

Е.С. Куркина^{1,2}, Э.М. Кольцова²

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ РАСПРОСТРАНЕНИЯ ВОЛН ЭПИДЕМИИ КОРОНАВИРУСА COVID-19 В РАЗНЫХ СТРАНАХ МИРА

Введение

Эпидемия коронавируса показала, что человеческое сообщество не готово к серьезным вызовам, связанным с распространением эпидемий. Ни экономика, ни здравоохранение, ни административно-управленческие органы в богатых и бедных странах не справились с распространением не самого опасного заболевания. Были парализованы целые отрасли экономики, рухнули международные производственно-экономические связи. Это самая глубокая рецессия за столетие, и она будет усугубляться, поскольку пандемия еще не побеждена. На конец февраля 2021 г. мы имеем более 110 млн заболевших и около 2.5 млн умерших от коронавируса. Косвенные потери – гораздо больше, их еще предстоит оценить.

Для того, чтобы успешно противостоять распространению инфекции, необходимо анализировать динамику роста заболеваемости, рассчитывать нагрузку на здравоохранение и делать реалистичные прогнозы. Все это требует разработки адекватной математической модели распространения эпидемии, параметры которой легко подбираются из статистических данных и легко корректируются при изменении условий распространения.

Математическое моделирование эпидемий, распространяющихся среди людей, имеет долгую историю, как в изучении математических свойств и возможностей моделей, так и в моделировании конкретных эпидемий и контроле их эволюции [1-11]. Математическое моделирование распространения эпидемии COVID-19 в разных странах и регионах мира началось с момента первой вспышки болезни в Китае и интенсивно проводится в настоящее время. Для этой цели используются различные модели: модели логистического роста Ферхюльста и его модификации [12 - 15], компланарные модели типа SIR [10,16,17,18], SEIR, SIRS, SAIR [19-22] и др. более сложные [23-30], которые учитывают различные уточняющие факторы [17, 19, 20, 21, 25,27]. Все эти модели применимы в

¹ Московский государственный университет им. М.В. Ломоносова, факультет ВМК

² Российский химико-технологический университет им. Д.И. Менделеева

однородных изолированных системах, когда имеет место хорошее перемешивание в популяции и выполняется закон действующих масс. Более-менее это выполняется в больших городах, в которых не вводятся ограничительные меры. Для распространения эпидемии COVID-19 компланарные модели с постоянными параметрами не применимы [21]. Они описывают распространение инфекции в виде одной эпидемиологической волны с одним пиком.

Реальные системы являются открытыми и неоднородными и демонстрируют очень сложную динамику распространения инфекции. Если сгладить графики ежедневного прироста заболевших, то во многих случаях получаются кривые, имеющие несколько локальных максимумов (пиков), или содержат «плечи», или «плато», что не укладывается в классическую модель с одним пиком. Такое поведение можно объяснить возникновением нескольких локальных волн распространения инфекции, сдвинутых по времени. Локальные волны могут возникать в разное время в разных регионах страны, или в одном регионе, но в разное время, когда инфекция приносится извне, вызывая новую вспышку, или, когда в данном регионе некоторые жители начинают игнорировать карантинные меры. В общем случае распространение инфекции является суперпозицией очень многих локальных волн эпидемии.

Ни одна математическая *модель с постоянными параметрами* не способна описать такую сложную динамику. Нужно использовать модели с переменными параметрами.

Было предпринято много усилий по адаптации моделей типа SIR к описанию распространения коронавируса [18, 21, 24] и определению параметров моделей. Однако почти нет работ, которые описывают распространение инфекции как суперпозицию волн. Как правило, рассматривается только одна волна. Попытка описать две волны в Иране сделана в работе [31]. В работе [15] автор моделировал распространение эпидемии в ряде стран как суперпозицию многих волн. Но параметры волн в модели выбирались окончательно на спаде ближе к концу очередной волны. Так что модель годилась для описания, но не для прогноза.

Для моделирования распространения коронавируса рассматривались также сетевые модели [32], модели с запаздыванием [33], модели, использующие фрактальные производные, [34, 35] и др. Все эти модели носят в большей степени теоретический характер. Они раскрывают детали механизмов распространения инфекций и помогают понять влияние тех или иных факторов на динамику эпидемии. Однако ни одна из них не годится для описания сложной реальной динамики распространения эпидемии и прогнозирования.

Кроме того, хотя в эпоху цифровых технологий мы имеем беспрецедентную в истории базу данных по статистике распространения

заболевания во всем мире, многие исследования показывают, что данные не точны, и по разным оценкам ошибка составляет от 20 до 80%. Неточность статистических данных связана и с погрешностью тестов, и с недостаточностью тестирования в некоторых странах, и с нежеланием людей обратиться к врачу при слабом и умеренном течении болезни.

Определение параметров модели по неточно заданным статистическим данным является обратной некорректно поставленной задачей, которая имеет неединственное решение. Для выделения правильного решения в таких задачах проводят метод регуляризации. Точность определения параметров модели согласуется с погрешностью статистических данных. Поскольку погрешность большая, то бессмысленно использовать сложные модели, содержащие много параметров, для определения которых требуется много дополнительной информации. Проблема описания сложной динамики развития эпидемии состоит не в сложности модели, а в большой погрешности статистических данных, в открытости и неоднородности моделируемых систем (стран, городов, регионов). В этих условиях нужна простая математическая модель, которая имеет минимум параметров, которые легко определяются и подстраиваются под изменяющиеся условия, модель, которая легко выявляет новые локальные волны и уточняет прогнозы.

Такой моделью является дискретное логистическое уравнение Фейгенбаума, которое содержит всего два параметра: емкость системы (связанное с потенциальным числом людей, которые могут заболеть), и коэффициент скорости роста заболевших. Оба они влияют на окончательную общую численность инфицированных.

Мы разработали методику выявления локальных волн и определения их параметров в модели. Локальные волны вводятся в модель последовательно. Новая волна добавляется, когда расчетные данные начинают отставать от фактических.

Используя дискретное логистическое уравнение, мы исследовали распространение эпидемии COVID-19 в ряде стран в разных регионах мира с самого начала распространения эпидемии. В работе представлены результаты математического моделирования для стран: Индии, Бразилии, Швеции, Испании, Соединенных Штатов Америки (США), Японии, Израиле, Москвы и всего Мира. Первое, что можно отметить, это то, что динамика распространения инфекции и параметры модели сильно различаются в разных странах. В некоторых странах распространение инфекции долгое время неплохо описывалось одной волной с одним пиком (например, в Китае, Индии, Бразилии). В других странах математическое описание динамики развития эпидемии потребовало введения нескольких волн. Так, для описания распространения эпидемии в течение года в Японии и в Испании было использовано 6 волн, в Швеции – 9 волн, в США

и в Москве – 11 локальных волн. По каждой волне после прохождения пика делается окончательный прогноз на каждый следующий день; рассчитывается, когда волна закончится и сколько заболевших она даст. До пика точно определить емкость волны нельзя, поэтому до пика строятся разные сценарии развития волны, отвечающие разным предполагаемым значениям емкости, и прогнозируется, когда будет пик и каковы будут его параметры.

Показано, что выбор дискретного логистического уравнения для моделирования распространения инфекций в неоднородных открытых системах при неточно заданных статистических данных является оптимальным. Модель имеет минимальное число параметров, которые относительно просто определяются из статистических данных. Модель хорошо описывает фактические данные и делает реалистичные прогнозы. Горизонт прогноза оказывается большим, когда условия распространения не меняются и закрываются границы, например, как в случае первой волны в Китае и во Вьетнаме. В подавляющем большинстве стран и регионов границы оставались условно закрытые, карантинные меры не были жесткими, в результате возникали все новые и новые очаги заражения, условия распространения изменялись. В этих условиях модель делала хороший прогноз на 7-14 дней (ошибка не больше 1%).

Показано, что для того, чтобы держать эпидемию под контролем, необходимо все время проводить мониторинг, уточнять параметры модели и делать прогнозы.

1. Математическая модель распространения эпидемии

1.1 Модель распространения одной волны эпидемии

Для описания и прогнозирования сложной динамики распространения эпидемий типа коронавируса COVID-19 мы использовали дискретное логистическое уравнение, предложенное Мэйем в 1976 г. [36] и ставшее популярным после работ Фейгенбаума [37], Шарковского и других ученых. Оно имеет вид:

$$y_{n+1} = \alpha y_n (1 - y_n/N), \quad y_1 = y_0 - \text{задано}, \quad (1)$$

где y_n – общее число рассматриваемых случаев (или кумулятивная численность популяции) в n -й день; y_1 – начальное число случаев в первый день $t_1 = t_0$, с которого начался отсчет; α – параметр, характеризующий скорость роста популяции; N – емкость, или нормировочный множитель. Решением уравнения (1) является последовательность чисел $\{y_n\}$.

Это уравнение широко использовалось для моделирования различных процессов в биологических, физических, экономических и др. системах, но для моделирования распространения эпидемий оно

практически не применялось. Оно известно, как простое уравнение, демонстрирующее сложную хаотическую динамику.

В зависимости от значения коэффициента α поведение системы может сильно изменяться. При значениях $0 < \alpha \leq 1$ независимо от начального значения y_0 численность популяции заболевших стремится к нулю. Колебательную динамику и хаос это уравнение демонстрирует при $\alpha > 3$. В диапазоне изменения параметра $1 < \alpha < 2$ оно демонстрирует логистический рост – это рост численности популяции в среде с ограниченными ресурсами. Именно этот диапазон изменения параметра мы используем для описания распространения эпидемий. Здесь дискретное логистическое уравнение описывает все фазы развития эпидемии: стадию экспоненциального роста на ранней стадии распространения при $y_n \ll N$, затем стадию замедления скорости роста заболеваемости, пик заболеваемости, характеризуемый наибольшим приростом, и заключительную стадию спада эпидемии. При этом к концу эпидемии общая численность заболевших асимптотически стремится к стационарному значению, равному:

$$\bar{y} = \left(\frac{\alpha-1}{\alpha}\right) \cdot N; \quad y_n \rightarrow \bar{y}, \quad \text{при } n \rightarrow \infty \quad (2)$$

Соотношение (2) верно, если α остается постоянным на всем протяжении эпидемии. Если коэффициент α изменяется, то в формулу (2), надо подставить последнее значение параметра.

Из выражения (2) видно, что окончательное число заболевших зависит не только от параметра α , но и от емкости системы N . Чем больше значение α , тем больше окончательное число заразившихся. При $\alpha = 2$ оно максимальное и равно половине емкости $\bar{y} = N/2$. При значениях α , близких к единице (то есть при медленном распространении инфекции), окончательное число заболевших может быть намного меньше (на порядок), чем потенциальное число восприимчивых индивидов N .

Значение параметров N и α зависит от целого ряда факторов, таких как открытость страны к вирусной инфекции (приток людей из зон очагов эпидемии), скученность и плотность населения, наличие мегаполисов, устойчивость к заболеванию, дисциплинированность населения во время карантинных мероприятий и др. Проводимые карантинные меры снижают вероятность заражения, и показатель α уменьшается, что приводит к снижению общего числа заболевших \bar{y} .

1.2 Дискретное уравнение для приростов

Дискретное логистическое уравнение (1) непосредственно описывает изменение со временем суммарного числа заболевших $y_n = y(n)$, то есть задает *кумулятивную кривую*. Выведем уравнение для ежедневных новых

случаев заболевания, или ежедневных приростов. Это уравнение описывает *эпидемическую кривую*.

$y_{n+1} = y_n + \Delta_{n+1}$, Δ_{n+1} – это прирост заболевших в $n+1$ -ый день

$y_n = y_{n-1} + \Delta_n$, Δ_n – это прирост заболевших в n -ый день, и т.д.

То есть кумулятивная функция y_n равна сумме всех ежедневных приростов вплоть до n дня.

Рассмотрим уравнение для приростов.

$$\Delta_{n+1} = y_{n+1} - y_n = \alpha y_n \left(1 - \frac{y_n}{N}\right) - y_n = b y_n \left(1 - \frac{y_n}{\bar{y}}\right), \quad (3)$$

где $b = (a - 1) < 1$, $\bar{y} = \left(\frac{\alpha-1}{\alpha}\right) N$ – стационарное значение (2).

Мы видим, что уравнение для приростов (3) – это такое же дискретное логистическое уравнение, как и (1), только со своими параметрами.

Поскольку $y_n \rightarrow \bar{y}$, при $n \rightarrow \infty$ из уравнения (3) получаем, что

$$\Delta_n \rightarrow 0 \text{ при } n \rightarrow \infty.$$

Из выражения (3) следует, что при относительно небольшом числе инфицированных приросты изо дня в день увеличиваются, пока не достигнут максимума, затем они начинают уменьшаться. Максимальное значение Δ_{max} найдем, продифференцировав (3). Оно достигается при:

$$y^* = \frac{\bar{y}}{2} = \left(\frac{\alpha-1}{2\alpha}\right) N, \quad \Delta_{max}(y^*) = \frac{(\alpha-1)^2}{4\alpha} N. \quad (4)$$

Если приросты уменьшились до нуля, то эпидемия закончилась. Однако бывает, что ежедневные приросты не падают до нуля, как в случае первой волны эпидемии коронавируса, а держатся довольно долго на каком-то небольшом уровне. Это означает, что эпидемия перешла в эндемическую стадию. Тогда возможна новая вспышка эпидемии спустя некоторое время.

Уравнению (3) можно дать такую интерпретацию. В начале эпидемии число ежедневных новых случаев заражения пропорционально числу инфицированных, и прирост новых случаев идет со скоростью b . Однако со временем растет число переболевших, и вырабатывается коллективный иммунитет, коэффициент скорости распространения уменьшается:

$$b(y_n) = b \left(1 - \frac{y_n}{\bar{y}}\right). \quad (5)$$

Из этого выражения следует, что на первой фазе эпидемии главный вклад в скорость распространения дает первый член, а в конце эпидемии – второй.

1.3 Сравнение дискретного логистического уравнения с уравнением Ферхюльста и моделью SIR

Распространение эпидемий подчиняется закону логистического роста. Многие модели описывают логистический рост. Наиболее употребительными моделями для описания и прогнозирования распространения эпидемии коронавируса COVID-19 являются модели, основанные на дифференциальных уравнениях, такие как уравнение Ферхюльста и его различные модификации [12-15] и модель SIR и ее расширения [16-30]. Сравним дискретное логистическое уравнение (1) с уравнением Ферхюльста и классической моделью SIR.

Сразу отметим, что дискретное логистическое уравнение (1) *не является разностным аналогом* дифференциального логистического уравнения Ферхюльста:

$$\frac{dx}{dt} = \lambda x \left(1 - \frac{x}{K}\right), \quad \lambda > 0, \quad x(t_0) = x_0. \quad (6)$$

Параметры этих уравнений имеют разный смысл. В дифференциальном уравнении (6) K – это число людей, которое точно заболеет к концу эпидемии независимо от значения коэффициента скорости распространения λ и начального значения x_0 . Коэффициент λ качественно не влияет на динамику системы. В дискретном уравнении (1), как мы видели, N – это емкость системы, связанная с потенциальным числом людей, которое может заболеть. В *открытых* системах емкость N может охватывать число восприимчивых к заболеванию людей, превосходящее численность города, страны или региона. На окончательное число заболевших влияет коэффициент α . Если проводятся карантинные мероприятия, то коэффициент α уменьшается, и уменьшается число людей, которые заболеют.

В отличие от уравнения (6), которое надо проинтегрировать, чтобы найти решение уравнения, описывающее логистический рост, уравнение (1) *непосредственно* описывает логистическую функцию с помощью последовательности. При соответствующем выборе параметров уравнения (1) и (6) *дают близкие результаты*. Решением уравнения (6) является функция:

$$x(t) = \frac{x_0 K}{x_0 + (K - x_0) e^{-\lambda(t-t_0)}} \quad (7)$$

На стадии экспоненциального роста решения этих уравнений абсолютно *совпадают*. Действительно, при $x \ll K$ решением уравнения (6) является экспоненциальная функция:

$$x(t) = x_0 e^{\lambda(t-t_0)}. \quad (8)$$

Уравнение (1) описывает экспоненту (8), как геометрическую прогрессию:

$$y_n = \alpha y_{n-1}. \quad (9)$$

Параметры уравнений (1) и (6) связаны формулой:

$$\alpha = e^{\lambda dt}, \quad dt = t_n - t_{n-1}. \quad (10)$$

Прохождение пика и выход на стационар уравнения (1) и (6) описывают немного по-разному, но в целом результаты близки (см. ниже).

Сравним теперь дискретную модель (1) с классической SIR-моделью, предложенной Кермаком и Маккендриком в 1927 г. [10]. В ней рассматривается три группы индивидов — восприимчивые к заболеванию (Susceptible), инфицированные (Infected) и переболевшие/умершие (Recovered/Removed). Передача инфекции осуществляется от инфицированных индивидов к восприимчивым. Считается, что переболевшие индивиды приобретают иммунитет и не могут быть заражены вторично. Модель описывается системой 3-х дифференциальных уравнений:

$$\begin{aligned} \frac{dS}{dt} &= -rSI, \\ \frac{dI}{dt} &= rSI - \nu I, \\ \frac{dR}{dt} &= \nu I, \end{aligned} \quad (11)$$

$$S(t_0) = S_0, \quad I(t_0) = I_0, \quad R(t_0) = R_0.$$

Здесь $S(t), I(t), R(t)$ — численности восприимчивых, инфицированных и переболевших индивидов в момент времени t соответственно; r — коэффициент скорости передачи инфекции; ν — коэффициент интенсивности выздоровления инфицированных индивидов.

Из системы уравнений видно, что выполняется закон сохранения общей численности популяции M :

$$\frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = 0.$$

$$M(t) = S(t) + I(t) + R(t) = M = const.$$

Общее число заболевших в модели SIR определяется по формуле:

$$Z(t) = M - S(t) = I(t) + R(t),$$

Модель SIR применима для моделирования распространения заболеваний в однородных или слабо структурированных популяциях с большим количеством индивидов в условиях хорошего перемешивания. Более-менее эти условия соблюдаются в больших городах.

Модель SIR имеет устойчивый стационар при $r > v$, описывающий завершающую стадию эпидемии. При $t \rightarrow \infty$ все фазовые переменные асимптотически стремятся к своим стационарным значениям. При этом число инфицированных стремится к нулю. В модели SIR часто в качестве параметра M берут начальное число восприимчивых: $M = S_0$. На окончательное число заболевших влияют скорости r и v , аналогично, как в дискретном уравнении (1). Физический смысл параметра M почти такой же, как у параметра K в уравнении Ферхюльста (6), это не емкость системы N в дискретном логистическом уравнении, это реальное число людей, которые должны заболеть. Емкость N имеет более глубокий физический смысл. Она характеризует численность популяции, в которой распространяется инфекция, и даже при максимально возможном показателе скорости распространения $\alpha = 2$ число заболевших достигает только половины емкости. В случае эпидемии COVID-19 реальное число заболевших много меньше, чем емкость N .

Для стационара в модели SIR нет аналитической формулы. Стационар \bar{S} для восприимчивых, которые так и не заболеют, находится из трансцендентного уравнения [16]:

$$\frac{r}{v}((I_0 + S_0) - \bar{S}) = -\ln\left(\frac{\bar{S}}{S_0}\right). \quad (12)$$

Решая уравнение (12) и вычисляя значение \bar{S} , найдем общее число выздоровевших \bar{R} и заболевших \bar{Z} .

$$\bar{R} = \bar{Z} = M - \bar{S}. \quad (13)$$

Сравним динамику рассмотренных моделей (1), (6) и (11) на конкретном примере. Все модели на ранних стадиях описывают экспоненциальный рост, затем пик, а на поздних стадиях – выход на стационар общего числа заболевших. Подберем параметры моделей так, чтобы: 1) они описывали одинаково экспоненциальный рост в начале эпидемии; 2) выходили на один и тот же стационар в конце эпидемии. Пусть заданы параметры в SIR. Тогда, чтобы выполнялись пункты 1 и 2, дифференциальное логистическое уравнение (6) должно иметь параметры:

$$\lambda = r - v; \quad K = M - \bar{S}. \quad (14)$$

Соответствующие параметры дискретного логистического уравнения найдем из соотношений (14) и (2):

$$\alpha = e^{\lambda dt}, \quad dt = 1 \text{ (день)}; \quad N = (M - \bar{S}) \frac{\alpha}{\alpha - 1}.$$

Пример. Рассмотрим модель SIR с некоторыми характерными для COVID-19 параметрами:

$$r = 0.3, \quad v = 0.1, \quad I_0 = 1, \quad S_0 = 10^5 - 1, \quad M = 10^5$$

$$\lambda = r - v = 0.2, \quad \alpha = e^{b/2} = 1.1052. \quad N = 988290$$

Из трансцендентного уравнения (12) находим стационар \bar{S} :

$$\bar{S} = 5951.9 \quad \text{Тогда} \quad K = 94048.$$

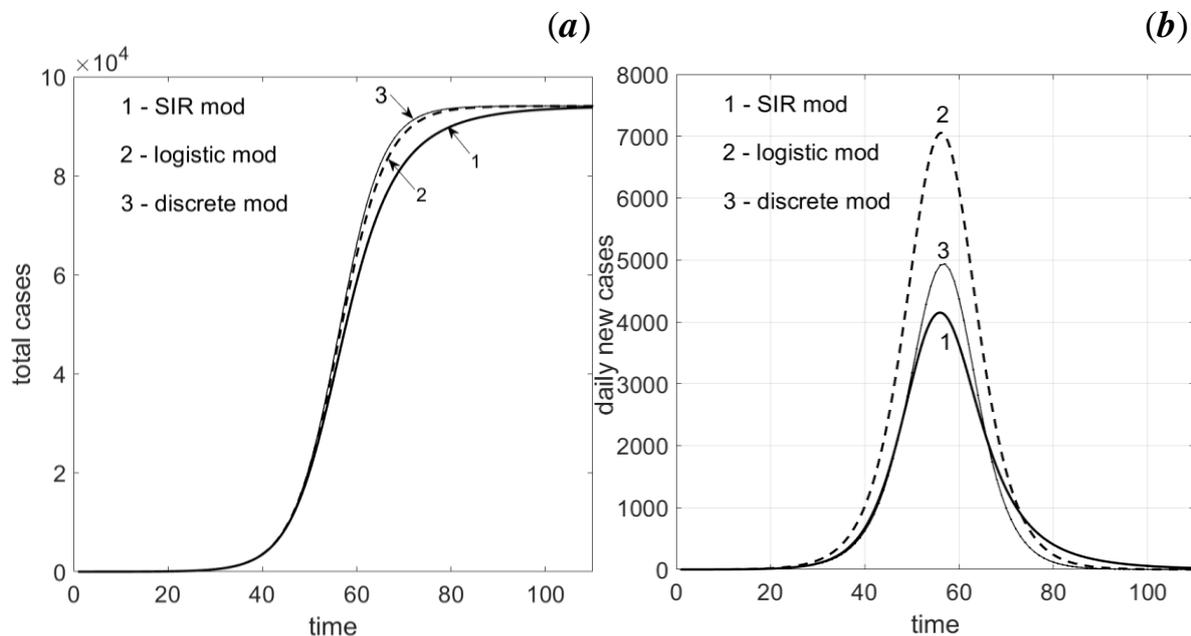


Рис. 1 Сравнение описания логистического роста в моделях SIR, Ферхюльста и дискретного уравнения (1). *a)* кумулятивные кривые; *b)* эпидемические кривые

Начальные данные подберем так, чтобы решения моделей совпадали на стадии экспоненциального роста. На рис. 1 а) и б) показаны кумулятивные и эпидемические кривые, описываемые тремя моделями. Мы видим, что на стадии экспоненциального роста все три модели совпадают, а далее немного расходятся. Все модели выходят на один и тот же стационар (так подобраны параметры). Модель Ферхюльста демонстрирует самый быстрый рост и самый высокий пик. Модель SIR имеет самый пологий спад эпидемии, поскольку учет выздоровевших уменьшает скорость распространения. Дискретная логистическая модель для данных значений параметров ближе к SIR-модели.

1.4 Определение параметров дискретного логистического уравнения по статистическим данным

Как было отмечено во введении, реальные статистические данные демонстрируют очень сложное динамическое поведение. Сглаженные эпидемические кривые имеют несколько локальных максимумов и содержат плато. Можно сделать вывод, что они представляют собой суперпозицию нескольких локальных волн. Причина сложной динамики понятна. Дело в том, что страны, города и отдельные регионы являются

открытыми неоднородными системами, в которых время от времени случайно возникают новые очаги эпидемии, порождающие новые цепочки передачи инфекции от инфицированных к восприимчивым. В неоднородных системах вспышки эпидемии случаются в разное время в разных местах и распространяются с разной скоростью. Все это приводит к возникновению локальных волн эпидемии, сдвинутых по времени. В результате вспышка эпидемии в том или ином регионе складывается из многих локальных волн. Некоторые из них сливаются в одну общую волну. Другие мощные локальные волны эпидемии, сильно сдвинутые по времени, не вкладываются в общую волну, и, когда та идет на спад, они дают новую отдельную вспышку эпидемии со своим локальным пиком.

Кроме того, статистические данные по распространению эпидемии имеют большую погрешность, а определение параметров модели по неточно известным данным является некорректной задачей, которая имеет неединственное решение. *И от этой проблемы не уйти*, поскольку найденное решение не может быть точнее статистических данных, а ошибка статистических данных составляет десятки процентов. Именно поэтому мы выбрали самую простую дискретную модель логистического роста с переменными параметрами, которая может описать любую динамику. Мы находим одно из возможных решений, хорошо описывающих статистические данные. Можно по-разному выделять локальные волны, и всегда есть люфт в определении параметров модели (1): коэффициента α и емкости N .

Рассмотрим сначала калибровку модели для описания одной волны.

Приведем модель (1) к безразмерному виду, сделав замену переменных:

$$y_n = x_n N, \quad 0 < x_n \leq 1, \quad \alpha = \alpha \quad (15)$$

Тогда она примет вид:

$$x_{n+1} = \alpha x_n (1 - x_n), \quad x_0 = y_0 / N \quad (16)$$

Как правило, начальное значение y_0 задавалось равным 1 ($y_0 = 1$).

Важной характеристикой уравнения также является *выбор шага дискретного временного интервала*, на котором рассчитывается численность населения, подвергшегося эпидемии. Мы взяли временной интервал в 12 часов, опираясь на расчеты, сделанные Майклом Левиттом для Китая [38]. Число случаев мы пересчитываем каждые 12 часов, а с фактическими данными сверяем через 24 часа, то есть раз день. Временной интервал связан со временем, в течение которого один человек в среднем может заразить другого. Оптимальность выбранного интервала подтвердили расчеты, проведенные для разных стран в течение всего прошедшего года.

Параметр α определяется в начале волны на стадии экспоненциального роста по статистическим данным.

$$\frac{X_{s,n+1}}{X_{s,n}} \cong \alpha_s \quad (17)$$

где $X_{s,n}$ – статистические данные, число случаев в n -ый день; α_s – среднее значение отношения (17) за несколько дней. Поскольку в модели число случаев x_n мы пересчитываем через 12 часов, то коэффициент α равен:

$$\frac{x_{n+1}}{x_n} = \alpha = \sqrt{\alpha_s}, \quad x_{n+1} = \alpha x_n, \quad (18)$$

Со временем показатель α может уменьшаться, и мы это отслеживаем из анализа статистических данных и связываем с информацией о введении ограничительных мер. При моделировании распространения эпидемии в разных странах коэффициент α для разных локальных волн или не менялся, или мы скачком изменяли несколько раз (обычно 2-4 раза), чтобы наилучшим образом описать фактические данные. Как правило, он последний раз менялся до середины пика, а после этого в модели оставался постоянным. Последнее значение α использовалось для прогнозирования.

Параметр N определяется в пике эпидемии из формул (4):

$$N \cong X_s^* \left(\frac{2\alpha}{\alpha-1} \right), \quad \text{или} \quad N \cong \Delta_s^* \frac{4\alpha}{(\alpha-1)^2}, \quad (19)$$

где X_s^* и Δ_s^* – средние значения числа случаев и ежедневных приростов в окрестности пика, соответственно.

Иногда пик эпидемии представляет собой достаточно длинное плато, тогда значение N приходится изменять, уточнять. До пика, начиная со стадии экспоненциального роста, мы задаем разные предполагаемые значения емкости N . Они берутся из разных оценок, из предшествующего опыта и из сравнительного анализа. Для разных значений N мы строим разные сценарии развития эпидемии. Для каждого сценария мы рассчитываем, когда будет пик, каков будет прирост в пике, сколько будет случаев каждый день, когда волна закончится и сколько заболевших она даст. Если коэффициент α уменьшился (ввели ограничительные меры), то снова пересчитываются сценарии. На стадии экспоненциального роста кумулятивные кривые для разных значений N не различаются. Расщепление этих кривых наблюдается ближе к пику. В пике окончательно выбирается значение N из формул (19), и строится окончательный прогноз по данной волне на каждый последующий день вплоть до окончательного спада волны.

1.5 Модель распространения нескольких волн эпидемии

Мы рассматриваем страну, город, регион и др., в которых распространяется инфекция, как *открытую систему*. Открытые границы,

умеренные ограничительные меры приводят к тому, что время от времени возникают новые очаги заражения, которые порождают новые цепочки последовательного заражения людей, тем самым увеличивая потенциальное количество людей, которое может заболеть. Это означает, что появляются новые емкости восприимчивых индивидов, и возникают новые локальные волны. Происходит суперпозиция локальных волн.

Для расчета распространения инфекции в виде нескольких волн используется несколько дискретных уравнений вида (1) (или (16)), описывающих каждая свою волну, со своей емкостью $N^{(i)}$, набором показателей $\alpha^{(i)}$, началом отсчета времени $t_0^{(i)}$ и начальным значением заболевших $x_0^{(i)} = y_0^{(i)}/N^{(i)}$:

$$x_{n+1}^{(i)} = \alpha^{(i)} x_n^{(i)} (1 - x_n^{(i)}), \quad x_0^{(i)} = y_0^{(i)}/N^{(i)}. \quad (20)$$

$$y_n^{(i)} = x_n^{(i)} N^{(i)}, \quad i = 1, \dots, M, \quad (21)$$

где M – число волн.

Общее число заболевших в момент времени t_n определяется суммой:

$$Y_n^{(i)} = \sum_{i=1}^M x_n^{(i)} N^{(i)} \quad (22)$$

Параметры уравнений подбираются так, чтобы наилучшим образом описать статистические данные.

Кратко опишем алгоритм определения параметров очередной локальной волны. Пусть в модели мы уже использовали j волн для описания распространения инфекции. Суперпозиция этих волн до какого-то момента времени t_n хорошо описывала статистические данные $X_{s,n}$:

$$Y_n = \sum_{i=1}^j x_n^{(i)} N^{(i)} \approx X_{s,n}.$$

А затем суммарная j -я волна пошла на спад, и статистические данные стали опережать расчетные данные. Тогда разница между статистическими данными и расчетными дает последовательность значений Z_n для определения следующей волны:

$$Z_n = X_{s,n} - Y_n, \quad n = n_1, n_1 + 1, n_1 + 2, \dots \quad (23)$$

Параметры уравнения (16) для описания $j+1$ -й волны подбираем так же, как для первой волны: сначала коэффициент α и несколько предполагаемых значений N , описывающих разные сценарии развития $j+1$ -й волны. Затем в окрестности пика уточняем значение емкости N и делаем окончательный прогноз по дальнейшему развитию этой волны.

2. Результаты моделирования

Статистические данные о подтвержденных случаях коронавируса (COVID-19) для рассматриваемых стран (Индия, Бразилия, Швеция,

Испания, Соединенные Штаты Америки (США), Япония, Израиль) и всего Мира мы брали с сайта Worldometer [39]. База данных случаев заболевания коронавирусом по Москве была взята с сайта [40].

Распространение инфекции можно описать в модели с помощью одной волны в двух случаях: 1) при слабых ограничительных мерах, которые почти не препятствуют свободным контактам людей, 2) при очень жестких ограничительных мерах, которые приводят к полной изоляции зараженной части населения. К первому случаю можно отнести распространение эпидемии в Индии и Бразилии до конца октября. Вторым случаем имел место в Китае и Вьетнаме. В большинстве других стран математические модели дают описание распространения эпидемии в виде нескольких волн. Локальные волны проявляются на сглаженных графиках ежедневных новых случаев заболевания в виде локальных пиков и плато. Отметим, что в силу того, что статистические данные неточные, а обратная задача выделения волн является некорректной задачей, то точно определить количество локальных волн и их характеристики невозможно. Поэтому мы рассматривали эту задачу как феноменологическую, и параметры локальных волн подбирали из требования как можно лучше описать имеющиеся на рассматриваемый момент данные по COVID-19.

Для всех рассматриваемых стран результаты моделирования представлены на двух рисунках. На рис. 2-10 а) показаны графики роста общего числа заболевших – кумулятивные кривые. На рис. 2-10 б) показаны ежедневные приросты новых случаев заболевания – эпидемические кривые. Статистические данные отмечены ромбиками, а результаты расчетов – кружками. Жирной сплошной линией отмечен график скользящего среднего по семи дням для ежедневных новых случаев заболевания. Описание результатов моделирования развития эпидемии коронавируса начнем с Бразилии.

Бразилия. Бразилия является третьей страной мира по общему количеству заболевших – 10.5 млн человек в конце февраля 2021 г. Статистические данные по ежедневным приростам заболевших демонстрируют сильную волатильность (на рис. 2 б) отмечены точками), сильнее, чем в других странах, возможно, это связано с проблемой массового тестирования. Результаты моделирования представлены на рис. 2 а) и б). Эпидемия в Бразилии начала разгораться во второй половине марта 2020 г., когда прирост заболевших в день стал составлять сотни человек. Локальный пик эпидемии в Бразилии наблюдался в начале июля 2020 г., после которого эпидемия пошла на спад (с 7 июля по 19 июля). Но с 20-го июля начался новый скачок в росте числа заболевших. Абсолютный максимум ежедневного прироста заболевших (70869 человек) имел место 29 июля, после чего эпидемия в среднем пошла на спад. Мы вели расчеты по модели для Бразилии, начиная с мая. Нормировочный множитель

уточнялся несколько раз. Емкости всех волн приведены на рис. 2 а). Показатель скорости роста по мере распространения эпидемии уменьшался несколько раз. Это можно объяснить и некоторыми ограничительными мерами, и уменьшением вероятности контактов с заболевшими. Модель хорошо описала локальную волну с пиком 7 июля. Но в августе пришлось заново подбирать параметры. До конца октября эпидемия в Бразилии находилась на спаде, и модель с уточненными в сентябре параметрами ($N = 450$ млн, $\alpha_7 = 1.0135$ с 16.07) хорошо описывала статистические данные до конца октября. С 29 июля по 26 октября расхождение модели со статистическими данными по общему числу заболевших составляло меньше 1 % (за исключением 3-х дней, в которых расхождение было чуть больше 1 процента). В ноябре в Бразилии начала формироваться новая вспышка эпидемии, и для ее описания в модели мы ввели вторую волну; предложили два сценария. По первому сценарию вторая волна в пике должна иметь значения параметров порядка первой волны. По второму сценарию вторая вспышка эпидемии должна была быть мощнее первой, как во многих европейских странах. Реализовался первый сценарий для второй волны, ее пик пришелся на 14 декабря. По расчетам первая волна в Бразилии закончится в июне 2021 (когда прирост заболевших будет исчисляться десятками), она даст порядка 6 млн инфицированных человек. Вторая волна закончится в апреле, она добавит 2.2 млн инфицированных.

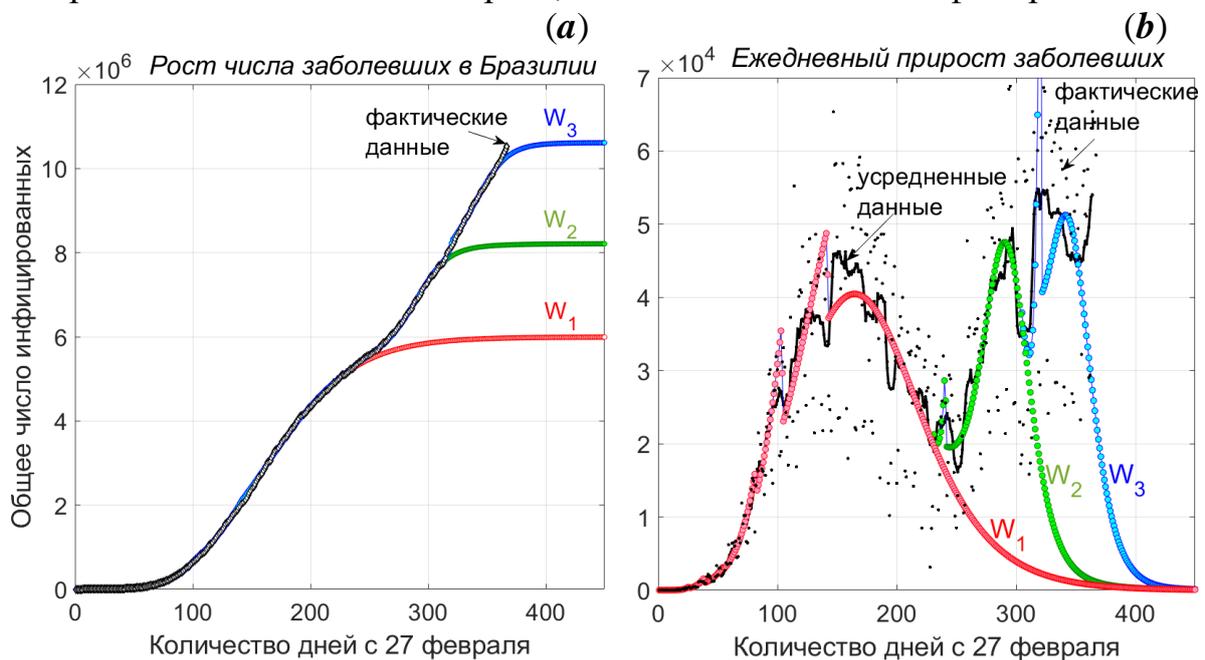


Рис. 2 Динамика распространения эпидемии коронавируса в Бразилии между 27.02.2020 и 27.02.2021. Фактические данные и результаты моделирования: W_1 , W_2 , W_3 – расчетные волны

Третья локальная волна в Бразилии начала формироваться в декабре. Ее пик наблюдался в начале февраля, после чего она пошла на спад. Третья

волна по расчетам закончится в мае, она принесет 2.4. млн заболевших. В конце февраля начался подъем новой четвертой локальной волны, параметры которой еще нельзя определить в модели.

Индия – вторая по численности населения страна в мире (≈ 1.4 млрд чел.), для которой характерна скученность населения в местах проживания. Несмотря на попытку ввести жесткие ограничительные меры на начальной стадии распространения инфекции, инфекция вырвалась из-под контроля и стала быстро распространяться в густонаселенной стране, обогнав по числу зараженных Бразилию и выйдя на второе место в мире после США. На рис. 3 представлены по дням статистические данные по общему числу заболевших (а) и ежедневным новым случаям заболевания (б), а также показаны результаты расчетов по модели. Мы видим, что вплоть до начала ноября модель в одну волну хорошо описывала учтенные данные. Пик этой волны был пройден в середине сентября 2020 г. Уточненный в сентябре нормировочный множитель первой волны равнялся $N = 550$ млн. Показатель скорости роста уменьшался по мере распространения эпидемии, последний (шестой) показатель скорости роста $\alpha_6 \cong 1.018$ и не изменялся с 3 августа. Но в ноябре скорость спада эпидемии замедлилась, и модель стала отставать от статистических данных, поэтому была введена вторая локальная волна.

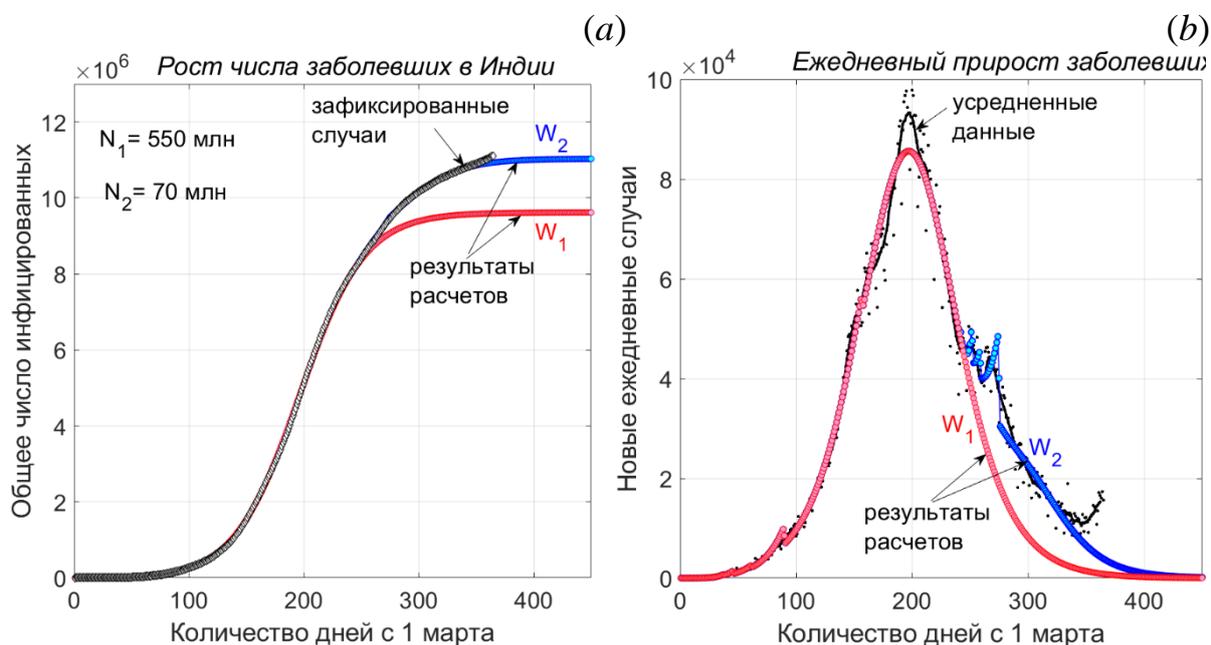


Рис. 3 Динамика распространения эпидемии коронавируса в Индии между 01.03.2020 и 28.01.2021. Фактические данные и результаты моделирования

Вторая волна оказалась гораздо слабее первой. Не успев развиваться, она прошла свой пик в конце ноября и далее пошла на спад. Прогнозируемые расчетные данные хорошо описывали статистические данные до середины февраля. Но во второй половине февраля 2021 г. стала подниматься третья

волна. Затухнет ли она быстро или разовьется в новую мощную вспышку, можно будет понять через некоторое время, описать ее в модели и сделать прогнозы по дальнейшему развитию. По расчетам первая волна в Индии закончится в июне 2021, она даст примерно 9.6 млн. заболевших. Вторая волна закончится примерно в это же время и добавит еще 1.4 млн человек.

США – огромная страна, в которую входят разные регионы, сильно различающиеся по образу жизни. В плане распространения инфекции здесь явно выделяются большие города, такие как Нью-Йорк, Лос-Анжелес, Чикаго, Сан-Франциско и др., с большой плотностью населения и бешеным ритмом жизни. При анализе статистических данных можно выделить три большие вспышки эпидемии в США (рис. 4 *а*) и *б*): весеннюю, летнюю 2020 г. и осеннюю, которая сейчас в феврале 2021 г. идет на спад. Как хорошо видно из сглаженных графиков ежедневных приростов (рис. 4 *б*), эти три большие волны имеют несколько локальных максимумов и на подъемах, и на спусках, что говорит о том, что они представляют собой суперпозицию многих локальных волн. Первая вспышка началась год назад в феврале 2020 г. Занимаясь моделированием распространения эпидемии в США с самого начала, мы последовательно вводили в модели локальные волны и давали по каждой волне прогноз. Ситуация в США из-за плохой организации карантинных мер, протестных движений жителей, открытых границ и т.п. все время непредсказуемо изменялась. В результате прогнозы по развитию очередной волны хорошо работали неделю, максимум – две. Потом расчетные данные начинали отставать от фактических, и приходилось вводить очередную волну. За прошедших двенадцать месяцев мы ввели одиннадцать локальных волн для описания распространения эпидемии в США, то есть в среднем локальная волна вводилась в модель раз в месяц. Расчётные данные прекрасно описали статистические данные, расхождение составляет меньше 1% за исключением нескольких дней. Еще раз подчеркнем, что задача определения параметров модели по неточно заданным статистическим данным имеет неединственное решение. Вначале мы хорошо описали весеннюю вспышку шестью волнами, но задним числом объединили их парами и получили всего три волны, как показано на рис. 4. Первая весенняя вспышка закончилась только в августе 2020 г., по расчетам она дала примерно 2.185 тыс. заболевших.

В то время, как в большинстве стран Европы в июне-июле 2020 уровень заболеваемости коронавирусом существенно снизился, в июне в США на фоне уличных протестов начала формироваться вторая большая волна эпидемии. Пик этой волны пришелся на конец июля. Она оказалась мощнее первой большой волны более чем в два раза. По расчетам она закончилась только в ноябре и добавила чуть более 5 млн заболевших. В модели мы ее описали тремя локальными волнами: четвертой, пятой и

шестой по счету с начала эпидемии (рис. 4 а), б)). Не успела летняя волна заметно спасть, как стала формироваться следующая волна, как и во многих странах Европы. По-видимому, она связана с игнорированием значительной частью населения ограничительных мер и открытием границ. Осенне-зимняя вспышка эпидемии оказалась намного мощнее, чем весенняя и летняя вспышки. Она уже растянулась на полгода и сейчас идет на спад, но пока прирост заболевших примерно такой, каким он был в пике летней вспышки. Осенняя вспышка, как хорошо видно из графика ежедневных приростов, (рис. 4 б)) имеет три ярко выраженных максимума. То есть три раза она уже было пошла на спад, но новая очередная локальная волна потянула ее вверх. Максимальные приросты наблюдались в начале января 2021 г. В модели для описания этой вспышки мы ввели последовательно пять локальных волн: одна небольшая на подъеме, три большие в пике и одна небольшая – на спаде. Но уже из графиков видно, что последняя одиннадцатая локальная волна стала отставать от фактических данных, и стала проявляться следующая локальная волна. По расчетам одиннадцатая волна закончится в апреле. Пять волн (7ая - 11ая), описывающих осенне-зимнюю вспышку, в сумме дадут примерно 21.6 млн заболевших.

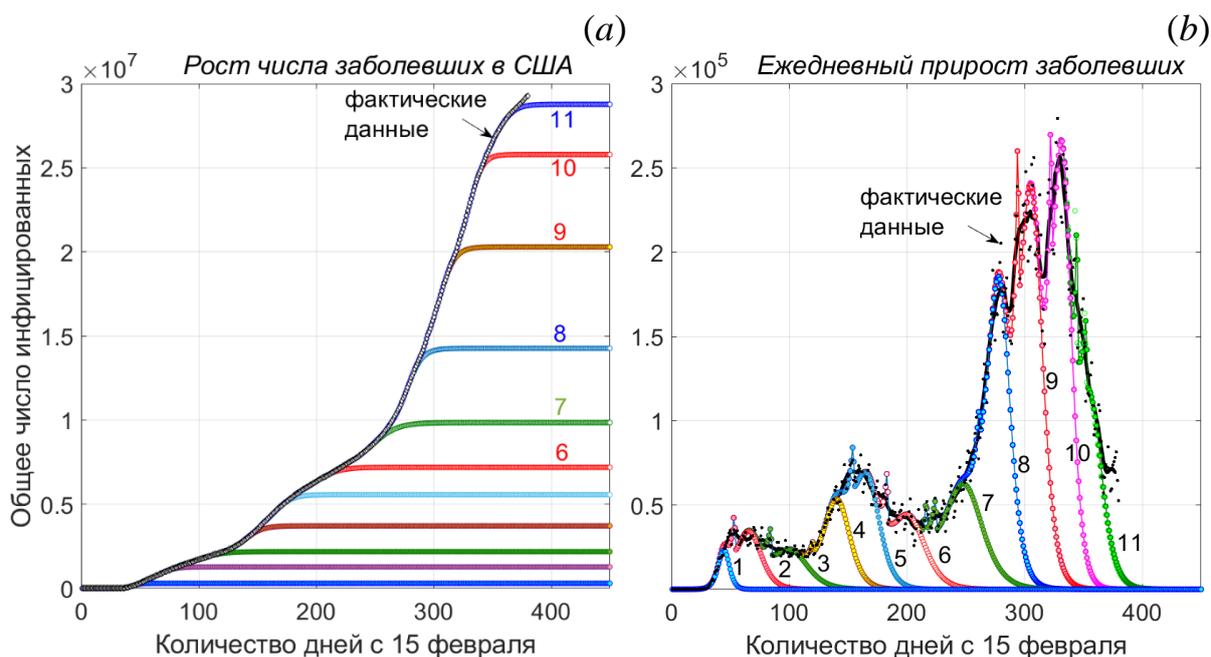


Рис. 4 Динамика распространения эпидемии коронавируса в США между 15/02/2020 и 01/03/2021. Фактические данные и результаты моделирования. Цифрами обозначены расчетные волны в модели

Москва. Москва – один из крупнейших мегаполисов мира, центр международной деловой активности и главный транспортный узел России. Это открытая система: внутреннее сообщение со всей страной не прекращалось даже при самых жестких ограничительных мерах. В

результате в Москве все время случайно возникали новые очаги заражения, которые запускали новые локальные волны. Мониторинг ситуации с коронавирусом в Москве мы проводим еженедельно, начиная с конца марта. На сегодняшний день Москва внесла чуть меньше четверти от общего число заболевших страны, в то время как ее население в 12.7 млн человек составляет примерно одиннадцатую часть жителей России. На рис. 5 а), б) представлены статистические данные и результаты моделирования. Мы видим, что распространение коронавируса в Москве проходит в виде двух больших волн: весенней вспышки и осенне-зимней, которые соединяются через нижнее плато. То есть весенняя вспышка в Москве не затухла до нуля, а перешла в эндемическую стадию, длившуюся два месяца. С начала июля до середины сентября средний прирост заболевших держался на уровне 600-700 человек в сутки. Для моделирования распространения первой вспышки заболевания мы использовали семь локальных волн: две волны на подъеме, три волны на спаде и две широкие волны для описания эндемической стадии.

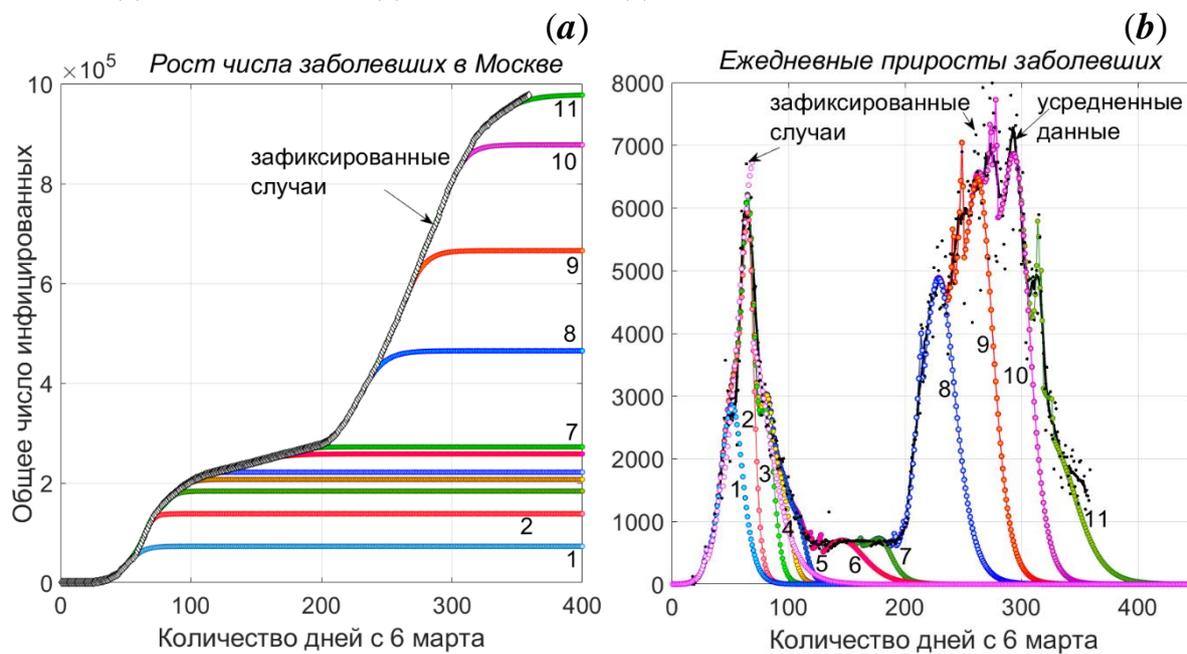


Рис. 5 Динамика распространения эпидемии коронавируса в Москве между 06.03.2020 и 27.02.2021. Фактические данные и результаты моделирования

Первая большая волна (суперпозиция пяти волн) по расчетам закончилась в июле 2020 г. Она дала 226 тысяч заболевших. Отголоски эндемической стадии окончательно затухли в октябре. Она добавила 50 тысяч заболевших. Показатель скорости роста α первой локальной волны уменьшался два раза, и его уменьшение коррелировало с введением ограничительных мер. Показатели скорости роста остальных пяти локальных волн были на уровне показателя скорости первой локальной волны после введения ограничительных мер.

Вторая большая вспышка эпидемии в Москве начала развиваться в середине сентября. Во второй половине октября ежедневные приросты инфицированных превзошли 5 тысяч. В Москве с каждым днем стали ужесточать ограничительные меры, которые затормозили распространение инфекции. В результате две недели с 16 по 29 октября ежедневные приросты держались примерно на одном уровне; среднее значение приростов составляло 4713 человек. Но затем приросты поползли вверх, и во второй половине ноября перевалили за семь тысяч. Очередной локальный максимум наблюдался в конце декабря, после чего, осциллируя, эпидемия пошла на спад. В модели осенне-весенняя вспышка коронавируса в Москве хорошо описывается четырьмя локальными волнами: одна – на подъеме (восьмая локальная волна по счету), одна – на спаде (одиннадцатая локальная волна) и две волны (девятая и десятая) описывают два самых больших локальных максимума. По расчетам она закончится в апреле 2021 г. и в сумме по четырем волнам (8-11) принесет 706 тысяч заболевших. Всего 11 волн дадут 982 тысячи заболевших. Очевидно, что 11-ая волна - не последняя в модели. Уже сейчас она стала отставать от статистических данных. 20.02.21 расхождение расчетных данных со статистическими составляло всего 0.02%, после этого расхождение стало монотонно возрастать и достигло 0.49% 27 февраля.

Япония. За 12 месяцев распространения заболевания коронавирусом в Японии наблюдается три большие вспышки, которые в модели представлены в виде суперпозиции шести локальных волн. Расчетные и фактические данные показаны на рис. 6 а), б). Мы видим, что модель хорошо описывает статистические данные. Япония – более закрытая и более однородная страна, чем США, Франция, Испания и другие европейские страны. Первая волна в Японии хорошо была описана с помощью одной волны в модели, характерной для замкнутой системы. Целый месяц после пика статистические данные изменялись в соответствие с расчетными прогнозными данными.

В Японии емкость 1-ой волны составляет $N=270$ тыс. человек, а последний показатель роста численности $\alpha = 1,065$. Теоретическое стационарное значение общей численности инфицированных (2) равно:

$$\bar{y} = \left(\frac{\alpha-1}{\alpha}\right) \cdot N = \left(\frac{1.065-1}{1.065}\right) \cdot 270\,000 = 16478 \text{ инфицированных.}$$

Таким образом, по расчету первая волна должна была закончиться в конце мая 2020 г. и выйти на стационар, равный около 16.5 тысяч. Фактические данные по Японии к концу 1-ой волны (23 мая) выявили 16536 инфицированных, что является убедительным подтверждением правильности выбранной модели.

Вторая и третья волны вспыхнули, по-видимому, из-за ослабления ограничительных мер. Как следует из расчетов по модели, вторая волна начала формироваться в конце мая, а давать существенные приросты (более 100 чел.) - с первых чисел июля. Она оказалась мощнее первой волны почти в четыре раза и добавила около 60.5 тыс. инфицированных. Ее пик пришелся на 10 августа 2020 г. В модели нормировочный множитель второй волны равняется $N = 1.5$ млн, а показатели роста α порядка значений этих показателей в первой волне.

Осенне-зимняя вспышка началась с возникновения третьей волны в начале сентября, когда еще не спала вторая волна. Третья локальная волна была относительно небольшой. Она прошла пик в середине октября 2020 г. и начала спадать. Если бы были введены жесткие ограничительные меры, как в Китае, то был бы шанс, что эта волна затухнет. Но слабые ограничения и открытые границы привели к новому мощному росту числа инфицированных, как во многих других странах. В модели этот рост описывается с помощью двух локальных волн: четвертой и пятой, самой большой (см. рис. 6 б)). Пик пятой волны наблюдался в середине января, потом она резко пошла на спад вместе с фактическими данными.

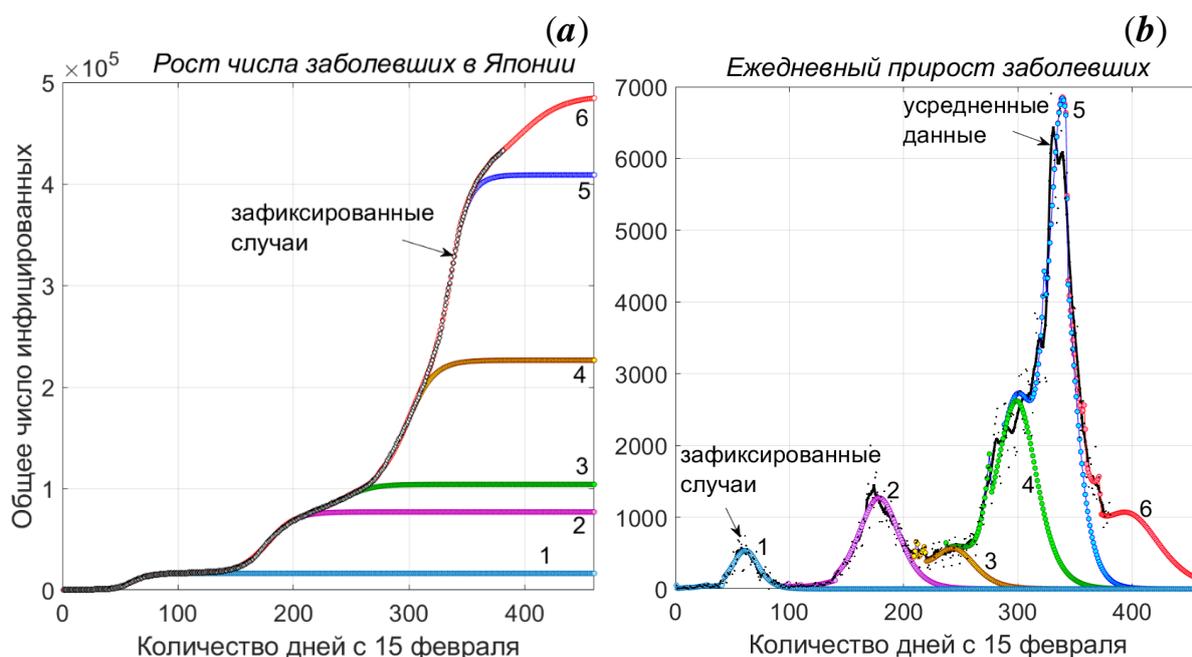


Рис. 6 Динамика распространения эпидемии коронавируса в Японии между 15.02.2020 и 01.03.2021. Фактические данные и результаты моделирования.

Цифрами обозначены расчетные волны в модели

К концу первой декады февраля скорость спада замедлилась, и фактические данные стали опережать расчетные, поэтому в модели была введена шестая волна. По расчетам шестая волна закончится в июне 2021 г, тогда суммарное количество заболевших по всем волнам будет 486 тысяч.

Израиль. В Израиле за год с начала эпидемии наблюдалось четыре вспышки заболевания коронавирусом, которые в модели описываются в виде четырех волн с емкостями $N_1=340$ тыс., $N_2=2$ млн, $N_3=4.4$ мл. и $N_4=17$ млн соответственно. Расчетные и фактические данные о количестве инфицированных представлены на рис. 7 а), б). Мы видим, что модель хорошо описывает статистические данные, особенно - первые две волны, которые развивались, как в замкнутой системе.

Первая волна в Израиле прошла свой пик 1 апреля 2020 г. Именно тогда в самом начале апреля была уточнена емкость первой волны в 340 тыс. человек. Последний показатель роста был $\alpha = 1,052$. Первая волна по расчетам должна была закончиться в конце мая и выйти на теоретическое стационарное значение, вычисленное по формуле (2):

$$\bar{y} = \left(\frac{\alpha-1}{\alpha}\right) \cdot N = \left(\frac{1,052-1}{1,052}\right) \cdot 340\,000 = 16806 \text{ инфицированных.}$$

Это полностью совпало с расчетными данными по модели (16804 инфицированных в конце мая) и прекрасно подтвердилось статистическими данными, которые 27 мая дали общее число инфицированных 16872 человек. Таким образом, прогноз, сделанный по модели, здесь работал целых два месяца, пока не началась новая летняя вспышка заболевания коронавирусом.

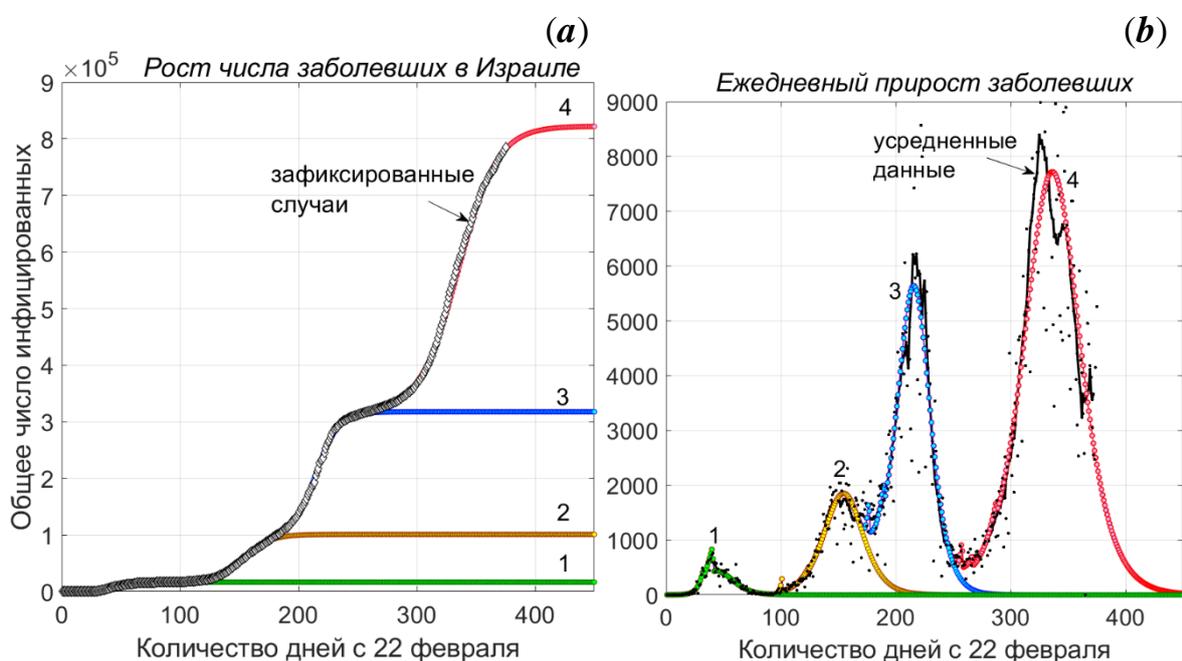


Рис. 7 Динамика распространения эпидемии коронавируса в Израиле между 2.02.2020 и 01.03.2021. Фактические данные и результаты моделирования. Цифрами обозначены расчетные волны в модели

Волны коронавируса в Израиле развивались по возрастающей. Вторая волна закончилась в начале октября 2020 г. и добавила 84.2 тыс. заболевших. Третья волна возникла, когда вторая только начала затухать.

Пик третьей волны наблюдался в конце сентября. Правительство Израиля в сентябре ввело сильные ограничительные меры, что позволило затормозить распространение инфекции. В модели это отразилось в уменьшении показателя скорости роста с 18 и 29 сентября. Третья волна закончилась в начале декабря и добавила 216.6 тыс. заболевших. Мы посчитали, что если бы коэффициент скорости роста был такой же, как до 18 сентября, то прогнозируемое общее число заболевших в результате третьей волны должно было быть примерно на 30 тыс. больше.

Четвертая волна – самая мощная, имеет самую большую емкость, превосходящую численность страны. Это указывает на открытость системы, в которую входят представители других стран. Четвертая волна имеет один ярко выраженный пик и еще один небольшой локальный пик на спаде, указывающий на локальную волну. Для более точного описания в модели можно было бы ввести эту локальную волну, но поскольку пик небольшой и прижат к основной волне, мы в модели описали эти две волны усредненной одной волной. Расхождение расчетных и фактических данных видно на графике приростов. Ошибка с общей численностью заболевших в окрестности заметного расхождения составляет в среднем 2 - 2.5 %, но не больше 3%. В подавляющих остальных случаях она меньше 1%. Пик четвертой волны наблюдался в середине января; сейчас, в конце февраля, 2021 она находится на спаде. Ежедневные приросты демонстрируют сильную волатильность. По расчетам она закончится в конце июня 2021 г. и добавит 504.7 тыс. заболевших, а суммарное количество заболевших в результате всех волн будет 822.5 тысяч.

Испания. Испания – одна из европейских стран, наиболее сильно пострадавших от пандемии коронавируса. Расчетные и статистические данные об изменении числа инфицированных людей представлены на рис. 8 а), б). Динамика распространения заболевания коронавирусом в Испании схожа с динамикой распространения эпидемии в Израиле. В Испании также с начала заболевания было 4 больших волны, и каждая следующая была мощнее предыдущей. Первая весенняя вспышка прошла пик в 20-х числах марта 2020 г. и далее пошла на спад. Но затухание волны длилось долго, имели место отдельные небольшие всплески. Если подъем волны занял примерно месяц, то спад длился три месяца. Для более точного описания медленного затухания первой волны в модели мы ввели две дополнительные локальные волны с небольшими емкостями: $N_2 = 0.52$ млн и $N_3 = 0.5$ млн соответственно; основная первая локальная волна имеет емкость $N_1 = 4.3$ млн человек. Суммарное количество инфицированных, принесенное первой волной, составляет примерно 271.7 тысяч человек.

Вторая большая волна стала формироваться в Испании по нашим расчетам в 20-х числа июня 2020 г. Ее возникновение, как и в других странах, связано с ослаблением ограничительных мер и открытием границ.

В модели она характеризуется емкостью $N_2 = 23$ млн, которая существенно превышает емкость первой волны. Пик этой волны по модели был пройден 10 сентября со значением прироста зараженных в максимуме 10855. Вторая большая волна в модели закончилась в январе, она добавила около 696 тыс. инфицированных. Ограничительные меры, которые стали вводиться правительством Испании в августе и сентябре, в модели отразились в снижении показателя скорости роста α и привели к уменьшению скорости распространения заболеваемости. Однако эти меры оказались недостаточными, новую волну остановить не удалось, и после небольшого спада в Испании снова пошел вверх прирост новых заболевших. В модели новая третья вспышка описывается пятой локальной волной. Она началась в конце августа 2020 г. Ее пик наблюдался в самом конце октября. По расчетам она окончательно затухла в феврале 2021 г. и принесла 773 тыс. заболевших.

Четвертая вспышка заболевания в Испании описывается в модели шестой волной с емкостью $N_6 = 37$ млн. Ее пик был пройден в 20-х числах января 2021 г., а сейчас она находится на спаде. Закончится шестая волна в конце мая 2021 г. и добавит к имеющимся еще 1.423 млн заболевших, в сумме эти волны покажут 3.164 млн заболевших.

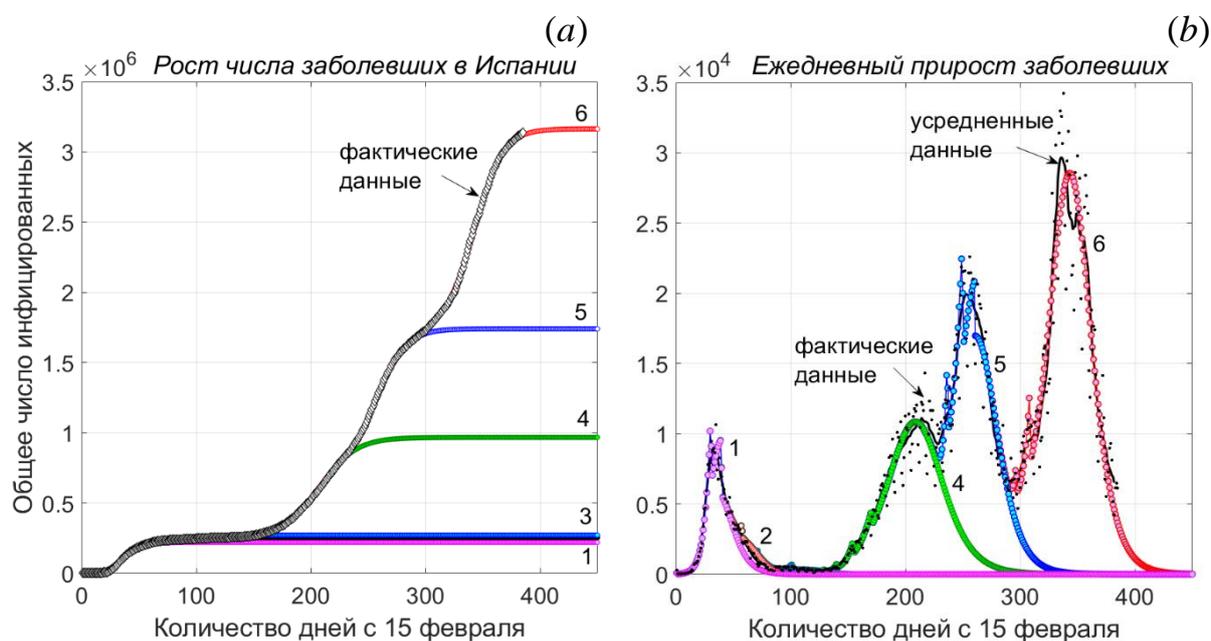


Рис. 8 Динамика распространения эпидемии коронавируса в Испании между 15.02.2020 и 03.03.2021. Фактические данные и результаты моделирования. Цифрами обозначены расчетные волны в модели

Швеция. Исследуем, как протекает эпидемия коронавируса в Швеции, единственной европейской стране, не считая Белоруссии, где не вводились жесткие ограничительные меры и работали все предприятия, учебные заведения, кафе, рестораны, театры и др. На рис. 8 представлены

фактические данные и результаты расчетов по математической модели для общего числа инфицированных и суточного прироста числа инфицированных COVID-19 в Швеции. Мы видим, что статистические данные демонстрируют сильную волатильность, сильнее, чем в других странах. Первая вспышка эпидемии в Швеции началась в конце февраля 2020 г. и вышла на плато в начале апреля, которое продержалось два месяца, до конца мая. Ежедневные приросты инфицированных в этот период колебались от 261 до 808, а среднесуточный прирост составил 560 человек. В июне 2020 г. наблюдалась локальная вспышка заражения коронавирусом, которая имела пик 24 июня со значением 1698 заболевших за день. Далее эпидемия быстро пошла на спад и достигла своего минимума 12 июля, когда было всего 106 заболевших за день. На нижнем плато с относительно небольшими локальными максимумами эпидемия продержалась до 20-х чисел сентября. Среднее значение инфицированных в день составило 236 человек.

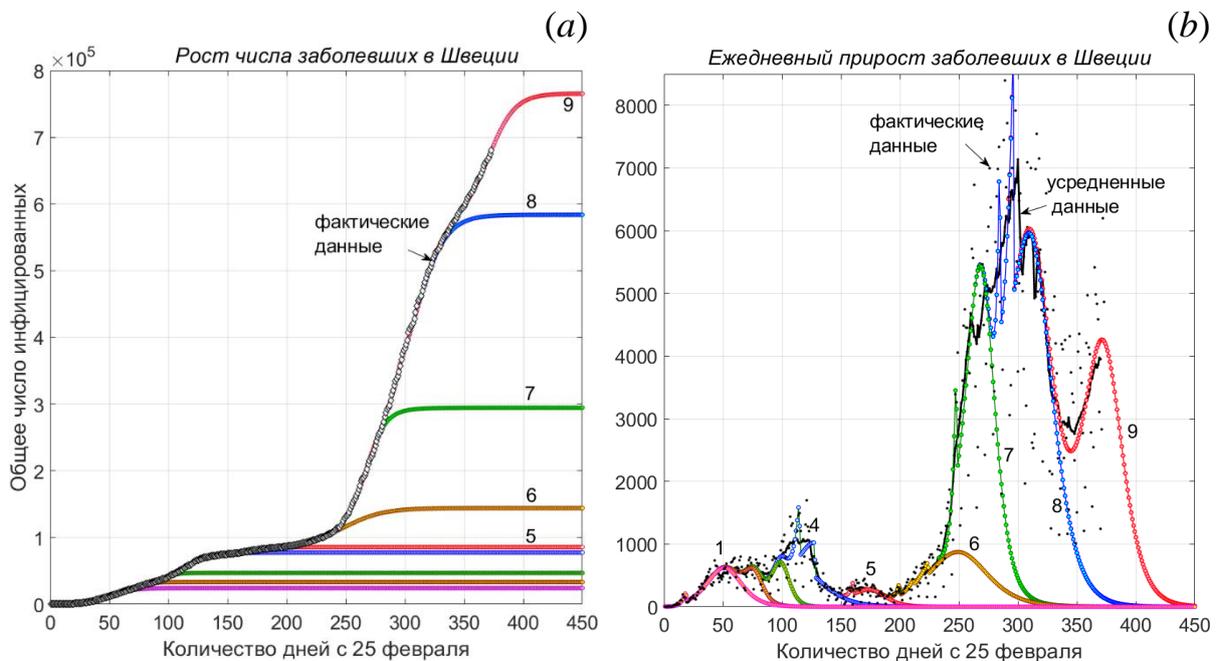


Рис. 9 Динамика распространения эпидемии коронавируса в Швеции между 25.02.2020 и 03.03.2021. Фактические данные и результаты моделирования. Цифрами обозначены расчетные волны в модели

Чтобы корректно описать в модели длинные плато и локальные максимумы весенней вспышки заболевания коронавирусом 2020 г. в Швеции, мы использовали суперпозицию пяти локальных волн с суммарной емкостью 1 млн 780 тысяч человек. Последняя локальная волна, относящаяся к весенней вспышке, сформировалась во второй половине июля и закончилась в конце сентября. Весенняя вспышка заболевания коронавирусом в Швеции протекала дольше и тяжелее и унесла больше жизней, чем в соседних европейских странах, в которых проводились

карантинные мероприятия. По расчетам она дала около 86 тысяч заболевших.

Осенне-зимняя вспышка заболевания коронавирусом в Швеции оказалась намного мощнее, чем весенне-летняя, как и в других европейских странах. Она начала нарастать в Швеции во второй половине сентября, а своего абсолютного максимума достигла 23 декабря со средним значением по 7 дням в 7136 заразившихся в сутки. В модели мы последовательно ввели 3 локальные волны (седьмую, восьмую и девятую) для описания осенне-зимней вспышки, но можно было ввести и большее количество локальных волн для более точного описания всех локальных максимумов. В середине января 2021 г. эпидемия в Швеции пошла на спад, но он длился недолго. Уже во второй декаде февраля ежедневные приросты в среднем начали расти. Средние приросты увеличились с 2796 заболевших в сутки 8 февраля до 3962 человек в пике 4 марта. В модели мы ввели девятую локальную волну для описания этой вспышки.

Седьмая волна закончилась в середине января 2021 г., она принесла 150 тыс. заболевших, восьмая по расчетам закончится в середине апреля 2021 г. Она добавит 290 тыс. заболевших. Девятая волна закончится в начале июня, она добавит еще 182 тысячи заболевших, а всего девять волн в сумме будут иметь примерно 765 тысяч заболевших. Пойдет ли эпидемия на спад, будет ли долго осциллировать в окрестности пика, или начнет снова разгораться, сегодня спрогнозировать невозможно.

Мир в целом. На рис. 10 а), б) представлены фактические данные и результаты расчетов по математической модели для общего числа инфицированных COVID-19 во всем Мире. Распространение эпидемии в Мире представляет собой суперпозицию многих эпидемиологических волн, идущих из разных очагов, расположенных в разных его частях. Из анализа статистических данных даже на беглый взгляд видно, что распространение коронавируса в Мире происходило в виде трех глобальных волн, поднимающих эпидемию на новый уровень. Первая глобальная волна подняла скорость заражения до 90 тысяч человек в сутки. Ее пик пришелся на 10 апреля 2020 г. со значением 90180 заболевших в день (фактические данные). Далее установилось плато, которое продержалось больше месяца до середины мая. Средний прирост инфицированных в день в этот период составил 80669 человек. В середине мая началась новая глобальная волна эпидемии, и суточные приросты стали расти. Они достигли пиковых значений - примерно в 290 тысяч зараженных за сутки - в самом конце июля. Далее эпидемия не пошла на спад, а установилось плато, которое продержалось полтора месяца, до середины сентября 2020 г. Ежедневный прирост заразившихся в этот период не превысил пикового значения, а средний прирост в день составил порядка 265 тысяч.

В сентябре началась новая глобальная осенне-зимняя вспышка эпидемии коронавируса в Мире, которая оказалась намного мощнее первых двух. Средние ежедневные приросты сначала медленно, а потом все быстрее начинали расти. Пиковых значений они достигли в середине ноября (примерно 590 тысяч заболевших в сутки). Два дальнейших месяца эпидемия развивалась медленно. Абсолютный максимум наблюдался 11 января 2021 г. со средним значением 745 тыс. человек в сутки. Далее эпидемия пошла на спад, который продолжался примерно 40 дней. Ежедневные приросты упали до значения 360 тыс. человек в сутки. В последние три недели наблюдается медленный подъем новой волны эпидемии коронавируса в Мире.

В нашей модели первая глобальная вспышка коронавируса в Мире хорошо описывается суммой четырех локальных волн, вторая глобальная летняя вспышка – одной волной (пятой локальной волной), третья осенне-зимняя вспышка – двумя локальными волнами (шестой и седьмой по счету) (см. рис. 10). Емкости волн, даты их начал и даты их окончаний, а также общее число заразившихся, которые дают эти волны, указаны в Таблице 1.

Таблица 1

Значения параметров локальных волн, описывающих распространение эпидемии коронавируса в Мире

	1	2	3	4	5	6	7
	волна	волна	волна	волна	волна	волна	волна
Емк. N	0.5 млн	0.25 млн	10 млн	90 млн	1.3 млрд	2.6 млрд	1 млрд
Дата начала	Декабрь 2019	Конец января 2020	Февраль 2020	Середина марта 2020	Конец апреля 2020	Август 2020	Октябрь 2020
Дата окончания	Конец февраля 2020	Начало марта 2020	Середина мая 2020	Конец июля 2020	Конец марта 2021	Август 2021	Август 2021
число заболев.	54 тысяч	27.4 млн	1.232 млн	3.876 млн	26.1 млн	55.2 млн	29.6 млн

Первые две волны описывают распространение эпидемии в Китае. В начале марта они в сумме дали около 81 тысячи зараженных, то есть чуть больше (на 500 человек), чем в это время было инфицированных в Китае.

Третья и четвертая волны главным образом описывают весеннюю вспышку заболевания коронавирусом в Европе, США и России. Третья волна, сформировавшаяся в середине февраля и закончившаяся в середине мая 2020 г., принесла миру 1 млн 232 тыс. заболевших. Пик четвертой волны пришелся на начало мая, а окончательно затухла она в конце июля. Четвертая волна по расчетам принесла 3.876 млн заболевших.

Пятая расчетная волна описывает вторую глобальную вспышку заболевания коронавирусом, когда эпидемия проникла и стала быстро

распространяться в странах Латинской Америки (Бразилии, Перу, Чили и др.), в странах Южной Азии (Индии, Пакистане, Бангладеш и др.) и Африке. В эту волну также большой вклад внесла летняя вспышка заболевания в США, связанная с демонстрациями протестующих. Пик суточного прироста числа инфицированных пятой волны пришелся по модели на 3 августа и составил около 268 тысяч человек. Окончание пятой волны по расчетам приходится на конец марта 2021. 5-я волна добавит примерно 26.1 млн инфицированных людей.

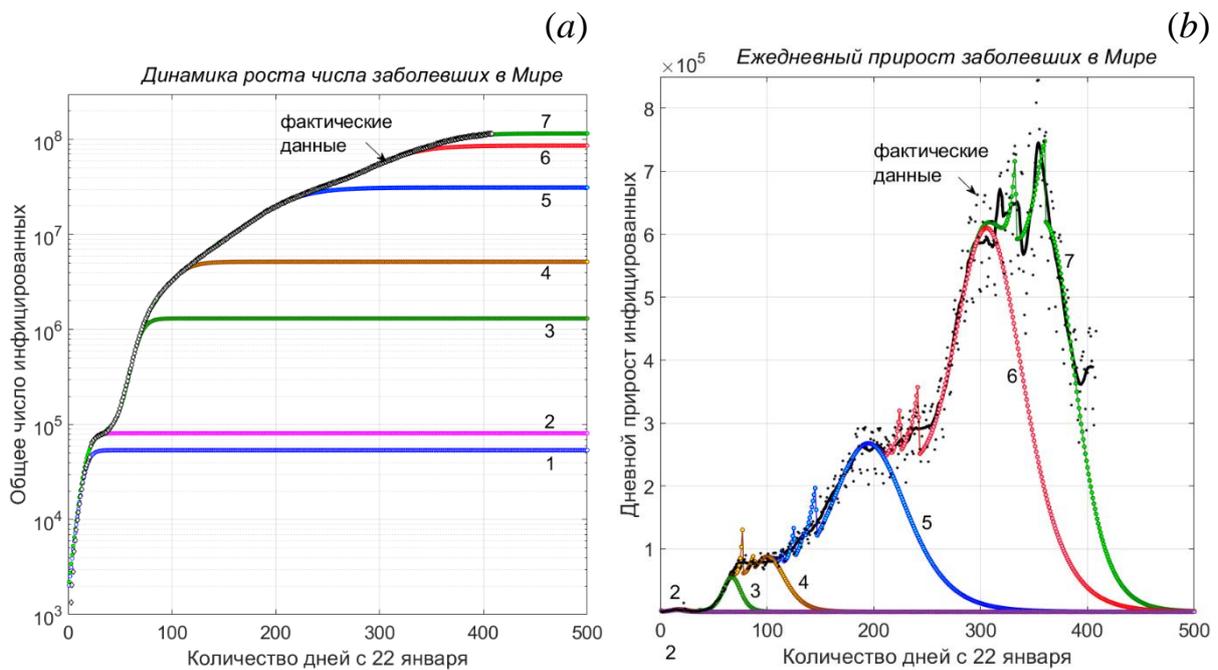


Рис. 10 Динамика распространения эпидемии коронавируса в Мире между 22.01.2020 и 03.03.2021. Фактические данные и результаты моделирования. *b* – линейная шкала; *a* – логарифмическая шкала

Шестая расчетная волна, описывающая осеннюю глобальную вспышку роста заболеваемости коронавирусом в Мире, охватывает распространение эпидемии во многих странах, в том числе ускорившийся рост заболеваемости в Европе, России и США. В октябре мы рассматривали два сценария развития шестой волны с емкостью $N = 2$ млрд и $N = 3$ млрд соответственно. Оба они хорошо описывали фактические данные. Во второй половине ноября стало ясно, что реализовался промежуточный сценарий с емкостью $N = 2.6$ млрд. Пик шестой волны был пройден 20 ноября 2020 г. По расчетам она закончится в августе 2021 г. и принесет 55.2 млн заболевших.

Седьмая волна начала формироваться в октябре. Она описывает широкий пик с несколькими локальными максимумами зимней вспышки эпидемии (рис. 10 б)). Абсолютный максимум 7-ой волны по модели наблюдался 14 января 2021 г. со значением 747.7 тыс. заболевших в день,

и это хорошо совпадает со статистическими данными, которые имели пик 12 января со значением 741.4 чел. в сутки (среднее значение). Седьмая волна сейчас находится на спаде. Она закончится в августе 2021 г. и добавит еще 29.6 млн заболевших.

Вот уже третью неделю подряд растут ежедневные приросты заболевших в Мире, и седьмая волна начала отставать от статистических данных. Это означает, что начала набирать силу новая вспышка эпидемии. На сколько мощной она будет, покажет время. В модели она будет описана восьмой локальной волной, параметры которой можно будет определить чуть позже.

Заключение

Вот уже более года человечество ведет наблюдение за распространением коронавируса COVID-19 в разных странах и регионах мира, собирает данные и пытается строить прогнозы на основе математических моделей. Динамика распространения эпидемии – очень сложная. Сглаженные графики эпидемических кривых во многих странах и регионах имеют несколько локальных максимумов и содержат плато. Это связано с тем, что города, страны, регионы являются открытыми неоднородными системами, в которых время от времени случайно в том или ином месте возникают новые источники заражения, запускающие новые цепочки передачи инфекции от зараженных людей к восприимчивым. Если они охватывают большое количество людей и сильно сдвинуты по времени, то они проявляются как отдельные волны эпидемии, распространяющиеся по-разному в неоднородной среде. В общем случае статистические данные, собранные в виде кумулятивных и эпидемических кривых, представляют собой суперпозицию многих различных локальных волн эпидемии.

Кроме того, при моделировании надо иметь в виду, что статистические данные имеют большую погрешность. Считается, что неточность может составлять несколько десятков процентов. Нахождение параметров модели по неточно заданным статистическим данным является обратной, некорректно поставленной задачей, которая имеет неединственное решение.

Для моделирования столь сложной динамики распространения инфекции в неоднородных открытых системах и при неточно заданных статистических данных мы использовали относительно простое дискретное логистическое уравнение Фейгенбаума, которое в определенном диапазоне изменения параметров описывает логистический рост популяции в среде с ограниченными ресурсами. Это уравнение имеет всего два параметра: емкость системы и коэффициент скорости роста заболевших. Оба они влияют на окончательную общую численность инфицированных. Они подбираются так, чтобы наилучшим образом

описать фактические статистические данные. Коэффициент скорости роста является переменной величиной. Первое его значение подбирается на стадии экспоненциального роста. В дальнейшем он может уменьшаться, например, при введении ограничительных мер. Емкость системы окончательно определяется в пике эпидемии, до этого из разных соображений выбирается несколько значений емкости и рассматривается несколько сценариев развития эпидемии.

Мы разработали методику выявления локальных волн и подбора параметров для их описания в модели. Локальные волны добавляются в модель последовательно со временем, когда расчетные данные начинают отставать от фактических. Разница между расчетными и статистическими данными дает последовательность значений для новой локальной волны, по которой определяются параметры.

Важно отметить, что локальные волны зачастую затухают медленно, и их влияние на общую картину распространения инфекции сказывается долго. Это означает, что моделирование развития эпидемии надо начинать с самого начала. Нельзя начать описывать в модели не первую волну эпидемии, если предыдущая не затухла окончательно.

Мы проводили моделирование и делали прогнозы по динамике распространения эпидемии в разных странах и регионах, начиная с марта 2020 года [41-44] по конец февраля 2021 года. В настоящей работе мы представили результаты моделирования распространения эпидемии коронавируса в Бразилии, Индии, США, Японии, Израиле, Испании, Швеции, в мегаполисе Москве и во всем Мире. Так, математическое описание роста общего числа инфицированных коронавирусом в Индии хорошо уложилось в одну волну, в Бразилии – в 3 волны, Израиле – в 4 волны, а в Японии и в Испании – в 6 волн, в Швеции – в 9 волн, в США и в Москве – в 11 локальных волн, в Мире в целом – в 7 волн. По каждой волне мы делали прогнозы на каждый день и уточняли параметры волны до пика. Сразу после прохождения пика окончательно уточнялась емкость волны при необходимости. После пика делался прогноз по волне, рассчитывалось, когда волна закончится и сколько заболевших она покажет.

Показано, что выбранная модель адекватно описывает динамику распространения инфекций в условиях открытости и неопределенности. С помощью нее можно делать реалистичные прогнозы. Горизонт прогноза зависит от степени замкнутости и однородности системы. Так, в Китае, Японии и Израиле, где во время первой волны были введены жесткие ограничительные меры, прогнозы давали ошибку меньше 1 % в течение полутора-двух месяцев. Хорошо совпали теоретические, расчетные и статистические данные. В Индии и Бразилии, где распространение инфекции долгое время описывалось с помощью одной волны в модели,

прогнозы тоже были длительными. В среднем прогнозы по модели работали 10-14 дней, в худшем случае – неделю. Проблема состоит не в точности модели, а в открытости и непредсказуемости системы.

Эпидемия заражения коронавирусом не заканчивается. Если в одних странах она пошла на спад, в других – наблюдается новый всплеск заболеваемости. Общий прирост числа заболевших в Мире последние три недели демонстрирует рост. Возможно – это является началом новой вспышки заболевания коронавирусом. Чтобы держать ситуацию под контролем, необходимо делать прогнозы на основе модели, сравнивать расчетные данные со статистическими, корректировать параметры модели или вводить новые локальные волны и делать прогнозы снова. На основании прогнозов принимать решение об усилении или ослаблении ограничительных мер.

Литература

1. *M.J. Keeling, P. Rohani* Modeling infectious diseases in humans and animals. / NJ: Princeton University Press, 2008. 408 pp.
2. *J.D. Murray* Mathematical Biology: An introduction Springer. 2002.
3. *W. Liu, H.W. Hethcote, S.A. Levin* Dynamical behavior of epidemiological models with nonlinear incidence rates //Journal of mathematical biology. – 1987. – Т. 25. – №. 4. – С. 359-380.
4. Infectious diseases of humans: Dynamics and control. By R.M. Anderson and R.M. May, 757 pp. Oxford: Oxford University Press, 1991.
5. *W.H. Hamer* Epidemic disease in England. Lancet I, 733- 739. (1906)
6. *W. Liu, S. Tang, Y. Xiao* Model selection and evaluation based on emerging infectious disease data sets including A/H1N1 and Ebola //Computational and mathematical methods in medicine. – 2015. – Т. 2015.
7. *Y.H. Hsieh, J.Y. Lee, H.L. Chang* SARS epidemiology modeling //Emerging infectious diseases. – 2004. – Т. 10. – №. 6. – С. 1165.
8. *S. Kingsland* The refractory model: The logistic curve and the history of population ecology //The Quarterly Review of Biology. – 1982. – Т. 57. – №. 1. – С. 29-52.
9. *H.W. Hethcote* The mathematics of infectious diseases //SIAM review. – 2000. – Т. 42. – №. 4. – С. 599-653
10. *W.O. Kermack, A.G. McKendrick* A contribution to the mathematical theory of epidemics. Proceedings of the Royal Society of London Series A, Containing Papers of a Mathematical and Physical Character 1927; 115(772):700–21.

11. *M. Y. Li, J.S. Muldowney* Global stability for the SEIR model in epidemiology //Mathematical biosciences. – 1995. – T. 125. – №. 2. – C. 155-164.
12. *Gabriele Martelloni, Gianluca Martelloni* Analysis of the evolution of the Sars-Cov-2 in Italy, the role of the asymptomatics and the success of Logistic model// Chaos, Solitons and Fractals 140 (2020) 110150
13. *P. Wang, X. Zheng, J. Li, B. Zhu* Prediction of epidemic trends in COVID-19 with logistic model and machine learning technics // Chaos, Solitons and Fractals 139 (2020) 110058 doi.org/10.1016/j.chaos.2020.110058.
14. *J. Guo* Theoretical Epidemic Laws Based on Data of COVID-19 Pandemic //medRxiv. – 2020.
15. *Ke Wu, D. Darcet, Q. Wang, D. Sornette* Generalized logistic growth modeling of the COVID-19 outbreak in 29 provinces in China and in the rest of the world//
medRxiv preprint doi: <https://doi.org/10.1101/2020.03.11.20034363>
16. *N.S. Barlow, S.J. Weinstein* Accurate closed-form solution of the SIR epidemic model // Physica D 408 (2020) 132540
17. *M. Cadoni, G. Gaeta* Size and timescale of epidemics in the SIR framework// Physica D 411 (2020) 132626
18. *T.T. Marinov, R.S. Marinova* Dynamics of COVID-19 using inverse problem for coefficient identification in SIR epidemic models // Chaos, Solitons & Fractals: X 5 (2020) 100041
19. *T. Odagaki* Analysis of the outbreak of COVID-19 in Japan by SIQR model //Infectious Disease Modelling 5 (2020) 691e698
20. *A.G.M. Neves, G. Guerrero* Predicting the evolution of the COVID-19 epidemic with the A-SIR model: Lombardy, Italy and São Paulo state, Brazil// Physica D 413 (2020) 132693
21. *A. Comunian, R. Gaburro, M. Giudici* Inversion of a SIR-based model: A critical analysis about the application to COVID-19 epidemic // Physica D 413 (2020) 132674
22. *M. Agrawal, M. Kanitkar, M. Vidyasagar* SUTRA: An Approach to Modelling Pandemics with Asymptomatic Patients, and Applications to COVID-19. <https://arxiv.org/abs/2101.09158>
23. *Y. Ding, L. Gao* An evaluation of COVID-19 in Italy: A data-driven modeling Analysis// Infectious Disease Modelling 5 (2020) 495-501

24. *A.M. Ramos, M.R. Ferrández, M. Vela-Pérez, A.B. Kubik, B. Ivorra* A simple but complex enough θ -SIR type model to be used with COVID-19 real data. Application to the case of Italy// *Physica D* 412 (2020) 132839
25. *C. H. Nkwayep et al.* Short-term forecasts of the COVID-19 pandemic: a study case of Cameroon // *Chaos, Solitons & Fractals*. – 2020. – T. 140. – C. 110106.
26. *H.P.C. Cintra, F.N. Fontinele* Estimative of real number of infections by COVID-19 in Brazil and possible scenarios // *Infectious Disease Modelling*. – 2020. – T. 5. – C. 720-736.
27. *S. Contreras, H.A. Villavicencio et al* A multi-group SEIRA model for the spread of COVID-19 among heterogeneous populations // *Chaos, Solitons and Fractals* 136 (2020) 109925
28. *U. Avila-Ponce de León, Á.G.C. Pérez, E. Avila-Vales* An SEIARD epidemic model for COVID-19 in Mexico: Mathematical analysis and state-level forecast// *Chaos, Solitons and Fractals* 140 (2020) 110165
29. *M. Higazy* Novel fractional order SIDARTHE mathematical model of COVID-19 pandemic// *Chaos, Solitons and Fractals* 138 (2020) 110007
30. *P. Wang, X. Zheng, G. Ai, D. Liu, B. Zhu* Time series prediction for the epidemic trends of COVID-19 using the improved LSTM deep learning method: Case studies in Russia, Peru and Iran// *Chaos, Solitons and Fractals* 140 (2020) 110214.
31. *B. Ghanbari* On forecasting the spread of the COVID-19 in Iran: The second wave // *Chaos, Solitons and Fractals* 140 (2020) 110176
32. *L.F.S. Scabini, L.C. Ribas, M.B. Neiva et al.* Social interaction layers in complex networks for the dynamical epidemic modeling of COVID-19 in Brazil // *Physica A* 564 (2021) 125498.
33. *K. Yew Ng, M. Mei Gui* COVID-19: Development of a robust mathematical model and simulation package with consideration for ageing population and time delay for control action and susceptibility// *Physica D* 411 (2020) 132599
34. *U. Sahin, T. Sahin* Forecasting the cumulative number of confirmed cases of COVID-19 in Italy, UK and USA using fractional nonlinear grey Bernoulli model// *Chaos, Solitons and Fractals* 138 (2020) 109948
35. *B.S. T. Alkahtani, S.S. Alzaid* A novel mathematics model of covid-19 with fractional derivative. Stability and numerical analysis // *Chaos, Solitons and Fractals* 138 (2020) 110006
36. *R. May* Simple Mathematical Models With Very Complicated Dynamics // *Nature*, V.261, 1976, pp. 459-467

37. *M.J. Feigenbaum* The universal metric properties of nonlinear transformations // *Journal of Statistical Physics.* — 1979. — Vol. 21. — P. 669-706.
38. <https://www.calcalistech.com/ctech/articles/0,7340,L-3800632,00.html>
39. <https://www.worldometers.info/coronavirus/>
40. <https://coronavirus-monitor.ru/coronavirus-v-rossii>
41. *Koltsova E.M, Kurkina E.S, Vasetsky A.M*, Mathematical Modeling of COVID-19 coronavirus spread in Moscow// *Computational nanotechnology*, Vol 7. № 1 (2020) 99-105.
42. *Koltsova E.M, Kurkina E.S, Vasetsky A.M*, Mathematical Modeling of COVID-19 coronavirus spread in several European, Asian countries, Israel and Russia, *Problems of economics and law practice*. 2020. Moscow. №. 2. p.154-165.
43. *Koltsova E.M, Kurkina E.S, Vasetsky A.M*, Mathematical Modeling of the Spread of COVID-19 in Moscow and Russian Regions, arXiv preprint arXiv:2004.10118, Submitted on 21 Apr 2020.
44. *Koltsova E.M, Kurkina E.S, Vasetsky A.M*, Mathematical modeling of the spread of waves of the COVID-19 coronavirus epidemic in different regions of the world, arXiv preprint arXiv:2007.02283, submitted 5 Jul 2020.