

УДК 578.087.1

ИМИТАЦИЯ ЭКОЛОГИЧЕСКИХ ОБЪЕКТОВ  
В СРЕДЕ ПАКЕТА MICROSOFT EXCEL

© 2002 г. А. В. Коросов

*Петрозаводский государственный университет  
185640 Республика Карелия, г. Петрозаводск, просп. Ленина, 33*

Поступила в редакцию 26.05.2000 г.

*Ключевые слова:* динамика биосистем, имитационные модели, табличное программирование, среда пакета Excel.

Количественное описание природных феноменов (биометрическое, статистическое, многомерное) давно стало нормой экологического исследования, чего пока нельзя сказать о количественном описании механизма наблюдаемых явлений, о моделировании. На наш взгляд, этому препятствует широко распространенное мнение о том, что “полноценными” могут быть лишь дающие прогноз аналитические модели; сопряженные с этим сложности построения системы дифференциальных уравнений и их решения оказываются серьезным испытанием для большинства биологов. Однако изучаемые экологические явления сначала нужно понять, дать им объяснение, а уже затем, при необходимости, и прогнозировать. Цель количественного объяснения состоит в вычислении неких параметров, определяющих динамику наблюдаемого процесса; их значения могут быть недоступны для прямого наблюдения или статистической оценки, но информация о них содержится в собранных данных. Решать такую задачу лучше с помощью имитационного моделирования.

Имитационные модели много проще прогностических аналитических моделей, поскольку представляют собой компьютерные программы (“Методы...”, 1980: с. 74), которые могут строиться на базе простейших (линейных) алгебраических уравнений. Переменная “время”, необходимый компонент аналитических моделей, в компьютерной программе выводится за рамки уравнений, т.е. счетчик времени организуется отдельно (Яковлев, 1975), модель становится дискретной. Второе упрощение – это отсутствие процедуры решения дифференциальных уравнений. Значения параметров имитационных моделей подбираются с помощью внешних процедур оптимизации (Пэнтл, 1979). Вместо составления и решения дифференциальных уравнений предлагается составлять программы и настраивать параметры имитационных моделей.

До последнего времени эта замена не сильно упрощала ситуацию. Ведь для создания даже простых

модельных программ требуются глубокие знания какого-либо языка программирования (Pascal, Basic, C++), не говоря уж о программировании процедуры оптимизации. Такие специализированные языки моделирования, как ЭКО-САПФИР (Иванищев и др., 1989) и другие (Страшкраба, Гнаука, 1989), в основном труднодоступны и также требуют особой подготовки. Ситуация в корне изменилась с распространением пакета Microsoft Excel. Обе заявленные проблемы решаются в его среде очень просто (Коросов, 1999).

Рассмотрим этапы моделирования в среде Excel на примере. В течение последних 10 лет на о-ве Кижы (Онежское озеро, Карелия) изучалась популяция обыкновенной гадюки (Коросов и др., 1999). Животных метили, затем определяли встречаемость меченых особей ( $m$ ) в повторных пробах разного объема ( $n$ ). Так, из 158 гадюк, помеченных в 1994 г., в пробе 1995 г. ( $n = 365$  экз.) содержалось  $m = 18$  особей (табл. 1).

Положим целью моделирования определение ежегодной численности ( $N$ ) и смертности ( $Nd$ ) в островной популяции гадюк (при отсутствии массовых миграций). Обычные методы расчетов (Коли, 1979) здесь не работают, так как в данном случае не выполняются важные требования (отсутствие смертности, только троекратный отлов и т.д.). Для иллюстрации работы метода имитации покажем решение упрощенной задачи, приняв ежегодную численность и смертность в островной популяции гадюки неизменной:

$$N = N_i = \text{const} \quad (i = 1994, \dots, 1998),$$

$$Nd = \text{const}.$$

Главный момент имитационного моделирования состоит в том, чтобы выразить известные переменные через неизвестные параметры. Имитационная модель должна вычислять те же величины, что наблюдаются в природе, опыте. Тогда появляется возможность, перебирая возможные значения параметров, найти такие, при которых

**Таблица 1.** Фрагмент листа Excel: имитационная модель снижения числа меченых гадюк ( $M'$ ) в островной популяции гадюки в предположении постоянства численности ( $N$ ), объемов отходов ( $Nd$ ) и пополнения ( $Nb$ ) до настройки параметров

	A	B	C	D	E	F	G	H	I
1	Год	$n$	$m$	$N'$	$d'$	$M'$	$m'$		$\Phi$
2	1994	158		5000		158			
3	1995	365	18	5000	0.1	142	10		58
4	1996	273	10	5000	0.1	128	7		9.1
5	1997	214	10	5000	0.1	115	5		26
6	1998	238	9	5000	0.1	104	5		17
7									
8		$S =$	4.2		$N =$	5000		$S_{\text{Ост}} =$	109
9		$df =$	3		$Nd =$	500		$D_{\text{Ост}} =$	36
10		$S_{\text{общ}} =$	53		$Nb =$	500		$F =$	-1.6
11		$D_{\text{Мод}} =$	-57		$d\% =$	10			

модельные значения переменных совпадут с реальными. В этом случае можно обсуждать найденные значения параметров как характеристику механизма наблюдаемого явления. Для популяции гадюки нам известны следующие переменные: число одноразово меченых животных ( $M$ ), объемы повторных отловов ( $n$ ), число повторно отловленных особей в каждой новой пробе ( $m$ ). Неизвестными остаются общая численность ( $N$ ), число ежегодно гибнущих особей ( $Nd$ ) и объем пополнения ( $Nb$ ) популяции. Три последних значения и требуют оценки, но их необходимо задать сразу же в первом приближении. Разместим их на электронном листе Excel (табл. 1) в отдельном блоке: [F8] = 5000, [F9] = 500, [F10] = F9, [D2] = F8.

В реальной популяции численность ежегодно поддерживается балансом процессов гибели и пополнения:  $N_{i+1} = N_i - Nd + Nb$ . Эта динамика в формате Excel примет вид: [D3] = D2 - \$F\$9 + \$F\$10, [D4] = D3 - \$F\$9 + \$F\$10, ..., [D6] = D5 - \$F\$9 + \$F\$10 (табл. 1, столбец D).

Несмотря на множество формул, их ввод не составляет проблемы: достаточно одну формулу ввести вручную, а остальные – с помощью операции “автозаполнение” (см. инструкцию к Excel). При этом важно следить за тем, чтобы ссылки на общие параметры были абсолютными, т.е. содержали префиксы \$, например \$F\$9. После ввода всех формул в таблица Excel отображаются результаты расчетов; в данном случае численность сохраняется неизменной  $N'_i = 5000$  экз. (табл. 1, графа  $N'$ ). Ежегодная смертность, в том числе среди меченых, составит  $d'_i = Nd/N'_i$ , или в формате Excel: [E3] = \$F\$9/D3... (табл. 1, графа  $d'$ ). Число погибших меченых особей составит  $dM = d'_i M$ , а число вы-

живших меченых будет равно:  $M'_{i+1} = M'_i - d'_i$ , или [F3] = F2 - F2\*E2, ... (табл. 1, графа  $M'$ ).

Как видно из табл. 1, число меченых гадюк со временем сокращается. Сокращаться должно и число повторно отловленных меток ( $m'$ ). Поскольку концентрация меченых особей равна  $pM'_i = M'_i/N'_i$ , то число меченых в пробе объемом  $n$  составит  $m'_i = n_i pM'_i = n_i M'_i/N'_i$ , или [G3] = B3\*F3/D3, ..., [G6] = B6\*F6/D6 (табл. 1, графа  $m'$ ).

Модельные объемы повторно отловленных гадюк ( $m'$ ) уменьшаются, но сильно отличаются от наблюдаемых значений ( $m$ ). Это говорит о том, что произвольно взятые величины  $N$  и  $Nd$  не соответствуют реальности.

Для расчета степени отличия модели от натурных наблюдений используем формулу  $d_i = (m_i - m'_i)^2$  (Пэнтл, 1979), или [I3] = (C3-G3)^2, ... (табл. 1, графа  $\Phi$ ). Общее отличие есть сумма всех частных отличий: [I8] = СУММ(I3 : I6). В нашем случае это обобщенное отличие (функция невязки) равно  $\Phi = 109$ . Понятно, что если бы модель абсолютно точно описывала реальность, то функция невязки была бы равно нулю.

Отсюда вытекает вторая главная задача моделирования – настройка модели, определение таких значений параметров ( $N$  и  $Nd$ ), которые нивелировали бы отличия модели от реальности. Эта очень трудоемкая операция оформлена в среде Excel в виде макроса “Поиск решения” (меню “Сервис”) с очень простым интерфейсом. Процедура настройки модели в среде Excel доступна любому пользователю. (Ответственное отношение к моделированию требует понимания существа процедуры настройки!) После вызова макроса остается заполнить его окно, т.е. указать, что це-

**Таблица 2.** Фрагмент листа Excel: имитационная модель снижения числа меченых гадюк ( $M'$ ) в островной популяции гадюки в предположении постоянства численности ( $N$ ), объемов отхода ( $Nd$ ) и пополнения ( $Nb$ ) после настройки параметров

	A	B	C	D	E	F	G	H	I
1	Год	$n$	$m$	$N'$	$d'$	$M'$	$m'$		$\Phi$
2	1994	158		3086		158			
3	1995	365	18	3086	0.07	142	17		0
4	1996	273	10	3086	0.07	135	12		4
5	1997	214	10	3086	0.07	125	9		2
6	1998	238	9	3086	0.07	116	9		0
7									
8		$S =$	4.2		$N =$	3086		$S_{\text{Oст}} =$	6
9		$df =$	3		$Nd =$	228		$D_{\text{Oст}} =$	2
10		$S_{\text{общ}} =$	53		$Nb =$	228		$F =$	23
11		$D_{\text{Мод}} =$	47		$d\% =$	7.4			

левой ячейкой выступает ячейка I8 (со значением функции невязки), что она должна быть равной значению 0, что для этого можно изменять значения в ячейках F8 : F9. После этого следует нажать кнопку “Выполнить” и в окне “Результаты поиска решения”, появившемся вслед за этим, нужно выбрать “Сохранить результаты”. Как видно из табл. 2, при численности островной популяции обыкновенной гадюки, равной  $N = 3086$  экз., и смертности  $d = 7.4\%$  модельная динамика снижения числа меченных животных оказалась почти такой же, что наблюдалась и в поле. “Почти”, потому что функция невязки так и не обнулилась; после настройки  $\Phi = 6$ .

Для решения вопроса, соответствует ли модель реальности, предлагается три способа: 1) проверка работы модели на независимо полученных данных (Розенберг, 1984); 2) оценка статистической ошибки найденных параметров путем рандомизации (Безель, 1987); 3) оценка адекватности модели – реальности с помощью дисперсионного анализа (Ивантер, Коросов, 1992). Рассмотрим лишь последний метод.

В соответствии со схемой дисперсионного анализа линейной регрессии общая сумма квадратов по всем наблюдениям представлена остаточной и модельной суммами квадратов. Функция невязки есть по существу остаточная сумма квадратов; остаточная дисперсия определяется из отношения  $D_{\text{Oст}} = S_{\text{Oст}}/(n - 1)$ , или [I9] = I8/C9. Общую сумму квадратов просто рассчитать по функции стандартного отклонения ([C8] = СТАНДОТ-КЛОН(C3 : C6)), возведя ее в квадрат и умножив на число степеней свободы: [C9] = СЧЕТ(C3 : C6)–1; [C10] = C9\*C8^2. Модельная сумма квадратов есть разность между общей и остаточной:  $S_{\text{Мод}} = S_{\text{Общ}} - S_{\text{Oст}}$ , или [C11] = C10 – I8. Это же значение равно модельной дисперсии, поскольку число сте-

пеней свободы  $df_{\text{Мод}} = 1$ . Величина критерия Фишера составит  $F = D_{\text{Мод}}/D_{\text{Oст}}$ , или [H10] = C11/I9. В нашем случае критерий превышает табличное значение  $F_{(0.05, 1, 3)} = 6.6$ ; модель в целом адекватна наблюдаемым данным. Видимо, численность наблюдаемой островной популяции гадюки действительно приближается к 3000 экз.

Обобщая рассмотренный пример, важно отметить, что для построения любой имитационной модели требуется конструирование имитационной системы (Моисеев, 1981), содержащей средства ввода данных, программирования, отладки, настройки модели и презентации результатов. Все эти функции оптимально выполняются в среде пакета Excel. В состав имитационной системы входят следующие основные компоненты: блок исходных данных, зачастую состоящий из массива независимых и зависимых переменных; блок расчета модельных данных, собственно имитационная модель, состоящая из уравнений; осуществляет расчет явных переменных (число повторно пойманных особей,  $m'$ ) и скрытых переменных (число меченых особей,  $M'$ ); блок параметров, участвующих в расчете модельных данных и изменяемых в процессе настройки; блок расчета отличий реальных и расчетных значений переменных; значение суммы отличий между моделью и реальностью (значение функции невязки), оно минимизируется в процессе настройки; блок процедуры настройки (окно “Поиск решения”); блок графического представления результатов.

Не менее существенно, что способ моделирования на листе Excel отличается от традиционного способов программирования (алгоритмического, структурного или объектного), это – табличное программирование. Главная особенность его состоит в предельном упрощении переменной “время”, в ликвидации счетчика времени; вместо цик-

лической жизни модели становится пространственной, построчной. Каждая отдельная строка модели представляет собой один временной шаг существования моделируемой системы. Гигантское число строк листа Excel (65 536) достаточно для имитации “жизни” любой модели. Важно лишь помнить, что единицы размерности параметров модели следует явно привязывать к величине выбранного шага. В нашем случае это был один год.

Рассмотренные две особенности среды Excel – простота программирования и встроенный модуль настройки – позволяют надеяться на широкое распространение имитационного моделирования в практике предметных экологических исследований.

#### СПИСОК ЛИТЕРАТУРЫ

- Безель В.С.* Популяционная экотоксикология млекопитающих. М.: Наука, 1987. 130 с.
- Иванищев В.В., Михайлов В.В., Тубольцева В.В.* Инженерная экология. Л.: Наука, 1989. 144 с.
- Ивантер Э.В., Коросов А.В.* Основы биометрии. Петрозаводск: “Карелия”, 1992. 168 с.
- Коли Г.* Анализ популяций животных. М.: Мир, 1979. 3645 с.
- Коросов А.В.* Имитация экологических объектов в среде Excel // Биологические основы изучения, освоения и охраны животного и растительного мира, почвенного покрова Восточной Фенноскандии. Петрозаводск: Изд-во КНЦ РАН, 1999. С. 83–84.
- Коросов А.В., Хилков Т.Н., Фомичев С.Н.* Кижы – “гажья мекка” // Острова Кижского архипелага: Биогеографическая характеристика. Петрозаводск: Изд-во КНЦ РАН, 1999. С. 91–95.
- Методы математической биологии. Математические решения задач биологии и медицины на ЭВМ. Киев: Высшая школа, 1980. Т. 8. 344 с.
- Моисеев Н.Н.* Математические задачи системного анализа. М.: Наука, 1981. 487 с.
- Пэнтл Р.* Методы системного анализа окружающей среды. М.: Мир, 1979. 214 с.
- Розенберг Г.С.* Модели в фитоценологии. М.: Наука, 1984. 265 с.
- Страшкраба М., Гнаук А.* Пресноводные экосистемы. Математическое моделирование. М.: Мир, 1989. 376 с.
- Яковлев Е.И.* Машинная имитация. М., Наука, 1975. 158 с.