



УДК 597.08

ОЦЕНИВАНИЕ ПАРАМЕТРОВ ПОЛОВОГО СОЗРЕВАНИЯ ГИДРОБИОНТОВ ПРИ ОТСУТСТВИИ ВИЗУАЛЬНЫХ ПРИЗНАКОВ ЗРЕЛОСТИ У ЧАСТИ ПОЛОВОЗРЕЛЫХ ОСОБЕЙ

© 2020 А. Н. Михайлюк, М. М. Пятинский

*Всероссийский научно-исследовательский институт рыбного хозяйства и океанографии (ФГБНУ «ВНИРО»),
Азово-Черноморский филиал ФГБНУ «ВНИРО» («АзНИИРХ»), Ростов-на-Дону 344002, Россия
E-mail: a.mikhaylyuk@mail.ru*

Аннотация. Рассмотрена проблема определения истинного значения длины рыб и ракообразных, при которой 50 % особей являются половозрелыми, в ситуациях, когда признаки половозрелости отсутствуют у части половозрелых особей. В качестве решения проблемы предлагается использование вместо традиционных двухпараметрических моделей, заданных логистической функцией или интегралом вероятностей нормального распределения, аналогичных трехпараметрических моделей. При этом третий параметр является оценкой доли половозрелых особей, у которых признаки половозрелости выражены внешне. На основе литературных данных показано, что для оценки параметров в таких случаях может быть использована трехпараметрическая логистическая модель в пакете расширения *drc* для среды открытого доступа R. В качестве материалов для проведения расчетов были использованы опубликованные данные по самкам таких видов живых организмов, как тигровая креветка *Penaeus semisulcatus* (De Naan, 1844) и средиземноморский ботус *Bothus podas* (Delaroche, 1809). В некоторых случаях для сравнения результатов расчеты выполнялись также по традиционной схеме, предусматривающей линеаризацию логистической функции, с использованием пакета прикладных программ SPSS. Был сделан вывод, что, когда задающий верхнюю асимптоту параметр статистически значимо отличается от 1, целесообразно использовать вышеуказанную трехпараметрическую модель. Но когда этот параметр не отличается статистически значимо от 1, рекомендуется использовать двухпараметрическую модель, задавая фиксированное значение данного параметра равным 1.

Ключевые слова: рыбы, ракообразные, половая зрелость, пропуск нереста, длина тела

ESTIMATION OF THE PARAMETERS OF SEXUAL MATURITY IN HYDROBIONTS IN THE ABSENCE OF VISIBLE MATURITY INDICATORS IN A PART OF MATURE INDIVIDUALS

A. N. Mikhaylyuk, M. M. Piatinskii

*Russian Federal Research Institute of Fisheries and Oceanography (FSBSI “VNIRO”),
Azov-Black Sea Branch of the FSBSI “VNIRO” (“AzNIIRKH”), Rostov-on-Don 344002, Russia
E-mail: a.mikhaylyuk@mail.ru*

Abstract. The problem of estimation of the true value of the length at which 50 % of fish or crustacean individuals are sexually mature is considered for the case when maturity indicators are absent in a part of mature individuals. As a means to solve this problem, it is proposed to use three-parameter models similar to traditionally used two-parameter ones, defined by a logistic function or the normal probability integral. In this case, the third parameter is the estimation of the fraction of mature individuals that present discernible maturity indicators. Based on the data from literature sources, it is shown that, in such cases, the three-parameter logistic model in the add-on package *drc* for the open-source environment R can be used. As the input data for the estimation, the published data on female individuals of such species of aquatic living organisms as tiger prawn *Penaeus semisulcatus* (De Haan, 1844) and wide-eyed flounder *Bothus podas* (Delaroche, 1809) have been used. In some cases, in order to compare the results, the estimation has also been done using the conventional method, involving linear approximation of the logistic function with application of the software package SPSS. It was concluded that, when the parameter defining the upper asymptote deviates from 1 in a statistically significant manner, it is expedient to use the aforementioned three-parameter model. However, when this parameter does not deviate from 1 in such a way, it is advisable to use the two-parameter model, defining this parameter as a fixed value equal to 1.

Keywords: fish, crustaceans, sexual maturity, body size, skipped spawning, body length

ВВЕДЕНИЕ

Важной характеристикой популяций у рыб и морских беспозвоночных является длина, при которой 50 % особей являются половозрелыми [1]. Эта характеристика представляет как чисто научный, так и практический интерес. Последнее обусловлено тем, что она используется в регулировании рыболовства при определении минимальной длины особей, допустимых для добычи. Для оценки этой характеристики принято применять методы, разработанные в токсикологии для анализа зависимостей «доза – эффект». При этом длина особей рассматривается как «доза», а «эффектом» является доля половозрелых особей в соответствующей размерной группе. Как правило, используются методы, основывающиеся на описании кривой созревания интегралом вероятностей нормального распределения или, чаще, логистической функцией, и при этом предусматривающие преобразование указанных функций к линейному виду.

Применение этих методов затруднено в тех случаях, когда у особей, являющихся на самом деле половозрелыми, отсутствуют признаки, по которым особи идентифицируются как половозрелые. Это имеет место, когда отдельные производители пропускают нерест в данном году (причем их гонады остаются незрелыми) или когда особи данного вида вообще нерестятся не каждый год, а также у самок ракообразных, когда из предыдущей порции икры произошел полный выклев, а новая порция икры еще не была отложена. М. Кинг [1] рекомендует в таких случаях предварительно корректировать данные путем умножения доли особей с признаками

зрелых на некий коэффициент. Этот коэффициент он рассчитывает как обратный той доле особей с признаками зрелых, которая имеет максимальное значение (т. е. $k=1/p_{\max}$). Конечно, применение такой корректировки предпочтительнее, чем непосредственное выполнение обычных расчетов, игнорирующее природную ущербность данных. Однако предложенный прием все-таки представляется слишком грубым, поскольку для расчета в нем используется только одно, а именно максимальное, значение доли. Очевидно, более правильным было бы оценивать истинное значение длины, при которой 50 % особей являются половозрелыми, а заодно и долю половозрелых особей, у которых не выражены признаки половозрелости, по всей имеющейся совокупности данных с использованием методов математической статистики. Это представляется актуальной задачей, поскольку ситуация с наличием половозрелых особей, у которых отсутствуют признаки половозрелости, является нередкой.

Для решения поставленной задачи полезно воспользоваться опытом, накопленным в токсикологии. В ней к настоящему времени было предложено большое количество различных нелинейных моделей для выполнения анализа данных типа «доза – эффект». К. Ритцем [2] была предпринята попытка выработки единого подхода к использованию таких моделей в экотоксикологии. Он рассматривает функции f следующего общего вида:

$$f(x; b, c, d, e, \dots) = c + (d - c)h(x; b, e, \dots).$$

В рамках данного подхода мы предлагаем использовать следующую трехпараметрическую функцию, описывающую зависимость доли особей с

визуальными признаками половозрелости от их длины: $f(x; b, d, e) = dh(x; b, e)$, где

$h(x; b, e)$ — обычная двухпараметрическая функция (т. е. логистическая функция или интеграл вероятностей нормального распределения), описывающая зависимость доли в действительности половозрелых особей от их длины;

d — корректирующий коэффициент, характеризующий долю половозрелых особей, у которых признаки половозрелости выражены внешне.

При этом доля половозрелых особей, у которых признаки половозрелости не выражены, будет равна $(1-d)$.

Для практического использования указанной формулы необходимо наличие алгоритма расчета ее параметров или соответствующей компьютерной программы. Нам не удалось найти ничего подобного для модели, основывающейся на интеграле вероятностей нормального распределения. Однако для модели, основывающейся на логистической функции, такая программа существует в пакете расширения *drc* для статистической программной среды открытого доступа R. Целями настоящей работы являются обоснование использования трехпараметрической логистической модели для оценки параметров полового созревания гидробионтов при наличии среди них особей, у которых отсутствуют визуальные признаки половозрелости, а также проверка возможности использования для указанной оценки пакета *drc*.

МАТЕРИАЛЫ И МЕТОДЫ

Статистическая программная среда открытого доступа R в настоящее время широко распространена, а работа в ней описана в большом количестве руководств (например, в [3]). Функциональные возможности *drc* изложены в [4]. К сожалению, предлагаемая авторами модель в данном описании отсутствует, хотя в самом пакете возможности для расчета ее параметров имеются.

Для установки пакета расширения *drc* в среде R необходимо выполнить команду *install.packages("drc")*. Установка выполняется единоразово. Для приведения *drc* в рабочее состояние необходимо после запуска R выполнить команду *library("drc")*. Данные, по которым должны оцениваться параметры модели, должны быть предварительно введены средствами среды R и сохранены под произвольным названием. В нашем случае эти данные представляют собой таблицу, в одном из столбцов кото-

рой указана длина гидробионтов (для сгруппированных данных — средняя длина в размерных группах), а в двух других — общее количество особей и количество особей с признаками половозрелости, имеющих соответствующую длину.

Перед оценкой параметров выбираем интересующую нас модель:

$formulae1 = L.3()$,

где $L.3()$ — код логистической функции с тремя параметрами: 1) e соответствует длине, при которой в действительности половозрелыми являются 50 % особей; 2) b — коэффициент, характеризующий наклон касательной к функции в точке ее перегиба; 3) d — доля половозрелых особей, у которых признаки половозрелости выражены внешне.

Для оценки параметров модели задаем:

$result1 = drm(affect/total \sim length, weights = total, fct = formulae1, type = "binomial", data = Data1)$,

где $result1$ — произвольное название записи, содержащей результаты данной оценки параметров модели;

$length$, $total$, $affect$ и $Data1$ — произвольные названия заголовков столбцов таблицы с данными и самой таблицы;

выражение $affect/total \sim length$ указывает, что аппроксимируется отношение количества особей, у которых признаки половозрелости выражены внешне ($affect$), к общему количеству особей в данной размерной группе ($total$) в зависимости от длины ($length$);

$weights = total$ задает выполнение «взвешенных» расчетов с учетом численности размерных групп;

fct — параметр, указывающий выбранную модель;

выражение $type = "binomial"$ указывает на биномиальный тип данных, имеющий место в случае анализа зависимости доли половозрелых особей от их длины.

Вывод результатов оценки параметров модели с указанием их стандартных ошибок осуществляется командой: *summary(result1)*.

Для оценки качества подгонки модели в *drc* используется критерий хи-квадрат, расчет которого осуществляется командой: *modelFit(result1)*.

При необходимости выполнения расчетов параметров для традиционной двухпараметрической модели (т. е. в случае наличия признаков половозрелости у всех половозрелых особей) в *drc* может быть использована указанная выше трехпараметрическая логистическая модель, в которой параметр

d задается равным 1. Задается такая модель следующим выражением: $formulae2 = L.3(fixed = c(NA, 1, NA), names = c("b", "d", "e"))$.

В таких случаях для сравнения результатов соответствующие расчеты выполнялись также по традиционной схеме, предусматривающей линеаризацию логистической функции, с использованием пакета прикладных программ SPSS.

Пакет *drc* позволяет рассчитывать длину особей, при которой доля производителей равна заданной исследователем величине (ED), а также определять доверительный интервал для полученной оценки.

Такой расчет (например, для $ED=10\%$) задается командой: $ED(result1, c(10), interval = "delta", level = 0,90)$, в которой $c(10)$ задает интересующее значение доли, в %;

$interval = "delta"$ задает метод расчета доверительного интервала (при задании $interval=none$ доверительный интервал не рассчитывается);

$level = 0,90$ задает доверительную вероятность, для которой рассчитывается доверительный интервал.

Задания $interval = none$ и $level = 0,95$ можно не указывать, в этом случае они будут выполнены по умолчанию. $ED=50\%$ тождественно вышеупомянутому параметру e . Расчет доверительного интервала по методу "delta" соответствует расчету по t -критерию Стьюдента при бесконечном числе степеней свободы.

В качестве материалов для проведения расчетов были использованы опубликованные данные [1, 5].

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Рекомендуемая нами для использования трехпараметрической логистической функция имеет вид: $f(x; b, d, e) = d / (1 + \exp(b(x - e)))$, где

x — аргумент функции, которым является длина особей;

b — параметр, характеризующий «растянутость» кривой вдоль оси координат, на которой отложена длина особей;

d — верхняя асимптота логистической функции, являющаяся оценкой доли половозрелых особей, у которых признаки половозрелости выражены внешне;

e — параметр, являющийся оценкой истинного значения длины, при которой 50 % особей являются половозрелыми.

Эта функция соответствует четырехпараметрической логистической функции $f(x; b, c, d, e)$ [2] при $c=0$. Параметр c является нижней асимптотой логистической функции, и в токсикологических

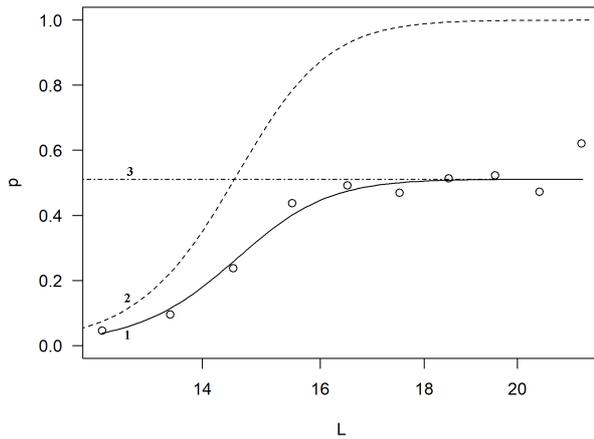
исследованиях он показывает долю особей, погибающих в эксперименте даже при отсутствии воздействия исследуемого препарата. При исследованиях полового созревания гидробионтов параметр c указывал бы на долю особей, которые половозрелы даже при самой малой длине. Поскольку такого не бывает, следовательно, в нашем случае должно быть $c=0$.

При расчете параметров полового созревания самок тигровой креветки *Penaeus semisulcatus* (De Naan, 1844) по литературным данным [1] результат был следующим: $b=-1,27$, $d=0,511$, $e=14,485$, а стандартные ошибки этих оценок, соответственно, равны 0,223, 0,012 и 0,233.

Оценка качества подгонки показала, что $\chi^2=5,856$ (при числе степеней свободы $df=7$); соответствующая вероятность $P=0,556$. Это показывает, что данные по созреванию самок тигровой креветки хорошо соответствуют трехпараметрической логистической модели. Вследствие этого указанная модель вполне допустима для использования в анализе этих данных. Результаты расчетов в пакете *drc* показывают, что только у половины половозрелых самок ($d \approx 0,5$) наблюдаются визуальные признаки половой зрелости (наличие икры или личинок). Разница между 1 и d гораздо больше стандартной ошибки d ; это свидетельствует о том, что отличие d от 1 не является случайным.

График этой функции, полученный в пакете *drc*, приводится на рисунке (кривая 1). Нами добавлена к нему прямая $p=d=0,511$, являющаяся горизонтальной асимптотой к кривой, заданной указанной функцией. Для сравнения на этом же рисунке приводится рассчитанный авторами график двухпараметрической логистической функции с параметрами $b=-1,27$ и $e=14,485$. Соответствующая кривая показывает долю в действительности половозрелых самок в зависимости от их длины. Горизонтальной асимптотой к кривой, заданной этой функцией, является прямая $p=1$, которая не приводится на рисунке во избежание его излишней детализации.

Для самок тигровой креветки была также выполнена попытка расчета параметров по скорректированным [1] данным при фиксированном значении $d=1$. Однако оценка качества подгонки при этом показала, что $\chi^2=118,17$ (при $df=8$), а соответствующая вероятность $P=0$. При расчете по традиционной схеме с использованием SPSS также оказалось, что $\chi^2=116,127$, а вероятность $P=0$. Вследствие того, что скорректированные данные не соответствуют



Доля (p) самок тигровой креветки *Penaeus semisulcatus* с наличием визуальных признаков половой зрелости (1) и в действительности являющихся половозрелыми (2) в зависимости от их длины (L). Горизонтальная асимптота (3) соответствует параметру $d=0,511$

The portion (p) of female individuals of the green tiger prawn *Penaeus semisulcatus* with visible indicators of maturity (1) that are actually mature (2) in relation to their length (L). Horizontal asymptote (3) corresponds to the parameter $d=0.511$

модели, расчет ее параметров был бы некорректен. Результаты проверки качества подгонки откорректированных данных показывают, что и после корректировки данные не стали соответствовать двухпараметрической логистической модели. Из этого следует, что для анализа данных в случае с неполной выраженностью признаков половозрелости целесообразно использовать трехпараметрическую модель, а не искусственно корректировать данные для последующего использования традиционных двухпараметрических моделей.

При расчете параметров полового созревания самок средиземноморского ботуса *Bothus podas* (Delaroche, 1809) по данным [5] результат оказался следующим: $b=-1,07$, $d=0,735$, $e=18,868$, а стандартные ошибки этих оценок, соответственно, равны 0,464, 0,203 и 0,784.

При оценке качества подгонки оказалось, что $\chi^2=2,14$ (при $df=5$), а вероятность $P=0,83$, что указывает на очень хорошее соответствие данных трехпараметрической логистической модели. Однако проверка значимости отличия d от 1 показала по t -критерию Стьюдента (при $df=5$), что это отличие не значимо даже на уровне 10 % (использовался односторонний критерий, т. к. d заведомо не более 1). Таким образом, несмотря на очень хорошее

соответствие данных трехпараметрической модели, нет оснований для утверждения о наличии половозрелых особей с отсутствующими признаками половозрелости. Исходя из этого, был выполнен расчет параметров при фиксированном значении $d=1$. Результат оказался следующим: $b=-0,743$, $e=19,823$, а стандартные ошибки этих оценок, соответственно, равны 0,145 и 0,312. Оценка качества подгонки показала, что $\chi^2=2,849$ (при $df=6$), а вероятность $P=0,827$. Таким образом, оказалось, что данные о половом созревании средиземноморского ботуса одинаково хорошо описываются как трехпараметрической, так и двухпараметрической логистической функцией (при $d=1$). Поэтому из того, что данные хорошо описываются трехпараметрической моделью, не следует делать вывод о том, что часть половозрелых особей в действительности пропускает нерест. Последнее можно утверждать только в том случае, если d статистически значимо отличается от 1. В тех случаях, когда отличие d от 1 не является статистически значимым (как в случае с ботусом), можно рекомендовать производить расчеты при фиксированном значении $d=1$. Это позволит обеспечить сопоставимость результатов с традиционно используемыми для оценки параметров полового созревания двухпараметрическими моделями.

Для средиземноморского ботуса был также выполнен расчет по традиционной схеме с использованием SPSS. Его результаты были практически такими же, как и с использованием пакета *drc* при $d=1$: $\chi^2=2,849$ (при $df=6$), а вероятность $P=0,828$; длина, при которой 50 % особей являются половозрелыми, также оказалась равной 19,823. Возможно, совпадение результатов объясняется тем, что в пакете *drc* в случае задания фиксированного значения параметра используется традиционный подход, предусматривающий линеаризацию логистической функции (как и в SPSS). Как указано в [4], в пакете *drc* используются модели зависимости «доза – эффект» в виде как нелинейной регрессии, так и обобщенной линейной регрессии (предусматривающей линеаризацию соответствующих функций).

В токсикологии обычно используются логарифмы величин доз. В пакете *drc* имеются как лог-логистическая, так и логнормальная модели, в которых вместо величин доз используются их логарифмы. Поскольку при изучении полового созревания гидробионтов значения их длины не принято логарифмировать, то авторы данной статьи рекомен-

дуют использовать именно логистическую модель, а не лог-логистическую или логнормальную.

ВЫВОДЫ

1. Для описания кривой полового созревания рыб и ракообразных в случаях, когда по каким-либо причинам у особей, являющихся на самом деле половозрелыми, отсутствуют признаки, по которым особи идентифицируются как половозрелые, могут быть использованы трехпараметрические логистическая функция или интеграл нормального распределения. Третий параметр является оценкой доли половозрелых особей, у которых признаки половозрелости выражены внешне, и задает верхнюю асимптоту кривой созревания.
2. Для оценки параметров кривой полового созревания гидробионтов в указанном выше случае рекомендуется использовать трехпараметрическую логистическую модель в пакете расширения *drc* для статистической программной среды открытого доступа R.
3. В тех случаях, когда задающий верхнюю асимптоту параметр не отличается статистически значимо от 1, рекомендуется использовать двухпараметрическую модель, задавая фиксированное значение данного параметра равным 1.

БЛАГОДАРНОСТИ

Авторы очень признательны К. Ритцу (Christian Ritz) за его любезные разъяснения по поводу деталей использования пакета *drc*, а также благодарят С.Н. Кульбу за ценную помощь в поиске литературы.

СПИСОК ЛИТЕРАТУРЫ

1. King M. Fisheries biology, assessment and management. The 2nd ed. Oxford: Blackwell Scientific Publications Ltd., 2007. 382 p.
2. Ritz C. Towards a unified approach to dose-response modeling in ecotoxicology // *Environmental Toxicology and Chemistry*. 2010. Vol. 29, no. 1. Pp. 220–229.
3. Венэблз У.Н., Смит Д.М. Введение в R версия 3.5.2 (2018-12-20). Заметки по R: среда программирования для анализа данных и графики / Пер. с англ. А.А. Фоменко. М.: Ridero, 2019. 150 с.
4. Ritz C., Baty F., Streibig J.C., Gerhard D. Dose-response analysis using R // *PLoS ONE*. 2015. Vol. 10, no. 12. e0146021. doi:10.1371/journal.pone.0146021.
5. Abid S., Ouannes-Ghorbel A., Jarboui O., Bouain A. Contribution to the study of the reproductive cycle of the wide-eyed flounder *Bothus podas podas* in the Gulf of Gabès (Tunisia) // *Journal of the Marine Biological Association of the United Kingdom*. 2010. Vol. 90, issue 3. Pp. 519–526.

REFERENCES

1. King M. Fisheries biology, assessment and management. The 2nd ed. Oxford: Blackwell Scientific Publications Ltd., 2007, 382 p.
2. Ritz C. Towards a unified approach to dose-response modeling in ecotoxicology. *Environmental Toxicology and Chemistry*, 2010, vol. 29, no. 1, pp. 220–229.
3. Venables W.N., Smith D.M. An introduction to R. Notes on R: a programming environment for data analysis and graphics Version 3.5.2 (2018-12-20). Bristol: Network Theory, 2018, 144 p.
4. Ritz C., Baty F., Streibig J.C., Gerhard D. Dose-response analysis using R. *PLoS ONE*, 2015, vol. 10, no. 12, e0146021. doi:10.1371/journal.pone.0146021.
5. Abid S., Ouannes-Ghorbel A., Jarboui O., Bouain A. Contribution to the study of the reproductive cycle of the wide-eyed flounder *Bothus podas podas* in the Gulf of Gabès (Tunisia). *Journal of the Marine Biological Association of the United Kingdom*, 2010, vol. 90, issue 3, pp. 519–526.

Поступила 12.12.2019

Принята к печати 17.02.2020