

УДК 519.233.5:578.7

## МОДЕЛИРОВАНИЕ ЭПИДЕМИИ КОРОНАВИРУСА КАК ФАЗОВОГО ПЕРЕХОДА

© 2020 г. В. Г. Суховольский<sup>1, 2, \*</sup>, А. В. Ковалев<sup>2, \*\*</sup>

<sup>1</sup>Институт леса им. В.Н. Сукачева СО РАН  
Академгородок, 50/28, Красноярск, 660036 Россия

<sup>2</sup>Федеральный исследовательский центр КНЦ СО РАН  
Академгородок, 50, Красноярск, 660036 Россия

\*E-mail: soukhovolsky@yandex.ru

\*\*E-mail: sunhi.prime@gmail.com

Поступила в редакцию 27.04.2020 г.

После доработки 19.06.2020 г.

Принята к публикации 25.06.2020 г.

Рассмотрена задача моделирования развития пандемии коронавируса на планете. Проведен анализ, направленный на выявление трех типов закономерностей в развитии коронавируса: существование критических плотностей заболевания, особенности распространения заболевания по странам и территориям в отдельной стране, характер относительной динамики заболеваемости в отдельных странах. Для моделирования предложено использовать аналоги моделей фазовых переходов первого и второго рода в физических системах и авторегрессионные модели развития заболевания. Выполнены расчеты критического числа заболевших, по достижению которого начинается экспоненциальная стадия развития заболевания. Предложена оценка скорости развития заболевания в разных странах и на разных территориях.

DOI: 10.31857/S0044459620050073

Существует большое число моделей развития инфекционных заболеваний, в которых с помощью систем дифференциальных уравнений (непрерывное время) или разностных уравнений (дискретное время) описывается динамика групп населения с разными характеристиками по отношению к инфекции (Allman, Rhodes, 2003; Edelstein-Keshet, 2005; Brauer et al., 2008; Iacoviello, Lizzi, 2008; Bolzoni et al., 2017; Bonyah et al., 2017; Zakary et al., 2017). В простейшей SIR-модели, используемой уже в течение почти 100 лет, вся популяция делится на три группы: восприимчивые (susceptible –  $S$ ), инфицированные (infected –  $I$ ), выздоровевшие (recovered –  $R$ ), – размеры которых подчиняются уравнениям (Kermack, McKendrick, 1927):

$$\begin{cases} \frac{dS(t)}{dt} = -\beta \frac{S(t)I(t)}{N}, \\ \frac{dI(t)}{dt} = \beta \frac{S(t)I(t)}{N} - \gamma I(t), \\ \frac{dR(t)}{dt} = \gamma I(t), \end{cases} \quad (1)$$

где  $\beta$ ,  $\gamma$  – константы, характеризующие скорость переходов из состояния в состояние. Размер популяции считается постоянным во время вспыш-

ки заболевания, поэтому в любой момент  $t$  численность популяции равна  $N$ .

Для верификации SIR-модели необходимо оценить по данным медицинской статистики значения ее двух свободных параметров  $\beta$  и  $\gamma$  и начальных условий.

Более детальные модели SIS, SIRS, SEIRD, SEIIFR – это модели того же типа, что и SIR-модель, однако включающие большее по сравнению с SIR-моделью число состояний. Так, в MSEIR-модели популяция делится на пять классов: с пассивным иммунитетом (maternally-derived immunity), восприимчивые, латентные (exposed), инфицированные, невосприимчивые (removed) (Hethcote, 2000). Математическая модель распространения лихорадки Эбола среди населения разделяет население на шесть классов: восприимчивые, латентные, инфицированные, госпитализированные (hospitalized), умершие и непохороненные (funeral), невосприимчивые или иммунные (Legrand et al., 2007). Для моделирования динамики заболевания необходимо параметризовать правые части уравнений и ввести значения параметров и начальные значения численности особей в каждом из состояний. Однако при расширении числа постулируемых состояний очень трудно получить

по данным медицинской статистики информацию о динамике численности соответствующих групп пациентов.

Модели развития эпидемии коронавируса в разных странах предложены многими авторами (Chen et al., 2020; Distant et al., 2020; Gomez et al., 2020; Hackl, 2020; Jin et al., 2020; Ediriweera et al., 2020; Lyra et al., 2020; Nishiura et al., 2020; Shen, 2020; Tuomisto et al., 2020; Zhao et al., 2020; Wu, McGoogan, 2020; Yue et al., 2020). В большинстве этих работ использованы SIR- и производные от нее модели со всеми сопутствующими факторами неопределенности, в частности, в оценке скоростей переходов из состояния в состояние, в учете влияния внешних факторов и в том, как описывать циклические вспышки инфекции.

Поскольку на начальных этапах развития эпидемии число  $S$  восприимчивых много больше числа  $I$  заболевших (миллионы и десятки миллионов против десятков тысяч) и  $S$  убывает очень медленно, на начальном этапе заболевания величину  $S$  можно считать константой. Если к тому же скорость выздоровления меньше скорости заражения, то SIR-модель развития эпидемии на ее начальном этапе можно переписать в следующем виде:

$$\frac{dI}{dt} = mI \quad \text{или} \quad I(t) = I_0 \exp(mt), \quad (2)$$

где  $m = \beta S = \text{const}$ .

Однако модель (2) не описывает ни лаг-фазу эпидемии, в ходе которой число заболевших практически не изменяется во времени, ни процесс перелома динамики числа заболевших (например, вследствие карантинных мероприятий), ни ее затухание. В настоящей работе мы пытаемся получить более общую модель динамики заболевания, в которой удалось обойти названные трудности SIR-моделей и их модификаций.

## ОБЪЕКТЫ И МЕТОДЫ ИССЛЕДОВАНИЯ

### *Объекты исследования*

Для анализа доступны данные о заболеваемости коронавирусом в любой стране мира. Однако в связи с ограниченностью времени мы провели анализ лишь для ряда стран: России, Италии, США (и отдельно для Нью-Йорка), Ирана (как страны с сильной эпидемией коронавируса), Швеции (как страны, в которой не был введен карантин) и Норвегии (страны с карантином в сопоставлении со Швецией), Вьетнама (как страны с отличными от стран Европы и Северной Америки климатическими условиями). Для расчетов использовались данные с начала января по середину апреля 2020 г., взятые с сайтов ВОЗ (WHO, 2020), Роспотребнадзора (2020) и Университета Дж. Хопкинса (Data Repository..., 2020).

### *Методы моделирования*

Модель SIR и производные от нее модели можно характеризовать как беспороговые. Согласно этим моделям, эпидемия может начаться после того, как заболит один пациент. Однако такое состояние популяции будет крайне неустойчивым, и эпидемии в этом случае будут возникать очень часто. В альтернативном подходе к моделированию эпидемии рассматривается возможность существования порога числа заболевших, по достижению которого начинается эпидемия. Тогда важным показателем риска эпидемии будет величина критической пороговой численности заболевших. Пороговую модель эпидемии будем рассматривать как аналог таких процессов в физических системах, как фазовый переход первого рода. Характерным примером фазового перехода первого рода является, например, кипение жидкости, происходящее при достижении критической температуры (для воды это  $100^\circ\text{C}$  при нормальном атмосферном давлении).

При описании пороговой модели эпидемии будем рассматривать процесс ее развития как реализацию во времени состояний  $x$  (например, числа зараженных) в некоторой взаимодействующей системе (стране или городе) и введем понятие функции  $F(x)$  состояния системы. Функцию  $F(x)$  определим как величину, обратную функции плотности распределения  $f(x)$  значений  $x$  в течение длительного периода  $T$ , достаточного для того, чтобы все возможные состояния системы могли реализоваться. Если вероятность реализации некоторого состояния  $x(m)$  за период  $T$  мала, то значение функции  $F(x)$  будет велико. Если в течение периода  $T$  состояние  $x(m)$  реализуется достаточно часто, то значение  $F(x)$  мало. Функция  $F(x)$  потенциально зависит от большого числа разнообразных факторов, и записать зависимости  $F(x)$  от этих факторов затруднительно. Однако  $F(x)$  можно классифицировать в зависимости от числа локальных минимумов и максимумов. Функцию  $F(x)$  с одним глобальным минимумом  $F(x_1)$  можно характеризовать как функцию с одним устойчивым состоянием  $x_1$ . Функцию с двумя локальными минимумами  $F(x_1)$  и  $F(x_2)$  и одним локальным максимумом  $F(x_r)$  между локальными минимумами можно характеризовать как бистабильную функцию с двумя возможными состояниями. Если значение  $F(x_r)$  велико, это означает, что состояние  $x_r$  встречается очень редко и система перескакивает из состояния  $x_1$  в  $x_2$  и обратно. Поведение системы такого рода хорошо известно для физических систем и носит название *фазового перехода первого рода* (Ландау, Лифшиц, 1964).

Выявить внешние факторы и признаки системы, для которой возможны фазовые переходы первого рода, достаточно трудно. Однако для описания фазовых переходов можно использовать подход,

предложенный Л. Д. Ландау (1937), согласно которому функция состояния  $F_1(q)$  зависит от некоторой переменной — *параметра порядка*  $q$  ( $0 \leq q \leq 1$ ), характеризующего интегральные свойства системы, — и  $F_1(q)$  можно выразить как степенной ряд по  $q$ :

$$F_1(q) = F_{00} + aq^2 + cq^3 + bq^4 + \dots \quad (3)$$

Если  $q < 1$ , то члены с высокими степенями  $q$  будут малы и их можно отбросить. Тогда локальные минимумы функции  $F_1(q)$  можно найти из обычных условий  $\frac{dF_1}{dq} = 0$ ;  $\frac{d^2F_1}{dq^2} > 0$ .

Для фазовых переходов первого рода функция состояния хорошо описывается полиномом шестой степени (Брус, Каули, 1984).

Если заболевание характеризуется существованием двух состояний системы: состоянием с низкой заболеваемостью  $x_1$  и состоянием с высокой заболеваемостью  $x_2$  — и может быть классифицировано как фазовый переход первого рода, то переход из состояния  $x_1$  в состояние  $x_2$  происходит при достижении критического состояния  $x_c$ .

Если время, в течение которого происходит фазовый переход, мало, то состояния вблизи значения  $x_c$  будут встречаться редко, и функция  $F_1(x)$  вблизи  $x_c$  будет иметь максимум. Тогда состояние системы, близкое к  $x_c$ , можно рассматривать как начало эпидемии заболевания. Вид функции  $F_1(x)$  будем оценивать, используя данные временной динамики заболеваемости.

Важным показателем развития пандемии коронавируса является его регистрация в разных странах. Так как заболевание развивается в разных странах и на разных территориях, в отдельной стране с разной скоростью и в разные периоды времени, необходима модель, описывающая динамику развития заболевания по территории отдельной страны и по планете в целом. Для описания такого процесса в настоящей работе использована модель *фазового перехода второго рода*. Для такого типа фазовых переходов введем функцию состояния  $F_2(w)$ , связывающую некоторый интегральный показатель распространения заболевания (параметр порядка  $w$ ) со временем  $t$ , прошедшем с начала заболевания (с регистрации первого заболевшего). Распространенность заболевания по стране на период времени с начала заболевания будем характеризовать значением *параметра порядка*  $w(t)$  — доли стран или регионов в пределах страны, в которых на момент  $t$  заболевание не наблюдается. Если в стране или в одном регионе страны на момент  $t$  нет заболевших, то  $w(t) = 1$ . Если на момент  $t$  заболевшие есть во всех регионах или во всех странах планеты, то  $w(t) = 0$ . По аналогии с фазовыми переходами второго ро-

да в физических системах зависимость функции состояния  $F_2(w, t)$  от параметра порядка можно описать следующим уравнением:

$$F_2(w) = F_{20} + (t - t_c)w^2 + bw^4. \quad (4)$$

Решение уравнения (4) имеет следующий вид:

$$w^2 = \begin{cases} 0, & t > t_c \\ 1 - Bt, & t \leq t_c \end{cases}. \quad (5)$$

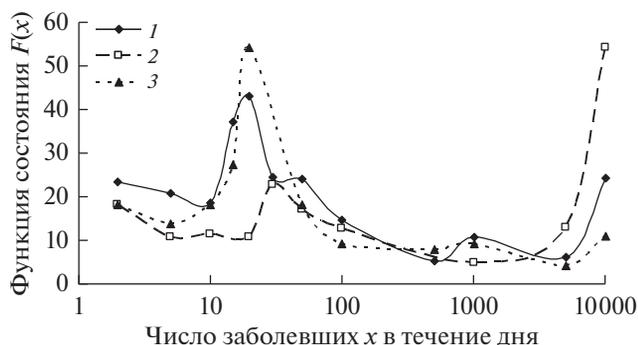
Согласно уравнению (5), изменение квадрата доли регионов, в которых заболевание еще не зарегистрировано, линейно убывает со скоростью  $B$  с момента  $t_0$  начала заболевания. Величина  $t_c$  характеризует дату, по достижению которой во всех странах планеты регистрируются заболевания коронавирусом, и это действительно *пандемия*. Модель (4) легко проверить, используя данные о регистрации первого заболевшего на территории страны или региона.

Для моделирования процессов динамики коронавируса будем использовать данные о текущей суточной заболеваемости  $[I(t) - I(t-1)]$ . Так как численность  $N$  населения в разных странах может очень сильно отличаться — от десятков тысяч до миллиарда жителей, — для получения сопоставимых данных для каждой страны рассматривали показатели относительной текущей заболеваемости в расчете на одного жителя страны

$p(t) = \frac{I(t) - I(t-1)}{N}$ . Для уменьшения дисперсии вместо величины  $p(t)$  используем величину  $\ln p(t)$ .

Развитие коронавируса характеризуется наличием инкубационного периода  $k$  между инфицированием и проявлением видимых признаков заболевания. Для коронавируса этот инкубационный период  $k$  составляет около семи дней (Lauer et al., 2020; Yang et al., 2020). Тогда можно ожидать, что текущая заболеваемость на дату  $t$  будет зависеть от числа заболевших на дату  $(t-k)$ . Процесс развития эпидемии можно характеризовать как авторегрессионный, в котором текущая заболеваемость коррелирует с заболеваемостью за  $k$  дней до того. Влияние продолжительности инкубационного периода на текущую заболеваемость можно оценить с помощью парциальной автокорреляционной функции (PACF) временного ряда заболеваемости, характеризующей запаздывание отклика системы на изменение ее характеристик. В случае наличия запаздывания на кривой PACF будет наблюдаться пик при значении сдвига, близкого к времени инкубационного периода.

Для описания динамики процесса развития заболевания можно записать ADL( $k, 1$ ) (autoregressive distributed lag)-модель (Сток, Уотсон, 2015):



**Рис. 1.** Функция состояния  $F(x)$  для США (1), планеты в целом (2) и России (3) по данным с начала января по середину апреля 2020 г.

$$\ln p(t) = a_0 + \sum_{j=1}^k a_j \ln p(t - j) + bt + \varepsilon(t), \quad (6)$$

где  $p(t)$  — относительная заболеваемость,  $t$  — продолжительность экспоненциальной фазы эпидемии (номер дня с ее начала),  $a_1, a_2, \dots, a_k, b$  — коэффициенты,  $k$  — порядок авторегрессии,  $\varepsilon(t)$  — случайный шум.

При известных суточных данных ВОЗ о числе зараженных коронавирусом уравнение (6) можно рассматривать как регрессионное с известными значениями текущей относительной заболеваемости, относительных заболеваемости в течение  $k$  предыдущих дней и продолжительности времени с начала экспоненциальной стадии эпидемии. Для расчетов коэффициентов регрессионного уравнения необходимо оценить порядок  $k$  авторегрессии. По-видимому, самый простой способ оценки  $k$  — вычисление парциальной автокорреляционной функции моделируемого временного ряда и определение числа членов PACF, значимо отличных от нуля (Бокс, Дженкинс, 1974).

Качество приближения ADL-модели можно определить с помощью двух показателей — степени согласия данных учетов и модельного числа заболевших и величины фазового сдвига между рядом данных учетов и модельным временным рядом. Оценку согласия амплитуд модельных и натуральных данных производили по показателю коэффициента детерминации  $R^2$  регрессионного уравнения (6). Для определения фазового сдвига модельного и натурального рядов заболеваемости возможно вычислить кросс-корреляционную функцию (CCF) модельного временного ряда и ряда натуральных данных. Если максимум взаимной корреляционной функции приходился на сдвиг  $\Delta = 0$ , и при этом значение взаимной корреляционной функции  $f(\Delta = 0)$  близко к 1, это указывало на то, что модельный и натуральный временные ряды изменялись синхронно и сдвиг по фазе между этими рядами не наблюдалось.

## РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

В развитии эпидемии в отдельной стране можно выделить следующие фазы:

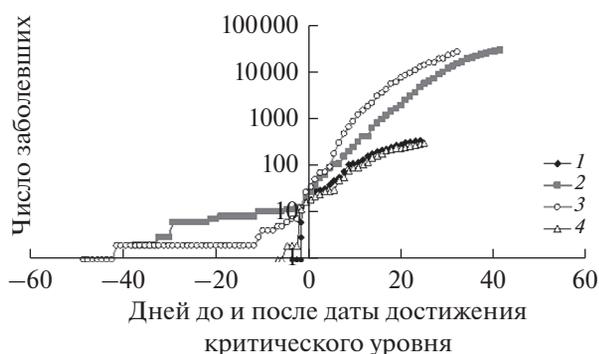
- *лаг-фазу*, когда в течение достаточно продолжительного времени регистрируется стабильно низкое число заболевших (один–два человека в отдельной стране);
- *экспоненциальную фазу*, когда относительное число заболевших (число заболевших в расчете на одного жителя страны) растет со временем по экспоненте;
- *фазу стабилизации и затухания*, когда относительное число заболевших не меняется во времени или начинает уменьшаться.

В настоящей работе проведен анализ, направленный на выявление трех типов закономерностей в развитии коронавируса: существование критических плотностей заболевания, особенности распространения заболевания по странам и территориям в отдельной стране, характер относительной динамики заболеваемости в отдельных странах.

1. На первом этапе анализа рассматривался феномен существования критической плотности заболевших, по достижению которой начиналась экспоненциальная стадия эпидемии. Для этого проводился расчет функции состояния  $F(x)$  для планеты в целом и для таких стран, как Россия и США (рис. 1). Как видно из рис. 1, критическая численность ежедневно заболевших в США, России и в среднем по планете близка к 20–30 пациентам. Если критическое число заболевших существует, то после достижения этого критического значения на территории должен наблюдаться скачок численности ежедневно заболевших и переход в экспоненциальную стадию заболевания. Для разных территорий этот скачок происходит в разное время, но если выровнять данные по дате достижения критической плотности на территории, то можно получить картину изменения заболеваемости, показанную на рис. 2.

Расчеты по развитию заболевания и переход через критическое состояние в различных штатах США показаны в табл. 1. Из нее видно, что при разной численности населения и различном числе заболевших величина критического числа заболевших, по достижению которого начинается экспоненциальная стадия эпидемии, примерно одинакова для всех штатов.

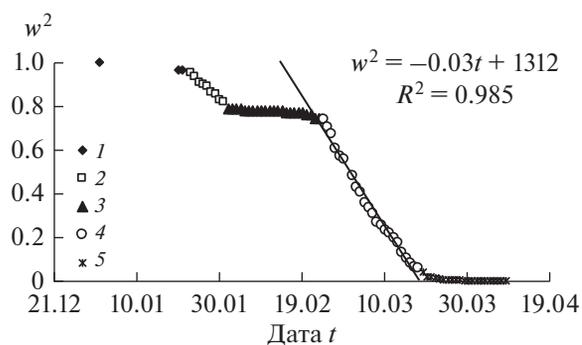
2. Важным показателем, характеризующим развитие пандемии, является пространственная распространенность заболевания по планете. В качестве характеристик распространения коронавируса вводим показатель  $w(t)$  — долю стран из общего числа  $n_0 = 202$  стран в реестре ВОЗ, в которых на момент  $t$  не было зарегистрировано коронавируса. Для описания развития заболевания



**Рис. 2.** Скачки заболеваемости в различных штатах США: 1 – Иллинойс, 2 – Северная Дакота, 3 – Калифорния, 4 – Вайоминг.

на планете в целом и в отдельных странах использована модель фазового перехода второго рода, согласно которой до достижения времени  $t_c$  полного распространения болезни на планете связь между квадратом параметра порядка  $w = \frac{n}{n_0}$  (где  $n$  – число стран, в которых не наблюдалось заболевание) и временем развития заболевания описывается уравнением (5). До начала заболевания коронавирусом в Китае, когда он не наблюдался на планете и  $n = 202$ , значение  $w$  будет равно 1. При нарастании эпидемии, когда заболевание не зарегистрировано в  $m$  ( $m < 202$ ) странах  $w = m/202 < 1$ . Состояние полной пандемии, когда стран без коронавируса нет и  $n = 0$ , будет характеризоваться значением  $w = 0$ .

На рис. 3–5 приведены данные о развитии пандемии на планете в целом, в США и России. Как видно из рис. 3–5, после прохождения начального этапа пандемии скорости  $V$  появления заболеваний в отдельных административных единицах России и в разных странах на планете в целом близки (0.043 и 0.030), тогда как в США заболевание распространялось по территории страны со скоростью более чем в 2 раза большей (0.067). При этом продолжительность стадии пространственного развития пандемии в целом на планете и в России составила 59 дней, а в США – 47 дней. По всей видимости, эти различия связаны с интен-



**Рис. 3.** Распространенность заболевания в целом по планете (по данным ВОЗ; WHO, 2020): 1 – начало пандемии (только одна страна – Китай), 2 – начальная стадия инфекции на планете, 3 – остановка развития инфекционного процесса, 4 – стадия массового инфицирования, 5 – полная пандемия на планете.

сивностью перемещения жителей на разных территориях.

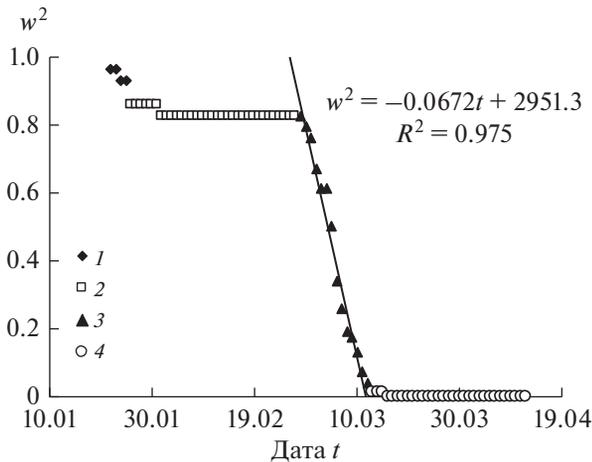
3. Для описания динамики развития заболевания использовали ADL-модель. Так как точность данных медицинской статистики неясна, перед моделированием выполняли высокочастотную фильтрацию данных статистики с помощью фильтра Ганна  $x(t) = 0.24x(t-1) + 0.52x(t) + 0.24x(t+1)$  (Хемминг, 1987). Для оценки порядка авторегрессии и характеристик запаздывания инфекционного процесса были вычислены PACF для Италии (рис. 6) и Калифорнии (рис. 7).

Расчеты парциальной автокорреляционной функции для разных стран дали неожиданный результат: вместо ожидаемого значения порядка авторегрессии  $k \approx 7$  величины PACF оказались значимыми только для  $k = 2$ . Отсюда следует, что существует связь между относительной заболеваемостью в момент  $t$  и относительными заболеваемостями в моменты  $(t-1)$  и  $(t-2)$ . Не пытаясь объяснить этот неожиданный эффект, для описания динамики развития заболевания мы выбрали ADL-модель в следующем виде

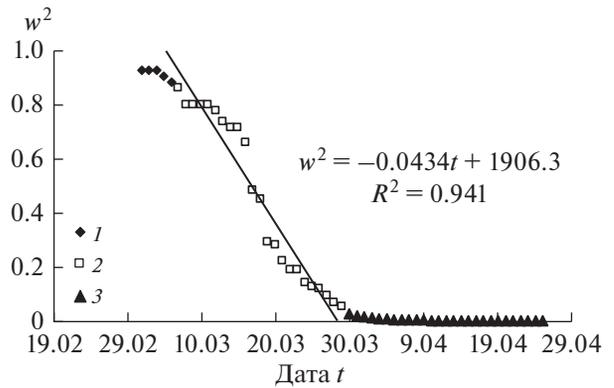
$$\ln p(t) = a_0 + a_1 \ln p(t-j) + a_2 \ln p(t-2) + bt + \varepsilon(t). \quad (7)$$

**Таблица 1.** Оценка критического числа заболевших в отдельных штатах США

Штат	Численность населения	Число заболевших на 08.04.2020	Пороговое число заболевших
Иллинойс	12671821	20852	25
Северная Дакота	762062	308	19
Калифорния	39512223	22795	21
Вайоминг	578759	270	26



**Рис. 4.** Распространенность заболевания по штатам и территориям США (по данным Университета Дж. Хопкинса; Data Repository..., 2020): 1 – начальная стадия эпидемии (три штата), 2 – остановка эпидемии, 3 – стадия массового заражения, 4 – полная эпидемия в США.



**Рис. 5.** Распространенность заболевания по регионам России (по данным Роспотребнадзора (2020)): 1 – начальный этап пандемии, 2 – развитие пандемии в течение марта 2020 г., 3 – полное освоение коронавирусом регионов России.

Исходя из значений известного ряда  $\{\ln p(t)\}$ , коэффициенты модели (7) вычисляли как параметры линейного регрессионного уравнения. На рис. 8 представлены данные о заболеваемости в штате Нью-Йорк и модельные расчеты. Схожие кривые были получены для динамики развития заболевания в России (рис. 9). В отдельных регионах России динамика заболеваемости была различной по абсолютным значениям  $\ln p(t)$ , однако характер развития эпидемии был схож (рис. 10).

В табл. 2 приведены параметры ADL-моделей для ряда стран с различными режимами противоэпидемических ограничений. Так, низкий уровень карантинных мероприятий характерен для Швеции и Бразилии, высокий – для Вьетнама и Норвегии. Нарастание уровня карантина в ходе развития эпидемии коронавируса происходило в Италии, Иране и Японии.

В ряде стран интенсивность заболевания начинает уменьшаться, и этот процесс невозможно описать простой экспоненциальной моделью (2). Так, снижение уровня заболеваемости характерно для Вьетнама, тогда как в Бразилии, стране со схожими климатическими условиями, до последнего времени наблюдался рост заболеваемости (рис. 11). Точно такие же различия – рост заболеваемости в Швеции и его падение в Норвегии – характерны для этих двух также сходных по климату стран.

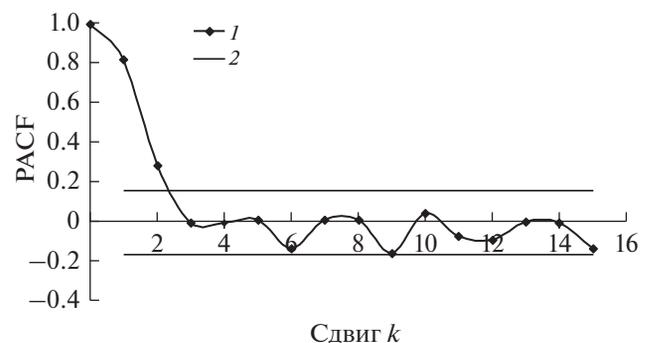
Проведение жестких карантинных мероприятий позволило очень быстро снизить уровень заболеваемости коронавирусом в Республике Корея (рис. 12). Как видно, уже после 29.02.2020 интенсивность заболевания в Южной Корее начала уменьшаться. В табл. 3 приведены расчеты параметров

ADL-модели для описания эпидемии на территории Республики Корея.

Используя рассчитанные значения ADL-модели для отдельной страны, можно на основе этих данных перейти к краткосрочному прогнозу динамики заболеваемости (рис. 13 и 14). Как видно из рис. 13, данные медицинской статистики и расчеты по ADL-модели для России практически совпадают. Такое же согласие характерно для данных и модели по США (рис. 14).

Насколько общей является предложенная модель и можно ли с ее помощью моделировать другие вирусные заболевания? Для проверки возможности модели мы рассмотрели динамику заболевания гриппом штамма H3N2 в европейских странах, используя базу данных Европейского центра профилактики и контроля заболеваний (<https://www.ecdc.europa.eu/en/seasonal-influenza/surveillance-and-disease-data/atlas>).

На рис. 15 приведен график функции состояния для заболевания гриппом H3N2 в Италии с



**Рис. 6.** PACF для Италии по данным с 22.02.2020: 1 – PACF, 2 – доверительный интервал для стандартной ошибки PACF.

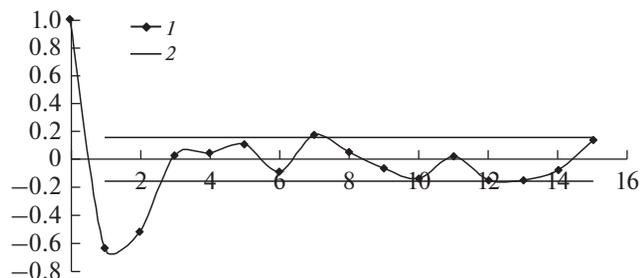


Рис. 7. PACF для Калифорнии по данным с 16.03.2020: 1 – PACF, 2 – доверительный интервал для стандартной ошибки PACF.

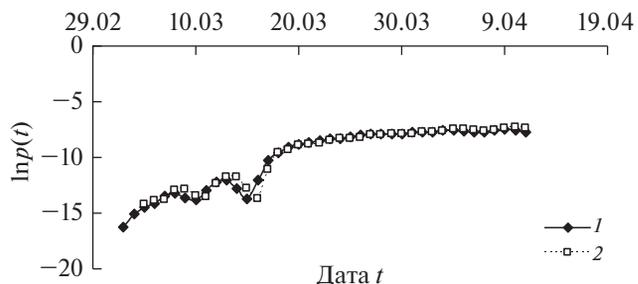


Рис. 8. Развитие заболевания в штате Нью-Йорк: 1 – данные с сайта Университета Дж. Хопкинса (Data Repository..., 2020), 2 – ADL-модель.

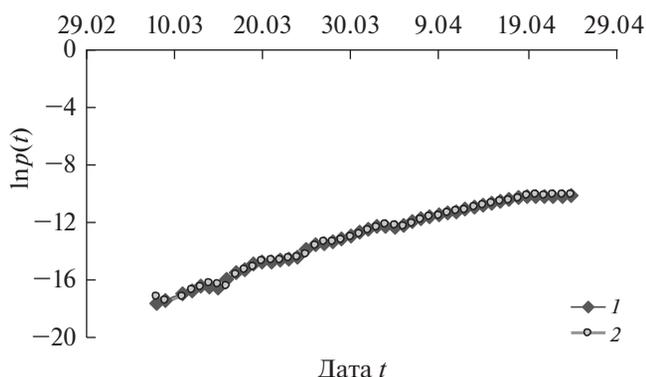


Рис. 9. Развитие заболевания в России: 1 – данные Роспотребнадзора (2020), 2 – ADL-модель.

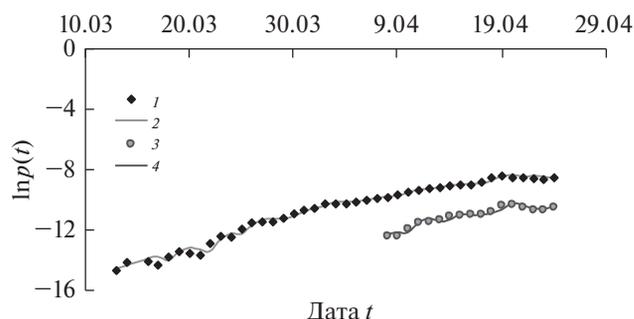


Рис. 10. Развитие заболевания в Москве и Нижегородской области: 1 – данные о заболевании по Москве, 2 – ADL-модель для Москвы, 3 – данные заболевания для Нижегородской области, 4 – ADL-модель для Нижегородской области.

2014 по 2020 год. Как следует из рисунка, критическое число заболевших для Италии составляет 20 человек в неделю.

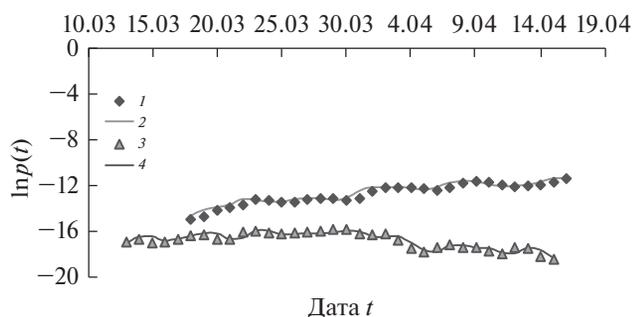
На рис. 16 приведена PACF для вирусного гриппа H3N2 в Италии в 2014–2020 гг. Как видно, парциальная автокорреляционная функция для гриппа H3N2 в Италии характеризуется порядком авторегрессии  $k = 4$ , что несколько больше, чем для коронавируса. Авторегрессионное уравнение, описывающее заболеваемость гриппом в Италии, запишется следующим образом:

$$\ln x(i) = a_0 + \sum_{j=1}^4 a_j \ln x(i - j). \quad (8)$$

Так как значения членов ряда  $\{\ln x(i)\}$  известны, уравнение (8) можно рассматривать как линейное регрессионное уравнение относительно неизвестных коэффициентов  $a_0, \dots, a_4$ . Для расчета коэффициентов использовались стандартные методы регрессионного анализа. Значения коэффициентов уравнения (8) приведены в табл. 4.

Таблица 2. Значения параметров уравнений ADL-модели для ряда стран

Коэффициент	Вьетнам	Италия	Норвегия	Бразилия	Швеция	Иран	Япония
$a_0$	703.34	97.02	55.04	-1209.43	-751.13	36.075	-1129
$b$	-0.016	-0.0022	-0.0013	0.03	0.017	-0.0008	0.03
$a_2$	-0.219	-0.247	-0.338	-0.35	-0.279	-0.4848	-0.35
$a_1$	1.022	1.215	1.229	1.05	0.999	1.450	1.01
$R^2$	0.851	0.989	0.94	0.964	0.973	0.994	0.97
$F$	57.1	1483.5	217.3	231.6	490.5	2689.6	524.6

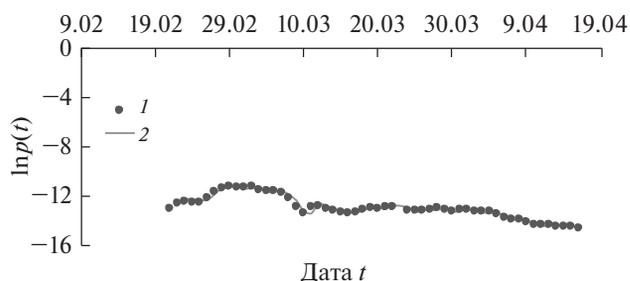


**Рис. 11.** Динамика заболеваемости коронавирусом в Бразилии (1) и Вьетнаме (3), ADL-модели для Бразилии (2) и Вьетнама (4).

На рис. 17 приведены трансформированный временной ряд статистики заболеваемости и модельный временной ряд, построенный по данным из табл. 4. Как видно, имеет место очень хорошее соответствие данных медицинской статистики и модельных расчетов. Аналогичные расчеты были выполнены для заболевания гриппом H3N2 в Великобритании. На рис. 18 приведена PACF для этой страны. В этом случае порядок авторегрессии  $k = 4$  и, значит, для описания динамики заболевания в Великобритании можно использовать модель (7). В табл. 5 приведены параметры модели AR(4) для Великобритании.

На рис. 19 приведены трансформированный временной ряд статистики заболеваемости гриппом в Великобритании и модельный временной ряд, построенный по данным из табл. 5. Расчет кросс-корреляционной функции временных рядов 1 и 2 на рис. 19 показывает, что они синхронны. Значение кросс-корреляционной функции при сдвиге  $k = 0$  равно 0.98. Сравнение данных табл. 4 и 5 показывает, что модельные коэффициенты для этих двух стран близки по значениям.

Таким образом, предложенная методика моделирования вирусных заболеваний оказывается эффективной и позволяет моделировать как ди-



**Рис. 12.** Динамика заболеваемости коронавирусом в Республике Корея (1) и ADL-модель (2).

намику коронавируса, так и динамику заболевания гриппом H3N2. Различие между гриппом и коронавирусом заключается в порядке авторегрессии. Если текущая заболеваемость коронавирусом зависит от числа заболевших в предыдущие два дня, то текущая заболеваемость гриппом зависит от заболеваемости в течение предшествующих четырех недель.

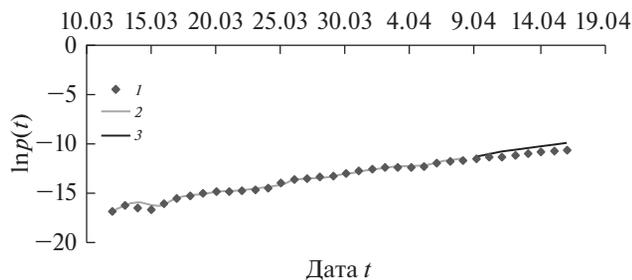
### ЗАКЛЮЧЕНИЕ

Выполненные исследования показали, что для описания развития пандемии коронавируса можно использовать модель, характеризуемую следующими свойствами: наличием порога в развитии критических явлений, качественным сходством кривых заболеваемости, что позволяет описывать их с помощью моделей общего типа, и схожестью процессов распространения эпидемии по территории.

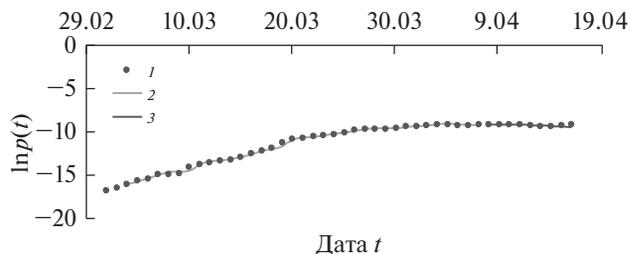
Показано, что для описания критических явлений в эпидемиологии возможно использовать подходы, применяемые для исследования физических систем. На наш взгляд, причиной эффективности использования моделей фазовых переходов в эпидемиологии является то, что в рамках теории фазовых переходов вводится принцип универсальности, согласно которому процессы фазовых переходов зависят только от некоторых

**Таблица 3.** Параметры ADL-модели развития эпидемии на территории Республики Корея

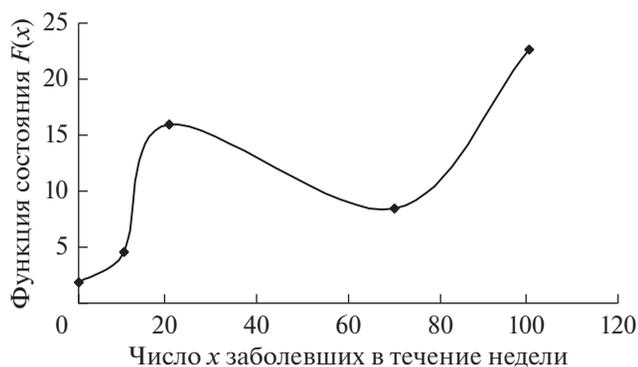
Коэффициент	Значение	Стандартная ошибка	$t$ (49)	Доверительный уровень
$a_0$	349.695	130.178	2.686	0.010
$b$	-0.008	0.003	-2.689	0.010
$a_2$	-0.437	0.116	-3.755	0.000
$a_1$	1.297	0.125	10.394	0.000
$R^2$	0.966			
$F$	458.9			



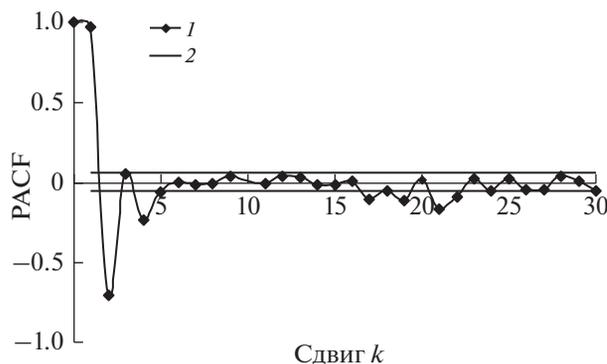
**Рис. 13.** Данные медицинской статистики ( $I$ ) до 18.04.2020, ADL-модель (2) по данным до 08.04.2020 и прогноз (3) на 9–18 апреля 2020 г. по ADL-модели для России.



**Рис. 14.** Данные медицинской статистики ( $I$ ) до 18.04.2020, ADL-модель (2) по данным до 08.04.2020 и прогноз (3) на 9–18 апреля 2020 г. по ADL-модели для США.



**Рис. 15.** Функция состояния  $F(x)$  для Италии.



**Рис. 16.** PACF для гриппа H3N2 в Италии с 46-й недели 2014 г. по 11-ю неделю 2020 г.: 1 – PACF, 2 – доверительный интервал стандартных ошибок PACF.

основных свойств систем, таких как размерность, число компонентов параметра порядка, зависимость взаимодействия в системе от расстояния (Брус, Каули, 1984). Характерной чертой простых статических моделей фазовых переходов является отсутствие в них зависимости от времени. Такое предположение существенно упрощает модель, в которой для описания системы требуется найти лишь устойчивые минимумы функции  $F$ .

Использование идеи универсальности фазовых переходов для систем с одним параметром порядка и однородной пространственной структурой упрощает и унифицирует построение моделей фазовых переходов. Даже не зная точного вида функции  $F$ , все же можно описать процесс фазового перехода и найти критические значения внешнего фактора и значения параметра порядка в устойчивых состояниях системы.

**Таблица 4.** Параметры AR(4)-модели развития эпидемии гриппа на территории Италии

Переменная	Коэффициент	Стандартная ошибка	$t(265)$	$p$ -level
$a_0$	0.060	0.021	2.875	0.004
$\ln x(i - 4)$	-0.307	0.059	-5.206	0.000
$\ln x(i - 3)$	0.661	0.120	5.487	0.000
$\ln x(i - 2)$	-1.238	0.120	-10.293	0.000
$\ln x(i - 1)$	1.829	0.059	31.125	0.000
$R^2$	0.979			
$F$	3114.000			

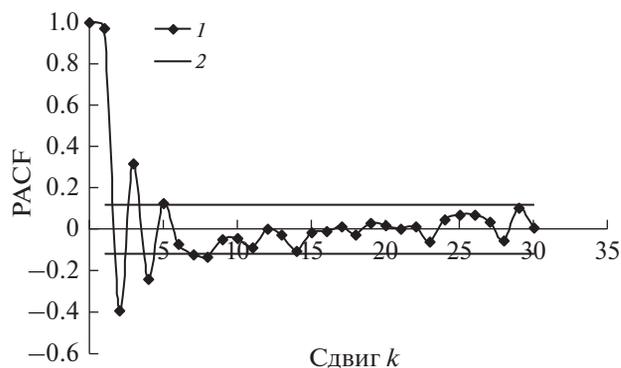


**Рис. 17.** Временные ряды динамики заболевания гриппом H3N2 в Италии: 1 – трансформированный временной ряд статистики заболеваемости, 2 – AR(4)-модель, 3 – критическое значение логарифма числа заболевших.

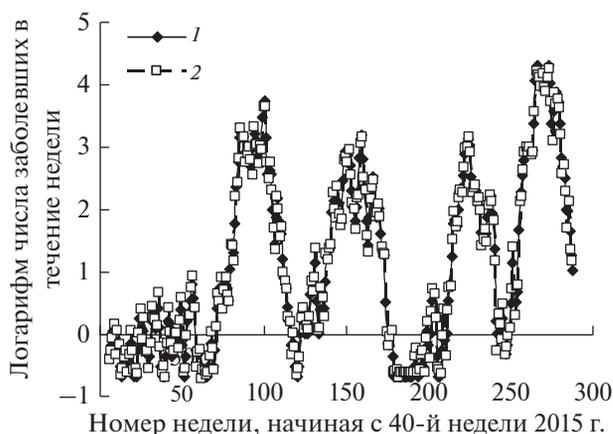
Важной деталью моделей фазовых переходов является возможность включения в модели (3) и (4) влияния внешних факторов, рассматриваемых как некоторое аддитивное поле, тогда как для учета влияния внешних полей в моделях типа SIR необходимо вводить дополнительные уравнения для каждого из параметров, что весьма затруднительно.

Использование ADL-моделей развития заболеваний, для построения которых необходимо фактически иметь информацию только об одной переменной в моделируемой системе, также существенно облегчает процесс построения модели и позволяет оценить с помощью расчета парциальной автокорреляционной функции для рассматриваемого ряда переменных порядок авторегрессии, существование и знаки обратных связей в моделируемой системе.

Представляется, что предлагаемые модели могут позволить описывать не только вирусные, но



**Рис. 18.** PACF для гриппа H3N2 в Великобритании с 40-й недели 2015 г. по 7-ю неделю 2020 г.: 1 – PACF, 2 – стандартные ошибки.



**Рис. 19.** Временные ряды динамики заболевания гриппом H3N2 в Великобритании: 1 – трансформированный временной ряд статистики заболеваемости, 2 – AR(4)-модель.

и бактериальные эпидемии, а также критические явления в экологических системах (Суховольский и др., 2020).

**Таблица 5.** Параметры AR(4)-модели развития эпидемии гриппа на территории Великобритании

Переменная	Коэффициент	Стандартная ошибка	<i>t</i> (278)	<i>p</i> -level
$a_0$	0.037	0.021	1.817	0.070
$\ln x(i - 4)$	-0.243	0.058	-4.208	0.000
$\ln x(i - 3)$	0.681	0.102	6.688	0.000
$\ln x(i - 2)$	-1.025	0.102	-10.046	0.000
$\ln x(i - 1)$	1.557	0.058	26.781	0.000
$R^2$	0.982			
$F$	1873			

## СПИСОК ЛИТЕРАТУРЫ

- Бокс Дж., Дженкинс Г., 1974. Анализ временных рядов. Прогноз и управление. М.: Мир. Вып. 1. 406 с.
- Брус А., Каули Р., 1984. Структурные фазовые переходы. М.: Мир. 408 с.
- Ландау Л.Д., 1937. К теории фазовых переходов // ЖЭТФ. Т. 7. С. 19–32.
- Ландау Л.Д., Лифшиц Е.М., 1964. Статистическая физика. М.: Наука. 567 с.
- Роспотребнадзор, 2020. <https://www.rosпотребнадзор.ru/>
- Сток Дж., Уотсон М., 2015. Введение в эконометрику. М.: “Дело” РАНХиГС. 864 с.
- Суховольский В.Г., Тарасова О.В., Ковалев А.В., 2020. Моделирование критических явлений в популяциях лесных насекомых // Журн. общ. биологии. Т. 81. № 5. С. 374–386.
- Хемминг Р.В., 1987. Цифровые фильтры. М.: Недра. 221 с.
- Allman E.S., Rhodes J.A., 2003. *Mathematical Models in Biology: An Introduction*. Cambridge: Cambridge Univ. Press. 384 p.
- Bolzoni L., Bonacini E., Soresina G., Groppi V., 2017. Time-optimal control strategies in SIR epidemic models // *Math. Biosci.* V. 292. P. 86–96.
- Bonyah E., Khan M.A., Okosun K.O., Islam S., 2017. A theoretical model for Zika virus transmission. A theoretical model for Zika virus transmission // *PLoS One*. V. 12. № 10. P. e0185540.
- Brauer F., Driessche P., van den, Wu J., 2008. *Mathematical Epidemiology*. Berlin-Heidelberg: Springer. 408 p.
- Chen Y., Cheng J., Jiang Y., Liu K., 2020. A time delay dynamical model for outbreak of 2019-nCoV and the parameter identification // *arXiv*: 2002.00418.
- Data Repository by Johns Hopkins University, 2020. <https://github.com/CSSEGISandData/COVID-19>.
- Distante C., Gadelha Pereira I., Garcia Gonçalves L.M., Piscitelli P., Miani A., 2020. Forecasting Covid19 outbreak progression in Italian Regions: A model based on neural network training from Chinese data // *medRxiv*. <https://doi.org/10.1101/2020.04.09.20059055>.
- Edelstein-Keshet L., 2005. *Mathematical Models in Biology*. Philadelphia: SIAM. 586 p.
- Ediriweera D., Silva N., de, Malavige N., Silva J., de, 2020. An epidemiological model to aid decision-making for COVID-19 control in Sri Lanka // *medRxiv*. <https://doi.org/10.1101/2020.04.11.20061481>.
- Gomez J., Prieto J., Leon E., Rodríguez A., 2020. INFEKTA: A general agent-based model for transmission of infectious diseases: Studying the COVID-19 propagation in Bogotá – Colombia // *medRxiv*. <https://doi.org/10.1101/2020.04.06.20056119>.
- Hackl K., 2020. Modeling the COVID-19 pandemic – parameter identification and reliability of predictions // *medRxiv*. <https://doi.org/10.1101/2020.04.07.20056937>.
- Hethcote H.W., 2000. The mathematics of infectious diseases // *SIAM Rev.* V. 42. № 4. P. 599–653.
- Iacoviello D., Liuzzi G., 2008. Fixed/free final time SIR epidemic models with multiple controls // *Int. J. Simul. Model.* V. 7. P. 81–92.
- Jin C., Yu J., Han L., Duan S., 2020. The impact of traffic isolation in Wuhan on the spread of 2019-nCoV // *medRxiv*. <https://doi.org/10.1101/2020.02.04.20020438>.
- Kermack W.O., McKendrick A., 1927. A contribution to the mathematical theory of epidemics // *Proc. Roy. Soc.* V. 115. № 772. P. 700–721.
- Lauer S.A., Grantz K.H., Qifang B.O., Forrest K., Zheng Q. et al., 2020. The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: Estimation and application // *Ann. Intern. Med.* V. 172. № 9. P. 577–582.
- Legrand J., Grais R.F., Boelle P.Y., Valleron A.J., Flahault A., 2007. Understanding the dynamics of Ebola epidemics // *Epidemiol. Infect.* V. 135. № 4. P. 610–621.
- Lyra W., Nascimento J.-D., de, Belkhiria J., Almeida L., de, Chrispim P.P.M., Andrade I., de, 2020. COVID-19 pandemics modeling with SEIR(+CAQH), social distancing, and age stratification. The effect of vertical confinement and release in Brazil // *medRxiv*. <https://doi.org/10.1101/2020.04.09.20060053>.
- Nishiura H., Linton N.M., Akhmetzhanov A.R., 2020. Serial interval of novel coronavirus (COVID-19) infections // *Int. J. Infect. Dis.* V. 93. P. 284–286.
- Shen J.A., 2020. Recursive bifurcation model for predicting the peak of COVID-19 virus spread in United States and Germany // *medRxiv*. <https://doi.org/10.1101/2020.04.09.20059329>.
- Tuomisto J.T., Yrjölä J., Kolehmainen M., Bonsdorff J., Pekkanen J., Tikkanen T., 2020. An agent-based epidemic model REINA for COVID-19 to identify destructive Policies // *medRxiv*. <https://doi.org/10.1101/2020.04.09.20047498>.
- WHO, 2020. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019>.
- Wu Z., McGoogan J.M., 2020. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: Summary of a report of 72314 cases from the Chinese Center for Disease Control and Prevention // *JAMA*. V. 323. № 13. P. 1239–1242.
- Yang X., Yu Y., Xu J., Shu H., Xia J. et al., 2020. Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: A single-centered, retrospective, observational study // *Lancet Respir. Med.* V. 8. P. 475–481.
- Yue Y., Chen Y., Liu K., Luo X., Xu B. et al., 2020. Modeling and prediction for the trend of outbreak of NCP based on a time-delay dynamic system // *Sci. Sin. Math.* V. 50. № 3. P. 1–8.
- Zakary O., Rachik M., Elmouki I., 2017. A new analysis of infection dynamics: Multi-regions discrete epidemic model with an extended optimal control approach // *Int. J. Dyn. Control.* V. 5. P. 1010–1019.
- Zhao S., Lin Q., Ran J., Musa S.S., Yang G. et al., 2020. Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: A data-driven analysis in the early phase of the outbreak // *Int. J. Infect. Dis.* V. 92. P. 214–217.

**Modeling the coronavirus epidemic as a phase transition****V. G. Soukhovolsky<sup>a, b, \*</sup>, A. V. Kovalev<sup>b, \*\*</sup>***<sup>a</sup>Sukachev Institute of Forest SB RAS  
Akademgorodok, 50/28, Krasnoyarsk, 660036 Russia**<sup>b</sup>Federal Research Center KSC SB RAD  
Akademgorodok, 50, Krasnoyarsk, 660036 Russia**\*E-mail: soukhovolsky@yandex.ru**\*\*E-mail: sunhi.prime@gmail.com*

The task of modeling the development of the coronavirus pandemic on the planet is considered. An analysis is conducted aimed at identifying the three types of patterns in the development of coronavirus: assessing the existence of critical disease densities, assessing the patterns of the disease spread by country and different territories within a particular country, modeling the relative dynamics of the incidence in individual countries. For modeling, it is proposed to use analogues of models of first- and second-order phase transitions in physical systems and autoregressive models of the disease development. Calculations are made of the critical number of patients, upon reaching which the exponential stage of the disease development begins, the rate of development of the disease in different countries and in different territories.