

УДК 512.542

ОБ ОДНОЙ СТОХАСТИЧЕСКОЙ МОДЕЛИ ДОЙНОГО СТАДА

А. Жерело¹, Г. Крылов¹, В. Оболонкин², М. Йоргенсен³, Р. Спелман⁴

¹Белорусский государственный университет, Минск, Беларусь

²Отдел общественного здравоохранения, Совет по здравоохранению округа Нортленд,
Новая Зеландия

³Университет Вайкато, Гамильтон, Новая Зеландия

⁴Livestock Improvement Corporation, Гамильтон, Новая Зеландия
e-mail: zherelo@bsu.by, krylov@bsu.by

Поступила: 12.04.2025

Исправлена: 19.05.2025

Принята: 23.05.2025

Ключевые слова: случайный процесс, стохастическое дифференциальное уравнение, интеграл Ито, математическое ожидание, процесс Пуассона, метод Монте-Карло, генная динамика.

Аннотация. В работе предложена стохастическая модель, описывающая динамику биаллельного полиморфизма в дойном стаде. Предложенная модель предполагает, что стадо формируется в условиях контролируемого спаривания. Система уравнений, описывающих модель основана на использовании случайных процессов, содержащих скачки. Построен дискретный аналог стохастической системы, для которого проведена Монте-Карло симуляция.

ABOUT ONE STOCHASTIC MODEL OF A DAIRY HERD

A. Zherelo¹, G. Krylov¹, V. Obolonkin², M. Jorgensen³, R. Spelman⁴

¹Belarusian State University, Minsk, Belarus

²Public Health Unit, Northland District Health Board, New Zealand

³University of Waikato, Hamilton, New Zealand

⁴Livestock Improvement Corporation, Hamilton, New Zealand
e-mail: zherelo@bsu.by, krylov@bsu.by

Received: 12.04.2025

Revised: 19.05.2025

Accepted: 23.05.2025

Keywords: random process, stochastic differential equation, Ito integral, mathematical expectation, approximate calculations, Poisson process, Monte-Carlo method, gene dynamics.

Abstract. The paper proposes a stochastic model, which describes the dynamics of biallelic polymorphism in a dairy herd. The proposed model assumes that the herd is formed under controlled mating conditions. The system of equations describing the model is based on the use of random processes containing jumps. A discrete analog of the stochastic system is constructed, for which a Monte Carlo simulation is performed.

1. Введение

Основополагающие работы генетики Г. Харди [1] и В. Вайнберга [2] по устойчивому распределению аллелей были основаны на законе Менделя.

Важнейшим результатом упомянутых исследований является равновесие Харди–Вайнберга, которое предполагает, что если аллель A встречается в популяции в некотором интересующем локусе с вероятностью p , а аллель B с вероятностью $q = 1 - p$, то наблюдаемые частоты генотипов в этом локусе равны p^2 для AA , $2pq$ для AB и q^2 для BB , причем выполняется данное соотношение в случае выполнения ряда условий (см. таблицу). Однако в популяциях домашних животных, используемых в современном сельском хозяйстве, указанные условия, как правило, не выполняются. В качестве примера можно привести дойное стадо Новой Зеландии, которое состоит из более чем $4 \cdot 10^6$ животных, сгруппированных приблизительно в $1,1 \cdot 10^4$ стад. Очевидно, что спаривание в этом национальном стаде далеко не случайно. В 2013–2014 гг. для осеменения было использовано чуть более 3700 быков, причем в 80 % спариваний всего национального стада было использовано менее 100 лучших быков. За исключением случайных мутаций популяции в целом далеки от пропорций

равновесия Харди–Вайнберга в любом локусе. Частоты генотипов могут быть различными даже для одного локуса в случае различных исследуемых стад или регионов. В таблице приведены условия, которые приводят к равновесию Харди–Вайнберга в некоторой идеальной модельной популяции и условия, существующие в молочной промышленности Новой Зеландии (как и в других странах, где используются промышленные подходы в сельском хозяйстве).

Условия Харди–Вайнберга и дойное стадо Новой Зеландии

Идеальная популяция	Индустриальное производство
Большая однородная популяция	Неоднородные стада и регионы
Случайное спаривание	Контролируемое спаривание
Отсутствие мутаций	Мутации случаются
Отсутствие отбора	Отбор существует
Отсутствие миграции	Экспорт/импорт генов

В действительности такие устойчивые результаты не представляют особого интереса для принятия решений по разведению животных, поскольку владелец стада стремится изменить производительные характеристики животных с целью повышения получаемой прибыли. Для начала упомянем наиболее распространенные модели в генетике, используемые для описания генной динамики в идеальной популяции.

Одной из моделей, оказавших существенное влияние на развитие моделирования в рассматриваемом в данной статье направлении, является генетическая модель эволюции (см. работы С. Райта [3] и Р. Фишера [4]). В данной модели предполагается, что размер исследуемой популяции является фиксированным, рассматривается динамика бинарных аллелей (A , B) и упомянутая модель является дискретной, участвующие особи – равно приспособленные, а моделируемые процессы – марковские. Модель является идеальной, поскольку фактически в ней полагается, что каждое из животных следующего поколения случайным образом выбирает родителей из предыдущего поколения, чего не происходит в реальных условиях. Значимым результатом исследования нейтральной модели Райта–Фишера без мутаций является то, что за конечное время произойдет либо закрепление, либо элиминация некоторого заданного аллеля.

Модель Райта–Фишера является слишком идеализированной и ее дальнейшее развитие осуществлялось за счет изменения ее основных предположений, приближающих модель к реальным условиям, например, за счет добавления большего количества типов аллелей (см. [5]) или использования плавающей общей численности популяции [6; 7]. Более усовершенствованные модели используют более сложные законы распределения (например, распределение Пуассона) для числа потомков особи (см. [5]), пытаются учесть наличие мутаций и/или отбора (см. [8]), добавить в модель процесс смертности (см. [9]) или являются моделями диффузионного типа (см. [10]). Ознакомиться с рядом из обозначенных моделей можно в [5; 11; 12].

Необходимо отметить, что одной из целей приведенных выше исследований было построение таких моделей, для которых возможно было бы получить аналитические представления для интересующих исследователей параметров или построить соответствующие асимптотические разложения для них (например, см. [13]), что неизбежно приводит к упрощению модели, и накладываемых на нее условий, и, как следствие, модель в меньшей степени соответствует процессу, для которого она была построена. В частности, при исследовании генной динамики зачастую предполагается, что популяция может иметь неограниченный размер, что, естественно, не соответствует случаю дойного стада на ферме, которое, как правило, не превышает нескольких сотен голов, и находится под управлением владельца стада как с точки зрения численности, так и с точки зрения показателей продуктивности особей. Далее мы будем рассматривать случай стада ограниченного размера в условиях стохастического управления.

В этой статье мы концентрируем внимание на коровах на конкретной ферме и, более того, в конкретном стаде, но упомянутый выше факт, что несколько десятков лучших быков оплодотворяют миллионы новозеландских коров, говорит о тесной связи всех фермерских стад. Такая ориентация на небольшое количество самцов еще больше контрастирует с ситуацией, которую предполагают классические популяционные модели.

Еще одной сложностью при рассмотрении моделей, ориентированных на получение аналитических представлений, является то, что они ориентированы прежде всего на получение аналитических представлений для описания эволюции средних значений интересующих переменных, в то время как в реальных приложениях существенное значение имеет также информация о, например, дисперсии или форме распределений исследуемых переменных. Это приводит к необходимости создания модели, близкой к реальным условиям, т. е., как следствие, к стохастической модели.

Благодаря успехам в области генетики и селекции разведение животных перестало быть процессом трудно управляемым и мало предсказуемым, а основным инструментом, используемым, в частности, в молочной промышленности для улучшения генетических качеств стада, стало искусственное оплодотворение, что позволяет осознанно выбрать родителей для следующего поколения. Именно с целью понимания последующего влияния вносимых в популяцию контролируемых человеком изменений с помощью искусственного разведения, строятся современные модели.

Для определения управляющего воздействия при искусственном разведении при выборе производителей чрезвычайно важными стали геномные методы [14]. Благодаря технологии статистического геномного прогнозирования, используемой для оценки животных, становятся все более доступными обширные наборы данных генотипирования.

Модель, которую мы разрабатываем в данной статье, является довольно общей в том смысле, что она допускает целый ряд корректировок. Например, можно рассмотреть другие виды, а также более сложные правила суммы генотипов, возможно, учитывающие зависящий от генотипа (геномный) отбор или другие отклонения от предположения о равной приспособленности. Представленная модель используется для описания динамики аллелей в одном локусе у каждого животного стада в условиях контролируемого спаривания.

При разработке моделей, представленных ниже, были приняты в расчет условия, являющиеся специфичными для сезонного управления новозеландским молочным стадом. Такие условия были учтены в модели потому, что одной из целей данного исследования была оценка времени, которое необходимо потратить на внедрение желательного геномного признака в рассматриваемое дойное стадо или на снижение доли нежелательного признака. В качестве примера желаемого признака можно выделить геномный вариант, благоприятно влияющий на состав молока, а в качестве нежелательного – ген «маленького теленка», обнаруженный усилиями Livestock Improvement Corporation (например, см. [15]), снижающий продуктивность, время жизни животного, а также приводящий к проблемам с фертильностью.

На данный момент существует достаточно много исследований в данном направлении (например, см. [16–18]). Однако очевидно, что число обнаруженных и описанных нежелательных генов будет возрастать в процессе новых знаний и понимания причин и механизмов развития распространенных заболеваний. Каждый бык, выбранный для искусственного осеменения, несет в себе некоторые нежелательные гены. Разработанная модель может помочь в анализе рисков путем моделирования различных сценариев для оптимизации стратегии искусственного осеменения в смысле закрепления генов, улучшающих производительность животных, и подавления вредоносных генов в популяции.

2. Описание модели

В данном разделе мы представляем стохастическую модель эволюции генотипа при условии ограничения размера стада некоторым максимальным числом особей и в предположении, что стадо изолировано. Для целей численного моделирования на основе указанной модели будет представлена ее модификация на случай дискретного времени, а также предложен вариант, допускающий приток животных из внешнего источника в предположении неограниченности внешнего источника по количеству доступных особей. В заключение, мы намечаем направления для возможного дальнейшего развития.

Исследуемое стадо фактически состоит из двух подстад: основного (производственного) стада и подменного стада.

Основное стадо состоит из взрослых коров, обеспечивающих производство молока. Подменное стадо состоит из молодых животных до момента их включения в основное стадо или удаления

по причине заболевания, смерти и по прочим причинам, а также включает некоторое количество взрослых животных, временно исключенных из основного стада.

Важно отметить, что использованные в уравнениях (2) и ниже индексы i и j не описывают какое-то конкретное животное, а используются для обозначения занимаемых животным в стадах мест (подобно номерам игроков в команде). Свойства животного в i -й в момент времени t будем считать функцией времени, которую будем обозначать $f_i(t)$. Новое животное добавляется в стадо либо в случае, когда стадо имеет размер меньше установленного ограничения, либо в качестве замены удаленному животному, если текущий размер стада является максимальным. В последнем случае вводимое в стадо животное займет первое незанятое положение. Если i -е животное заменяется в момент времени t_* , это обычно вызывает разрыв или «скачок» в функции $f_i(t)$ в $t = t_*$. Обращаем внимание читателей, что при таких предположениях стохастические процессы в нашей модели могут иметь скачки в случае замены животных.

Здесь мы будем полагать, что стохастические процессы со скачками могут быть представлены в виде интегральной формы стохастического дифференциального уравнения, опирающегося на использование обобщения интеграла Ито на случай стохастически разрывных процессов ([19; 20]).

Отметим, что скачки в случайных процессах часто возникают и в других областях, таких как, финансы, например, при наблюдении и моделировании поведения акций на рынке либо при моделировании поведения портфеля акций в случае принятия решения о замене одного или нескольких активов на другие.

Приведем далее ряд предположений, необходимых для построения модели стада.

2.1. Предположения для модели с непрерывным временем

Будем полагать, что:

1. Количество коров в основном стаде в начальный момент времени $t_0 = 0$ равно N_0 и никогда не превышает этого значения. $N(t)$ обозначает количество животных в основном стаде в момент времени t , $t \in [0, T]$, где $T \in \mathbb{R}$ – некоторое заданное значение.

2. Количество коров в подменном стаде обозначается $M(t)$, в момент времени $t_0 = 0$ равно M_0 и $M(t) \leq M_0$, $t \in [0, T]$.

3. В начальный момент времени $t_0 = 0$ возраст коров в основном стаде генерируется настраиваемым генератором случайных чисел.

4. Генотип потомства следует из генотипа его родителей, определенного через правило суммирования. В этой статье мы рассмотрим только случай суммирования по Менделю, задаваемый правилом (1).

5. С целью упрощения модели с точки зрения моделирования причин удаления животного из основного стада будем считать, что данный процесс подчиняется пуассоновскому процессу интенсивностью λ_D .

6. Аналогично удалению животного из основного стада, для подменного стада также будем полагать, что удаление подчиняется процессу Пуассона интенсивностью λ_d . Отметим, что удаление из подменного стада единообразно описывает и удаление животного по причине смерти либо заболевания, и по причине перевода его в основное стадо.

7. Перемещение животных из подменного стада в основное и удаление из стад происходят один раз в год и моделируются следующим образом: исходя из предположений 1–6, моделируется набор животных текущего года, которые будут удалены из стад; затем вакансии в основном стаде заполняются случайно выбранными животными подменного стада (переменная ξ), достигшими возраста t_{\min} , до заданного максимального размера основного стада; если это окажется невозможным из-за отсутствия животных подходящего возраста в подменном стаде, то заменяется максимально возможное количество животных. Очевидно, что может возникнуть ситуация, когда особей из подменного стада будет недостаточно для заполнения основного стада до максимального размера.

8. Замена животных в подменном стаде также производится один раз в год за счет добавления новорожденных телят, полученных от коров основного стада до определенного ранее размера. Очевидно, что если это окажется невозможным из-за недостатка телят, размер подменного стада уменьшится.

9. Для осеменения всего стада используется детерминированная последовательность известного генотипа, материалом, полученным от одного или нескольких быков-производителей.

10. Приток животных извне отсутствует.

Предположения 5 и 6 введены с целью унификации модели и упрощения процессов численного моделирования. Такой подход позволяет существенно снизить ресурсы, необходимые для проведения симуляции, однако, в случае необходимости, их можно заменить на процессы, которые более точно описывают управление стадом.

Параметры λ_D, λ_d выбираются путем оценки средней продолжительности жизни животных в основном и подменном стадах либо на основе общей практики, либо на основе детального анализа кривой выживаемости. В Новой Зеландии фактическая смерть животного как причина удаления его из любого стада является крайне редкой, но плохое состояние здоровья животного может привести к решению об удалении животного из любого стада.

2.2. Правило суммы одного генотипа

Для построения формальной математической модели представим генотип животного в интересующем локусе как число из $\{-1, 0, 1\}$ и введем термин «индекс генозиготы» (GZI). Здесь -1 и 1 обозначают две гомозиготы (AA и BB) и 0 для гетерозиготных (AB) генотипов соответственно.

Предполагается, что тип наследования в одном локусе менделевский, описываемый следующим правилом суммирования (P задает вероятность каждого результата):

$$\begin{aligned}
 (-1) \dot{+} (-1) &= -1, P = 1; \\
 0 \dot{+} (-1) &= \begin{cases} -1 & P = 0,5 \\ 0 & P = 0,5 \end{cases}; \\
 1 \dot{+} (-1) &= 0, P = 1; \\
 0 \dot{+} 0 &= \begin{cases} -1, & P = 0,25 \\ 0, & P = 0,5 \\ 1, & P = 0,25 \end{cases}; \\
 1 \dot{+} 0 &= \begin{cases} 0, & P = 0,5 \\ 1, & P = 0,5 \end{cases}; \\
 1 \dot{+} 1 &= 1, P = 1.
 \end{aligned} \tag{1}$$

Здесь $\dot{+}$ – коммутативная инфиксная операция, задающая случайное значение генотипа потомка как случайной функции двух переменных соответствующих родительских генотипов.

2.3. Модель с непрерывным временем: интегральная форма

Целью данной статьи было предложить максимально общую модель, которая позволит гибко представлять различные типы управляемых популяций животных. Чтобы выразить такую модель на языке стохастических дифференциальных уравнений, будем полагать, что время непрерывно.

Модели с непрерывным временем могут демонстрировать близость к наблюдаемой динамике стада, но требуют некоторой модификации предположений 6 и 7 в приведенном выше списке предположений относительно случайного времени скачка для процессов удаления из стада.

Будем использовать индекс j для вышеупомянутой формальной позиции в основном стаде и индекс i для позиции в подменном стаде. В нашей стохастической модели взаимодействия двух стад определим по два элементарных эволюционных процесса для каждого из индексов.

В модели процессы $D_j(t)$ или $d_i(t)$ описывают процесс изменения значения генотипа (genezygosis index – GZI) в позициях j или i в основном и подменном стаде. Процессы $A_j(t)$ и $a_i(t)$ отражают процесс изменения возраста животного в позициях. Отметим, что процессы участвуют и в принятии решений относительно характеристик животных в стаде.

Для моделирования процесса удаления животных из стада используются дополнительные независимые процессы Пуассона P_{D_j} и P_{d_i} . Изменения $D_j(t)$, $A_j(t)$ и $d_i(t)$, $a_i(t)$ происходят точно в моменты скачков соответствующего пуассоновского процесса.

Тогда, учитывая предыдущее обсуждение и условия, мы можем записать систему стохастических уравнений, описывающую эволюцию во времени ансамбля коров в основном и подменном стадах в следующем виде:

$$\begin{aligned} D_j(t) &= D_j(0) + \int_0^t (-D_j(s-) + d_{\xi_j}(s-)) dP_{D_j}(s + A_j(0)), \\ d_i(t) &= d_i(0) + \int_0^t (-d_i(s-) + f(D_{\eta_i}(s-), S_{\zeta_i})) dP_{d_i}(s + a_i(0)), \\ A_j(t) &= A_j(0) + t - \int_0^t (A_j(s-) - a_{\xi_j}(s-)) dP_{D_j}(s + A_j(0)), \\ a_i(t) &= a_i(0) + t - \int_0^t a_i(s-) dP_{d_i}(s + a_i(0)), \end{aligned} \quad (2)$$

где $t \in [0, T]$, а другие величины в уравнении (2) определяются следующим образом:

D_j – значение *GZI* (аллеля) для j -й коровы в основном стаде;

P_{D_j} – процесс Пуассона с параметром λ_D , который определяет показатель удаления коровы из основного стада;

A_j – возраст j -й коровы основного стада;

d_i – значение аллеля для i -й коровы подменного стада;

P_{d_i} – процесс Пуассона с параметром λ_d , который определяет показатель удаления коровы из подменного стада;

a_i – возраст i -й коровы подменного стада;

$f(\cdot, \cdot)$ – генотип теленка женского пола как случайная функция родительских генотипов.

Здесь $f(\cdot, \cdot)$ обычно задается правилом суммирования (1), но возможны и другие правила;

η_i – случайная величина, соответствующая случайному выбору коровы из основного стада, которая впоследствии родит самку, которая, в свою очередь, будет помещена в подменное стадо;

ξ – правило для выбора животного из подменного стада, в случае, если животное в основном стаде необходимо заменить;

$\{S_k\}$ – набор значений аллелей для быков-производителей;

величина ζ задает правило для выбора производителя для осеменения из набора $\{S_k\}$.

Опишем более детально систему (2) для j -й позиции основного стада. Пусть в момент времени s произойдет скачок процесса Пуассона для первого уравнения, что означает, что было принято решение об удалении из j -й позиции находящегося на ней животного. Тогда новое значение локуса формируется за счет суммы $-D_j(s-) + d_{\xi}(s-)$, которая позволяет установить новое значение локуса в соответствии с значением правилом ξ . В уравнение для возраста животного в основном стаде в (2) скачок того же процесса Пуассона приводит к изменению в j -й позиции переменной $A_j(s)$, с учетом естественного увеличения возраста (параметр t в этом уравнении). Аналогично в модель введены уравнения для подменного стада. Исключение составляет новое значение генотипа в локусе, которое определяется правилом суммирования (1). Для определения нового значения генотипа животного в подменном стаде используется функция суммирования генотипов f через аргумент $D_{\eta_i}(s-)$.

Для упрощения моделирования здесь и далее мы предполагаем, что все переменные и процессы взаимно независимы. Также полагаем, что интегралы в правой части системы это стохастические интегралы в смысле Ито (см., напр., [21]).

Важно отметить, что стохастический интеграл в смысле Ито по процессу Пуассона, вообще говоря, не определен, однако, принимая во внимание ограниченность выражений под знаком интеграла и тот факт, что $P(t) = \tilde{P}(t) + \lambda t$, где $P(t)$ – процесс Пуассона с параметром λ , а $\tilde{P}(t)$ – это компенсированный процесс Пуассона, соответствующий $P(t)$, являющийся мартингалом, для

которого можно определить интеграл Ито, можно говорить о том, что предложенная модель представляет собой строго определенную систему стохастических дифференциальных уравнений, представленных в соответствующем интегральном виде.

Существуют различные подходы к моделированию, например, аналогичные используемым в эволюционной теории игр [22], которые, как и текущая модель, позволяют использовать при моделировании методы Монте-Карло, однако использованная здесь форма в виде интегральных уравнений позволяет использовать дополнительно приближенные методы слабого типа (например, см. [23; 24]), которые в ряде случаев получают приближенные значения не только для средних, но и для функционалов, заданных на решении представленной системы. Возможность применения таких методов может стать предметом последующих исследований.

2.4. Модель в дискретной форме

При проведении компьютерного моделирования зачастую удобно работать с моделью, использующей дискретное время.

Здесь необходимо обратить внимание на то, что в рассматриваемой задаче обычно отсутствует информация о времени события в стаде, однако доступна информация, предоставляемая ежемесячно или ежегодно. Мы дискретизируем предложенную выше модель по последовательности фиксированных моментов времени $0 = \tau_0 < \tau_1 < \dots < \tau_L = T$.

Будем полагать, что в каждом из L интервалов $(\tau_{l-1}, \tau_l]$ вероятность более чем одного скачка процесса Пуассона (управляющее воздействие) в пределах интервала пренебрежимо мала, или, в качестве альтернативы, несколько скачков можно заменить одним скачком, величина которого равна сумме отдельных скачков. Тогда мы приходим к следующей дискретной модели:

$$\begin{aligned} D_j(t) &= D_j(0) + \sum_{l=1}^L (-D_j(\tau_{l-1}) + d_\xi(\tau_{l-1})) \text{sign}[\mathcal{P}(\lambda_D A_j(\tau_{l-1}))] I_{[0,t)}(\tau_l), \\ d_i(t) &= d_i(0) + \sum_{l=1}^L (-d_i(\tau_{l-1}) + f(D_\eta(\tau_{l-1}), S_\zeta)) \text{sign}[\mathcal{P}(\lambda_d a_i(\tau_{l-1}))] I_{[0,t)}(\tau_l), \\ A_j(t) &= A_j(0) + \sum_{l=1}^L (1 + (a_\xi(\tau_{l-1}) - A_j(\tau_{l-1})) \text{sign}[\mathcal{P}(\lambda_D A_j(\tau_{l-1}))]) I_{[0,t)}(\tau_l), \\ i(t) &= i(0) + \sum_{l=1}^L (1 - i(\tau_{l-1}) \text{sign}[\mathcal{P}(\lambda_d a_i(\tau_{l-1}))]) I_{[0,t)}(\tau_l). \end{aligned} \quad (3)$$

Здесь $I_{[0,t)}(u)$ равна 1, если $u \in [0, t)$ и 0 в противном случае, а $\mathcal{P}(k)$ – независимые случайными величинами Пуассона.

2.5. Модель с притоком животных

При проведении моделирования в описанной в предыдущем разделе модели наблюдается «эффекта вымирания», который упоминался в литературе. На практике такой эффект сокращения численности популяции не наблюдается, в связи с тем, что фермер компенсирует недостаток животных за счет их покупки на рынке.

Ниже мы предлагаем способ учета такого способа управления стадом в предложенной выше модели. Будем полагать, что новое животное всегда помещается в подменное стадо. Такое условие позволяет включить приток животных извне просто за счет изменения значения переменной η в уравнении (3) не внося существенных изменений в используемые уравнения. Для переменной η предположим, что ее нулевое значение будет соответствовать рождению самца животного, а для рожденных особей женского пола значение переменной по-прежнему равно индексу матери в основном стаде. Тогда уравнения для подменного стада примут вид:

$$\begin{aligned}
d_i(t) &= d_i(0) + \sum_{l=1}^L \left(-d_i(\tau_{l-1}) + \text{sign}(\eta) f(D_\eta(\tau_{l-1}), S_\zeta) + \right. \\
&\quad \left. + (1 - \text{sign}(\eta)) D_M \right) \text{sign}[P_d(\lambda_d a_i(\tau_{l-1}))] I_{[0,t)}(\tau_l), \\
a_i(t) &= a_i(0) + \sum_{l=1}^L \left(1 + ((1 - \text{sign}(\eta)) a_M - \right. \\
&\quad \left. - \text{sign}(\eta) a_i(\tau_{l-1})) \text{sign}[P_d(\lambda_d a_i(\tau_{l-1}))] \right) I_{[0,t)}(\tau_l). \tag{4}
\end{aligned}$$

В этих уравнениях мы вводим случайную величину D_M , которая задает распределение смоделированного аллеля у животных с рынка. Распределение для D_M необходимо определить экспериментально или каким-либо еще образом. a_M – это аналогичный предыдущему тип случайной величины, которая задает распределение возраста коров на рынке.

Из уравнений (4) следует, что при $\eta = 0$, т. е. в случае рождения самца, множитель $1 - \text{sign}$ – ненулевой вклад в члены с индексом M , что используется как признак для включения животных извне. В случае же, когда $\eta > 0$, данный множитель равен нулю и, соответственно, исключает влияние внешнего притока животных и приводит систему к исходному виду (3). Еще одним техническим преимуществом последней предложенной модели является то, что теперь размер подменного стада постоянен, что несколько упрощает численное моделирование.

В данной статье мы не намерены приводить результаты моделирования последней предложенной модели и ограничиваемся только ее формулировкой.

3. Результаты моделирования динамики генотипа

Приведем далее результаты моделирования в предположении, что суммирование генотипа осуществляется согласно (1) и при условии контролируемого спаривания. Как и ранее, будем использовать $-1, 0, 1$ при кодировании генотипа. При моделировании для каждого животного мы начинаем с коровы с генотипом -1 (гомозиготный с аллелем, который необходимо устранить). Самка, полученное от нее потомство и так далее осеменяются последовательностью производителей с известным генотипом. Зададим дискретизацию по времени с постоянным интервалом в один год, то есть полагаем $\tau_l - \tau_{l-1} = 1$ (год), $l = 0, 1, \dots, L$. В начальный момент времени τ_0 также предполагается, что распределение аллелей в подменном стаде такое же, как и в основном. Это приводит, как мы увидим позже, к двухлетней задержке при переключении к новому значению.

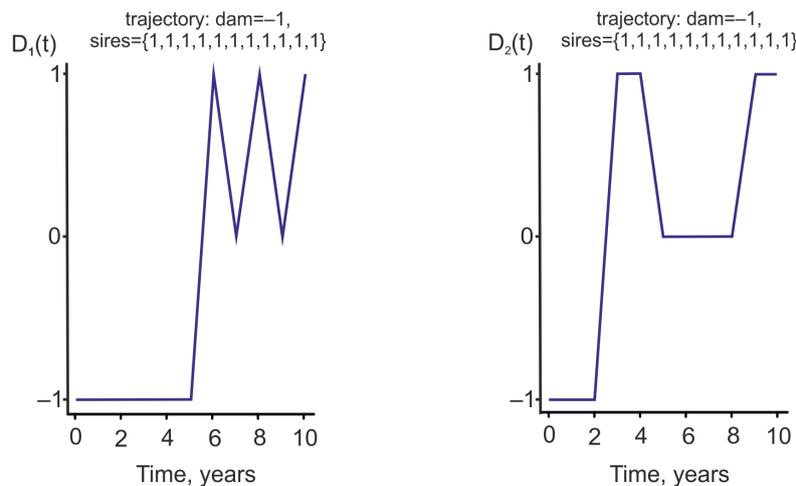


Рис. 1. Две случайно выбранные траектории стохастического процесса D_1, D_2 для заданной позиции в основном стаде

Сначала рассмотрим динамику безусловного переключения генотипа в состояние 1 (гомозиготный с аллелем, который будет введен). Для достижения цели будем использовать последовательность производителей, которые имеют генотип 1. На рис. 1 приведены две реализации единичной траектории случайного процесса для такого случая.

Стоит отметить, что каждый график показывает зависимость GZI в конкретной, скажем в j -й, позиции в основном стаде. На рис. 1 только прыжки можно интерпретировать как замену животного, тогда как горизонтальные линии могут указывать как на замену коровы на животное с тем же генотипом, так и на отсутствие замены.

Подчеркнем, что отдельные траектории (рис. 1) и пучки траекторий (рис. 2, 3) перехода из одного состояния в другое заметно отличаются от среднего (рис. 4).

Далее, на рис. 2 мы строим набор из 100 траекторий с небольшим возмущением, чтобы на графике можно было выделить вероятность каждой отдельной траектории. Плотность линий позволяет визуальнo оценить численность животных в отдельных состояниях $\{-1, 0, 1\}$.

Особое внимание стоит обратить на рис. 3. На нем, как и на рис. 2, рассмотрено 50 траекторий, и всего у одного из производителей значение GZI равно -1 , однако это оказывает существенное влияние на переход к желаемому значению в основном стаде. Напомним также, что почти исчезающее значение GZI , равное -1 , в конце периода моделирования на рис. 2 и 3 не означает, что значение -1 было полностью исключено из стада, т. е. существуют траектории, у которых значение по прошествии 10 лет будет равно -1 .

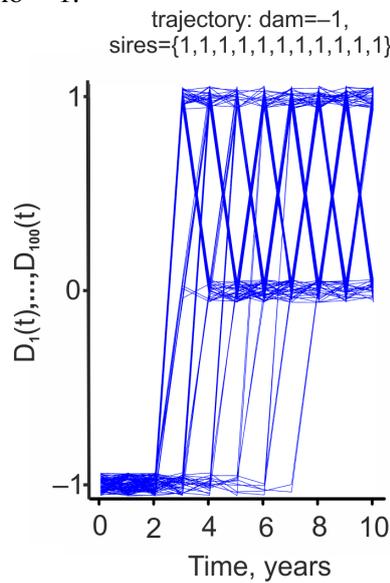


Рис. 2. Набор из 100 траекторий стохастического процесса D_1, \dots, D_{100} (небольшое случайное возмущение было добавлено на каждом временном шаге для лучшей визуализации индивидуальной траектории), все быки в последовательности с $GZI = 1$

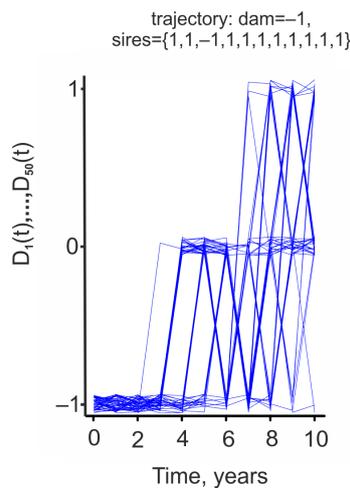


Рис. 3. Набор из 50 траекторий стохастического процесса D_1, \dots, D_{50} , один бык в последовательности (третий) с $GZI = -1$

На рис. 4 мы показываем временную эволюцию среднего значения и дисперсии для случая 1000 стад. Каждая симуляция начинается со стада из 200 гомозиготных самок в основном стаде (все начальные GZI равны -1) и затем стада развиваются независимо. Среднее значение D определяется

стандартным способом как $\text{mean } D(t) \equiv \bar{D}(t) = 1/N \sum_j^N D_j(t)$, где N – общее количество животных во всех стадах. Дисперсия σ^2 рассчитывается как дисперсия среднего для стад.

Чтобы записать формулу для дисперсии, добавим индекс k к $D_j(t)$ для подсчета заданного количества стад, тогда

$$\sigma^2 = \frac{1}{N_{\text{herds}} - 1} \times \sum_{k=1}^{N_{\text{herds}}} \left(\bar{D}(t) - \frac{1}{N_{\text{size}}} \sum_{j=1}^{N_{\text{size}}} D_j^k(t) \right)^2,$$

где N_{herds} – это количество моделируемых стад; N_{size} – это количество коров в стаде (при моделировании использовались стада размером 100, 200 или 400 животных).

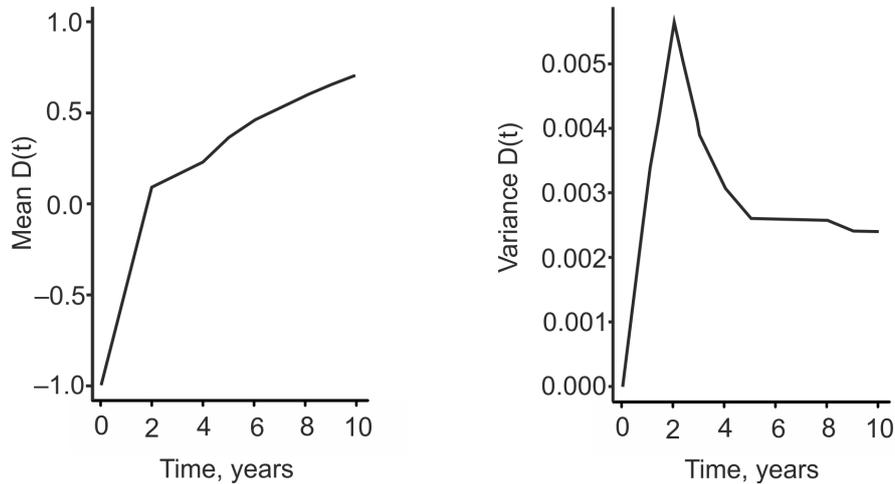


Рис. 4. Среднее (слева) и дисперсия (справа) значения GZI для 1000 стад, размер каждого основного стада составляет 200, размер подменного стада составляет 100 в начальный момент времени, $\lambda_D = 0,114$, $\lambda_d = 0,25$

Важной особенностью, обнаруженной в ходе этого моделирования, является то, что размер стада со временем достаточно быстро падает и стабилизируется на уровне, меньшем заданного в качестве верхней границы. На рис. 5 показан график для среднего значения размера стада, что похоже на «частичное вымирание».

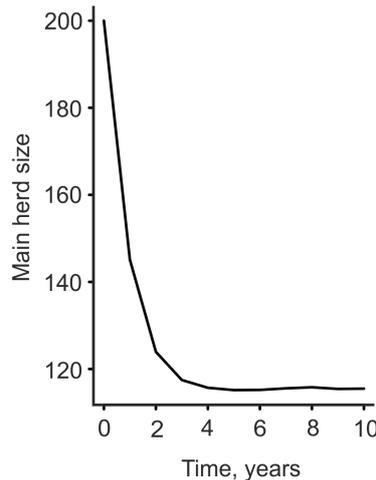


Рис. 5. Зависимость среднего размера основного стада от времени

Это можно объяснить тем фактом, что положительные колебания числа рожденных особей женского пола ограничиваются заданным предельным размером стада, в то время как редкие сильные отрицательные колебания численности рождающихся самок создают дефицит коров в подменном стаде и серьезно влияют на последующую динамику размера стада. Таким образом, использование подменного стада сильно влияет на основное стадо. Представленные результаты по основному стаду следует рассматривать только как демонстрацию тенденций, потому что в данном исследовании мы не стремились изучать влияние управления подменным стадом или пытаться как-то оптимизировать его.

В реальности, на первый взгляд неприятный «эффект вымирания» компенсируется притоком животных с рынка и может быть использован для более быстрого переключения всего стада на заданный аллель за счет введения в стадо животных, у которых интересующее значение уже присутствует, что, в свою очередь, отразится на последующем воспроизводстве потомства.

Интенсивность процесса Пуассона λ_D для основного стада напрямую влияет на уровень «частичного вымирания», поскольку, как это можно увидеть из графиков на рис. 6–10, на которых мы демонстрируем переключение от определенного знака (был выбран равным -1) аллеля в значение 1 , скорость переключения зависела от значений λ_D . На рисунках графики для λ_D , соответствующих вероятности, равной $0,8$, для коровы дожить в основном стаде до $4, 6, 8, 10, 12$ лет, обозначены красным, голубым, зеленым, синим, черным цветами и метками $1, 2, 3, 4, 5$ соответственно. Моделирование было проведено по 10000 стад.

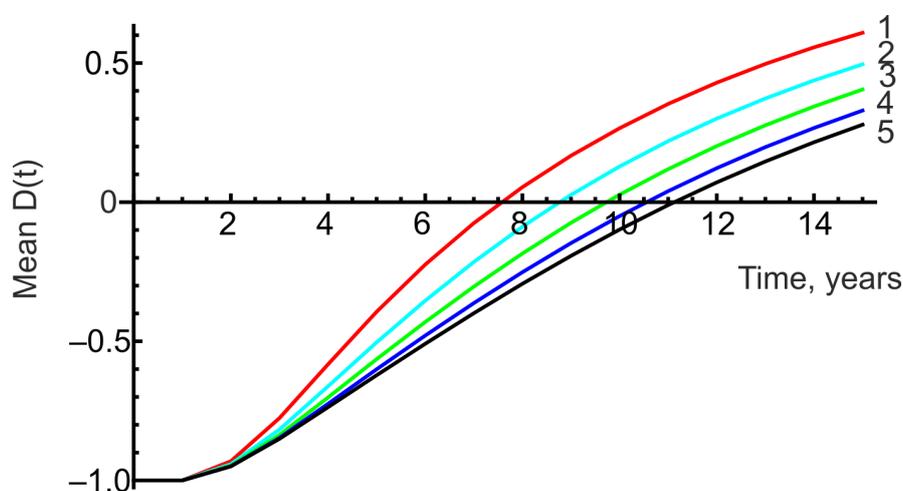


Рис. 6. Среднее значение GZI во времени для 10000 стад, размер каждого основного стада составляет 400 голов, размер подменного стада в начальный момент времени составляет 200 голов

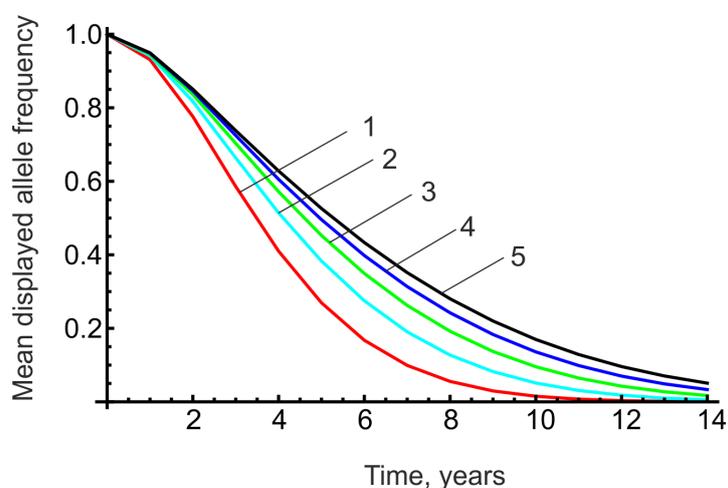


Рис. 7. Средняя частота в популяции замещаемого аллеля по времени для эволюции 10000 стад, размер каждого основного стада составляет 400 , размер подменного стада в начальный момент времени составляет 200 . Последовательность цветов (и числовых меток) соответствует разным значениям средней продолжительности жизни коровы в основном стаде (см. подробности в тексте)

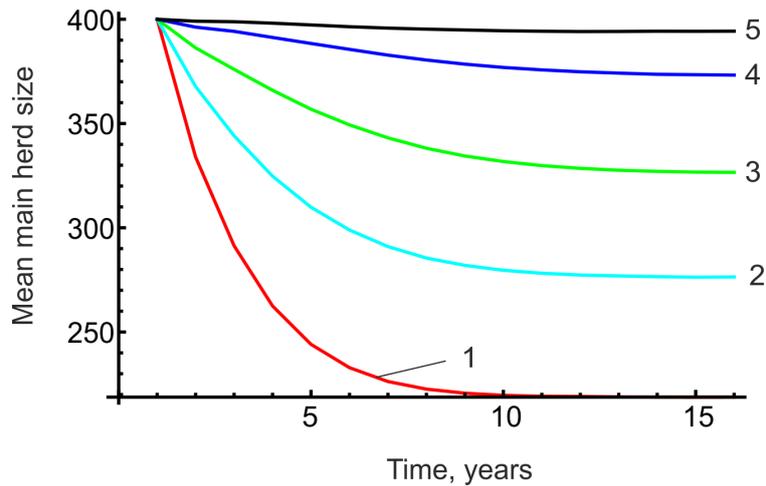


Рис. 8. Изменение среднего размера основного стада во времени для 10000 стад, размер каждого основного стада составляет 400 голов, размер подменного стада в начальный момент времени составляет 200 голов. Последовательность цветов (и числовых меток) соответствует разным значениям средней продолжительности жизни коровы в основном стаде (см. подробности в тексте)

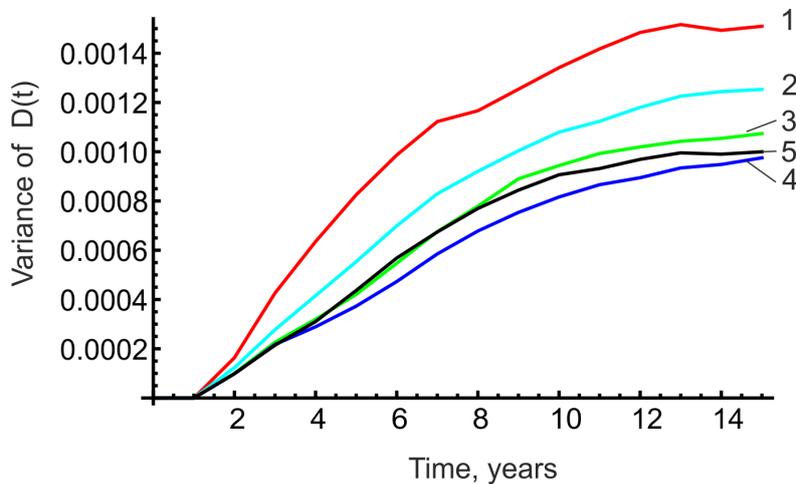


Рис. 9. Дисперсия среднего изменения значения GZI для 10000 стад, размер каждого основного стада составляет 400 голов, размер подменного стада в начальный момент времени составляет 200 голов. Последовательность цветов (и числовых меток) соответствует разным значениям средней продолжительности жизни коровы в основном стаде (см. подробности в тексте)

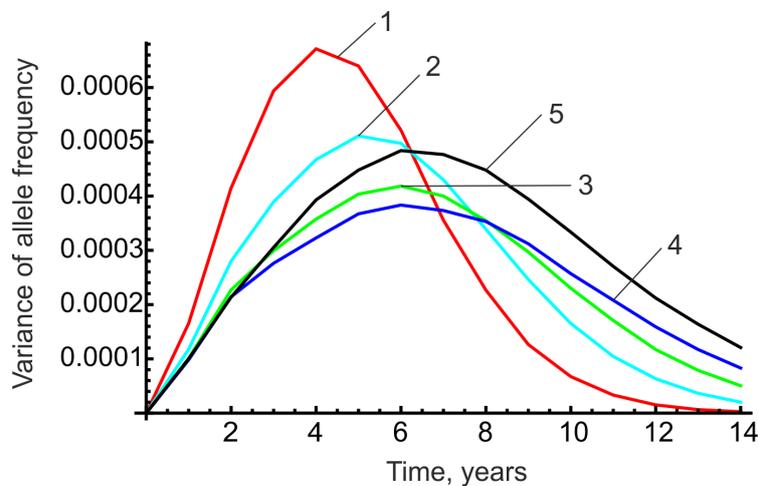


Рис. 10. Дисперсия частоты замещаемого аллеля для 10000 стад, размер каждого основного стада составляет 400 голов, размер подменного стада в начальный момент времени составляет 200 голов. Последовательность цветов (и числовых меток) соответствует разным значениям средней продолжительности жизни коровы в основном стаде (см. подробности в тексте)

Одной из интересных особенностей предложенной модели дойного стада является чувствительность к деталям перехода животного из подменного стада в основное. Рассмотрим две модели управления процессом исключения и замещения животных в стаде.

В модели 1 мы заполняем основное стадо до предельного размера сначала из подменного стада, затем случайным образом осуществляем удаление животных из подменного стада из числа оставшихся.

В модели 2 сначала делаем удаление коров из стада (и основного, и подменного), и лишь потом животных, удаленных из подменного стада, вводим в основное стадо.

Оказывается, что эти две немного отличающиеся процедуры дают заметно различное поведение, что показано на рис. 11 для средних значений, для дисперсии – рис. 12 и для среднего размера основного стада – рис. 13. Динамика среднего значения GZI очень близка, отклонения немного отличаются, но средний размер стада указывает на существенное влияние схемы управления.

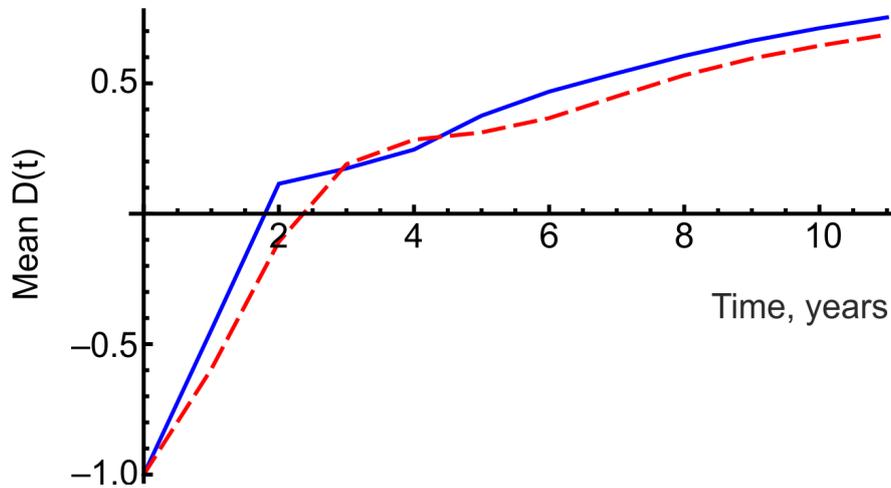


Рис. 11. Сравнение среднего значения аллеля основного стада GZI по времени для двух моделей. Пунктирная красная линия – первая модель, сплошная синяя линия – вторая модель. Все параметры такие же, как на рис. 4

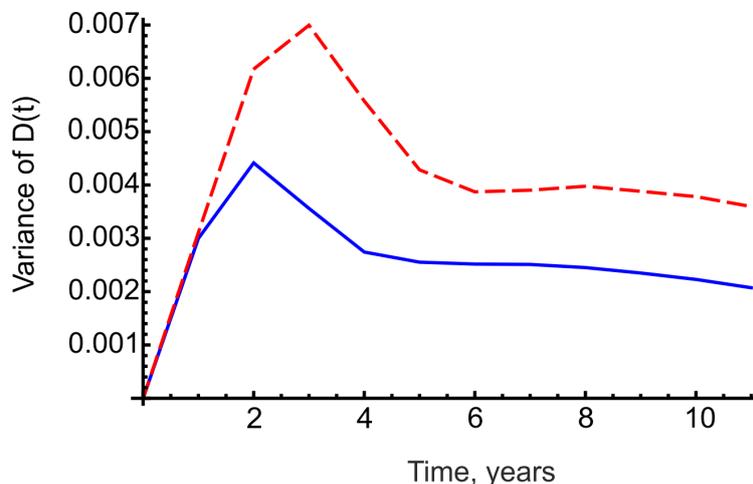


Рис. 12. Сравнение усредненной дисперсии аллеля основного стада GZI по времени для двух моделей. Пунктирная красная линия – первая модель, сплошная синяя линия – вторая модель

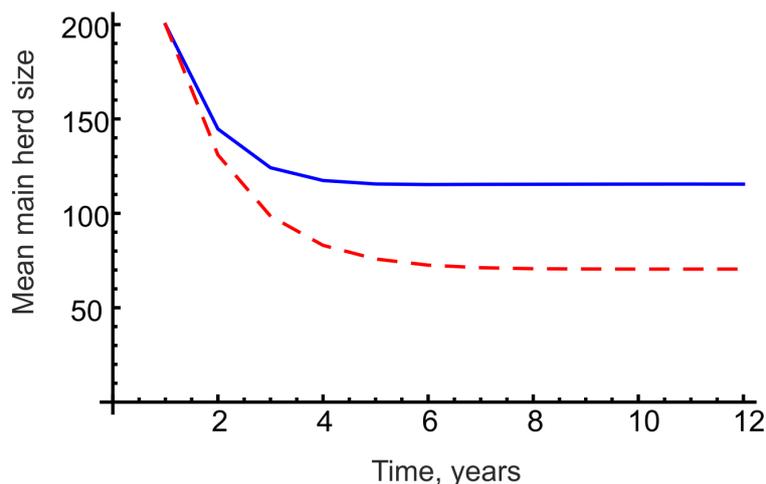


Рис. 13. Сравнение среднего размера основного стада для двух моделей. Пунктирная красная линия – первая модель, сплошная синяя линия – вторая модель

Рис. 12, 13 обращают внимание на то, что необходимо очень внимательно относиться к правилам, используемым в управлении стадом, поскольку на первый взгляд близкие правила приводят к существенным различиям.

4. Заключение

В данной работе для целей моделирования временной эволюции биаллельного полиморфизма дойного стада была предложена модель, основанная на использовании стохастических дифференциальных уравнений. Модель включает ряд дополнительных условий: разделение стада на основное и подменное стадо, типичное время жизни животного в стаде, ограничение размера основного и подменного стада, приток замены животных в основном стаде.

Моделирование показало, что при установлении максимального размера стада локальные колебания новорожденных животных будут сильно влиять на динамику системы и приводить к заметному уменьшению размера стада («частичному вымиранию»). Для устранения этого явления, которое не наблюдается в реальных фермерских ситуациях, была предложена модификация модели с учетом притока животных извне. Моделирование показало, что политика управления животными в стаде может иметь большое значение для достижения целей оптимизации.

Благодарности

Мы выражаем благодарность Джеку Хуперу из Livestock Improvement Corporation за прочтение черновиков этой статьи и за ценные замечания по существующей в Новой Зеландии практике управления стадом.

Литература

1. Hardy G. H. Mendelian proportions in a mixed population // *Science*. 1908. N 28. P. 49–50.
2. Weinberg W. Über den nachweis der vererbung beim menschen // *W. Jahreshefte des Vereins für vaterländische Naturkunde in Württemberg*. 1908. N 64. P. 368–382.
3. Wright S. Evolution in Mendelian populations // *Genetics*. 1931. N 16. P. 97–159.
4. Fisher R. A. *The Genetical Theory of Natural Selection*. Oxford: Clarendon Press, 1931.
5. Feng S. *Poisson-Dirichlet Distribution and Related Topics: Models and asymptotic behaviors*. Berlin, Heidelberg: Springer-Verlag, 2010.
6. Donnelly P., Kurtz T. G. Genealogical processes for Fleming–Viot models with selection and recombination // *Annals of Applied Probability*. 1999. Vol. 9, N 4. P. 1091–1148.
7. Kaj I., Krone S. M. The coalescent process in a population with stochastically varying size // *Journal of Applied Probability*. 2003. Vol. 40, N 1. P. 33–48.

8. Steinsaltz D., Evans S. N., Wachter K. W. A generalized model of mutation-selection balance with applications to aging // *Advances in Applied Mathematics*. 2005. Vol. 35, N 1. P. 16–33.
9. Moran P. A. P. Random processes in genetics // *Mathematical Proceedings of the Cambridge Philosophical Society*. 1958. Vol. 54. P. 60–71.
10. Feller W. Diffusion processes in genetics // *Proc. Second Berkeley Symp. Math. Statist. Prob.*, 1951. P. 227–246.
11. Dawson D. A. "Stochastic population dynamics. Lectures at Summer School in Probability, at PIMS-UBC" [Electronic resource]. – Mode of access: <https://personal.math.ubc.ca/~db5d/SummerSchool09/LectureNotes.html>. – Date of access: 12.04.2025.
12. Pfaffelhuber P., Pennings P., Hermisson J. "Tutorial: Population genetics. Mathematics and BioSciences Group, University of Vienna" [Electronic resource]. – Mode of access: https://www.mabs.at/fileadmin/user_upload/p_mabs/popgen2009.pdf. – Date of access: 12.04.2025.
13. Draghi J. A., Parsons T. L., Plotkin J. B. Epistasis increases the rate of conditionally neutral substitution in an adapting population // *Genetics*. 2011. Vol. 187. P. 1139–1152.
14. Harris B., Johnson D. Genomic predictions for New Zealand dairy bulls and integration with national genetic evaluation // *Journal of Dairy Science*. 2010. Vol. 93. P. 1243–1252.
15. Inbreeding and recessive genes [Electronic resource]. – Mode of access: <https://www.lic.co.nz/products-and-services/artificial-breeding/inbreeding-and-recessive-genes/>. – Date of access: 12.04.2025.
16. Grisart B., Farnir F., Karim L., Cambisano N., Kim J. J. et al. Genetic and functional confirmation of the causality of the DGAT1 K232A quantitative trait nucleotide in affecting milk yield and composition // *Proceedings of the National Academy of Sciences of the United States of America*. 2004. Vol. 101. P. 2398–2403.
17. Allais-Bonnet A., Grohs C., Medugorac I., Krebs S., Djari A. et al. Novel Insights into the Bovine Polled Phenotype and Horn Ontogenesis in Bovidae // *PLoS ONE*. 2013. Vol. 8. Art. e63512. <https://doi.org/10.1371/journal.pone.0063512>
18. Morris A. P., Zeggini E. An evaluation of statistical approaches to rare variant analysis in genetic association studies // *Genetic epidemiology*. 2010. Vol. 34. P. 188–193.
19. Øksendal B. K., Sulem A. Applied stochastic control of jump diffusions. Springer, 2005.
20. Applebaum D. Lévy processes and stochastic calculus. Cambridge University Press, 2009.
21. Gardiner C. W. Handbook of stochastic methods: for Physics, Chemistry and the Natural Sciences. Springer, 2004.
22. Smith M. Evolution and the Theory of Games. Cambridge: Cambridge University Press, 1982.
23. Egorov A. D., Sobolevsky P. I., Yanovich L. F. Functional Integrals: Approximate Evaluation and Applications. Springer, 1993.
24. Zherelo A. Approximate Calculation of Mathematical Expectations on Processes with a Drift // *Nonlinear Phenomena in Complex Systems*. 2015. Vol. 18, N 2. P. 207–214.

References

1. Hardy G. H. Mendelian proportions in a mixed population. *Science*, 1908, no. 28, pp. 49–50.
2. Weinberg W. Über den nachweis der vererbung beim menschen. *W. Jahreshefte des Vereins für vaterländische Naturkunde in Württemberg*, 1908, no. 64, pp. 368–382.
3. Wright S. Evolution in Mendelian populations. *Genetics*, 1931, no. 16, pp. 97–159.
4. Fisher R. A. *The Genetical Theory of Natural Selection*. Oxford, Clarendon Press, 1931.
5. Feng S. *Poisson-Dirichlet Distribution and Related Topics: Models and asymptotic behaviors*. Berlin, Heidelberg, Springer-Verlag, 2010.
6. Donnelly P., Kurtz T. G. Genealogical processes for Fleming–Viot models with selection and recombination. *Annals of Applied Probability*, 1999, vol. 9, no. 4, pp. 1091–1148.
7. Kaj I., Krone S. M. The coalescent process in a population with stochastically varying size. *Journal of Applied Probability*, 2003, vol. 40, no. 1, pp. 33–48.
8. Steinsaltz D., Evans S. N., Wachter K. W. A generalized model of mutation-selection balance with applications to aging. *Advances in Applied Mathematics*, 2005, vol. 35, no. 1, pp. 16–33.

9. Moran P. A. P. Random processes in genetics. *Mathematical Proceedings of the Cambridge Philosophical Society*, 1958, vol. 54, pp. 60–71.
10. Feller W. Diffusion processes in genetics. *Proc. Second Berkeley Symp. Math. Statist. Prob.*, 1951, pp. 227–246.
11. Dawson D. A. "Stochastic population dynamics. Lectures at Summer School in Probability, at PIMS-UBC". – URL: <https://personal.math.ubc.ca/~db5d/SummerSchool09/LectureNotes.html> (accessed 12.04.2025)
12. Pfaffelhuber P., Pennings P., Hermisson J. "Tutorial: Population genetics. Mathematics and BioSciences Group, University of Vienna". – URL: https://www.mabs.at/fileadmin/user_upload/p_mabs/popgen2009.pdf (accessed 12.04.2025).
13. Draghi J. A., Parsons T. L., Plotkin J. B. Epistasis increases the rate of conditionally neutral substitution in an adapting population. *Genetics*, 2011, vol. 187, pp. 1139–1152.
14. Harris B., Johnson D. Genomic predictions for New Zealand dairy bulls and integration with national genetic evaluation. *Journal of Dairy Science*, 2010, vol. 93, pp. 1243–1252.
15. *Inbreeding and recessive genes*. – URL: <https://www.lic.co.nz/products-and-services/artificial-breeding/inbreeding-and-recessive-genes/> (accessed 12.04.2025).
16. Grisart B., Farnir F., Karim L., Cambisano N., Kim J. J. et al. Genetic and functional confirmation of the causality of the DGAT1 K232A quantitative trait nucleotide in affecting milk yield and composition. *Proceedings of the National Academy of Sciences of the United States of America*, 2004, vol. 101, pp. 2398–2403.
17. Allais-Bonnet A., Grohs C., Medugorac I., Krebs S., Djari A. et al. Novel Insights into the Bovine Polled Phenotype and Horn Ontogenesis in Bovidae. *PLoS ONE*, 2013, vol. 8, art. e63512. <https://doi.org/10.1371/journal.pone.0063512>
18. Morris A. P., Zeggini E. An evaluation of statistical approaches to rare variant analysis in genetic association studies. *Genetic epidemiology*, 2010, vol. 34, pp. 188–193.
19. Øksendal B. K., Sulem A. *Applied stochastic control of jump diffusions*. Springer, 2005.
20. Applebaum D. *Lévy processes and stochastic calculus*. Cambridge University Press, 2009.
21. Gardiner C. W. *Handbook of stochastic methods: for Physics, Chemistry and the Natural Sciences*. Springer, 2004.
22. Smith M. *Evolution and the Theory of Games*. Cambridge, Cambridge University Press, 1982.
23. Egorov A. D., Sobolevsky P. I., Yanovich L. F. *Functional Integrals: Approximate Evaluation and Applications*. Springer, 1993.
24. Zherelo A. Approximate Calculation of Mathematical Expectations on Processes with a Drift. *Nonlinear Phenomena in Complex Systems*, 2015, vol. 18, no. 2, pp. 207–214.