

ДИНАМИКА СИСТЕМЫ «РЫСЬ–ЗАЯЦ»: ПРИМЕНЕНИЕ МОДЕЛИ ЛОТКИ–ВОЛЬТЕРРА

© 2016 г. Л.В. Недорезов

Центр междисциплинарных исследований по проблемам окружающей среды РАН,
191187, Санкт-Петербург, наб. Кутузова, 14

E-mail: l.v.nedorezov@gmail.com

Поступила в редакцию 06.09.15 г.

Для аппроксимации известных эмпирических временных рядов по изменению численностей рысей и зайцев в Канаде, собранных компанией Гудзонова залива за 1845–1935 гг., использовалась модель Лотки–Вольтерра. Предполагалось, что модель дает удовлетворительную аппроксимацию данных, если множества отклонений модельных и эмпирических данных для каждого из временных рядов удовлетворяют ряду статистических тестов (для выбранного уровня значимости). Плотности распределений отклонений проверялись на симметричность относительно оси ординат и монотонность поведения ветвей при положительных и отрицательных значениях (критерии Колмогорова–Смирнова и Лемана–Розенблатта); последовательности отклонений тестировались на наличие/отсутствие сериальной корреляции (тесты Сведенберга–Эйзенхарта и «скачков вверх – скачков вниз»). Численные расчеты показывают, что множество точек пространства параметров модели, при которых отклонения удовлетворяют статистическим критериям, не пусто и, следовательно, модель пригодна для описания эмпирических данных.

Ключевые слова: модель Лотки–Вольтерра, оценка параметров модели, анализ отклонений.

Одной из самых интересных (и важных для развития математической экологии) является задача аппроксимации известных данных, собранных компанией Гудзонова Залива [1–5] по колебаниям численностей зайцев (*Lepus americanus*) и рысей (*Lynx lynx*), с помощью траекторий модели Лотки–Вольтерра [6–9]:

$$\frac{dx}{dt} = ax - bxy, \quad \frac{dy}{dt} = -cy + dxy, \quad (1)$$

где $x(t)$ – численность жертв (зайцев), $y(t)$ – численность хищников (рысей) в момент времени t ; a – Мальтузианский параметр популяции жертв (в дальнейшем предполагается, что $a > 0$; если $a \leq 0$ система из двух взаимодействующих видов вырождается асимптотически при любых начальных значениях численностей), c – Мальтузианский параметр популяции хищников (интенсивность гибели особей), b, d – коэффициенты взаимодействия популяций. В модели (1) все коэффициенты неотрицательны, $a, b, c, d = \text{const} \geq 0$. На рис. 1 представлены данные по колебаниям численностей рысей и зайцев в Канаде (в тысячах сданных шкур за год).

Попытки аппроксимации представленных данных с помощью модели (1) препринимались неоднократно, но каждый раз оканчивались безуспешно. В качестве основных причин неудач указывались следующие: это и негрубость модели (1), и априорное несоответствие данных теоретическим результатам (в некоторые годы наблюдается «неправильное» расположение максимумов численностей – в соответствии с моделью (1) максимум численности хищников всегда должен наблюдаться после максимума численности жертв), «неправильное» движение эмпирической траектории на плоскости «численность жертв – численность хищников» (в некоторых случаях эмпирическая траектория «закручена» в обратную сторону) и другие причины. Кроме этого, указывались (возможные) особенности сбора данных [3–5].

В частности, М. Гилпин [3,4], оценивая параметры более общей, чем (1), модели (которая включает (1) как частный случай), следовал методике, предложенной Ф. Дж. Аялой [10,11]. В соответствии с этой методикой левые части уравнений (1) заменяются на приращения соответствующих численностей за год, после чего параметры оцениваются при минимизации суммы квадратов отклонений (квадратов разностей левых и правой частей уравнений модели).

Сокращение: МНК – метод наименьших квадратов.

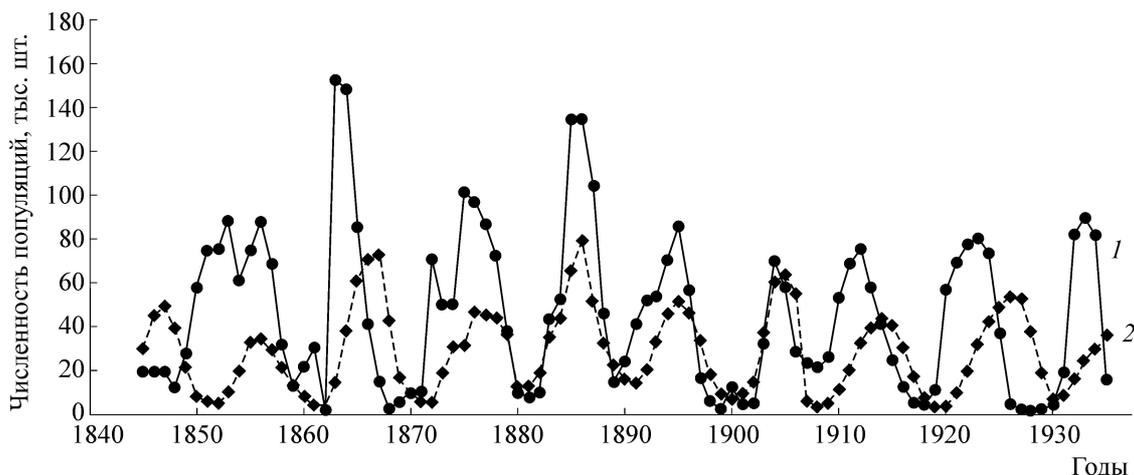


Рис. 1. Колебания численностей рысей и зайцев в Канаде с 1845 по 1935 года: 1 – кривая изменения численности зайцев, 2 – кривая изменения численности рысей.

Заметим, что при таком подходе к оценке параметров используется не сама исходная модель, а ее дискретный аналог, получаемый с помощью схемы Эйлера. С помощью формулы Адамса можно было бы получить счетное множество дискретных аналогов исходной модели, с помощью которых, не исключено, можно было бы получить более правдоподобные результаты. Кроме этого, при этом подходе эмпирическая траектория приближается вовсе не траекторией модели, которая «ближе» других к исходным данным. Наконец, метод наименьших квадратов (МНК) весьма часто приводит к неверным результатам: модель приходится признавать непригодной для аппроксимации данных, когда при близких к МНК-оценкам параметров модель дает удовлетворительное описание [12–16].

В настоящей работе исходные данные [1–4] аппроксимируются траекториями модели (1): система интегрируется методом Рунге–Кутты 4-го порядка с шагом 10^{-3} ; считается, что единица времени модели (тысяча шагов) равняется одному году. Параметры модели (1) и начальные значения численностей популяций выбирались случайным образом (равномерное распределение) в области $\Delta: x_0, y_0 \in [0, 220], a, b, d \in [0, 1], c \in [0, 20]$. После разыгрывания случайной точки и нахождения соответствующей траектории модели (1) вычислялись две последовательности отклонений эмпирических данных от модельной траектории. Для каждой из последовательностей отклонений проверялся ряд статистических тестов (для 5% уровня значимости). Только в том случае, если все используемые тесты для каждой из последовательностей отклонений показывали требуемый результат, точка простран-

ства параметров считалась принадлежащей допустимому множеству. Если один какой-либо тест показывал нетребуемый результат, соответствующая точка считалась не принадлежащей допустимому множеству.

Если для выбранного уровня значимости допустимое множество оказывается пустым, то только в этом случае модель следует считать непригодной для аппроксимации данных. Если же допустимое множество не пусто, то оснований для утверждения о непригодности модели нет. Цель настоящей работы состоит в том, чтобы показать, что широко распространенное мнение о непригодности модели (1) для аппроксимации каких-либо данных вообще не является состоятельным. Зайцам вовсе не нужно быть кровожадными и потреблять в пищу рысей – достаточно хорошее соответствие траекторий модели и эмпирической траектории наблюдается и при неотрицательных значениях параметров.

ИСПОЛЬЗУЕМЫЕ СТАТИСТИЧЕСКИЕ КРИТЕРИИ

Пусть дана выборка $\{(x_k, y_k)\}, k = 0, 1, \dots, N$, где x_k – численность популяции жертв (зайцев) в k -й момент времени, y_k – численность популяции хищников (рысей) в тот же момент времени, $N + 1$ – объем выборки ($N = 90$). Пусть также $x = x(t, x_0, y_0, \vec{\alpha}), y = y(t, x_0, y_0, \vec{\alpha})$ – решение системы (1) при заданных начальных значениях численностей x_0, y_0 и заданных значениях параметров модели $\vec{\alpha}$.

Пусть e_k^1 и e_k^2 – следующие величины:

$$e_k^1 = x_k - x(k, x_0, y_0, \vec{\alpha}), \quad e_k^2 = y_k - y(k, x_0, y_0, \vec{\alpha}).$$

Пусть также $\{e_k^{j+}\}$ – совокупность положительных отклонений, $\{e_k^{j-}\}$ – совокупность отрицательных отклонений, $j = 1, 2$.

Будем считать, что траектория модели дает удовлетворительную аппроксимацию данных, если:

1. Плотности распределений отклонений $\{e_k^1\}$ и $\{e_k^2\}$ симметричны относительно оси ординат. Симметричность относительно оси ординат означает, что выборки $\{e_k^{j+}\}$ и $\{e_k^{j-}\}$, $j = 1, 2$, имеют одну и ту же функцию распределения (для разных значений j эти функции различны). Поэтому для проверки симметричности можно использовать критерии однородности выборок Колмогорова–Смирнова, Вальда–Вольфовица, Манна–Уитни и Лемана–Розенблатта [17–19]. Однако разыгрывание 10^7 случайных точек в Δ не позволило выявить таких значений параметров, при которых две последовательности отклонений одновременно удовлетворяют всем этим критериям. Поэтому в дальнейших расчетах использовался «ослабленный» вариант – при тестировании были использованы только критерии Колмогорова–Смирнова и Лемана–Розенблатта.

2. Ветви плотности распределения должны быть монотонны: значение плотности должно возрастать при отрицательных значениях аргумента и снижаться в области положительных значений (большие отклонения должны наблюдаться с меньшими вероятностями). Для проверки монотонности поведения ветвей плотности распределения были использованы коэффициенты корреляции Спирмена и Кендалла [17].

Пусть $\{e_k^{j*+}\}$ – упорядоченная по возрастанию выборка $\{e_k^{j+}\}$: $\{e_1^{j*+}\} < \{e_2^{j*+}\} < \dots$. При монотонном снижении плотности распределения в идеальном случае длины интервалов $[0, e_1^{j*+}]$, $[e_1^{j*+}, e_2^{j*+}]$, ... также должны быть упорядочены по возрастанию (и могут быть ранжированы 1, 2, ...). Пусть ρ – ранговый коэффициент корреляции Спирмена. Сравнивая ранги длин интервалов для выборки $\{e_k^{j*+}\}$ с идеальным случаем, нулевая гипотеза $H_0: \rho = 0$ (при альтернативной гипотезе $H_1: \rho > 0$ и выбранном уровне значимости) должна быть отклонена. В дальнейших расчетах использовался «ослабленный» вариант – для каждой точки пространства параметров Нулевая гипотеза проверялась два раза: для выборок $\{e_1^{j+}\} \cup \{e_2^{j+}\}$, $j = 1, 2$.

3. В последовательности отклонений не должно быть сериальной корреляции. Для этих целей использовались тест серий Сведа–Эйзенхарта [20,21]. В некоторых случаях дополнительно использовался тест серий «скачков вверх – скачков вниз» [18].

Если какой-либо из используемых статистических критериев давал негативный результат, то соответствующая точка пространства параметров модели считалась не принадлежащей допустимому множеству. Если же все используемые критерии давали положительный результат (т.е. гипотеза о симметрии относительно оси ординат распределения отклонений не отвергалась, не отклонялась гипотеза об отсутствии сериальной корреляции и не отклонялась гипотеза о монотонном поведении ветвей функций плотностей распределений), то это означало, что при тестируемых значениях параметров модель вполне может использоваться для аппроксимации исходных данных и объяснения особенностей популяционной динамики, а соответствующая точка принадлежит допустимому множеству.

На следующем этапе в допустимом множестве выбираются точки с экстремальными характеристиками (метод экстремальных точек [14–16]), а именно такие точки, для которых требуемый результат может быть получен с максимальным (или в некоторых случаях с минимальным) уровнем значимости. Если, к примеру, гипотеза о симметрии распределения не может быть отклонена с уровнем значимости 5%, то это может служить основанием для принятия нулевой гипотезы; но существенно более сильный результат наблюдается в случае, когда та же гипотеза не может быть отклонена с 95% уровнем значимости. Увеличение уровня значимости (в некоторых случаях – уменьшение) характеризует более хорошее соответствие теоретических и эмпирических данных.

РЕЗУЛЬТАТЫ РАСЧЕТОВ

Некоторые авторы полагают [1–5], что колебания численностей взаимодействующих видов имеют примерно 10-летний цикл. Однако в последовательности максимумов численностей такая закономерность не прослеживается: в частности, максимумы могут повторяться и через два года (рис. 1). Конечно, если учитывать одни максимумы и не учитывать другие, то вполне можно обнаружить и 10-летний цикл.

Анализ элементов допустимого множества показывает, что среди возможных динамических режимов имеются 10-летние циклы (рис. 2). Режим, представленный на рис. 2, реализуется

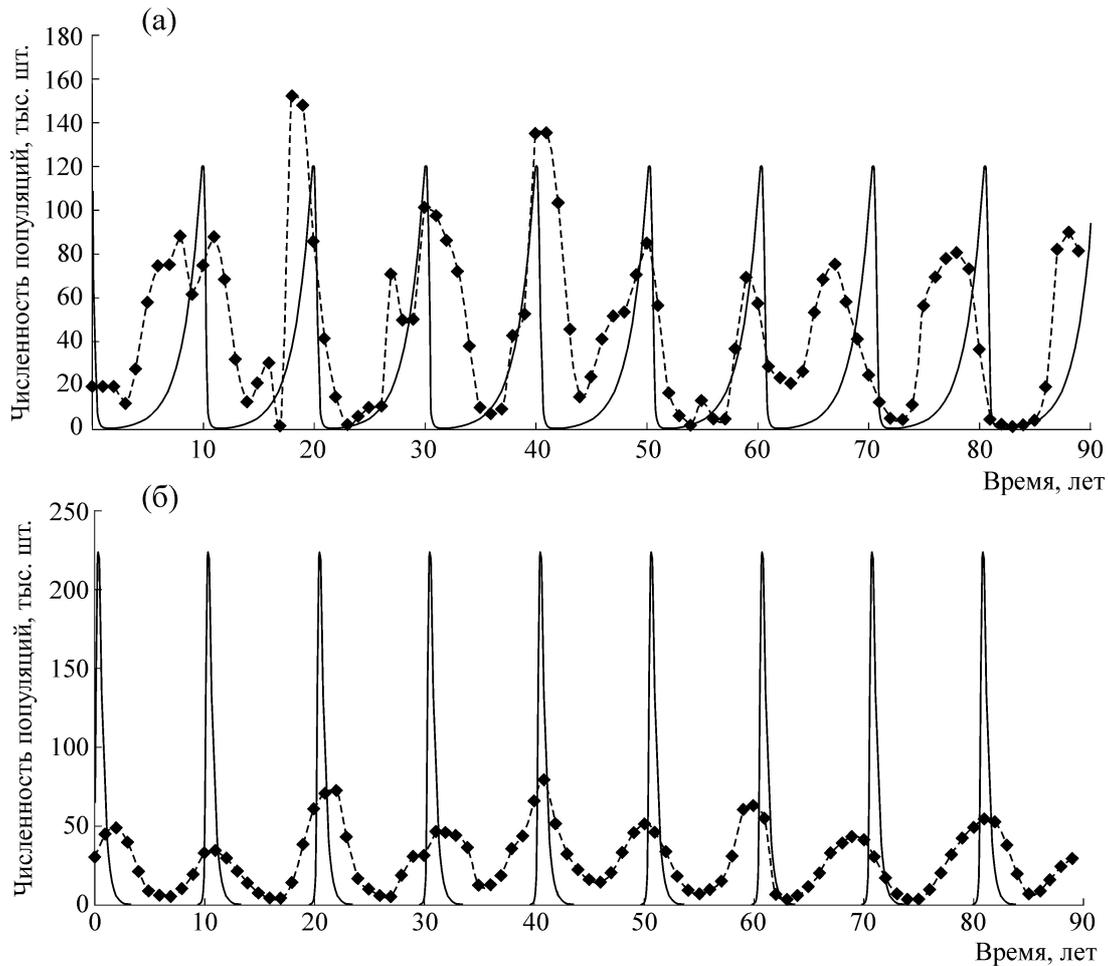


Рис. 2. Результаты аппроксимации данных (10-летний цикл): (а) – для численности зайцев, (б) – для численности рысей. Сплошная линия – траектория модели (1). Штриховая линия – эмпирические данные.

при $x_0 = 115,0546$, $y_0 = 45,7497$, $a = 0,631315$, $b = 0,036912$, $c = 2,49187$, $d = 0,104151$ и обладает следующими характеристиками: с уровнями значимости 43,57 и 10,19% (для $\{e_k^1\}$ и $\{e_k^2\}$ соответственно) гипотезы о симметричности распределений не могут быть отклонены по критерию Колмогорова–Смирнова, с 49,54 и 6,4% – по критерию Лемана–Розенблатта.

Коэффициенты корреляции Спирмена равны 0,4352 и 0,3113 с соответствующими p -уровнями $1,6 \cdot 10^{-5}$ и $2,67 \cdot 10^{-3}$. Коэффициенты корреляции Кенделла равны 0,3153 и 0,2181 с соответствующими p -уровнями $1,0 \cdot 10^{-5}$ и $2,2 \cdot 10^{-3}$. Нулевые гипотезы (о равенстве коэффициента нулю) отклоняются с весьма малыми уровнями значимости. С уровнями значимости 14,7 и 35,4% не отклоняется гипотеза об отсутствии сериальной корреляции по критерию Сведенборга–Эйзенхарта. Можно также отметить, что и визуально наблюдается достаточно хорошее соответствие теоретической кривой и эмпириче-

ских данных (рис. 2а,б): максимумы траектории модели располагаются весьма близко к максимальным значениям исходной выборки.

В допустимом множестве имеются элементы, для которых соответствующие отклонения удовлетворяют большему числу статистических тестов. В частности, точка с координатами $x_0 = 53,3244$, $y_0 = 107,0217$, $a = 0,80522$, $b = 0,025355$, $c = 6,476933$, $d = 0,128812$ обладает следующими характеристиками: с уровнями значимости 99,81 и 6,46% гипотезы о симметричности распределений не могут быть отклонены по критерию Колмогорова–Смирнова, с 97,43 и 9,83% – по критерию Лемана–Розенблатта, с 44,23 и 26,72% – по критерию Вальда–Вольфовица, с 87,1 и 14,08% – по критерию Манна–Уитни.

Коэффициенты корреляции Спирмена равны 0,27605 и 0,4459 с соответствующими p -уровнями $8,08 \cdot 10^{-3}$ и $9,0 \cdot 10^{-6}$. Коэффициенты корреляции Кенделла равны 0,1912 и 0,3118 с соответствующими p -уровнями $7,27 \cdot 10^{-3}$ и $1,2 \cdot 10^{-5}$.

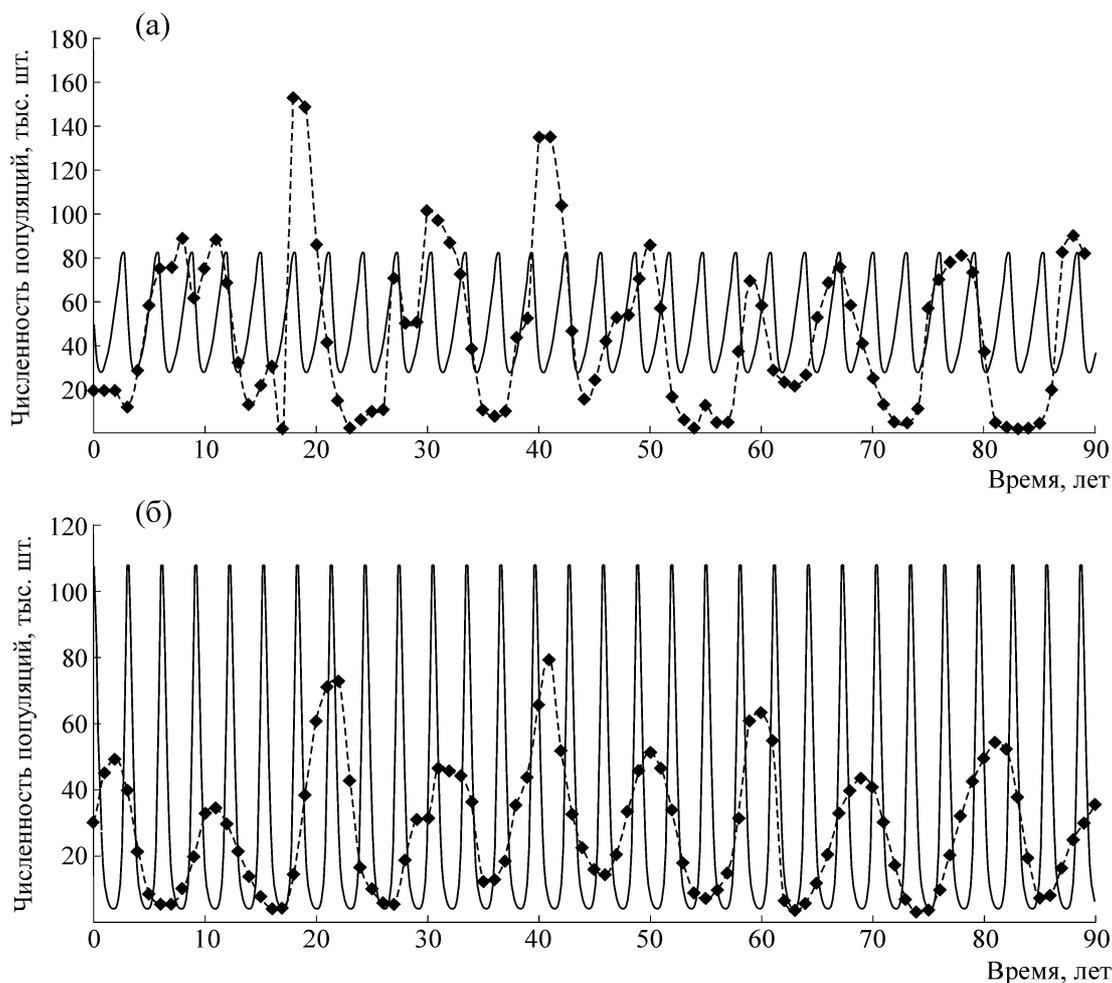


Рис. 3. Результаты аппроксимации данных (3-летний цикл): (а) – для численности зайцев, (б) – для численности рысей. Сплошная линия – траектория модели (1). Штриховая линия – эмпирические данные.

Нулевые гипотезы (о равенстве коэффициента нулю) отклоняются. С уровнями значимости 16,22 и 35,18% не отклоняется гипотеза об отсутствии сериальной корреляции по критерию Сведа–Эйзенхарта. Критические значения при 20%-м уровне значимости для критерия «скачков вверх – скачков вниз» [18] равны 54 и 66. Отклонения для временного ряда для зайцев имеют 56 серий, для другого временного ряда – 59. Таким образом, гипотезы о наличии сериальной корреляции в последовательностях отклонений должны быть отклонены.

На рис. 3 представлены эмпирические данные и траектория модели (1) для рассматриваемого варианта. Длина цикла – 3,06 (года). Визуально представленные на рис. 3 результаты несколько хуже представленных на рис. 2: и амплитуда колебаний явно не соответствует эмпирическим данным (для траектории изменения численности зайцев амплитуда значительно меньше, чем у временного ряда; для траектории

изменения численности рысей наоборот), и частота колебаний существенно меньше эмпирической.

Поиск элемента допустимого множества, соответствующего минимуму суммы квадратов отклонений для обоих временных рядов, показал, что это цикл длиной 2,06 лет, что также не соответствует априорным предположениям. Картина для временного ряда для численности зайцев сходна с представленной на рис. 3а (амплитуда колебаний имеет близкое значение, но частота, очевидно, больше). Амплитуда колебаний численности рысей существенно меньше эмпирической.

Анализ 30000 случайно выбранных точек из допустимого множества показывает, что с наибольшей вероятностью (0,37107) встречаются циклы длиной из полуоткрытого интервала [1,2); с вероятностью 0,2972 – из [2,3), с вероятностью 0,1595 – из [3,4). Остальные вероятности существенно меньше. В частности, циклы

длиной из интервала [10,11) встречаются с вероятностью 0,001439. Циклы длиной более 12 не были выявлены.

ЗАКЛЮЧЕНИЕ

Если в рамках метода наименьших квадратов предполагается, что первоначально найдутся некие оценки параметров модели и только после этого проверяются свойства последовательностей отклонений теоретических и эмпирических данных, то в рамках используемого в настоящей работе подхода все осуществляется с точностью до наоборот. Сначала находятся точки пространства параметров, при которых последовательности отклонений теоретических и эмпирических данных обладают необходимыми свойствами (допустимое множество), и только после этого из найденных выбираются точки, обладающие экстремальными свойствами.

Отказ от метода наименьших квадратов обусловлен целым рядом обстоятельств. Это и отсутствие критериев для выбора вида минимизируемого функционала (который, как правило, не имеет никакого отношения ни к биологической проблеме, ни к имеющимся данным, ни к модели); это и финальный результат – вердикт о пригодности или непригодности модели к аппроксимации данных выносится на основе свойств одной-единственной точки пространства параметров (глобального минимума функционала). И, как результат, обилие неверных результатов при использовании нелинейных моделей. Один из таких примеров как раз и рассмотрен в настоящей работе.

Анализ соответствия модели Лотки–Вольтерра (без саморегуляции в популяциях) эмпирическим данным по колебаниям численностей зайцев и рысей в Канаде показывает, что существует множество точек в пространстве параметров модели, дающих достаточно хорошую аппроксимацию данных. Причем удовлетворяется не только ряд статистических критериев, но и в некоторых случаях визуально наблюдается соответствие теоретических и эмпирических данных.

При использовании метода наименьших квадратов для получения некоего соответствия данным требуется поменять знаки у коэффициентов, описывающих взаимодействие видов.

Иными словами, соответствие возможно, если зайцы станут кровожадными и будут потреблять рысей в пищу. Однако, как показывает проведенный анализ, в действительности ничего подобного (ни от зайцев, ни от модели) не требуется.

СПИСОК ЛИТЕРАТУРЫ

1. D. A. MacLulich, *Fluctuations in the numbers of varying hare (Lepus americanus)* (The University of Toronto press, Toronto, 1937).
2. C. S. Elton and M. Nicholson, *J. Anim. Ecol.* **11**, 215 (1942).
3. M. E. Gilpin, *Am. Nat.* **107** (957), 727 (1973).
4. М. Гилпин, Самарская Лука: проблемы региональной и глобальной экологии **19** (3), 177 (2010).
5. Г. С. Розенберг, Самарская Лука: проблемы региональной и глобальной экологии **19** (3), 180 (2010).
6. В. Вольтерра, *Математическая теория борьбы за существование* (Наука, М., 1976).
7. V. Volterra, *Lecons sur la theorie mathematique de la lutte pour la vie* (Gauthiers-Villars, Paris, 1931).
8. A. I. Lotka, *J. Amer. Chem. Soc.* **42** (8), 1595 (1920).
9. A. I. Lotka, *Elements of physical biology* (Williams & Wilkins, Baltimore, 1925).
10. F. J. Ayala, M. E. Gilpin, and J. G. Ehrenfeld, *Theor. Pop. Biol.* **4**, 331 (1973).
11. M. E. Gilpin and F. J. Ayala, *Proc. Nat. Acad. Sci. USA* **70**(12), 3590 (1973).
12. Л. В. Недорезов, *Журн. общ. биологии* **73** (2), 114 (2012).
13. L. V. Nedorezov, *Population Dynamics: Analysis, Modelling, Forecast* **1** (1), 47 (2012).
14. Л. В. Недорезов, *Хаос и порядок в популяционной динамике: моделирование, анализ, прогноз* (LAP Lambert Acad. Publ., Саарбрюкен, 2012).
15. Л. В. Недорезов, *Биофизика* **60** (3), 564 (2015).
16. Л. В. Недорезов, *Биофизика* **60** (5), 1039 (2015).
17. Л. Н. Большев и Н. В. Смирнов, *Таблицы математической статистики* (Наука, М., 1983).
18. И. Ликеш и Й. Ляга, *Основные таблицы математической статистики* (Финансы и статистика, М., 1985).
19. М. Холлендер и Д. Вулф, *Непараметрические методы статистики* (Финансы и статистика, М., 1983).
20. Н. Дрейпер и Г. Смит, *Прикладной регрессионный анализ* (Финансы и статистика, М., 1986), т. 1.
21. Н. Дрейпер и Г. Смит, *Прикладной регрессионный анализ* (Финансы и статистика, М., 1987), т. 2.

Dynamics of a “Lynx–Hare” System: an Application of the Lotka–Volterra Model

L.V. Nedorezov

*Research Center for Interdisciplinary Environmental Cooperation, Russian Academy of Sciences,
nab. Kutuzova 14, St. Petersburg, 191187 Russia*

For approximation of well-known empirical time series on lynx–hares system dynamics in Canada, collected by the Hudson’s Bay Company in 1845–1935, the Lotka–Volterra model of predator–prey system dynamics was used. As it was assumed the model demonstrates good fitting of datasets if a group of statistical criterions show required results for both considering time series. Sets of deviations between theoretical (model) trajectories and empirical datasets were tested for symmetry (with respect to y-axis; the Kolmogorov–Smirnov and Lehmann–Rosenblatt tests) and the existence/absence of serial correlation (Swed–Eisenhart test and “jumps up–jumps down test). For testing of monotonic behavior of branches of density function of distribution of residuals the rank correlation coefficients were used. Provided analysis showed that it is possible to point out values of model parameters when model trajectory gives a good fitting of dataset, and all statistical criterions give required results.

Key words: the Lotka–Volterra model, model parameter estimation, analysis of deviations