y(t) = y_0(t) e^{D(t-9)^{2.5}},$$

где $y_0(t)$ - численность Cyclopoida. При этом t – номер эксперимента, он изменяется от 1 до 17, то есть представляет собой строку

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17

поэтому $t-9$ изменяется от -8 до 8:

-8 -7 -6 -5 -4 -3 -2 -1 0 1 2 3 4 5 6 7 8

При этом множитель $e^{D(t-9)^{2.5}}$ (обозначен в R как Exp) убывает от 25,487097 до единицы и затем снова возрастает до 25,487097. Точнее, он представляет собой строку (показываем, как она выводится в R)

[1] 25.487097 10.166554 4.842548 2.718282
1.772575 1.321615 1.106490
[8] 1.018050 1.000000 1.018050 1.106490
1.321615 1.772575 2.718282
[15] 4.842548 10.166554 25.487097

Предположения относительно $w(t)$ давали возможность применить метод наименьших квадратов (МНК).

Заметим, что умножение численности Cyclopoida на строку $e^{D(t-9)^{2.5}}$ обосновано. Если не умножать, то рассмотренные ниже модели (для самой численности) будут неадекватными. Как известно, количественная оценка адекватности модели системы в большинстве практически важных случаев может быть определена лишь косвенно [11]. В данном случае авторегрессионные модели для величины $y_0(t)$, например, дадут отрицательные значения численности либо дадут явное увеличение численности в конце сезона, что не соответствует изучаемому явлению. Причиной этой неэффективности, скорее всего, является гетероскедастичность (непостоянство дисперсии отклонений). Как известно, одной из ключевых предпосылок МНК является условие постоянства дисперсии случайных отклонений, то есть, эта дисперсия должна быть постоянна для любых наблюдений (в данном случае их 17, и это условие не выполняется – численность циклопов сильнее меняется летом, чем весной или осенью).

Как известно, вопрос обеспечения гомоскедастичности, или хотя бы уменьшения гетероскедастичности, является одним из главных при моделировании. Поэтому вводится множитель $e^{D(t-9)^{2.5}}$ – он служит уменьшению гетероскедастичности.

Модели численности циклопов и применение языка R. Мы рассмотрели пять классов моделей:

$C_i, i = 0, 1, 2, 3, 4.$

Соответствующие уравнения приведены ниже.

Пусть $y_1 = \nabla \ln y(t) = \ln y(t) - \ln y(t-1)$,

$y_2 = \ln y(t)$. Тогда

$$L_0 : y(t) = w(t) + \theta_0 + \theta_1 y(t-1),$$

$$L_1 : y_1(t) = w(t) + \theta_0 + \theta_1 y_1(t-1),$$

$$L_2 :$$

$$y(t) = w(t) + \theta_0 + \theta_1 y(t-1) + \theta_2 y(t-2),$$

$$L_3 : y_2(t) = w(t) + \theta_0 + \theta_1 y_2(t-1),$$

$$L_4 :$$

$$y_2(t) = w(t) + \theta_0 + \theta_1 y_2(t-1) + \theta_2 y_2(t-2).$$

Уравнение L_1 есть модель типа IAR, оказавшаяся очень успешной при моделировании численности населения США в 1954-1984 годах и в исследованиях популяции американского журавля [3, с. 343, 346]. Уравнение L_0 — обычный авторегрессионный процесс первого порядка — использовалось в ранних исследованиях популяции американского журавля [4]. Уравнение L_2 — обычный авторегрессионный процесс второго порядка. Уравнение L_4 есть обобщение на стохастический случай модели Гомпертца.

Уравнение L_3 – авторегрессионный процесс второго порядка с заменой переменных их логарифмами.

Для оценки параметров различных уравнений мы используем наблюдения за период с 24 апреля по 25 августа. Наиболее успешным (в первую очередь учитывая коэффициент детерминации) оказались уравнения L_0 и L_3 . Соответствующие модели $M_i, i = 0, 3$ задаются уравнениями

$$M_0 : y(t) = w(t) + 0,3666 + 0,8341 y(t - 1),$$

$$M_3 : y_2(t) = w(t) + 0,07819 + 0,7465 y_2(t - 1).$$

При этом прогнозировалась численность взрослых особей лишь примерно на 10 дней (на дату следующего замера), в дальнейшем модель уточнялась с учётом нового замера. То есть коэффициенты были уже другие – хотя несильно отличались от этих. Это вызвано большой нестабильностью численности рачков по сравнению с позвоночными, объективной невозможностью точного предсказания на длительный срок.

Для практического применения разработанных моделей было необходимо использование автоматизированных средств анализа, моделирования и прогнозирования. Была использована статистическая программа R [12, 13], с помощью которой осуществлялась параметризация предложенных моделей и оценка их точности.

R – мощный функциональный язык программирования и среда анализа наборов статистических данных. Будучи средой анализа, R позволяет создавать различные графические представления данных из командной строки.

Опишем примерную схему применения R в нашем случае.

В частности, в модели AR(1) для ввода строки численности циклопов применялась команда

```
Cyclopoida_mN<-c(0.06, 0.43, 1.99,
3.08, 3.57, 3.36, 1.91, 2.11,
```

```
0.75, 0.07, 0.15, 0.86, 0.45,
0.26, 0.13, 0.14, 0.04)
```

для их преобразованной («гомоскедастичной») численности

```
Cyclopoida_mNExp<- Cyclopoida_mN*Exp
```

для предыдущего значения этой численности:

```
Cyclopoida_mNExp_k_1<- c(0,
```

```
Cyclopoida_mNExp[1:16])
```

для рассмотрения первых 12 замеров (с учётом дальнейшего применения метода МНК – то, что было известно на момент 25 августа) использовалась строка:

```
Cyclopoida_mNExp_12<-
```

```
c(0,0,Cyclopoida_mNExp [3:12],0,0,0,0,0)
```

(либо в случае модели второго порядка, строка Cyclopoida_mNExp_12<- c(0,Cyclopoida_mNExp [2:12],0,0,0,0,0))

для получения коэффициентов методом МНК:

```
fmt3<- lm(Cyclopoida_mNExp_12 ~
```

```
Cyclopoida_mNExp_k_1_12)
```

```
summary(fmt3)
```

полученные коэффициенты 0,3666 и 0,8341 подставлялись в формулу

```
Cyclopoida_mNExp_Prob<- 0.3666+ 0.8341*
Cyclopoida_mNExp_k_1
```

и далее, после пересчётов и ввода вспомогательных переменных, формировалась строка вычисленных значений.

Подстановка полученных оценок параметров в рассмотренные модели позволило рассчитать сглаженные модельные значения уровней временного ряда $y(t)$. Вывод в пользу той или иной модели временного ряда можно делать по большему значению меры точности моделирования – коэффициенту детерминации R^2 . Коэффициент принимает значения из интервала [0; 1]. Чем ближе значение к 1, тем ближе модель к эмпирическим наблюдениям. При незначительном статистическом различии коэффициентов детерминации для рассматриваемых моделей предпочтительнее выбрать более простую.

При этом авторегрессионная (и более простая) модель M_0 лучше описывает численность циклопов, чем логарифмическая модель M_3 . В данном случае для модели M_0 коэффициент R^2 меняется от 0,6922 до 0,6626, а для модели M_3 коэффициент R^2 меняется от 0,5561 до 0,5477. То есть в обоих случаях сила связи заметная, но в первом случае она выше (почти высокая).

Кроме того, модель M_0 лучше предсказывает численность циклопов, чем авторегрессионная модель второго порядка M_2 , несмотря на то, что

во втором случае коэффициент детерминации равен 0,816.

То есть в обоих случаях более сложная модель обладает худшими предсказательными свойствами и меньшим коэффициентом детерминации, что не удивительно. Известно, что модели, лежащие за пределами определенного

уровня сложности, часто оказываются более низкими по качеству, чем некоторые более простые модели.

Данные прогноза численности циклопов по моделям M_0 , M_2 и M_3 приведены в табл. 2.

Таблица 2. Прогноз численности циклопов по моделям M_0 , M_2 и M_3

	4.09	17.09	26.09	11.10	25.10
Cyclopoida	0,45	0,26	0,13	0,14	0,04
Прогноз Cyclopoida по модели M_0 (AR(1))	0,74164928	0,3040788	0,1555374	0,068038	0,056294
Прогноз Cyclopoida по модели M_3	0,67121379	0,32917635	0,16647387	0,07148187	0,05519078
Прогноз Cyclopoida по модели M_2 (AR(2))	0,97869663	0,28969454	0,177483	0,08076815	0,07564871

Учитывая размах колебаний численности циклопов в течение весны-осени, точность предсказания хорошая. Результаты моделирования численности циклопов в сравнении с реальной численностью представлены на рисунке.

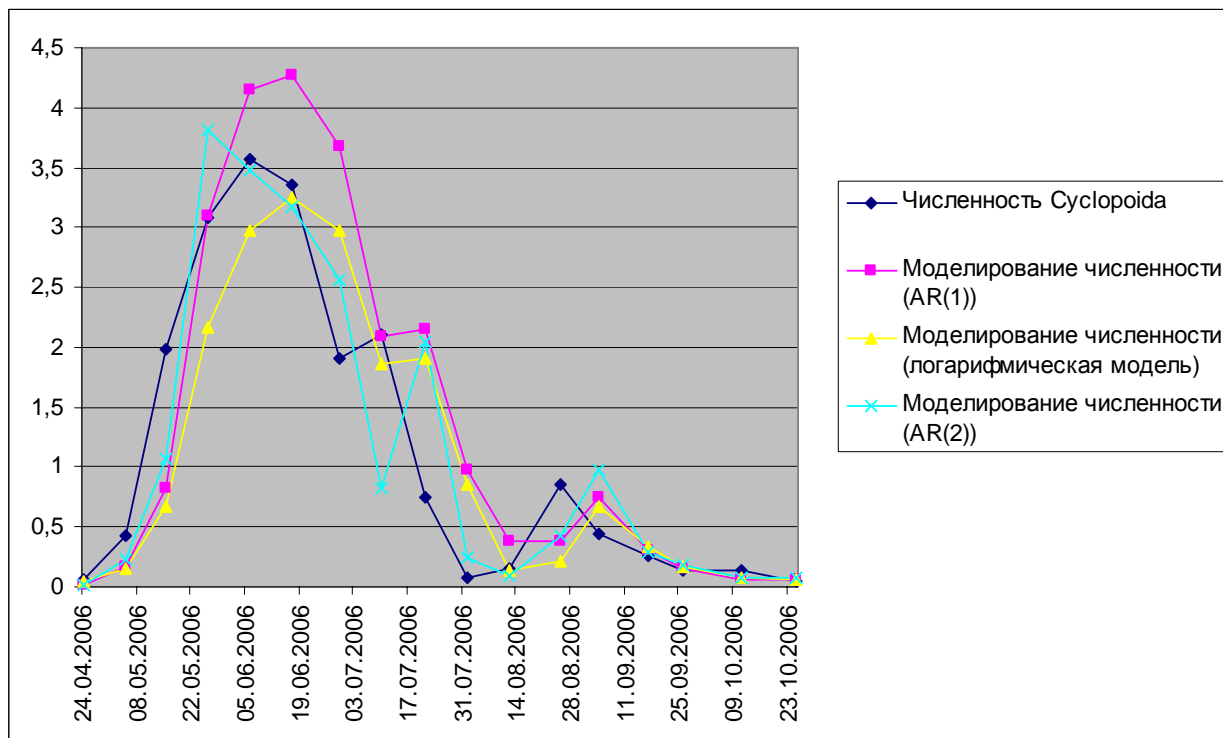


Рисунок 1. Сезонная динамика плотности популяции циклопов и их моделирование при помощи авторегрессионных и логарифмической моделей (в тыс.экз/м³).

Как видно из рис.1, модель выражает общий характер колебаний численности, в первую очередь – колебаний её порядка. При разработке модели использовали только данные о численности половозрелых циклопов и копепоидитов, без учёта количества науплиев. Для ракообразных из семейств подкласса Cladocera учет численности личинок возможен. В модели не

учитывали также условия существования ракообразных в изучавшемся водоёме – температуру, гидрохимические показатели, изменения проточности. Не вводились в модель данные о количестве пищевых объектов, наличии и численности беспозвоночных и позвоночных хищников разных видов, сильно влияющих на зоопланктон [14].

Не исключено, что возможна стохастическая (более общая, чем авторегрессионная) модель, которая будет учитывать все эти факторы, т.е. такая модель будет точнее, чем полученная нами. Однако в этом случае такая модель будет иметь значительно более громоздкую функциональную форму, она также будет требовать большой точности вводимых факторов. Всё же мы считаем, что результаты нашей работы показывают перспективность применения авторегрессионной модели M_0 для описания изменений численности ракообразных в непроточном водоёме.

СПИСОК ЛИТЕРАТУРЫ

1. *Гиляров А.М.* Динамика численности пресноводных планктонных ракообразных. М.: Наука, 1987. 190 с
10. Определитель пресноводных беспозвоночных России и сопредельных территорий. Т. 2. Ракообразные. СПб: ЗИН, 1995. 627 с.
11. *Балдин К.В., Быстров О.Ф., Соколов М.М.* Эконометрика: Учеб. пособие для вузов. 2-е изд., перераб. и доп. М.: ЮНИТИ-ДАНА, 2004. 254 с.
12. Основные понятия языка R / Сост. Савельев А.А., Мухарамова С.С., Пилюгин А.Г. Учебно-методическое пособие. Казань, Изд-во Казан. ун-та. 2007. 30 с.
13. Использование языка R для статистической обработки данных / Сост. Савельев А.А., Мухарамова С.С., Пилюгин А.Г. Учебно-методическое пособие. Казань, Изд-во Казан. ун-та. 2007. 30 с.
2. *Lewis W.M.* (jun.) Zooplankton community analysts: studies on a tropical system. N.Y.: Springer, 1979. 163 p.
3. *Кашиял Р.Л., Рао А.Р.* Построение динамических стохастических моделей по экспериментальным данным. Пер. с англ. М.: Наука. Главная редакция физико-математической литературы, 1983. 384 с.
4. *Миллер, Боткин* (Miller R.B., Botkin D.B.). Endangered species: models and predictions. Amer. Sci., 62, 1974. 172-181 pp.
5. *Герасимов Ю.Л.* Зоопланктон водоема в Самарском заречье в 2005 г. // Вест. Самарск. гос. ун-та. 2006, №7 (47). С. 31-36.
6. *Герасимов Ю.Л.* Ракообразные пойменного водоёма реки Самары в черте города Самары // Вестн. Мордовского ун-та, 2010, № 1. С. 118-123.
7. *Рылов В.М.* Cyclopoidea пресных вод. Фауна СССР, Ракообразные. Т. 3. Вып. 3. М.-Л.: Изд-во АН СССР. 1948. 318 с.
8. Методика изучения биогеоценозов внутренних водоемов. М., 1975. 240 с.
9. Руководство по гидробиологическому мониторингу пресноводных экосистем. СПб.: Гидрометеиздат, 1992. 246 с.
14. *Gliwicz Z.M., Ghilarov A.M., Pijanowska J.* Food and predation as major factors limiting two natural populations of Daphnia. Hydrobiologia, 1981. V. 80, № 3. P. 205-218.

FORECASTING OF CYCLOPOIDA NUMBER IN THE SAMARA RIVER FLOODPLAIN LAKE IN 2006 BY MEANS OF AUTOREGRESSION MODELS

© 2013 V.B. Dmitriev¹, Yu.L. Gherasimov²

¹Samara State College of Technology and Design Service, Samara

²Samara State University, Samara

Seasonal dynamics of number of populations of planktonic Copepoda in the Samara River floodplain lake in 2006 is studied. Stochastic models for modeling and forecasting of number of a Cyclops were constructed. It turns out that among these models AR(1) model forecasts better than the other ones do.

Key words: autoregression model, floodplain lake, Cyclops, number.