

УДК 519.837

ББК 22.18

ДВЕ ЭПИДЕМИОЛОГИЧЕСКИЕ МОДЕЛИ МАЛЯРИИ И ИХ ПРАКТИЧЕСКОЕ ПРИМЕНЕНИЕ*

ВИКТОР В. ЗАХАРОВ

СЕРИНЬ М. НДИАЙЕ

Санкт-Петербургский государственный университет
199034, Санкт-Петербург, Университетская наб., 7-9
e-mail: mcvictor@mail.ru, mamounepourtoi@gmail.com

В статье на основе данных о ежегодной заболеваемости малярией в Сенегале в период с 2000 по 2021 год демонстрируются возможности прогнозирования ежегодной динамики эпидемии с использованием модели SIR и модели CIR . Строится модифицированная модель SIR с постоянными коэффициентами, приводится описание балансовой модели CIR со стохастическими параметрами. Исследуется вопрос о точности прогнозирования ежегодных статистических показателей эпидемии при использовании указанных моделей. Как показывают численные эксперименты, средняя ошибка прогнозирования ежегодного количества болеющих людей по отношению к фактическим статистическим данным при использовании модели SIR является достаточно большой, в то время как модель CIR при сравнительном анализе генерирует более точные прогнозы.

Ключевые слова: эпидемическая модель малярии, модель SIR , прогнозирование активных случаев болезни, балансовая модель эпидемии, принцип динамического баланса, модель CIR .

Поступила в редакцию: 17.05.23 *После доработки:* 01.06.23 *Принята к публикации:* 15.06.23

©2023 В.В. Захаров, С.М. Ндиайе

* Исследование выполнено за счет гранта Российского научного фонда № 23-21-10049 (<https://rscf.ru/project/23-21-10049/>) и гранта Санкт-Петербургского научного фонда.

1. Введение

Прогнозирование динамики заболеваемости малярией является серьезной проблемой из-за отсутствия систематически собираемых достоверных статистических данных в течение продолжительного периода времени, что не позволяет с достаточной степенью точности моделировать и прогнозировать развитие эпидемии в реальном времени при небольшом временном шаге дискретности. Однако данные о ежегодных характеристиках эпидемий можно считать вполне достоверными. Поэтому в рамках этой статьи мы обсудим возможности прогнозирования динамики ежегодных показателей развития малярии в Сенегале в период с 2001 по 2021 год. Исследование, результаты которого представлены в данной статье, опиралось как на модели развития как эпидемии *Covid-19* в России и других странах, так и малярии в странах Африки [1–7, 12, 14, 18–20]. Многие исследовательские группы работают над созданием эффективных моделей и методов прогнозирования распространения вирусов, но малярия остается пока непривлекательной темой для большинства ученых. Модели, позволяющие понять динамику распространения новых вирусов, таких как *Covid-19*, уже были представлены в широко известных периодических научных журналах [7, 14, 19].

Существующие многокамерные детерминированные модели типа восприимчивый — инфицированный — выздоровевший (*SIR*) или восприимчивый — бессимптомный носитель — инфицированный — выздоровевший (*SEIR*) [6, 16] построены на основе механизмов распространения вируса от человека к человеку. Помимо различных эпидемиологических моделей, предлагаются различные модели и методы анализа временных рядов [18]. Кроме того, стали появляться модели, основанные на методах машинного обучения. Например, авторы исследования [21] оценили эффективность динамической байесовской сети в эпиднадзоре за инфекционными заболеваниями. В статье [17] сравнивается статистический подход, часто используемый на практике, и предложенный авторами метод типа *case-based*, в большей степени опирающийся на текущие данные. Они убедительно продемонстрировали преимущество своего метода в прогнозировании динамики вспышек эпидемий, при котором волны характеризуются нерегулярными циклами, которые трудно предсказать с использова-

нием более ранней эпидемической статистики.

Одной из наиболее популярных математических моделей эпидемий является модель SIR . Причем первые методы анализа смертности были проведены еще в 1766 г. Д. Бернулли [15] с целью повлиять на государственную политику в области вакцинации, чтобы остановить распространение оспы. Стоит отметить вклад Р. Росса, У.Х. Хамера, У.О. Кермака и А.Г. Маккендрика, чьи работы в области эпидемиологии внесли значительный вклад (см. [9, 11]). Предложенные вышеупомянутыми авторами модели носят детерминированный характер и для них можно рассчитать базовое репродуктивное число R_0 , которое определяет порог между эпидемическим ростом и спадом эпидемии. Предложенные модели, включая модель SIR доказали свою эффективность при моделировании распространения туберкулеза, СПИДа, гриппа, Эболы и малярии [8]. Но пандемия коронавируса, вызванная вирусом $SARS-CoV-2$, стала толчком к развитию математического моделирования эпидемий и созданию более сложных моделей [10].

В статьях [4, 5] представлены модели развития малярии в популяции при наличии и без вакцинации, в основе которых лежит $SEIR$ модель. В основу данной статьи легла работа [1], в которой описана новая модель развития эпидемии $Covid-19$, а также строятся прогнозы на основе имеющихся статистических данных и сформулированного авторами принципа динамического баланса эпидемиологического процесса. В нашей работе мы используем две модели для построения ежегодных прогнозов развития эпидемии малярии: модифицированная модель SIR и балансовая модель CIR .

Статья имеет следующую структуру. В разделе 2 дается описание выборочных данных по заболеваемости малярией в Сенегале. В разделе 3 приводится описание модели эпидемии малярии на основе модели SIR и ее модификация. Также в этом разделе проводится численное моделирование и построение прогноза динамики популяции на основе предложенной модели, проводится сравнительный анализ полученных прогнозов с имеющимися статистическими данными. Модель эпидемии малярии, основанная на балансовой модели [1] описана в разделе 4, где также строятся прогнозы с использованием данной модели. В конце работы дается заключение (см. раздел 5) и

список литературы.

2. Описание выборочных данных¹

Имеется выборка из 22 наблюдений о ежегодной заболеваемости малярией в Африке (на примере Сенегала) в период с 2020 по 2021 год. Каждое наблюдение представляет собой трехмерный вектор, соответствующий определенному году и включающий в себя количество людей в восприимчивой части населения, количество инфицированных, подтвержденных тестированием, а также количество выздоровевших людей в каждом году. Имеющиеся данные представлены в таблице. На основе ежегодных значений показателей формируется новая таблица, в которой данные приведены нарастающим итогом.

3. Построение модифицированной модели SIR на основе статистических данных

Предполагается, что вся популяция людей делится на несколько субпопуляции. Любой индивид принадлежит единственной субпопуляции. Несмотря на то, что модель SIR является одной из простых эпидемических моделей, она активно применяется на практике и содержит небольшое количество параметров, которые необходимо оценить по имеющимся статистическим данным. Пусть популяция состоит из N человек, структура которой включает три непересекающиеся подгруппы: восприимчивые S , инфицированные I и выздоровевшие R .

Количество человек в каждой из этих подгрупп не является постоянным и меняется в зависимости от года. Будем использовать следующие обозначения:

$S(t)$ — размер субпопуляции восприимчивых в момент времени t ;

$I(t)$ — размер субпопуляции инфицированных в момент времени t ;

$R(t)$ — размер субпопуляции выздоровевших (людей с иммунитетом) в момент времени t ;

α — скорость роста числа инфицированных;

β — скорость роста числа выздоровевших людей.

Запишем систему дифференциальных уравнений, соответствующую

¹ Авторы благодарны Фэлисигэ Дупа за предоставленные данные

щую модели SIR [2]:

$$\begin{cases} \frac{dS(t)}{dt} = -\alpha \frac{S(t)I(t)}{N_0}, \\ \frac{dI(t)}{dt} = \alpha \frac{S(t)I(t)}{N_0} - \beta I(t), \\ \frac{dR(t)}{dt} = \beta I(t). \end{cases} \quad (3.1)$$

Эта система дифференциальных уравнений обычно изучается в области допустимых значений: $\{(S, I, R) \in \mathbb{R}_+^3 : S + I + R = N\}$.

Предложенная система SIR может быть использована для моделирования эпидемии малярии при известных в момент времени $t = 0$ начальных условиях на любом заданном промежутке времени. Траектории системы, построенные с учетом наилучшего подбора коэффициентов α и β , обычно представляют собой детерминированные по времени функции, которые, как показано на многих примерах в научной литературе, часто являются довольно хорошими приближениями к траекториям изменения структуры популяции в реальном времени. Однако известно, что длина промежутка времени, на котором ошибка аппроксимации является небольшой, как правило, мала. При предположении, что число людей, восприимчивых к вирусу в начале вспышки, пропорционально величине популяции N , то есть $S \approx \gamma N$ и $\gamma \in [0, 1]$, получим модифицированную систему дифференциальных уравнений типа SIR :

$$\begin{cases} \frac{dS(t)}{dt} = -\alpha\gamma I(t), \\ \frac{dI(t)}{dt} = \alpha\gamma I(t) - \beta I(t), \\ \frac{dR(t)}{dt} = \beta I(t). \end{cases} \quad (3.2)$$

Решение системы дифференциальных уравнений, описывающей динамику популяции восприимчивых (S), инфицированных (I) и выздоровевших (R), следующее:

$$\begin{aligned} I(t) &= I_0 e^{(\alpha\gamma - \beta)t}, \\ R(t) &= \frac{\beta I_0}{\alpha\gamma - \beta} (e^{(\alpha\gamma - \beta)t} - 1) + R_0, \\ S(t) &= \frac{\alpha\gamma I_0}{(\alpha\gamma - \beta)} (1 - e^{(\alpha\gamma - \beta)t}) + \frac{\gamma}{1 - \gamma} (I_0 + R_0). \end{aligned} \quad (3.3)$$

Используя заданные начальные данные для переменных системы (3.3), построим прогнозные траектории количества инфицированных и количества выздоровевших с использованием модифицированной модели SIR (см. табл. 1 и табл. 2).

Как правило, система (3.2) используется для описания динамики эпидемии на промежутках, длина которых несущественно превышает продолжительность болезни. В рассматриваемом нами случае прогнозирования годовых значений количества инфицированных и количества выздоровевших требуется по-другому интерпретировать полученные траектории системы (3.2). Учитывая возрастающий характер решений этой системы, мы предлагаем сравнивать полученные прогнозные траектории с интегральными фактическими данными, которые суммируют соответствующие значения фактических ежегодных показателей за все предыдущие годы. Таким образом, во втором и третьем столбцах мы внесли именно интегральные фактические данные.

Значения параметров α , β и γ выбирались из условия минимизации ошибок аппроксимации построенных при использовании модифицированной модели траекторий $\hat{I}(t)$ и $\hat{R}(t)$ от фактически реализовавшихся на промежутке от 2000 до 2021 года траекторий $I(t)$ и $R(t)$. Прогноз этих значений представлял наибольший интерес врачей-эпидемиологов по сравнению с прогнозом субпопуляции S . В результате, для проведения сравнения были выбраны следующие значения параметров для (см. табл. 1): $\alpha = 0,4$, $\beta = 0,06$ и $\gamma = 0,9$ и параметров $\alpha = 0,4$, $\beta = 0,1$ и $\gamma = 0,9$ для (см. табл. 2).

Графики 1 и 2 построены по данным, представленным в таблице 1, а графики 3 и 4 построены по данным, представленным в таблице 2, где синий цвет соответствует фактическим данным по малярии, полученным в больницах, а оранжевый цвет — прогнозным значениям, полученным при применении модифицированной модели SIR , представленной выше.

Средняя ошибка аппроксимации вычисляется по формуле:

$$M = \frac{1}{n} \sum_{i=1}^n \left| \frac{A_i - F_i}{A_i} \right|,$$

где A_i — фактическое значение, а F_i — прогнозное значение.

Таблица 1: Интегрированные фактические данные по малярии за 2000-2016 гг. и прогноз по модифицированной модели SIR

Год	$I(t)$	$R(t)$	$\hat{I}(t)$	$\hat{R}(t)$
2000	44959	44083	44959	44083
2001	57879	55861	44959	47229
2002	72304	69320	60688	51475
2003	99169	95138	81921	57207
2004	121403	116132	110581	64945
2005	154563	148030	149269	75390
2006	202633	194948	201492	89488
2007	320965	311845	271986	108520
2008	562891	553845	367143	134209
2009	728824	719853	495591	168887
2010	1059155	1050103	668978	215696
2011	1333274	1324303	903026	278883
2012	1613515	1604469	1218957	364175
2013	1980202	1971082	1645420	479308
2014	2249114	2241429	2221085	634721
2015	2741367	2734834	2998151	844507
2016	3090907	3085636	4047080	1127689

3.1. Прогнозирование эпидемии малярии в Сенегале с 2000 по 2016 год

Результаты расчетов модели SIR представлены на рисунке 2 и в таблице 1. Можно заметить, что с 2000 по 2007 год модель достаточно хорошо отражает динамику распространения эпидемии. Траектории, построенные для наилучшего образного коэффициентов такой системы, обычно представляют собой детерминированные по времени функции, которые, как показано на многих примерах в научной литературе, часто являются довольно хорошими приближениями к траекториям изменения популяции в реальном времени. Однако на точность оценки влияют многие факторы, такие как качество доступных статистических данных и изменчивость течения заболевания.

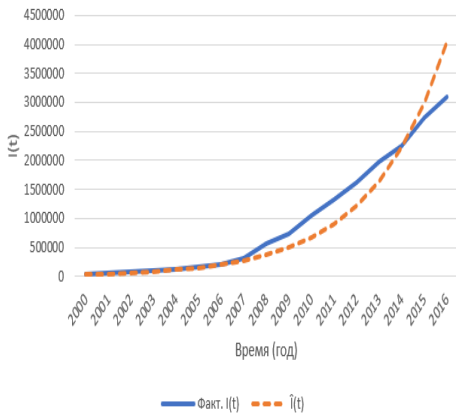


Рис. 1: Интегрированные значения количества инфицированных $I(t)$ (фактические данные и прогнозные данные, полученные с помощью модифицированной модели SIR)

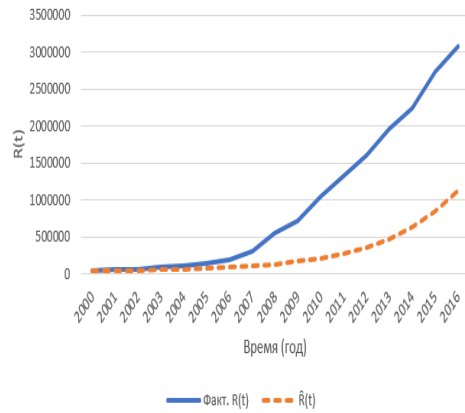


Рис. 2: Интегрированные значения количества выздоровевших $R(t)$ (фактические данные и данные, полученные с помощью модифицированной модели SIR)

Неизвестные параметры модели (α и β) могут существенным образом снизить качество построенных прогнозов динамики эпидемиологической ситуации. Коэффициенты α и β в общем случае являются случайными величинами. Функции распределения этих случайных величин заранее неизвестны. Эффективность модели, основанной на методе обучения, оценивается с помощью ошибки прогнозирования, известной также как средняя ошибка аппроксимации. Измерение этой эффективности очень важно, поскольку, с одной стороны, оно позволяет проводить выбор модели в семействе, связанном с используемым методом обучения, а с другой стороны, оно направляет выбор метода путем сравнения каждой из оптимизированных моделей с предыдущим этапом. Наконец, при любом сделанном выборе она обеспечивает показатель качества или даже достоверности, которые можно придать прогнозу. Выбор зависит от нескольких факторов, включая желаемую цель, размер исходной выборки, сложность рассматриваемой модели, дисперсию ошибок, сложность алгоритмов. В нашей модели с 2007 по 2016 год наблюдается существенное расхождение между расчетными и фактическими данными. Средняя ошиб-

ка аппроксимации расчетной траектории числа активных случаев за период с 2000 по 2016 год составила 17,81%.

На рисунке 2 и в таблице 1, представлены результаты расчетов по модифицированной модели SIR . Можно видеть, что с 2000 по 2006 год модель достаточно хорошо отражает динамику распространения эпидемии. С 2006 по 2016 год наблюдается нарастающее расхождение между расчетными и фактическими данными. Средняя ошибка аппроксимации расчетной траектории интегрированного числа активных случаев (инфицированных) за период с 2000 по 2016 год составила 56,56%.

3.1.1 Прогнозирование эпидемии малярии в Сенегале с 2000 по 2021 год.

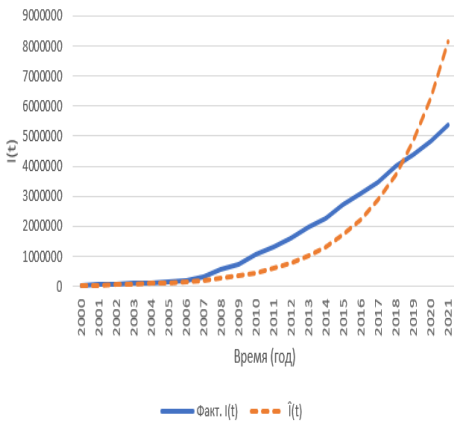


Рис. 3: Интегрированные значения количества инфицированных $I(t)$ (фактические данные и данные, полученные с помощью модифицированной модели SIR)

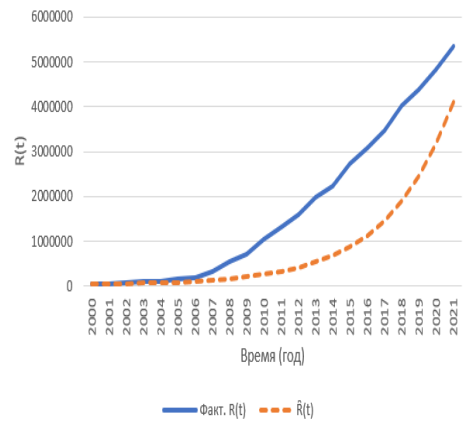


Рис. 4: Интегрированные значения количества выздоровевших $R(t)$ (фактические данные и данные, полученные с помощью модифицированной модели SIR)

Результаты расчетов модели SIR представлены на рисунках 3 и 4 и в таблице 2. Можно видеть, что с 2000 по 2007 год и с 2000 по 2006 год соответственно модель достаточно хорошо отражает динамику

Таблица 2: Фактические данные по малярии за 2000-2021 гг. и прогноз по модифицированной модели SIR

Год	$I(t)$	$R(t)$	$\hat{I}(t)$	$\hat{R}(t)$
2000	44959	44083	44959	44083
2001	57879	55861	44959	47229
2002	72304	69320	60688	51475
2003	99169	95138	81921	57207
2004	121403	116132	110581	64945
2005	154563	148030	149269	75390
2006	202633	194948	201492	89488
2007	320965	311845	271986	108520
2008	562891	553845	367143	134209
2009	728824	719853	495591	168887
2010	1059155	1050103	668978	215696
2011	1333274	1324303	903026	278883
2012	1613515	1604469	1218957	364175
2013	1980202	1971082	1645420	479308
2014	2249114	2241429	2221085	634721
2015	2741367	2734834	2998151	844507
2016	3090907	3085636	4047080	1127689
2017	3486613	3482582	5462987	1509943
2018	4017557	4014573	7374261	2025933
2019	4372265	4370247	9954211	2722447
2020	4817578	4816702	13436779	3662642
2021	5354428	5354428	18137755	4931773

распространения эпидемии. С 2007 по 2021 год и с 2006 по 2021 год соответственно наблюдается значительное расхождение между расчетными и фактическими данными. Средняя ошибка аппроксимации расчетной траектории числа активных случаев за период с 2000 по 2021 год составила 31,40% на рисунке 3 и 49,58% на рисунке 4.

4. Балансовая модель малярии

4.1. Описание балансовой модели эпидемии малярии на основе процентного прироста

Рассмотрим балансовую модель малярии, построенную на основе CIR модели, представленной в [1]. Обозначим через $C(t)$ общее количество зарегистрированных случаев заболевания малярией для $t = 0, 1, 2, \dots$. Число новых случаев заболевания $\Delta(t)$ в момент t можно вычислить по формуле

$$\Delta(t) = C(t) - C(t - 1).$$

Процентный прирост $r(t)$ величины $C(t)$ в момент t обозначим через $r(t)$ и запишем в виде

$$r(t) = 100 \frac{\Delta(t)}{C(t - 1)}.$$

Дискретное уравнение для переменной $C(t)$ тогда можно записать в виде

$$C(t) = \left(1 + \frac{r(t)}{100}\right) C(t - 1).$$

Обозначим через $R(t)$ общее количество выздоровевших или умерших пациентов к моменту времени t . Полагая, что продолжительность болезни в общем случае является конечной величиной, для любого заданного момента времени $T > 0$, рассмотрим неравенства

$$C(t) \geq R(T), \quad 0 < t \leq T.$$

С учетом неубывающего характера функции $C(t)$ относительно переменной t и ограниченности интервала $0 \leq t \leq T$ задача целочисленного программирования

$$\min_{0 < t \leq T} t,$$

$$C(t) \geq R(T)$$

имеет решение, которое обозначим $\tau(T)$.

Тогда в соответствии с принципом динамического баланса эпидемиологического процесса [21] для любого времени t , $0 < t \leq T$ справедливо следующее неравенство:

$$C(\tau(t)) \geq R(t) \geq C(\tau(t) - 1).$$

В этом случае значение $R(t)$ будет принадлежать промежутку $[C(\tau(t) - 1), C(\tau(t))]$ и может быть представлено в виде

$$R(t) = \lambda_t C(\tau(t) - 1) + (1 - \lambda_t) C(\tau(t)), \quad 0 \leq \lambda_t \leq 1.$$

Динамика количества болеющих людей (активных случаев болезни) $I(t)$ с учетом балансового соотношения $C(t) = I(t) + R(t)$ может быть задана следующим образом:

$$I(t) = \left(1 + \frac{r(t)}{100}\right) C(t - 1) - R(t).$$

Таким образом, систему дискретных уравнений, описывающих распространение эпидемии малярии в популяции хозяина, будет иметь вид

$$\begin{aligned} C(t) &= \left(1 + \frac{r(t)}{100}\right) C(t - 1), \\ I(t) &= \left(1 + \frac{r(t)}{100}\right) C(t - 1) - R(t), \\ R(t) &= \lambda_t C(\tau(t) - 1) + (1 - \lambda_t) C(\tau(t)). \end{aligned}$$

Величина $\theta(t) = t - \tau(t)$ называется характеристикой динамического баланса эпидемиологического процесса [1]. Процентный прирост и характеристика динамического баланса, как правило, являются нестационарными случайными величинами. Для прогнозирования значений процентного прироста $r(t)$ можно использовать сценарный или использовать "case-based rate reasoning" (CBRR) подход [1]. Прогнозирование значений характеристики динамического баланса на горизонте прогнозирования $[t, t + \theta(t)]$ можно проводить при предположении стационарности или не стационарности характеристики динамического баланса на этом промежутке.

4.2. Практическое применение стохастической балансовой модели

В этой части статьи изучим стохастические прогнозы малярии. Исследование основано на статье [20], в которой исследуется балансовая модель эпидемии новых вирусов на основе процентного прироста и характеристики динамического баланса, имеющими стохастическую природу. Высокая точность прогнозирования будущей динамики основных статистических показателей пандемии *COVID-19* в России подтверждена многочисленными численными экспериментами. Используемая база данных представляет ежегодную эволюцию малярии в Сенегале (Африка) с 2000 по 2021 год.

Данные в приведенных ниже таблицах структурированы следующим образом: столбец 2 содержит фактические данные общее количество зарегистрированных случаев заболевания $C(t)$, начиная с 2020 года, столбец 3 содержит интегрированные фактические данные о инфицированной популяции $I(t)$, столбец 4 содержит интегрированные фактические данные о выздоровевшей популяции $R(t)$, столбец 5 содержит значения характеристики динамического баланса $\theta(t)$, столбец 6 содержит расчетные данные по общему количеству зарегистрированных случаев заболевания $\hat{C}(t)$, столбец 7 содержит расчетные данные по инфицированной популяции $\hat{I}(t)$, столбец 8 содержит расчетные данные о выздоровевшей популяции $\hat{R}(t)$.

4.2.1 Модельный прогноз на 2011-2017 г.

Основываясь на данных, доступных с 2020 по 2011 годы, мы построим прогноз течения заболевания на период с 2011 по 2017 год. Прогнозирование проводилось на основе применения интеллектуальных алгоритмов построения кусочно-линейных трендов стохастических значений процентного прироста и характеристики динамического баланса (при предположении о ее стационарности или нестационарности). Графики фактических и прогнозных значений характеристики динамического баланса в стационарном и случае представлены на Рис. 5. Графики фактических и прогнозируемых траекторий переменных модели изображены на рисунках 6-8.

Таблица 3: Фактические данные по малярии за 2011-2017 гг. и прогноз по модифицированной модели CIR

Год	$C(t)$	$I(t)$	$R(t)$	$\theta(t)$	$\hat{C}(t)$	$\hat{I}(t)$	$\hat{R}(t)$
2000	89042	44959	44083				
2001	113740	57879	55861	1			
2002	141624	72304	69320	2			
2003	194307	99169	95138	2			
2004	237535	121403	116132	2			
2005	302593	154563	148030	2			
2006	397581	202633	194948	2			
2007	632810	320965	311845	1			
2008	1116736	562891	553845	1			
2009	1448677	728824	719853	1			
2010	2109258	1059155	1050103	2			
2011	2657577	1333274	1324303	2	2636484	1353778	1282707
2012	3217984	1613515	1604469	2	3191408	1908702	1282707
2013	3951284	1980202	1971082	3	3816514	2037546	1778968
2014	4490543	2249114	2241429	3	4544977	2161560	2383418
2015	5476201	2741367	2734834	3	5389759	2451978	2937781
2016	6176543	3090907	3085636	4	6364612	2779978	3584634
2017	6969195	3486613	3482582	4	7483965	3263051	4220914

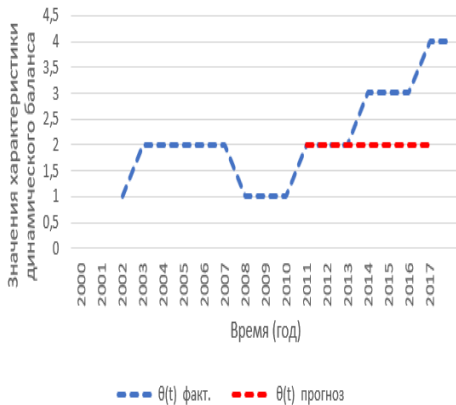


Рис. 5: Фактические значения характеристики динамического баланса на промежутке 2000 – 2017 годы и прогнозные (стационарные значения) на промежутке 2011 – 2017 годы.

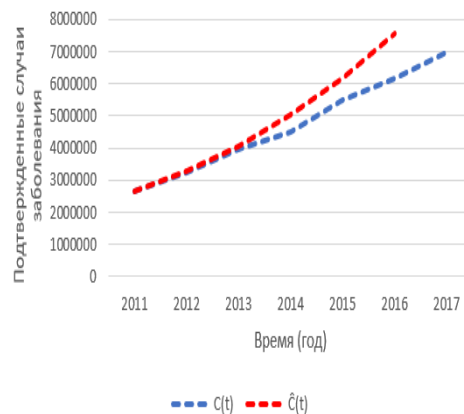


Рис. 6: Фактическая и прогнозируемая динамика общего количества заболевших на промежутке 2011 – 2017 годы

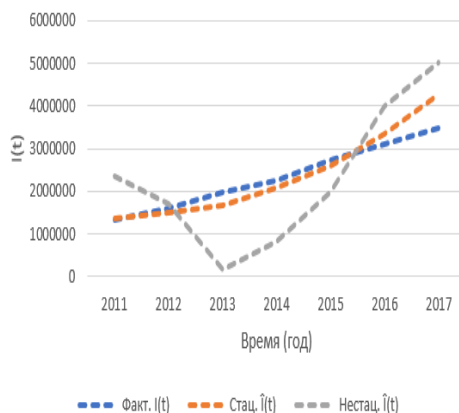


Рис. 7: Динамика $I(t)$, фактической и смоделированной траекторий в интервале с 2011 по 2017 год, в стационарном и нестационарном случаях характеристики динамического баланса

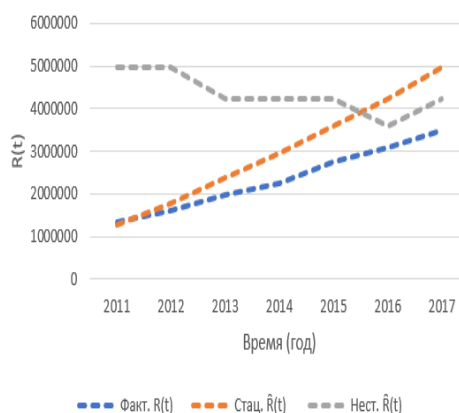


Рис. 8: Динамика $R(t)$, фактической и смоделированной траекторий в интервале с 2011 по 2017 год, в стационарном и нестационарном случаях характеристики динамического баланса

С 2011 по 2017 год в прогнозах использовалось значение $\theta = 2$. График расчетных данных по инфицированной популяции почти идеально отображает фактические данные по инфицированной популяции. Средняя ошибка аппроксимации составляет 9,44%.

4.2.2 Модельный прогноз на 2018-2021 г.

С 2018 по 2021 год в прогнозах использовалось значение $\theta = 4$ и предположение о незначительное убывание $r(t)$. Динамика характеристики динамического баланса и фактические и прогнозные траектории основных переменных графически представлены на рисунках 9-12. График расчетных данных по инфицированной популяции почти идеально отображает фактические данные по инфицированной популяции. Средняя ошибка аппроксимации составляет 18,12%.

Таблица 4: Фактические данные по малярии за 2018-2021 гг. и прогноз с использованием модели CIR и условия стационарности

Год	$C(t)$	$I(t)$	$R(t)$	$\theta(t)$	$\hat{C}(t)$	$\hat{I}(t)$	$\hat{R}(t)$
2000	89042	44959	44083				
2001	113740	57879	55861	1			
2002	141624	72304	69320	2			
2003	194307	99169	95138	2			
2004	237535	121403	116132	2			
2005	302593	154563	148030	2			
2006	397581	202633	194948	2			
2007	632810	320965	311845	1			
2008	1116736	562891	553845	1			
2009	1448677	728824	719853	1			
2010	2109258	1059155	1050103	2			
2011	2657577	1333274	1324303	2			
2012	3217984	1613515	1604469	2			
2013	3951284	1980202	1971082	3			
2014	4490543	2249114	2241429	3			
2015	5476201	2741367	2734834	3			
2016	6176543	3090907	3085636	4			
2017	6969195	3486613	3482582	4			
2018	8032130	4017557	4014573	4	7793878	3572965	4220914
2019	8742512	4372265	4370247	5	8677179	3693807	4983372
2020	9634280	4817578	4816702	5	9617200	3790828	5826372
2021	10708856	5354428	5354428	6	10610971	4038102	6572869

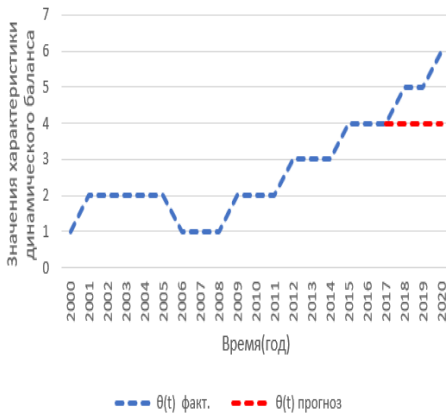


Рис. 9: Фактические и смоделированные значения характеристики динамического баланса $\theta(t)$ за период с 2000 по 2020 год.

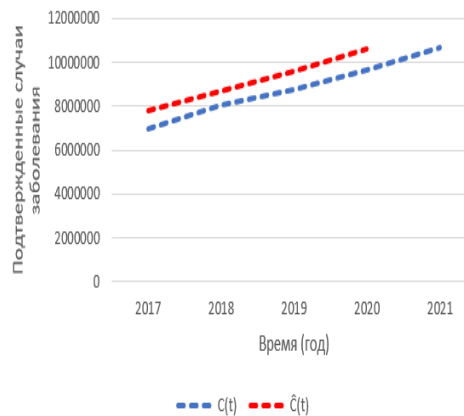


Рис. 10: Динамика $C(t)$, фактической и смоделированной траекторий в интервале с 2017 по 2021 год.

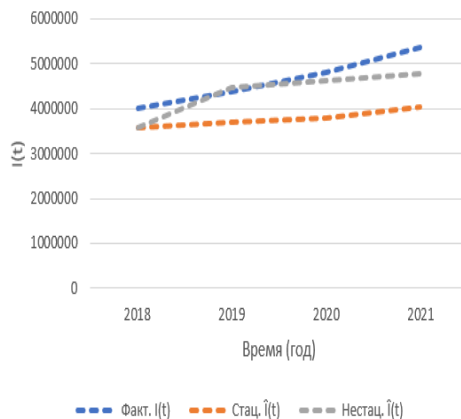


Рис. 11: Динамика $I(t)$, фактической и смоделированной траекторий в интервале с 2018 по 2021 год, в стационарном и нестационарном случаях характеристики динамического баланса

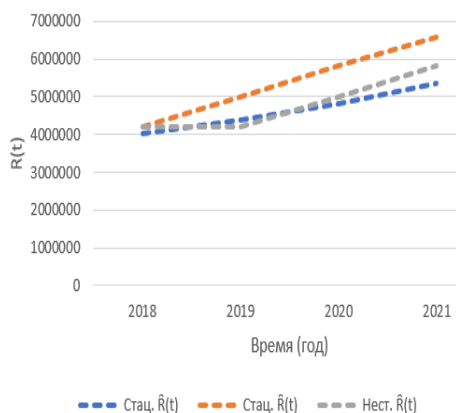


Рис. 12: Динамика $R(t)$, фактической и смоделированной траекторий в интервале с 2018 по 2021 год, в стационарном и нестационарном случаях характеристики динамического баланса

5. Заключение

В работе представлено две модели эпидемического процесса малярии. Первая модель основана на классической модели SIR и модифицирована для описания процесса развития малярии в Сенегале в формате ежегодных данных. Главной особенностью модифицированной модели SIR является то, что она подходит для моделирования на интервалах, продолжительность которых значительно превышает длительность болезни. Прогнозные значения сравниваются с фактическими данными, представленными в интегрированной форме, для различных временных интервалов получены средние ошибки аппроксимации, на основании которых делается вывод о возможности применения модели для прогнозирования числа активных случаев и общего количества выздоровевших. Вторая модель основана на принципе динамического баланса эпидемиологического процесса и учитывает при построении прогнозов сгенерированные динамические тренды стохастических значений процентного прироста общего количества заболевших и предположения о стационарном или нестационарном

ционарном характере изменения характеристики динамического баланса. Полученные для выбранных промежутков прогнозирования оценки демонстрируют более высокую точность прогнозов балансовой модели для оценки будущих значений активных случаев и размера выздоровевшей части популяции.

СПИСОК ЛИТЕРАТУРЫ

1. Захаров В.В., Балыкина Ю.Е. *Балансовая модель эпидемии COVID-19 на основе процентного прироста* // Информатика и автоматизация. 2021. Том 20, № 5. С. 1034–1064.
2. Захаров В. В., Балыкина Ю. Е. *Прогнозирование динамики эпидемии коронавируса (COVID-19) на основе метода прецедентов* // Вестн. С.-Петербург. ун-та. Сер. 10. Прикл. матем. Информ. Проц. упр. 2020. Т. 16, вып. 3. С. 249–259.
3. Кондратьев М. А. *Методы прогнозирования и модели распространения заболеваний* // Компьютерные исследования и моделирование. 2013 Т. 5. № 5 С. 863–882.
4. Ндияйе С. М., Парилина Е. М. *Эпидемическая модель малярии без вакцинации и при ее наличии. Ч. 1. Модель малярии без вакцинации* // Вестн. С.-Петербург. ун-та. Сер. 10. Прикл. матем. Информ. Проц. упр. 2022. Т. 18, вып. 2. С. 263–277.
5. Ндияйе С. М., Парилина Е. М. *Эпидемическая модель малярии без вакцинации и при ее наличии. Ч. 2. Модель малярии с вакцинацией* // Вестн. С.-Петербург. ун-та. Сер. 10. Прикл. матем. Информ. Проц. упр. 2022. Т. 18, вып. 4. С. 554–566.
6. Cooper I., Mondal A. Antonopoulos C.G. *A SIR model assumption for the spread of COVID-19 in different communities* // Chaos Solitons Fractals. 2020. Vol. 139. № 110057.
7. Fanelli D., Piazza F. *Analysis and forecast of COVID-19 spreading in China, Italy and France* // Chaos Solitons Fractals. 2020. Vol. 134. № 109761.

8. Fred B., Carlos C., Zhilan F. *Mathematical Models in Epidemiology*. Springer, 2019.
9. Fred B., Pauline V. D. D., Jianhong W. *Mathematical Epidemiology*. Springer-Verlag Berlin Heidelberg. 2008.
10. <https://statsandr.com/blog/covid-19-in-belgium/#more-sophisticated-models>.
11. Kermack W. O., McKendrick A. G. *Contribution to the Mathematical Theory of Epidemics* // The Royal Society. Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character. 1927. Vol. 115, no. 772. P. 700–721.
12. Layne S.P., Hyman J.M., Morens D.M., Taubenberger J.K. *New coronavirus outbreak: Framing questions for pandemic prevention* // Science Translational Medicine. 2020. Vol. 12(534). eabb1469.
13. Linda J.S. Allen. A. *A primer on stochastic epidemic models: Formulation, numerical simulation, and analysis* // Infectious Disease Modelling. 2017. Vol. 2, is. 2. P. 128–142.
14. Mandal M., Jana S., Nandi S., Khatua A., Adak S., Kar T.K. *A model based study on the dynamics of COVID-19: Prediction and control* // Chaos Solitons Fractals. 2020. Vol. 136. № 109889.
15. Nicholas F. B. *Infectious Diseases* // Essential Mathematical Biology. Springer-Verlag London, 2003.
16. Otunuga O.M. Ogunsolu M.O. *Qualitative analysis of a stochastic SEITR epidemic model with multiple stages of infection and treatment* // Infectious Disease Modelling. 2020. Vol.5. P. 61–90.
17. Schmidt R., Waligora T. *Influenza forecast: Case-based reasoning or statistics* // Lecture Notes in Computer Science. 2007. Vol. 4692. P. 287–294.
18. Singh R.K., Rani M., Bhagavathula A.S., Sah R., Rodriguez-Morales A.J., Kalita H., Nanda C., Sharma S., Sharma Y.D., Rabaan

- A.A. *Prediction of the COVID-19 pandemic for the top 15 affected countries: Advanced autoregressive integrated moving average (ARIMA) model* // JMIR Public Health Surveill. 2020. Vol. 6. P. 19115.
19. Wu J.T., Leung K., Leung G.M. *Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: A modelling study* // Lancet. 2020. Vol. 395. P. 689–697.
20. Zakharov, V., Balykina, Y., Ilin, I., Tick, A. *Forecasting a New Type of Virus Spread: A Case Study of COVID-19 with Stochastic Parameters* // Mathematics. 2022. Vol. 10. P. 3725.
21. Zhang T., Ma Y., Xiao X., Lin Y., Zhang X., Yin F., Li X. *Dynamic bayesian network in infectious diseases surveillance: A simulation study* // Sci. Rep. 2019. Vol. 9. № 10376.

TWO EPIDEMIC MODELS OF MALARIA AND THEIR PRACTICAL APPLICATIONS

Serigne Modou Ndiaye, Saint Petersburg State University
(mamounepourtoi@gmail.com),

Victor V. Zakharov, Saint Petersburg State University, Dr.Sc.,
professor (mevictor@mail.ru).

Abstract: Based on data on the annual incidence of malaria in Senegal in the period from 2000 to 2021, the article demonstrates the possibilities of predicting the annual dynamics of the epidemic using the *SIR* model and the *CIR* model. A modified *SIR* model with constant coefficients is constructed, and a description of the *CIR* balance model with stochastic parameters is given. The question of the accuracy of forecasting the annual statistical indicators of the epidemic when using these models is investigated. As numerical experiments show, the average error in predicting the annual number of sick people compared to actual statistical data when using the *SIR* model is quite large, while the *CIR* model generates more accurate forecasts when compared.

Keywords: epidemic model of malaria, *SIR* model, prediction of active cases of the disease, epidemic balance model, dynamic balance principle, *CIR* model.