

Динамика численности сосновой пяденицы в Нидерландах: оценка параметров обобщённой дискретной логистической модели

Dynamics of a pine looper moth population in The Netherlands: estimation of parameters of a generalized discrete logistic model

А.В. Недорезов

L.V. Nedorezov

Учреждение Российской Академии наук Центр междисциплинарных исследований по проблемам окружающей среды РАН, наб. Кутузова 14, Санкт-Петербург 191187 Россия. E-mail: l.v.nedorezov@gmail.com
Research Center for Interdisciplinary Environmental Cooperation RAS, Nab. Kutuzova 14, St.-Petersburg 191187 Russia.

Ключевые слова: сосновая пяденица, временные ряды, обобщённая дискретная логистическая модель, статистические критерии.

Key words: pine looper moth, time series, generalized discrete logistic model, statistical criterions

Резюме. В работе рассматриваются проблемы известных методов оценки параметров экологических моделей (в том числе, метода наименьших квадратов) и предлагается альтернативный подход, базирующийся на первоначальном определении границ и структуры допустимого множества. Для аппроксимации известных данных по колебаниям численности сосновой пяденицы в Нидерландах [Klomp, 1966] была использована обобщённая дискретная логистическая модель, параметры которой были оценены с помощью метода наименьших квадратов. Для этих же данных и модели были получены допустимые множества, для каждого элемента которых модель удовлетворяет целому ряду статистических критериев. Сравнение полученных результатов показывает, что во всех трёх случаях метод наименьших квадратов не позволил получить наилучший результат. Также показано, что существуют такие значения параметров, при которых дискретная логистическая модель позволяет получать долгосрочные прогнозы изменения численности пяденицы.

Abstract. A new alternative method of pine looper moth population parameter estimation, is based on an initial determination of structure and boundaries of feasible sets, is offered. For an approximation of well-known time series for pine looper moth dynamics in The Netherlands [Klomp, 1966] a generalized discrete logistic model was used. Parameters of this model were estimated with Least Squared Method (LSM), and for the same time series feasible sets were determined and the results were compared. In all three situations LSM-estimations belong to domains where the model has no prognostic properties, and an asymptotic dynamic regime for the population could not be determined. An alternative method allowed one to obtain sets of points (in a space of model parameters) which can be used to forecast and create hypotheses about dynamic regimes.

Введение

Для моделирования динамики численности насекомых часто используются модели следующего вида

[Кондаков, 1974; Голубев и др., 1980; Исаев и др., 1984, 2001; Корзухин, Семевский, 1992 и др.]:

$$x_{k+1} = F(x_k, \vec{\alpha}). \quad (1)$$

В модели (1) x_k — численность популяции (или её плотность) в некоторый момент времени k , $k = 0, 1, 2, \dots$ (время дискретно); $\vec{\alpha}$ — вектор параметров модели. При решении задач по оценке значений параметров модели (1) по эмпирическим данным предполагается [Недорезов, 2010, 2011, 2012a], что начальная численность популяции x_0 (в момент времени $k = 0$) также является неизвестным параметром, который требуется оценить по имеющейся выборке. Для экспериментальных данных эта величина может быть известной [Гаузе, 1934; Гаузе, Витт, 1934; Недорезов, 2012б; Nedorezov, 2011].

Пусть $\{x_k^*, k = 0, 1, \dots, N\}$ — имеющийся временной ряд, значения численности популяции в различные моменты времени (например, раз в год); $N + 1$ — размер выборки. Используя эти значения необходимо оценить значения параметров модели $\vec{\alpha}$ и x_0 .

Для оценки значений параметров моделей используются самые различные методы и приемы: от экспертных оценок до байесовского подхода [Бард, 1979; Большев, Смирнов, 1983; Боровков, 1984; McCallum, 2000; Wood, 2001a, 2001b]. Один из таких методов — это метод наименьших квадратов (МНК), который в настоящее время широко используется при анализе биологических данных [регрессионные задачи; Дрейпер, Смит, 1986, 1987; Рокицкий, 1973; Доспехов, 1985; Лакин, 1990]. Использование МНК для оценки параметров модели предполагает, прежде всего, выбор (или конструирование) некоей функции потерь, которая отражает уровень соответствия теоретических (модельных) и эмпирических данных.

В частности, функция потерь может быть выбрана в виде суммы квадратов отклонений эмпирических значений от *траектории модели* (1) [Глобальное приближение; Wood, 2001a, b]:

$$Q(\bar{\alpha}, x_0) = \sum_{k=0}^N (x_k(\bar{\alpha}, x_0) - x_k^*)^2 \quad (2)$$

В формуле (2) $(x_k(\bar{\alpha}, x_0))$ — решение уравнения (1) при фиксированных значениях параметров $\bar{\alpha}$ и начального значения численности x_0 . Известны различные модификации выражения (2), когда, например, вместо квадратов отклонений выбираются их абсолютные значения, или выбираются квадраты отклонений с какими-либо неотрицательными весами и так далее.

После определения вида функции потерь (2) находят значения величин параметров, минимизируя это выражение. Считается, что таким образом мы получаем наилучшие значения оценок параметров. На следующем шаге процесса анализируется множество отклонений теоретических значений от эмпирических. Пусть

$$e_k = x_k(\bar{\alpha}, x_0) - x_k^*$$

есть те самые отклонения, полученные при наилучших значениях параметров. В рамках традиционного подхода к оценке параметров считается, что являются значениями независимых нормально распределённых случайных величин с нулевым средним, что может быть проверено, например, с помощью тестов Колмогорова – Смирнова, Лиллифорса и Шапиро – Вилка [Shapiro et al., 1968; Lilliefors, 1967; Бард, 1979; Большев, Смирнов, 1983]. Также считается, что в последовательности отклонений не должно быть зависимостей (сериальной корреляции). Если подобная зависимость обнаруживается, то это даёт основание для утверждения о том, что в модели не учитывается какой-то существенный фактор и, соответственно, модель нуждается в модификации. Проверка на наличие /отсутствия сериальной корреляции в последовательности отклонений осуществляется обычно с помощью теста Дарбина – Уатсона и теста серий Свуда – Эйзенхарта [Дрейпер, Смит, 1986, 1987].

Если в последовательности отклонений обнаруживается сериальная корреляция или гипотеза о нормальности совокупности величин $\{e_k\}$ отклоняется (при выбранном уровне значимости), то это даёт основание для признания модели непригодной для аппроксимации рассматриваемых данных. Иными словами, окончательное решение о пригодности или непригодности модели для аппроксимации данных выносится на основе анализа поведения модели при *одной-единственной точке пространства параметров*.

Возникает целый ряд естественных вопросов, связанных с применением МНК к анализу биологических процессов. В частности, какое отношение имеет

функционал (2) к анализируемому биологическому процессу? Ответ очевиден: никакого отношения не имеет (как и все возможные его модификации). Но, если так, то и значения параметров, при которых этот функционал достигает минимум, не могут характеризовать возможности модели для аппроксимации данных?

Еще один вопрос представляется весьма важным — это вопрос о проверке нормальности распределения отклонений. Имеет ли смысл предположение о нормальности отклонений к биологическим задачам или нет? Единственно возможный ответ — нет, не имеет. И в подтверждение этого ответа можно привести большое количество различных примеров. В частности, если измеряется вес гусениц в миллиграммах, то предположение о нормальности ошибок измерений означает, что ошибка с *положительной вероятностью* может быть в несколько тонн, как и с положительной вероятностью можно получить гусеницу с отрицательным весом.

Сказанное выше позволяет сделать следующее заключение: МНК базируется на *логической ошибке* (выбор функции потерь и её минимизация до определения требований к модели и нахождения допустимых множеств в пространстве параметров), а также *небиологических требованиях* к биологическим данным. Под допустимым множеством понимается такая совокупность точек в пространстве параметров, при которых модель даёт удовлетворительное описание данных (для выбранного уровня значимости и совокупности выбранных статистических критериев).

В настоящей работе рассматривается альтернативный подход к оценке параметров экологических моделей: сначала определяются основные требования к модели и к отклонениям теоретических значений от эмпирических, определяются соответствующие статистические тесты, а после этого проводятся численные расчеты для выявления допустимых множеств в ограниченной части пространства параметров. Численные расчеты проведены для данных по колебаниям плотности сосновой пяденицы (*Bupalus piniarius* L.) в Нидерландах [Klomp, 1966] при аппроксимации данных с помощью обобщённой дискретной логистической модели [Недорезов, 2012a].

Основные требования к модели

Требования к модели — это, по сути, требования к свойствам отклонений теоретических (модельных) от эмпирических данных. Большая часть этих требований / ограничений хорошо известна и используется на практике (но только для анализа отклонений при «наилучших значениях параметров»). Все эти требования можно условно разбить на следующие три группы:

1. Отклонения между теоретическими и эмпирическими данными должны иметь симметричную относительно нуля плотность распределения (ошибки измерений плотности популяции в ту или другую сторону должны иметь одинаковые вероятности). Кроме этого, обе ветви плотности распределения (в отрицательной и положительной областях) должны быть монотонными функциями: большие ошибки должны иметь меньшую вероятность. Эти ограничения являются более слабыми, чем требование нормальности распределения: указанным ограничениям удовлетворяет и нормальное распределение, и многие другие [Корн, Корн, 1973].

Пусть $\{e_k^+\}$ — совокупность положительных отклонений, $\{e_k^-\}$ — совокупность отрицательных отклонений (взятых со знаком минус). Симметрия плотности распределения относительно оси ординат означает, что для выбранного уровня значимости гипотеза об однородности двух выборок $\{e_k^+\}$ и $\{e_k^-\}$ не может быть отклонена. Для проверки симметрии распределения использовались тесты Колмогорова – Смирнова, Лемана – Розенблатта и Манна – Уитни [Hollander, Wolfe, 1973; Большев, Смирнов, 1983; Ликеш, Ляга, 1985].

Проверка монотонности поведения ветвей плотности распределения проверялась с помощью рангового коэффициента корреляции Спирмена [Большев, Смирнов, 1983]. Пусть $\{e_k^{*+}\}$ — упорядоченная по возрастанию выборка $\{e_k^+\}$: $e_1^{*+} < e_2^{*+} < \dots$. Если ветви плотности распределения монотонны (и при этом правая ветвь монотонно убывает), то тогда в *идеальном варианте* длины интервалов $[0, e_1^{*+}]$, $[e_1^{*+}, e_2^{*+}]$, ... также должны образовывать упорядоченную по возрастанию последовательность. Этот идеальный вариант может быть сопоставим с той последовательностью рангов, которая определяется выборкой $\{e_k^{*+}\}$.

Пусть ρ — ранговый коэффициент корреляции Спирмена. Нулевая гипотеза $H_0: \rho = 0$ (при альтернативной гипотезе $H_1: \rho > 0$ и фиксированном уровне значимости) должна быть отклонена. Если нулевая гипотеза отклоняется, то это служит определённой гарантией того, что у ветвей плотности распределения наблюдается монотонное поведение.

2. Общепринятое ограничение: в последовательности отклонений не должно быть сериальной корреляции. Наличие сериальной корреляции в последовательности свидетельствует, например, об отсутствии в модели каких-то важных компонентов. Соответственно, это свидетельствует о том, что модель непригодна для аппроксимации данных. Для тестирования наличия/отсутствия сериальной корреляции использовались тест серий Свуда – Эйзенхарта [Дрейпер, Смит, 1986, 1987] и тест серий «вверх и вниз» [Ликеш, Ляга, 1985].

3. Выполнение ограничений, представленных выше, ещё не даёт достаточных оснований для заключения о пригодности модели для аппроксимации данных. В частности, нет никаких оснований для утверждения о пригодности модели, если в каждом случае, когда временной ряд показывает увеличение численности, модель демонстрирует её снижение или наоборот. Естественно считать, что модель можно считать пригодной для аппроксимации данных только в том случае, если число разнонаправленных изменений численности во временном ряду и модели относительно невелико.

Пусть q — доля случаев, когда модель и временной ряд показывают однонаправленное изменение численности популяции. Если модель соответствует имеющимся эмпирическим данным, то тогда нулевая гипотеза $H_1: q > 0,5$ (при альтернативной гипотезе $H_1: q > 0,5$ и выбранном уровне значимости) должна быть отклонена.

Проверка выполнения всех указанных выше статистических критериев (именно в том порядке, в каком они представлены) в точках пространства параметров (параметры меняются с относительно небольшими фиксированными шагами) позволяет выявить допустимое множество, в каждой точке которого все критерии показывают требуемый (положительный) результат. Если какой-либо один критерий показывает отрицательный результат или не может быть применён к анализу отклонений (в силу незначительного количества элементов в той или иной выборке), соответствующая точка считается непринадлежащей допустимому множеству. Ниже представлены допустимые множества для обобщённой дискретной логистической модели при её применении к аппроксимации данных по колебаниям численности сосновой пяденицы в Нидерландах [Klopp, 1966] в проекции на плоскость (a, b) . Если для каких-либо фиксированных значений параметров a и b находилось хотя бы одно значение x_0 такое, что точка (x_0, a, b) принадлежит допустимому множеству, в проекции точка отмечалась чёрным цветом. Если таких значений x_0 не находилось, то точка на плоскости (a, b) отмечалась белым цветом.

Заметим, что указанные требования являются своеобразным минимальным набором ограничений, которым должна удовлетворять каждая модель, если она предназначена для того, чтобы быть *законом популяционной динамики*.

Обобщённая дискретная логистическая модель

Для аппроксимации данных по колебаниям численности сосновой пяденицы использовалась обобщённая дискретная логистическая модель, которая имеет следующий вид [Недорезов, 2012a]:

$$x_{k+1} = \begin{cases} ax_k(b - x_k), & 0 \leq x_k \leq b, \\ 0, & b < x_k. \end{cases}, \quad a, b = \text{const} \geq 0 \quad (3)$$

В (3) a, b — максимальное значение коэффициента размножения, которое наблюдается лишь при достаточно низких значениях численности популяции; параметр b обычно трактуется как максимально возможная численность, но в рамках модели (3) допускается однократное превышение этого уровня (после чего численность считается тождественно равной нулю).

Эта модель обладает весьма богатым набором динамических режимов, включающем различные циклические режимы и хаос. На рис. 1 представлена бифуркационная диаграмма модели (3) при $b = 1$ (по оси ординат откладываются координаты устойчивых аттракторов). Как видно из этого рисунка, если $ab \leq 1$, популяция вырождается при любых начальных значениях численности. При $1 < ab \leq 3$ численность популяции стабилизируется на единственном ненулевом уровне. При пересечении уровня 3 сначала появляется цикл длины 2, потом 4 и так далее. При $ab > 4$ начало координат снова становится глобально устойчивым равновесием (популяция вырождается при любых начальных условиях). Все эти критические уровни $ab = 1, ab = 2, ab = 3$ и $ab = 4$ представлены на рисунках.

Заметим, что при $ab > 4$ модель (3) может не иметь прогностической ценности (так как через некоторое количество шагов траектория «заходит» в ноль). Эта ценность сохраняется только при $ab \leq 4$ (рис. 2). Однако и в этой области имеются подмножества (закрашенная область на рис. 2), где для популяции мо-

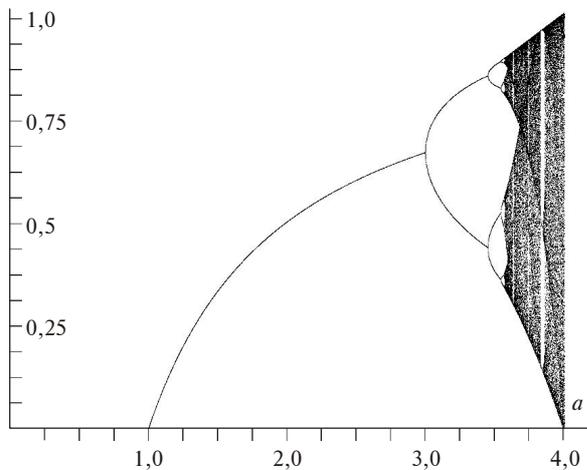


Рис. 1. Бифуркационная диаграмма для обобщённой дискретной логистической модели при параметре $b = 0$.

Fig. 1. Bifurcation diagram for generalized discrete logistic model with parameter $b = 0$.

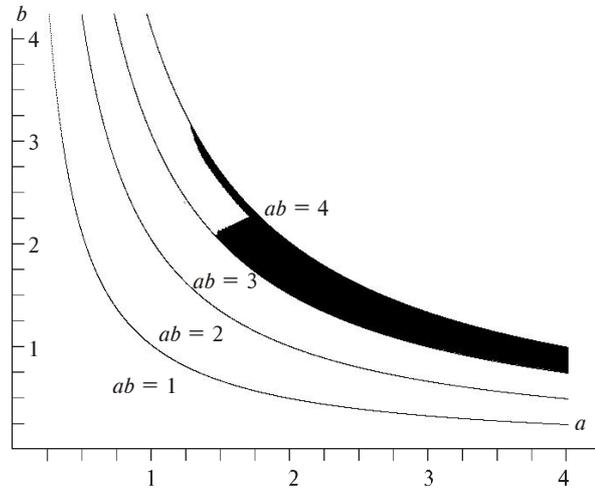


Рис. 2. Основные подмножества пространства параметров модели. В подмножестве $\{(a,b): ab < 1\}$ популяция вырождается при любых начальных значениях численности; в подмножестве $\{(a,b): 1 < ab < 3\}$ численность популяции асимптотически стабилизируется на единственном ненулевом уровне; в подмножестве $\{(a,b): 3 < ab < 4\}$ наблюдаются циклические и хаотические режимы. Закрашенная область включает значения параметров, для которых выполняются условия теоремы Даймонда [Diamond, 1976] (для популяции могут реализоваться хаотические режимы изменения численности).

Fig. 2. Basic subsets of the space of model parameters. In subset $\{(a,b): ab < 1\}$ population eliminates for all initial values of population size; in subset $\{(a,b): 1 < ab < 3\}$ population size stabilizes asymptotically at one non-zero level; in subset $\{(a,b): 3 < ab < 4\}$ cyclic and chaotic dynamic regimes can be realized for population. Painted over domain contains values of model parameters when conditions of Diamond's theorem [Diamond, 1976] are truthful (chaotic regimes can be realized for population).

жет реализоваться хаотический режим: в этом случае прогноз также невозможен (но уже по другой причине — малые отклонения могут привести к большой разности в результатах).

Временные ряды

В настоящей работе рассматриваются временные ряды по динамике численности сосновой пяденицы в Нидерландах [Klomp, 1966]. В числовом формате данные временные ряды представлены в Интернете в свободном доступе (NERC Centre for Population Biology, Imperial College (1999)) The Global Population Dynamics Database, №№ 2727, 2728 and 2729). Данные первого временного ряда (№ 2727) представлены в единицах «среднее число яиц на квадратный метр»; данные второго временного ряда (№ 2728) — в единицах «среднее число гусениц на квадратный метр»; третий ряд (№ 2729) — «среднее число куколок на квадратный метр». Размер первой выборки равен 15 (первое значение было получено в 1950); размер второй выборки — 14 (первое значе-

ние получено в 1950, но информация за 1962 год в выборке отсутствует); размер третьей выборки равен 14 (первое значение было получено в 1951 году).

Данные были собраны в национальном парке De Hoge Veluwe в Нидерландах. Как было показано в наших предыдущих публикациях [Недорезов, 2010, 2011], дискретная логистическая модель была единственной, которая в рамках традиционного подхода к оценке параметров модели позволила получить удовлетворительную аппроксимацию данных. В настоящей работе проводится сравнение результатов, которые были получены в рамках традиционного подхода (МНК), с результатами, которые были получены при нетрадиционном подходе, описанном выше.

Временной ряд № 2727

Для временного ряда № 2727 следующие МНК-оценки параметров модели (3) были получены [Недорезов, 2010, 2011]: $x_0 = 38,23$, $a = 0,0437$, $b = 92,158$, $Q_{\min} = 7290,89$ где Q_{\min} — минимальное значение функционала (2). Для этих значений параметров гипотеза о нормальности отклонений не может быть отклонена даже с 10 % уровнем значимости: вероятность (по критериям Колмогорова–Смирнова и Лиллифорса) и (по критерию Шапиро–Вилка) [Lilliefors, 1967; Shapiro et al., 1968; Hollander, Wolfe, 1973; Большев, Смирнов, 1983].

По критерию Дарбина–Ватсона $d = 1,628$, $d_U < d$, где критический уровень $d_U = 1,36$. Тест серий Свезда–Эйзенхарта показывает, что кумулятивная вероятность $p = 0,063$ [Дрейпер, Смит, 1986, 1987]. Следовательно, оба теста показывают, что с 5 % уровнем значимости сериальной корреляции в последовательности отклонений нет. Таким образом, по МНК-оценкам параметров обобщённую дискретную логистическую модель следует признать пригодной для аппроксимации временного ряда № 2727.

На рис. 3 представлена проекция допустимого множества на плоскость (a, b) : точка отмечена черным цветом только в том случае, если для данных параметров a и b найдется хоть одно значение начальной численности x_0 , при котором все статистические критерии показывают требуемый результат (при 5 % уровне значимости). Точка, соответствующая МНК-оценкам, располагается между двумя «облаками» черных точек в области $\{(a, b): ab > 4\}$.

Относительно небольшие облака выделенных точек располагаются в области $\{(a, b): 1 < ab \leq 3\}$ (рис. 3). Наличие этих «малых» облаков точек не позволяет отклонить гипотезу о том, что наблюдаемый динамический режим — это колебания численности около стационарного уровня. Но поскольку временные ряды связаны и имеют отношение к одной популяции в одном локальном местообитании, то любая гипотеза может быть признана состоятельной толь-

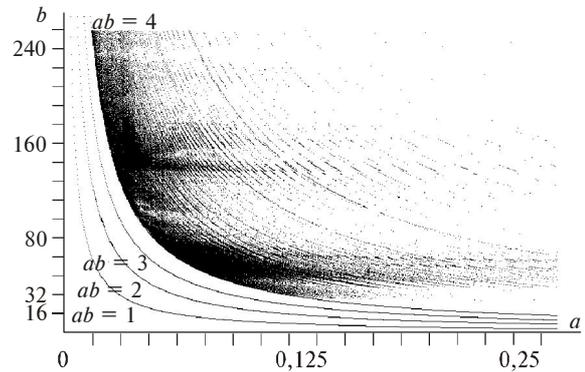


Рис. 3. Проекция допустимого множества (для временного ряда 2727) на плоскость (a, b) . Функции $ab = 1$, $ab = 2$, $ab = 3$ и $ab = 4$ — основные бифуркационные кривые для модели (3).

Fig. 3. Projection of feasible set (for time series 2727) onto the plane (a, b) . Functions $ab = 1$, $ab = 2$, $ab = 3$ and $ab = 4$ — are basic bifurcation curves for model (3).

ко в одном случае — если она имеет определенные основания для всех трёх рядов.

Относительно небольшое число выделенных точек наблюдается в области $\{(a, b): 3 < ab < 4\}$ вблизи бифуркационной кривой $ab = 4$. Поскольку эти точки лежат в области, где параметр a достаточно мал, и, наоборот, параметр b велик, то соответствующая точка лежит вне области, где выполняются условия теоремы Даймонда (рис. 2). Это означает, что для сосновой пяденицы могут реализоваться циклические режимы, но не хаотические.

Еще один важный результат позволяет получить рис. 3: нет ни одной выделенной точки в области $\{(a, b): ab < 1\}$. Таким образом, гипотеза о том, что наблюдаемый динамический режим является режимом асимптотического вырождения популяции, не имеет под собой оснований.

Временной ряд № 2728

Для временного ряда № 2728 следующие МНК-оценки параметров модели (3) были получены [Недорезов, 2010, 2011]: $x_0 = 6,5$, $a = 0,227$, $b = 19,33$, $Q_{\min} = 152,6$. Для этих значений параметров гипотеза о нормальности отклонений не может быть отклонена даже с 10 % уровнем значимости: вероятность $p > 0,2$ (по критериям Колмогорова–Смирнова и Лиллифорса) и $p = 0,13045$ (по критерию Шапиро–Вилка).

Критерий Дарбина–Ватсона $d = 1,319$, $d_U < d$, где критический уровень для 2,5 % уровня значимости $d_U = 1,18319$ (значения критерия определялись для первых 12 значений выборки из-за отсутствия значений за 1962 г.). Кумулятивная вероятность для критерия Свезда–Эйзенхарта $p = 0,762$. Следовательно, оба теста показывают, что с 5 % уровнем значимости сериальной корреляции в последовательности отклонений нет. Таким образом, по МНК-оценкам параметров обобщённую дискретную логистическую модель следует признать пригодной для аппроксими-

мации временного ряда № 2728. Но при этом модель будет непригодна для построения прогнозов, поскольку параметры лежат в области $\{(a,b): ab > 4\}$.

На рис. 4 представлена проекция допустимого множества на плоскость (a,b) . Как видно из этого рисунка, большие группы выделенных точек наблюдаются в зоне $\{(a,b): 1 < ab < 2\}$ (режим монотонной стабилизации численности на ненулевом уровне). Как и в предыдущем случае, выделенных точек в области $\{(a,b): ab < 1\}$ (вырождения популяции) нет.

Большая область выделенных точек располагается в зоне $\{(a,b): 3 < ab < 4\}$, где параметр a достаточно мал. Таким образом, как и в предыдущем случае с временным рядом 2727, можно высказать сходные гипотезы о типе популяционной динамики [Исаев и др., 1984, 2001]. И, кроме этого, имеется возможность указать такие значения параметров модели (рис. 4), при которых соответствующим образом удовлетворены все статистические критерии, и при этом модель обладает прогностическими возможностями. МНК-оценки такой возможности не дают.

Временной ряд № 2729

Для временного ряда № 2729 следующие МНК-оценки параметров модели (3) были получены [Недорезов, 2010, 2011]: $x_0 = 3,33$, $a = 1,64$, $b = 4,05$, $Q_{\min} = 15,4$. Для этих значений параметров гипотеза о нормальности отклонений не может быть отклонена даже с 10 % уровнем значимости: вероятность $p > 0,2$ (по критериям Колмогорова – Смирнова и Лиллифорса) и $p = 0,68609$ (по критерию Шапиро – Вилка).

Критерий Дарбина – Ватсона $d = 0,967$, кумулятивная вероятность для критерия Сведы – Эйзенхарта $p = 0,413$. Таким образом, для всех трёх временных рядов при МНК-оценках нет серийной корреляции, и модель (3) пригодна для аппроксимации вре-

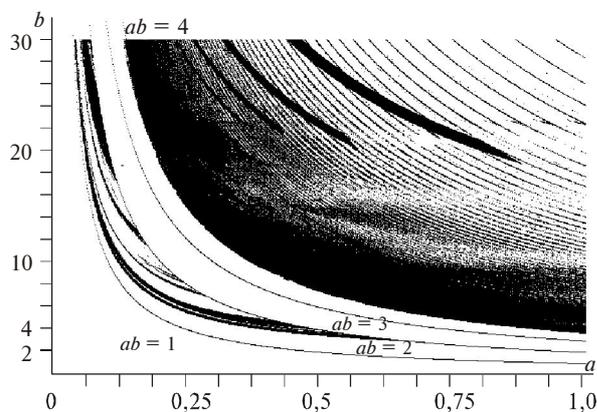


Рис. 4. Проекция допустимого множества (для временного ряда 2728) на плоскость (a,b) . Функции $ab = 1$, $ab = 2$, $ab = 3$ и $ab = 4$ — основные бифуркационные кривые для модели (3).

Fig. 4. Projection of feasible set (for time series 2728) on plane (a,b) . Functions $ab = 1$, $ab = 2$, $ab = 3$, and $ab = 4$ are basic bifurcation curves for model (3).

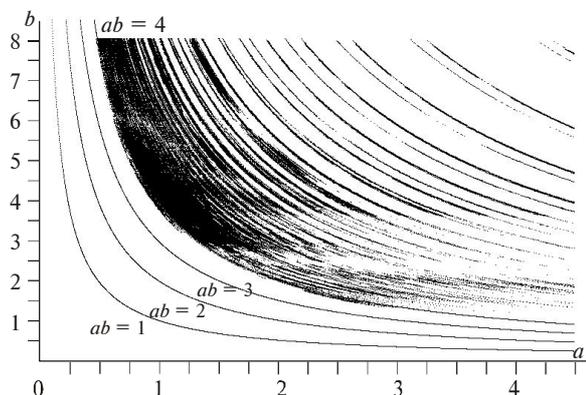


Рис. 5. Проекция допустимого множества (для временного ряда 2729) на плоскость (a,b) . Функции $ab = 1$, $ab = 2$, $ab = 3$ и $ab = 4$ — основные бифуркационные кривые для модели (3).

Fig. 5. Projection of feasible set (for time series 2729) on plane (a,b) . Functions $ab = 1$, $ab = 2$, $ab = 3$, and $ab = 4$ are basic bifurcation curves for model (3).

менных рядов. Как и в двух предыдущих случаях МНК-оценки лежат в области $\{(a,b): ab > 4\}$, но при этом не принадлежат областям с максимальной концентрацией выделенных точек.

На рис. 5 представлена проекция допустимого множества на плоскость (a,b) . Как видно из этого рисунка, выделенных точек в зоне $\{(a,b): 1 < ab < 2\}$ нет: гипотеза о реализации режима монотонной стабилизации численности на ненулевом уровне не имеет оснований. Как и в предыдущих случаях, выделенных точек в области $\{(a,b): ab < 1\}$ (вырождения популяции) нет. Во всех трёх случаях имеется полоса из выделенных точек, располагающаяся вблизи бифуркационной кривой (a,b) в области $\{(a,b): 3 < ab < 4\}$. Это дает основание для заключения о том, что наблюдаемый динамический режим сосновой пяденицы — циклический (заметим, что МНК-оценки не позволяют сделать такой вывод), а во множестве значений параметров можно выбрать такие, которые позволяют прогнозировать изменения численности.

Заключение

В настоящей работе проведено сравнение двух различных методов оценки параметров экологических моделей: классического метода наименьших квадратов (МНК) и нетрадиционного метода, не связанного с минимизацией какого-либо функционала. Сравнение проведено на временных рядах по колебаниям численности сосновой пяденицы в Нидерландах [Клюмп, 1966] для обобщённой дискретной логистической модели. Для всех трёх рассматриваемых рядов было получено, что МНК-оценки лежат в области, когда модель не обладает прогностическими свойствами и, следовательно, установить её реальный динамический режим (или высказать более-

менее правдоподобную гипотезу о нем) не представляется возможным.

В то же время при нетрадиционном подходе можно указать такие значения параметров модели, при которых все основные статистические критерии показывают, что модель может быть использована для аппроксимации данных, и, кроме того, модель позволяет прогнозировать изменения численности и выявлять её асимптотический динамический режим. Столь сильное расхождение результатов связано с тем, что МНК базируется на логической ошибке: вид минимизируемого функционала, который и позволяет получить *наилучшие оценки*, может быть самым различным, и во всех случаях он не имеет никакого отношения к рассматриваемой биологической проблеме. Поэтому применение МНК для оценки параметров нелинейных экологических моделей (и не только) следует проводить с большой осторожностью.

Поскольку исследователи в своем большинстве предпочитают работать не с *множествами хороших значений*, а с конечным множеством таких значений (и, лучше всего, только с одним), то можно использовать следующую процедуру (вместо минимизируемого функционала). Можно использовать статистические критерии (указанные в разделе Основные требования к модели) в качестве своеобразных фильтров, соответствующим образом изменяя уровень значимости. К примеру, если некая нулевая гипотеза не отклоняется с 5 % уровнем значимости и не отклоняется с 10 % уровнем значимости, то во втором случае мы имеем более сильный результат. И во втором случае допустимое множество будет меньше, чем в первом случае.

Литература

- Бард Й. 1979. Нелинейное оценивание параметров. М.: Статистика. 349 с.
- Большев Л.Н., Смирнов Н.В. 1983. Таблицы математической статистики. М.: Наука. 416 с.
- Боровков А.А. 1984. Математическая статистика. М.: Наука. 472 с.
- Гаузе Г.Ф. 1934. Математическая теория борьбы за существование и её применение к популяциям дрожжевых клеток // Бюллетень МОИП. Серия биологическая. Т.43. No.1. С.69–87.
- Гаузе Г.Ф., Витт А.А. 1934. О периодических колебаниях численности популяций: математическая теория релаксационного взаимодействия между хищниками и жертвами и её применение к популяциям двух простейших // Известия СО АН СССР. No.10. С.1551–1559.
- Голубев А.В., Инсаров Г.Э., Страхов В.В. 1980. Математические методы в лесозащите. М.: Лесная промышленность. 104 с.
- Доспехов Б.А. 1985. Методика полевого опыта (с основами статистической обработки результатов исследований). М.: Агропромиздат. 351 с.
- Дрейпер Н., Смит Г. 1986. Прикладной регрессионный анализ. Т.1. М.: Финансы и статистика. 366 с.
- Дрейпер Н., Смит Г. 1987. Прикладной регрессионный анализ. Т.2. М.: Финансы и статистика. 351 с.
- Исаев А.С., Хлебопрос Р.Г., Недорезов Л.В., Киселев В.В., Кондаков Ю.П. 1984. Динамика численности лесных насекомых. Новосибирск: Наука. 224 с.
- Исаев А.С., Хлебопрос Р.Г., Недорезов Л.В., Киселев В.В., Кондаков Ю.П., Суховольский В.Г. 2001. Популяционная динамика лесных насекомых. М.: Наука. 347 с.
- Кондаков Ю.П. 1974. Закономерности массовых размножений сибирского шелкопряда // Петренко Е.С. (ред.): Экология популяций лесных животных Сибири. Новосибирск: Наука. С.206–265.
- Корзухин М.Д., Семевский Ф.Н. 1992. Синэкология леса. СПб.: Гидрометеиздат. 192 с.
- Корн Г.А., Корн Т.М. 1973. Справочник по математике для научных работников и инженеров. М.: Наука. 832 с.
- Лакин Г.Ф. 1990. Биометрия. М.: Высшая школа. 352 с.
- Ликеш И., Ляга Й. 1985. Основные таблицы математической статистики. М.: Финансы и статистика. 356 с.
- Недорезов Л.В. 2010. Анализ динамики численности сосновой пяденицы с помощью дискретных математических моделей // Математическая биология и биоинформатика. Т.5. No.2. С.114–123.
- Недорезов Л.В. 2011. Динамика численности сосновой пяденицы: прогноз с помощью дискретных математических моделей // Евроазиатский энтомологический журнал. Т.10. No.3. С.280–288.
- Недорезов Л.В. 2012а. Хаос и порядок в популяционной динамике: моделирование, анализ, прогноз. Саарбрюкен: LAP Lambert Academic Publishing. 352 с.
- Недорезов Л.В. 2012б. Анализ экспериментальных временных рядов Г.Ф. Гаузе с помощью моделей с непрерывным временем // Журнал общей биологии. Т.73. No.2. С.114–123.
- Рокицкий П.Ф. 1973. Биологическая статистика. Минск: Высшая школа. 320 с.
- Diamond P. 1976. Chaotic behaviour of systems of difference // International Journal of Systems Science. Vol.7. No.8. P.953–956.
- Hollander M., Wolfe D.A. 1973. Nonparametric statistical methods. N.Y.: John Wiley and Sons. 503 p.
- Klomp H. 1966. The dynamics of a field population of the pine looper, *Bupalus piniarius* L. (Lep., Geom.) // Advances in Ecological Research. Vol.3. P.207–305.
- Lilliefors H.W. 1967. On the Kolmogorov–Smirnov test for normality with mean and variance unknown // Journal of the American Statistical Association. Vol.64. P.399–402.
- McCallum H. 2000. Population parameters estimation for ecological models. Brisbane: Blackwell Sciences Ltd. 224 p.
- Nedorezov L.V. 2011. Analysis of some experimental time series by Gause: Application of simple mathematical models // Computational Ecology and Software. Vol.1. No.1. P.25–36.
- Shapiro S.S., Wilk M.B., Chen H.J. 1968. A comparative study of various tests of normality // Journal of the American Statistical Association. Vol.63. P.1343–1372.
- Wood S.N. 2001a. Minimizing model fitting objectives that contain spurious local minima by bootstrap restarting // Biometrics. Vol.57. P.240–244.
- Wood S.N. 2001b. Partially specified ecological models // Ecological Monographs. Vol.71. P.1–25.