

МИНОБРАЗОВАНИЯ РОССИИ  
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ  
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ  
«НИЖЕГОРОДСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ  
УНИВЕРСИТЕТ им. Р.Е. АЛЕКСЕЕВА»

**ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА**

**Глазова Сергея Игоревича**

*(фамилия, имя, отчество)*

Институт (факультет) Институт экономики и управления

Кафедра Цифровая экономика

Группа 16-САИ

Дата защиты « 8 » июля

Индекс

МИНОБРНАУКИ РОССИИ  
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ  
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ  
«НИЖЕГОРОДСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ  
УНИВЕРСИТЕТ им. Р.Е.АЛЕКСЕЕВА»  
(НГТУ)

Институт \_\_\_\_\_ Институт экономики и управления (ИНЭУ)

Направление подготовки (специальность) \_\_ 01,04,02 Прикладная математика и информатика

Направленность (профиль) образовательной программы) Прикладная математика и информатика

Кафедра \_\_\_\_\_ «Цифровая экономика»

## ВЫПУСНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА БАКАЛАВРА

\_\_\_\_\_ (бакалавра, магистра, специалиста)

Студента \_ Глазова Сергея Игоревича \_\_\_\_\_ группы \_\_\_\_\_ 16-САИ

на тему \_\_\_\_\_ «Математическое моделирование распространения коронавируса»  
(наименование темы работы)

### СТУДЕНТ

Глазов С.И.

\_\_\_\_\_  
(подпись)

\_\_\_\_\_  
(фамилия, и.,о.)

\_\_\_\_\_  
(дата)

### РУКОВОДИТЕЛЬ

Митяков С.Н.

\_\_\_\_\_  
(подпись)

\_\_\_\_\_  
(фамилия, и.,о.)

\_\_\_\_\_  
(дата)

### РЕЦЕНЗЕНТ

\_\_\_\_\_  
(подпись)

\_\_\_\_\_  
(фамилия, и.,о.)

\_\_\_\_\_  
(дата)

### ЗАВЕДУЮЩИЙ КАФЕДРОЙ

Митяков С.Н.

\_\_\_\_\_  
(подпись)

\_\_\_\_\_  
(фамилия, и.,о.)

\_\_\_\_\_  
(дата)

### КОНСУЛЬТАНТЫ

1. По нормоконтролю

\_\_\_\_\_  
(подпись)

\_\_\_\_\_  
(фамилия, и.,о.)

\_\_\_\_\_  
(дата)

2. По \_\_\_\_\_

\_\_\_\_\_  
(подпись)

\_\_\_\_\_  
(фамилия, и.,о.)

\_\_\_\_\_  
(дата)

3. По \_\_\_\_\_

\_\_\_\_\_  
(подпись)

\_\_\_\_\_  
(фамилия, и.,о.)

\_\_\_\_\_  
(дата)

ВКР защищена \_\_\_\_\_ (дата)

протокол № \_\_\_\_\_  
с оценкой \_\_\_\_\_

2020 г.

## Содержание

Введение .....	4
1 Обзор существующих методов математического моделирования эпидемий.....	6
1.1 SIR- модель ( <i>Susceptible - Infectious - Recovered</i> ).....	6
1.2 SI- модель ( <i>Susceptible – Infectious</i> ).....	9
1.3 SIS- модель ( <i>Susceptible – Infectious – Susceptible</i> ).....	11
1.4 SEIR- модель ( <i>Susceptible - Exposed – Infectious - Recovered</i> ).....	12
1.5 SEIRS- модель ( <i>Susceptible - Exposed - Infectious - Recovered – Susceptible</i> ).....	16
1.6 Вывод из обзора.....	16
2 Разработка математической модели распространения эпидемии .....	18
2.1 Качественные особенности моделей заболеваемости и смертности.....	20
2.2 Количественные особенности моделей заболеваемости и смертности. ....	21
2.3 Идентификация модели .....	24
2.4 Описание алгоритма.....	24
3 Верификация алгоритмов.....	27
3.1 Верификация алгоритма (Италия).....	27
3.2 Верификация алгоритма (Россия) .....	32
3.3 Верификация алгоритма (Нижегородская область) .....	36
4 Сравнение моделей .....	40
Заключение .....	41
Список литературы.....	42

				<b>ВКР-НГТУ-01.03.02-16-САИ-010-2020</b>			
Изм.	Лист	№ докум.	Подп.	Дата			
Разраб.		Глазов С.И.			Лит.	Лист	Листов
Пров.					У	3	
Н. контр.					НГТУ им. Р.Е. Алексеева ИНЭУ кафедра «ЦЭ» группа 16-САИ		
Утв.		Митяков С.Н.					

Пояснительная записка  
к выпускной квалификационной работе на тему:

«Математическое моделирование  
распространения коронавируса»

## Введение

Объектом данного исследования является заболевание Covid-19 – самая сложная пандемия со времен испанского гриппа, которым более века назад заразилось около 30% населения планеты. Быстрая вспышка Covid-19 и его широкое распространение по всему миру превратили локальную болезнь, первоначально находившуюся в Китае, в глобальную проблему; таким образом, она приобрела статус пандемии. Предмет исследования - динамика распространения вирусного заболевания COVID-19.

Математические модели, описывающие динамику явления, о которых говорится в этой работе, позволяют понять эволюцию и ежедневно готовиться к постоянно растущей нагрузке. Официальные случаи смерти — это случаи, которые случаются с пациентами, находящимися в больницах после того, как они дают положительный результат на мазок из носа. Официальные случаи смерти важны, поскольку они измеряют один из двух возможных результатов лечения в стационаре: успех или неудачу. Врачи должны знать и каким-то образом прогнозировать не только ежедневное количество смертей, но и асимптотическое значение, прогнозируемое в конце пандемии. Такое число является большой психологической нагрузкой для тех, кто борется за жизнь (например, врачей, медсестер), но оно может подготовить их к тому, чтобы справиться с этими фатальными последствиями и каким-то образом найти верхнюю границу явления.

Математические модели могут быть использованы для получения информации о предпосылках прошлых событий, понимания текущей ситуации и прогнозирования развития пандемии на краткосрочные или долгосрочные периоды времени. Качественная и количественная оценка динамики пандемии позволяет лицам, принимающим решения в медицине, планировать чрезвычайную ситуацию, готовиться к пику заболеваемости, и, наконец, ослабить меры безопасности и вернуться к стандартному режиму работы, когда пандемия спадет. Модели могут очертить возможные отклонения пандемии от ожидаемых изменений и, следовательно, играть роль систем раннего предупреждения в случае новых вспышек.

Целью исследования является выявление динамики развития и прогнозирование распространения вирусного заболевания Covid-19.

### Задачи исследования

- Описание и анализ математических моделей распространения инфекционных заболеваний
- Обзор существующих методов математического моделирования эпидемий
- Разработка математических моделей распространения эпидемии

- Моделирование распространения заболевания с реальными статистическими данными
- Сравнение построенных моделей между собой

В этой работе будет осуществлен обзор существующих методов математического моделирования эпидемий, будут построены собственные альтернативные модели и будет выяснено, какая из построенных моделей отражает характеристики распространения заболевания лучше.

					ВКР-НГТУ-01.03.02-16-САИ-010-2020	Лист
Изм.	Лист	№ докум.	Подп.	Дата		5

# 1 Обзор существующих методов математического моделирования эпидемий

Первая математическая модель в эпидемиологии была разработана Бернулли для изучения вариации против оспы [3][4]. Следующий шаг сделал Уильям Фарр, применивший в 1840 году нормальное распределение к анализу смертности от оспы. Основу математической эпидемиологии заложил вклад нескольких биологов и врачей - П.Д. Энко, У.Х. Хамер, Р.А. Росс, А.Г. МакКендрик и А. С. Кермак. Работы Росса о малярии [31] и Кермака и МакКендрика[18][19] считается вехой в развитие математической эпидемиология. Росс, основываясь на его обширные исследования малярии в Индии, показал, что болезнь была распространена комарами и разработал модель описания передачи инфекции [31]. Он предсказал из этой модели, что сокращение популяции комаров будет эффективно контролировать эпидемию малярии в географическом районе. Дальше, было несколько исследований по моделированию конкретных заболеваний, в том числе корь, гонорея, СПИД, проказа [14][24].

Концепция базового репродуктивного числа была разработана в работах Кермака и МакКендрика [20]. Авторы проанализировали распространение заболевания при:

- заболеваниях, при которых зараженный человек выздоравливает и получает иммунитет от перенесенного заболевания (вирусные заболевания)
- заболевания с выздоровлением, но без приобретения иммунитета против возбудителя (бактериальный и венерические заболевания).

Базовое репродуктивное число, обозначаемый повсеместно как  $R_0$  – это безразмерный параметр, характеризующий заразность инфекционного заболевания. Чем больше  $R_0$ , тем интенсивнее передача болезни, следовательно, более серьезной является эпидемия. Например, если  $R_0 > 1$ , то на начальном этапе эпидемии, число инфицированных будет расти по экспоненте. Концепция  $R_0$  является центральной идеей в математической эпидемиологии и жизненно важна для предсказания или описания динамики передачи эпидемии. Далее будут рассмотрены основные виды моделей распространения эпидемий.

## 1.1 SIR- модель (Susceptible - Infectious - Recovered)

Первая математическая модель, используемая для описания эпидемии гриппа, была разработана Кермаком и МакКендриком. Аббревиатура SIR расшифровывается как:

- S - восприимчивый (без иммунитета к болезни)
- I - заразившийся и распространяющий вирус,

									Лист	
Изм.	Лист	№ докум.	Подп.	Дата	ВКР-НГТУ-01.03.02-16-САИ-010-2020					6

- R – выздоровевший и получивший иммунитет, хотя бы временный.

В этой модели люди последовательно перемещаются между тремя классами. Каждый зараженный человек передает восприимчивым членам населения инфекцию со средней скоростью передачи  $\beta$ . А после инфекционного периода, он считается выздоровевшим. Если средняя скорость выздоровления  $\alpha$ , то средний период передачи у любого индивида равна  $1/\alpha$ .

$$\frac{dS(t)}{dt} = -\beta S(t)I(t)$$

$$\frac{dI(t)}{dt} = \beta S(t)I(t) - \alpha I(t)$$

$$\frac{dR(t)}{dt} = \alpha I(t)$$

(1)

Здесь  $S(t)$  и  $I(t)$  обозначают число особей в восприимчивом и инфекционном состояниях соответственно в любой момент времени  $t$ . Скорости изменения  $S(t)$  и  $I(t)$  во времени обозначаются производными  $dS(t)/dt$  и  $dI(t)/dt$  соответственно. Общая численность населения считается постоянной и определяется как  $N = S(t) + I(t) + R(t)$ , когда никто не входит и не выходит из системы.

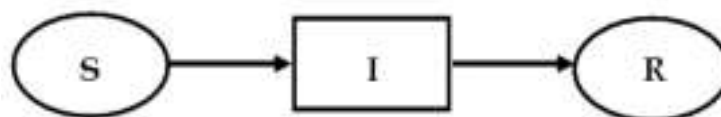


Рисунок 1. Схема принципа работы SIR-модели.

Число восприимчивых особей  $S(t)$  уменьшается с увеличением числа случаев (то есть инфекционных заболеваний  $I(t)$ ). Пики эпидемии снижаются, так как все больше людей выздоравливают и прекращают передачу болезни. Исходя из того, что каждый изначально был восприимчивым (то есть при  $t = 0$ ,  $S(t) = N$ ), вновь введенный инфицированный индивид может заразить в среднем  $\beta N / \alpha = R_0$  индивидуумов. Это базовое репродуктивное число,  $R_0$ . Другими словами,  $R_0$  описывает среднее число вторичных инфекций, генерируемых одним инфекционным индивидуумом при введении в полностью восприимчивую популяцию. Степень тяжести эпидемии и темпы ее роста зависят от значения базового репродуктивного числа. Если  $R_0 > 1$ , эпидемия продолжится. Если  $R_0 < 1$ , эпидемия погаснет.  $R_0$  можно рассчитать по темпам роста эпидемии ( $r$ ), полученным из данных совокупных случаев на начальной фазе роста

$$R_0 = \left(1 + \frac{r}{\alpha}\right) \text{ вспышки:}$$

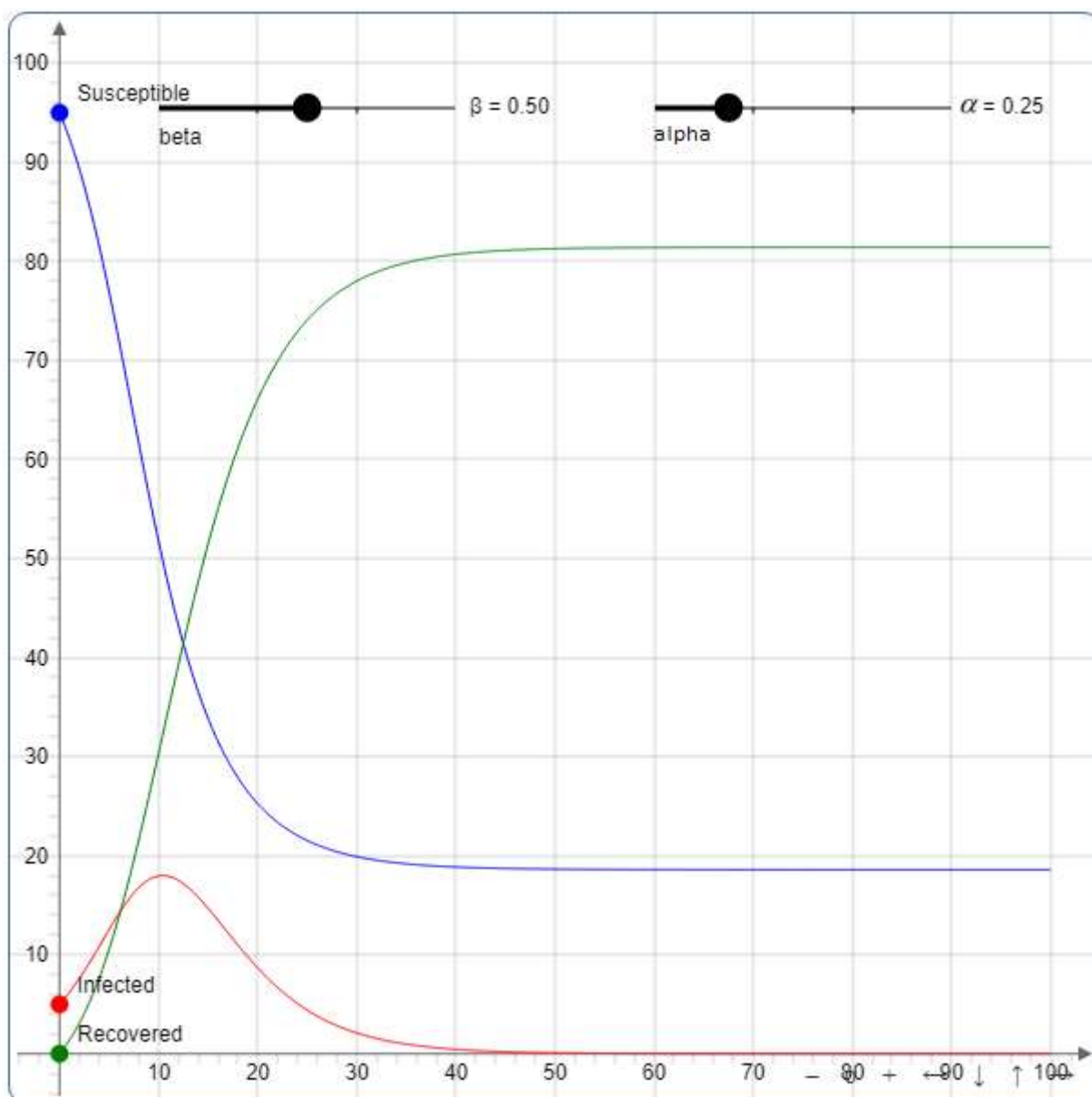


Рисунок 2. График решения SIR-уравнений. Здесь синяя линия — число восприимчивых индивидов, красная — инфицированных, зеленая — переболевших.

Красная линия - интенсивность эпидемии, показывает количество одновременно болеющих индивидов, определяется базовым репродуктивным числом

Численные решения обыкновенных дифференциальных уравнений (1) могут быть получены с подходящими граничными условиями (подходящими для болезни) с использованием компьютерного моделирования. Эта модель использовалась для объяснения передачи кори в Нью-Йорке в 1962 году, а также для повторных вспышек заболевания между 1930 и 1962 годами [2]. Модель SIR может быть расширена для объяснения возникновения повторных эпидемий в одном месте из-за возбудителя учитывая демографию, то есть добавление и удаление людей из популяции через рождение и смерть соответственно. Принимая во внимание, что В является



коэффициентом рождаемости в единицу времени и коэффициентом смертности (на душу населения)  $\mu$ , уравнение (1) можно изменить как:

$$\begin{aligned}\frac{dS(t)}{dt} &= B - \beta S(t)I(t) - \mu S(t) \\ \frac{dI(t)}{dt} &= \beta S(t)I(t) - \alpha I(t) - \mu I(t)\end{aligned}\tag{3}$$

Такая модификация базовой модели SIR была использована для объяснения возникновения кори [2]. Модифицированные модели SIR также использовались для объяснения динамики передачи различных заболеваний, таких как корь [1] и грипп [38]. Модель SIR также была соответствующим образом модифицирована, чтобы отражать или прогнозировать пространственно-временную динамику заболевания, особенно вспышки гриппа в СССР [32], а также учитывать влияние перелетов на пандемии гриппа [33].

### 1.2 SI- модель (Susceptible – Infectious)

Модель SI является самой простой формой из всех моделей заболеваний. Люди рождаются без иммунитета (восприимчивый). После заражения и без лечения люди остаются инфицированными и заразными в течение всей своей жизни и остаются в контакте с восприимчивым населением. Эта модель соответствует поведению таких заболеваний, как цитомегаловирус (ЦМВ) или герпес. Динамика I в модели SI также известна как логистический рост. Если нет жизненно важных процессов (рождение и смерть), каждый восприимчивый в конечном итоге заразится. Модель SI может быть записана в виде следующего обыкновенного дифференциального уравнения:

$$\begin{aligned}\frac{dS}{dt} &= -\frac{\beta SI}{N} \\ \frac{dI}{dt} &= \frac{\beta SI}{N} = \beta I \left(1 - \frac{I}{N}\right)\end{aligned}\tag{4}$$

где  $N = S + I$  общая численность населения.

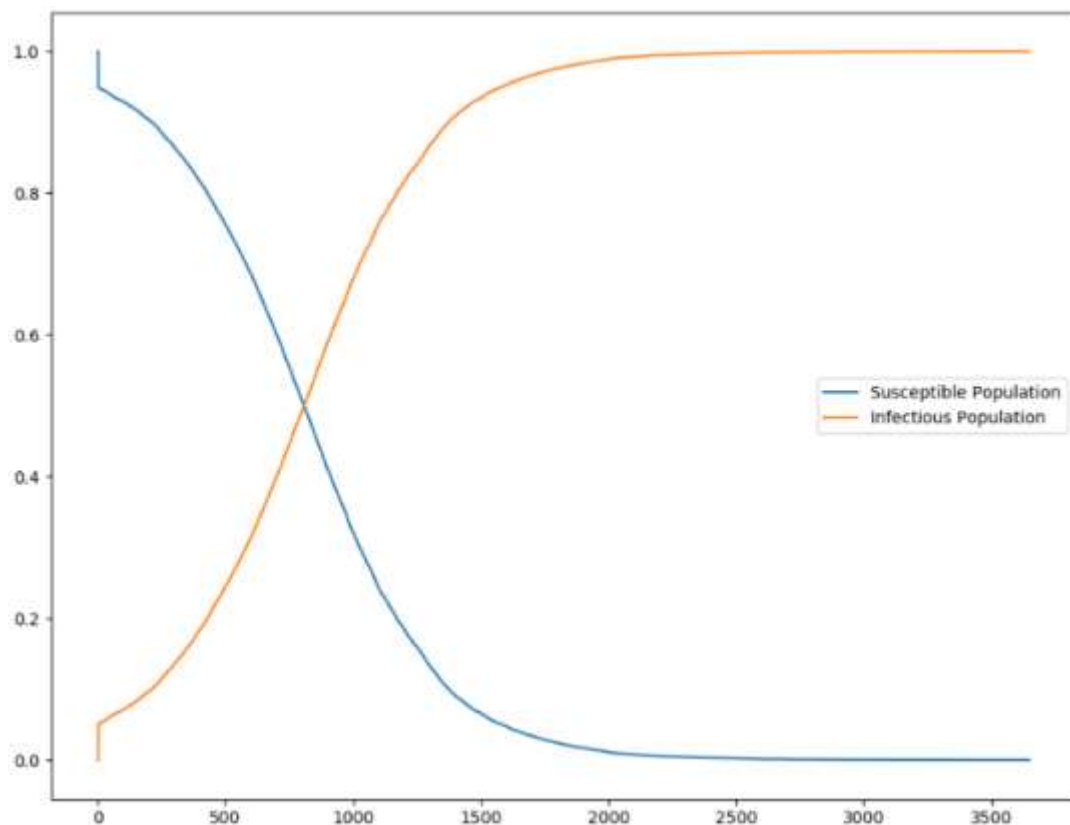


Рисунок 3. Пример работы SI- модели. Желтая линия- инфицированные индивиды, синяя линия- восприимчивые индивиды.

Можно добавить динамику жизни населения. Для этого добавим коэффициенты. Пусть  $\mu$  и  $\nu$  представляют соответственно коэффициент рождаемости и смертности для модели. Чтобы поддерживать постоянную популяцию, предположим, что  $\mu = \nu$ . Следовательно, ОДУ принимает вид:

$$\begin{aligned} \frac{dS}{dt} &= \mu N - \frac{\beta SI}{N} - \nu S \\ \frac{dI}{dt} &= \frac{\beta SI}{N} - \nu I \end{aligned}$$

(5)

где  $N = S + I$  общая численность населения.

Изм.	Лист	№ докум.	Подп.	Дата

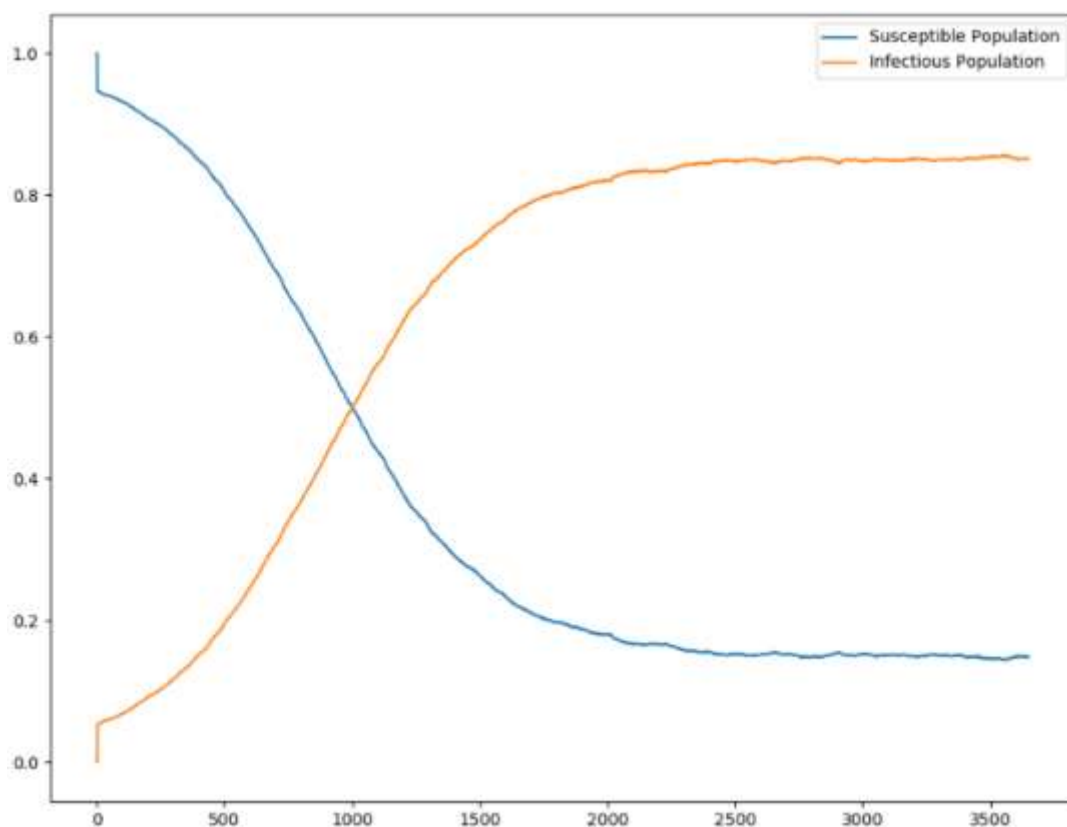


Рисунок 4. Пример работы SI- модели с параметрами рождаемости и смертности.

### 1.3 SIS- модель (Susceptible – Infectious – Susceptible)

Модель SIS подходит для заболеваний, к которым не вырабатывается иммунитет, например простуда или венерические заболевания, такие как гонорея или хламидиоз. Поскольку люди становятся снова восприимчивыми после переноса инфекции, заболевание достигает устойчивого состояния в популяции, даже без жизненной динамики. ОДУ для модели SIS без жизненной динамики может быть аналитически решено для понимания динамики заболевания. ОДУ выглядит следующим образом:

$$\begin{aligned} \frac{dS}{dt} &= -\frac{\beta SI}{N} + \gamma I \\ \frac{dI}{dt} &= \frac{\beta SI}{N} - \gamma I \end{aligned} \quad (6)$$

Эти уравнения показывают, что число инфицированных и здоровых суммарно не изменяется, а количество заражений пропорционально количеству контактов между инфицированными и здоровыми.

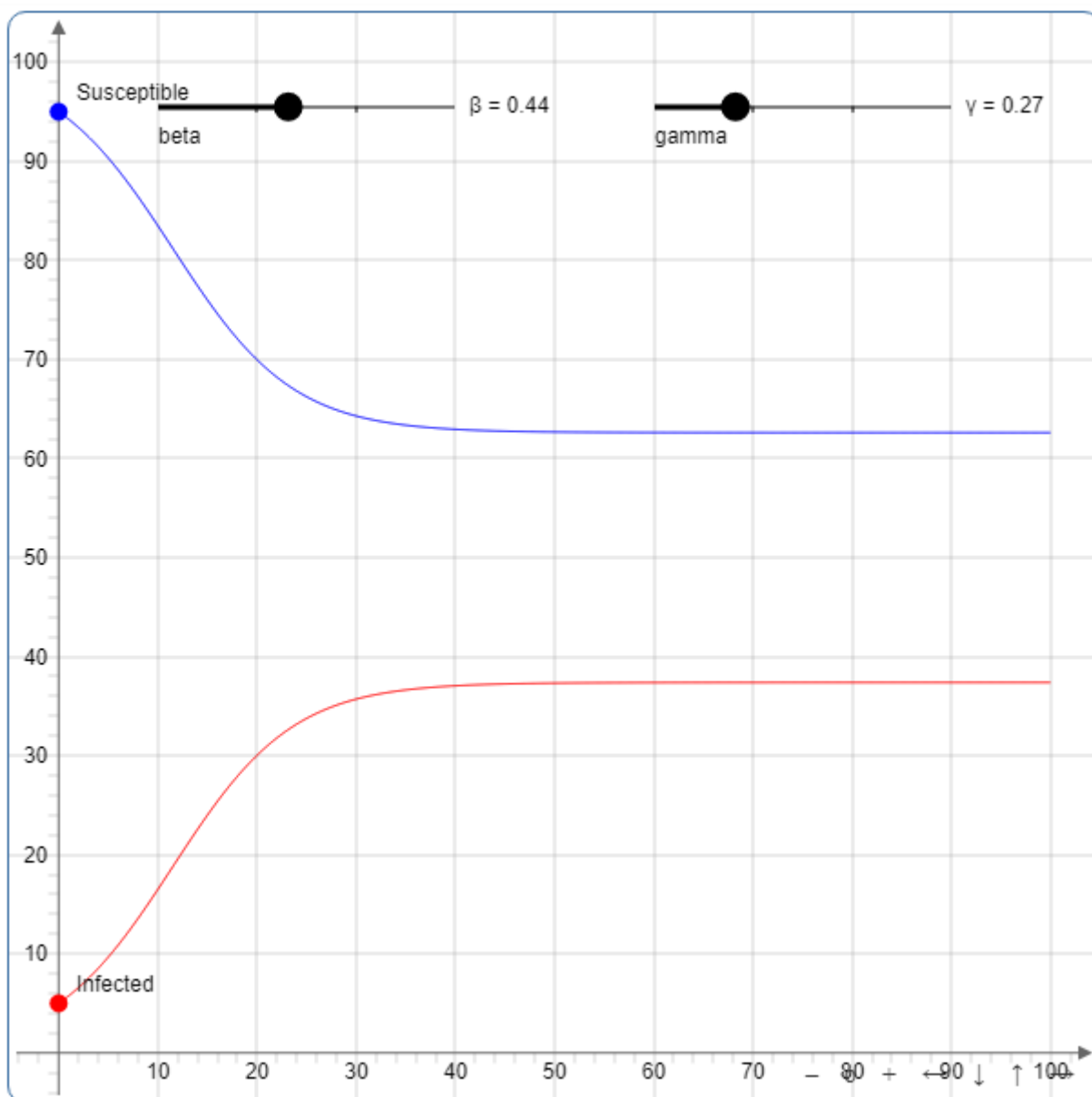


Рисунок 5. Пример работы SIS- модели. Красная линия- инфицированные индивиды, синяя линия- восприимчивые индивиды.

#### 1.4 SEIR- модель (Susceptible - Exposed – Infectious - Recovered)

В случае некоторых инфекционных заболеваний наблюдается инкубационный период у индивидуума после получение возбудителя и до появления симптомов (латентная фаза, во время которой человек заражен, но еще не заразен). Следовательно, простая модель SIR не может эффективно описать передачу таких заболеваний. Следовательно, математическая модель должна учитывать открытое состояние или скрытое состояние, что приводит к развитию модели восприимчивый – контактный – инфицированный – выздоровевший.

Модель SEIR также предполагает введение одного инфицированного индивида в популяцию, где члены ранее не подвергались воздействию патогена и, следовательно,

все подвержены (S). Каждый человек, получивший возбудителя (патоген), существует в состоянии «открыто» или «скрыто» (E), в течение которого он / она переносит вирус или бактерии, но не передает инфекцию никому. С появлением симптомов один и тот же индивид переходит в инфицированное состояние и считается инфицированным (I).

Если  $k$  - скорость перехода из Открытого состояния в Инфицированное состояние, то продолжительность среднего открытого периода или скрытой фазы составляет  $1 / k$ . Зараженный индивид передает вирус восприимчивым членам населения со средней скоростью передачи  $\beta$ . В конце инфекционного периода человек выздоравливает и считается выздоровевшим (R) членом населения. Если средняя скорость восстановления составляет  $\alpha$ , то средний период передачи для любого человека составляет  $1 / \alpha$ .

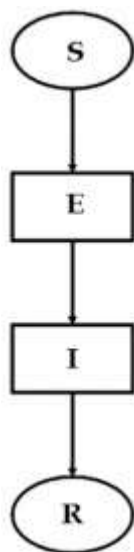


Рисунок 6. Схема принципа работы SEIR-модели.

Учитывая постоянный размер популяции  $N = S + E + I + R$ , набор дифференциальных уравнений, описывающих передачу в соответствии с базовой моделью SEIR, определяется как

$$\begin{aligned} \frac{dS(t)}{dt} &= -\beta S(t)I(t) \\ \frac{dE(t)}{dt} &= \beta S(t)I(t) - \kappa E(t) \\ \frac{dI(t)}{dt} &= \kappa E(t) - \alpha I(t) \\ \frac{dR(t)}{dt} &= \alpha I(t) \end{aligned} \quad (7)$$

Если мы предположим, что часть  $f$  людей, покидающих инфицированное состояние в момент времени  $t$ , выздоравливает, а часть  $(1-f)$  умирает из-за болезни, то уравнения (7) можно изменить так:

$$\begin{aligned} \frac{dS(t)}{dt} &= -\beta S(t)I(t) \\ \frac{dE(t)}{dt} &= \beta S(t)I(t) - \kappa E(t) \\ \frac{dI(t)}{dt} &= \kappa E(t) - \alpha I(t) \\ \frac{dN(t)}{dt} &= -(1-f)\alpha I(t) \end{aligned} \quad (8)$$

Следует отметить, что в этом случае популяция не постоянна, а уменьшается, так как все больше людей страдают от этой болезни. Принимая во внимание сценарий отсутствия смертельных исходов, базовое репродуктивное число может быть оценено на основе скорости роста начальной фазы вспышки для простой модели SEIR следующим образом. Скорость роста эпидемии ( $r$ ) может быть рассчитана из оценок совокупного числа подтвержденных инфекций ( $y$ ) и предполагаемой даты начала и размера вспышки ( $t$  и  $y$ ), соответственно, с использованием уравнения [9]

$$y = y_0 e^{r(t-t_0)} \quad (9)$$

Базовое репродуктивное число ( $R_0$ ) определяется по формуле:

$$R_0 = \left(1 + \frac{r}{\alpha}\right) \left(1 + \frac{r}{\kappa}\right) \quad (10)$$

Где средний инфекционный период  $1/\alpha$  и средний инкубационный период  $1/\kappa$ . Это дает более точную оценку  $R_0$  по сравнению с моделью SIR, где скрытая фаза не рассматривалась. Это лучше всего объяснить с помощью примера. Гурав и др. [15] сообщили о новой вспышке гриппа А / Н1N1 2009 (свиной грипп) в школе-интернате в Панчгани, Махараштра. На основании эпидемиологических данных о вспышке, Шил и др. [34] определили, что собственная экспоненциальная скорость роста ( $r$ ) составляет 0,2341 в день. Предполагая, что средний инкубационный период составляет 1,5 дня, а средний инфекционный период - 4 дня,  $R_0$  оценивается в 2,61 (согласно уравнению 10). Подобные более высокие значения  $R_0$  и интенсивная передача также наблюдались в разных странах для мест с тесной группировкой людей, таких как деревня и школы [13][37].

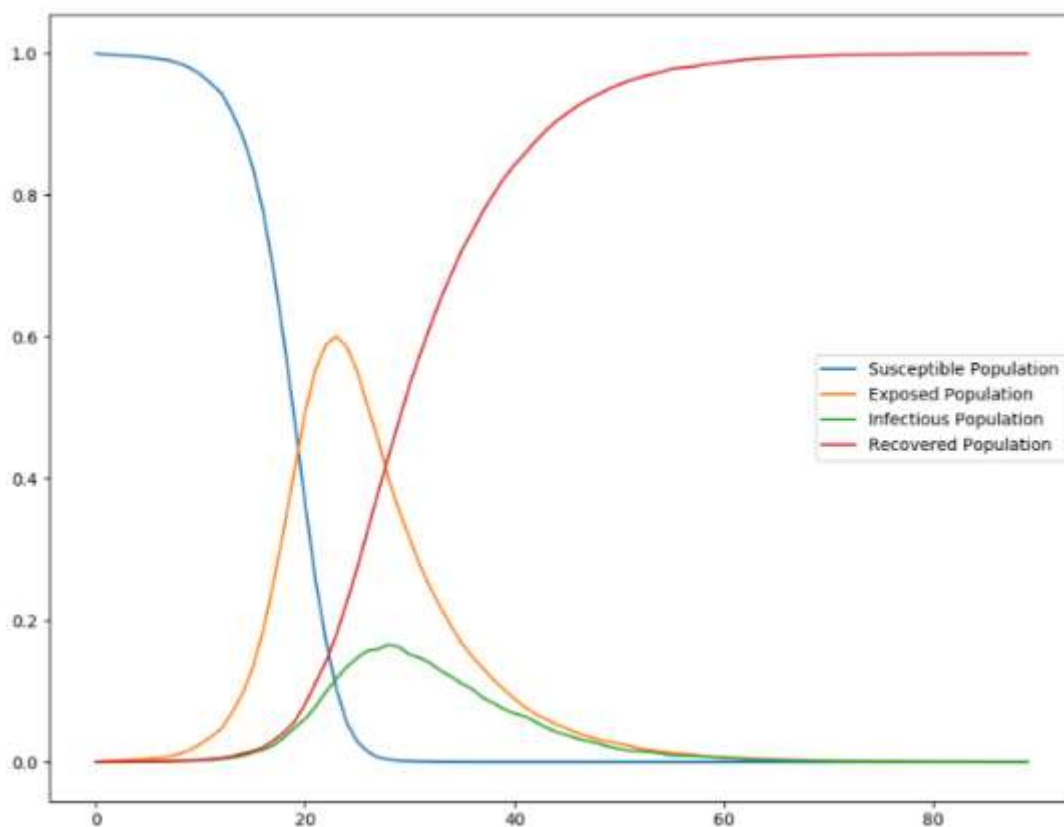


Рисунок 7. Пример работы SEIR- модели с 8-дневным инкубационным периодом. Синяя линия- восприимчивые индивиды, желтая линия- индивиды в инкубационном периоде, зеленая линия- инфицированные, красная линия- выздоровевшие индивиды.

Модель SEIR с подходящими адаптациями широко использовалась при различных заболеваниях, включая грипп, ветряную оспу и SARS [7]. Deguen и другие [7] проанализировали сезонный характер эпидемии ветряной оспы во Франции, сопоставив модель SEIR с периодической функцией частоты контакта с еженедельными данными о заболеваемости ветряной оспой, собранными в 1991-1996 гг. Обе модели, предполагая либо непрерывную, либо кусочно-постоянную периодическую функцию, давали разумное соответствие данным о заболеваемости и давали оценки инкубационного и инфекционного периодов в соответствии с клинически или серологически оцененными значениями. Ванг и соавторы (39) адаптировали модель SEIR с зависимой от времени скоростью передачи (число контактов на одного инфекционного человека в день) для описания вспышки атипичной пневмонии в городе Пекин. Решение SEIR точно соответствовало эпидемиологическим данным. Чтобы изучить динамику передачи вспышки атипичной пневмонии в Гонконге (2003 г.), Смолл и его коллеги [35][36] адаптировали концепцию SEIR в «модели маленького мира», где передача была разрешена внутри групп населения и между случайными количеством географически удаленных кластеров. Передача была разрешена только между связанными узлами /

кластерами. Эта концепция могла бы эффективно описать вспышку атипичной пневмонии в 2003 году, когда компьютерное моделирование соответствовало зарегистрированным данным.

### 1.5 SEIRS- модель (Susceptible - Exposed - Infectious - Recovered – Susceptible)

Модель SEIR предполагает, что люди получают пожизненный иммунитет к заболеванию после выздоровления, но для многих заболеваний иммунитет после заражения со временем снижается. В этом случае модель SEIRS используется, чтобы позволить выздоровевшим людям вернуться в восприимчивое состояние. В частности,  $\xi$  это скорость, с которой выздоровевшие люди возвращаются к восприимчивому состоянию из-за потери иммунитета. Если имеется достаточный приток в восприимчивую популяцию, в равновесии динамика будет в эндемичном состоянии с затухающими колебаниями. SEIRS- модель описывается следующими ОДУ:

$$\begin{aligned} \frac{dS}{dt} &= -\frac{\beta SI}{N} + \xi R \\ \frac{dE}{dt} &= \frac{\beta SI}{N} - \sigma E \\ \frac{dI}{dt} &= \sigma E - \gamma I \\ \frac{dR}{dt} &= \gamma I - \xi R \end{aligned} \tag{11}$$

где  $N = S + E + I + R$  общая численность населения.

### 1.6 Вывод из обзора

Исследование динамики передачи любого заболевания зависит от характера данных и разработки модели, которая наилучшим образом описывает сценарий вспышки. Подгонка эпидемиологических данных помогает оптимизировать параметры модели, особенно те, которые невозможно определить экспериментальным путем. Например, бессимптомные параметры (возможна ли передача вируса в бессимптомный период, сколько и как долго длятся бессимптомный период и т. д.) для гриппа у людей не могут быть оценены экспериментально или при помощи наблюдений, но могут быть оценены с помощью модельных исследований при условии, что известно общее количество бессимптомных индивидуумов. (по результатам серологического обследования) для конкретной вспышки [34]. Моделирование и имитационные исследования, основанные на эпидемиологических данных, также могут помочь оценить эффективность мер контроля и могут быть использованы для оценки эффективности



вакцины. Однако, несмотря на преимущества, моделирование эпидемий также имеет ограничения. Ограничения в моделировании заболеваний возникают из-за неправильной регистрации данных, особенно если это связано с отслеживанием контактов (методы и эффективность могут варьироваться в зависимости от страны) и / или предположениями для описания сценария вспышки.

					ВКР-НГТУ-01.03.02-16-САИ-010-2020	Лист
Изм.	Лист	№ докум.	Подп.	Дата		17

## 2 Разработка математической модели распространения эпидемии

Математическое моделирование может оказать реальную помощь в описании, понимании и, в конечном итоге, прогнозировании распространения вируса на определенной территории (например, город, область, регион, страна и континент). Математические модели содержат набор уравнений, которые включают ряд адаптивных параметров, которые могут быть определены численно на основе имеющихся реальных данных [28]. После уточнения этих моделей и вычисления адаптивного параметра можно использовать эти модели, чтобы понять, что произошло в прошлом (извлечь уроки), и, что еще более важно предсказать, что может произойти в будущем [29].

В этом разделе основное внимание уделяется описанию временной эволюции (то есть динамики) пандемии Covid-19, основанной только на двух переменных, которые являются наиболее надежными, а именно на количестве зараженных и на официальных смертях в больницах.

Я решил описать и проследить, а также предсказать динамику этих реальных значений с помощью простейших доступных математических моделей. Эти модели основаны на регрессии, они минимизируют расстояние между предсказаниями модели и реальными данными. Я выбрал модели, которые лучше всего описывают динамику явления и характеризуются наименьшим количеством адаптивных параметров. На самом деле, я использовал только два или три адаптивных параметра, чтобы сохранить простой подход, избежать чрезмерной параметризации и сохранить числовую устойчивость и стабильность [23]. Такой подход позволяет реализовать эти модели на любых компьютерах без необходимости использования специальных инструментов программирования. Намеренно я реализовал набор математических моделей, которые могут работать в Excel, поскольку он широко доступен в любых операционных системах (например, Windows, Mac, Linux) и работает на нескольких платформах (например, компьютеры, планшеты, мобильные телефоны). Кроме того, Excel требует более низкой кривой обучения по сравнению с другими инструментами программирования, такими как Matlab или Mathematica, или с языками программирования, такими как Fortran, C / C ++, Python (и это лишь некоторые из них).

Я намеренно не использовал никаких эпидемиологических моделей, таких как модели SIR, SEIR, SEIRS и т.д., которые основаны на наборе нескольких дифференциальных уравнений с начальными условиями и рядом адаптивных параметров, а также сильных предположениях и упрощениях. Семейство моделей SIR показывает более высокую детализацию описания явления. Однако, по моему мнению,

их использование оправдано для запуска параметрических прогнозирующих сценариев, основанных на ряде предположений и гипотез, которые достаточно чувствительны к выбору правильных значений адаптивных параметров и функционального описания. Надежные значения этих параметров будут доступны только в конце пандемии и будут в значительной степени зависеть от политических решений, принятых в разное время и по интенсивности мер по социальному дистанцированию в каждой стране (или даже в каждом регионе) и, в целом, от фармацевтического вмешательства [5].

И наоборот, математические модели, предлагаемые далее, могут использоваться ежедневно (путем обновления данных заболеваемости и смертности соответственно) и автоматически адаптироваться к эволюции этих значений.

Явления, лежащие в основе заражаемости и смертей, следуют двум различным эволюционным кривым в случае пандемии. Значения «количество зараженных» начинаются с нуля в самом начале пандемии и продолжают увеличиваться, пока пандемия развивает свою интенсивность до максимального значения. После плато количество зараженных начинает уменьшаться и достигает окончательного нулевого значения, когда пандемия заканчивается.

И наоборот, число смертей продолжает монотонно увеличиваться и достигает окончательного значения, то есть количества жертв, которые страна / регион должна заплатить за пандемию. Число смертей является накопленной величиной, которая увеличивается с ежедневными жертвами.

Что касается больших регионов и стран, то количество зараженных монотонно увеличивается до максимального плато. Это то, что происходит в Италии [26]. И наоборот, в небольших регионах с точки зрения населения и / или положительных случаев могут наблюдаться немонотонно растущие тенденции. Например, Умбрия - центральный регион Италии, где 882 000 жителей проживают в небольших муниципалитетах, распределенных по ее территории, что приводит к тому, что количество зараженных значительно отличается от наиболее вовлеченных итальянских регионов - северных. Действительно, в Умбрии часто количество зараженных оставалось постоянным в течение нескольких дней. Время от времени это число уменьшалось, а затем снова увеличивалось с довольно небольшими колебаниями (т.е. несколькими единицами в день) из-за крошечного числа инфицированных людей. Для такого рода регионов / стран модели и тенденции, обсуждаемые в далее, не рекомендуются, и их следует избегать.

## 2.1 Качественные особенности моделей заболеваемости и смертности

Сначала сосредоточимся на качественном тренде роста инфицированных. В первые дни пандемии наблюдается экспоненциальный рост [30]. Экспоненциальная кривая характеризуется периодом удвоения, который остается постоянным, если не происходит внешнего воздействия. Довольно легко показать, что явление экспоненциально, когда оно соответствует прямой линии на полулогарифмической диаграмме (то есть графике, где ось  $x$  является линейной (время, количество дней), а ось  $y$  представляет логарифм действительных значений (количество больных)). Можно измерить качество аппроксимирующей кривой (т. е. прямой линии) относительно реальных данных с помощью коэффициента детерминации ( $R^2$ ). Это безразмерный коэффициент в диапазоне от 0 до 1. Чем выше значение, тем лучше согласованность модели с реальными данными. В случае Италии, России и Нижегородской области в первые дни эпидемии значения  $R^2$  были достаточно близки к единице.

Экспоненциальный рост остается стабильным в течение 15–16 дней. Каждый день наклон прямой линии немного уменьшается. Это присуще динамике эпидемии и показывает, что феномен начинает уменьшать свой импульс, хотя и остается экспоненциальным.

Существует время, дни 16–22, когда новые значения по количеству заболевших начинают отходить от прямой линии и начинают следовать параболе на полулогарифмическом графике. Это случай так называемого экспоненциально модифицированного гауссова (EMG) тренда [10]. На стандартной декартовой плоскости явление хорошо описывается двумя различными областями, разделенными точкой перегиба. Первая часть практически не отличается от экспоненциальной кривой, а вторая (после точки перегиба) в конечном итоге выравнивается (постепенно становится линейной).

В точке перегиба происходит максимальная положительная скорость изменения функции. Иными словами, в точке перегиба первая производная (т. е. максимальное суточное изменение наблюдаемой переменной, количество зараженных) имеет самый высокий положительный наклон (т. е. максимальный прирост). Это самые трудные дни для медицинских работников, в которые, день за днем, наблюдается наибольший прирост пациентов. Кроме того, чем больше регион / страна с количеством инфицированных пациентов, тем выше их инерция. После точки перегиба феномен продолжает нарастать, но более медленными темпами. На этом этапе увеличение является монотонным, что означает, что количество зараженных продолжает расти, но ежедневный рост постепенно снижается до максимального плато (около 39–43 дня). Если

					ВКР-НГТУ-01.03.02-16-САИ-010-2020	Лист
Изм.	Лист	№ докум.	Подп.	Дата		20

инерция достаточно высока (большое количество заболевших), плато будет стабильным в течении нескольких дней). Для медицинских работников это максимальное плато имеет решающее значение, поскольку пребывание пациентов в реанимации достаточно длительное [25]. Обычно пациенты остаются в отделениях интенсивной терапии не менее пятнадцати дней [2]. На максимальном плато могут возникать небольшие колебания, но, наконец, количество заболевших за день начнет уменьшаться (дни 42–45) и постепенно стремится к нулю. Эти явления могут быть описаны при помощи нескольких моделей: логистической кривой и кривой Гомпертца. Все эти модели имеют в основе физические или биологические процессы. Логистическая модель [16] был первоначально предложен в 1838 году Пьером Франсуа Ферхюльстом для описания роста популяций, где скорость размножения зависит как от существующего населения, так и от количества доступных ресурсов. Модель Гомпертца [27] похожа на логистическую и была разработана Бенджамином Гомпертцем в 1825 году для описания закона человеческой смертности.

Так же, данные модели могут быть адаптированы для моделирования смертности населения от болезни. Кривая смертности только монотонно увеличивается и, наконец, приближается к максимальному плато после истечения пандемии. Количественно точка перегиба, когда происходят максимальные ежедневные смертельные случаи, наступает в дни 34–38, и для достижения 98% окончательного максимального плато требуется всего 100–120 дней. Математические модели Ферхюльста и модели Гомпертца предсказывают практическое исчезновение смертей через 150–200 дней. К предсказанным прогнозам нужно относиться с осторожностью и скептицизмом, поскольку 1) это просто прогнозы, на момент, когда работа была написана; 2) феномен смертельных исходов и заражаемости в значительной степени зависит от мер социального дистанцирования, принятых регионами и странами, а также от весьма неопределенных результатов, которые даст так называемая фаза 2, когда люди постепенно уменьшат социальное дистанцирование и снова начнут жить и работать, как раньше, до пандемии, хотя и с более высокой осведомленностью о рисках, связанных с безопасностью и здоровьем. Только время покажет, приближаются ли эти прогнозы к реальному явлению, или следует учитывать дальнейшие проблемы моделирования.

## 2.2 Количественные особенности моделей заболеваемости и смертности

Математическое описание экспоненциальной модели:

$$y = a 10^{bt} \quad (12)$$

					ВКР-НГТУ-01.03.02-16-САИ-010-2020	Лист
Изм.	Лист	№ докум.	Подп.	Дата		21

Математическое описание логистической модели:

$$y = \frac{a}{1+b \exp(-ct)} = \frac{a}{1+be^{-ct}} \quad (13)$$

Математическое описание модели Гомпертца:

$$y = a \exp(-b \exp(-ct)) = a e^{-be^{-ct}} \quad (14)$$

где  $t$  - время,  $a$ ,  $b$ ,  $c$  - адаптивные параметры,  $e = 2.71828182845\dots$  это число Эйлера,  $y$  — это зависимая переменная, которая прогнозируется моделью (т. е. зараженные или умершие).

Экспоненциальная модель имеет два параметра, в то время как логистическая модель и модель Гомпертца имеют три параметра. Как уже говорилось выше, подход к параметризации моделей является минимальным и соответствует принципу «keep-it-simple». Что касается независимой переменной, то  $t$  можно считать либо непрерывным, либо, что более правильно, прерывистым, поскольку оно соответствует количеству дней с начала пандемии в конкретном регионе / стране.

Уравнение (12) в полулогарифмических координатах принимает вид:

$$Y = A + bt \quad (15)$$

которое представляет собой уравнение прямой с наклоном  $b$ .

Логистическая кривая (уравнение 13) симметрична относительно точки перегиба, а сигмовидная функция является частным случаем логистической функции [16]. Кривая Гомперца качественно похожа на логистическую функцию, но она не симметрична и требует больше времени для достижения конечного асимптотического плато. Кроме того, плато модели Гомперца выше, чем плато логистической модели.

Далее, будет приведена аналитическая формула, позволяющая вычислить точку перегиба. Точка перегиба определяет время, когда скорость изменения функции является самой высокой, т.е. после нее ежедневных случаев заболевания будет меньше.

$$t_{infl} = \frac{\ln(b)}{c} \quad (16)$$

И логистическая модель, и модель Гомпертца имеют одинаковое максимальное плато при  $t \rightarrow \infty$ ,  $y = a$ .

Каждая кривая имеет свое собственное  $a$  значение, которое вычисляется путем минимизации суммы квадратов отклонений между реальными данными и моделью с помощью следующей процедуры нелинейной регрессии [17]:

$$\text{Min}_{a,b,c} \left\{ SSE = \sum_{i=1}^{NP} (y_i^{real} - y_i^{model})^2 \right\} \quad (17)$$

Минимизация уравнения (17) является многомерной и неограниченной, хотя процедуру оптимизации можно облегчить, указав, что степени свободы (то есть адаптивные параметры) положительны.

После того, как модель определена, можно ответить на следующие вопросы: сколько времени требуется феномену для достижения заданного процентного отношения к максимальному плато? Например, в какое время феномен смерти достигнет 98% или 99% асимптотического состояния (т. е. максимального конечного плато после истечения пандемии)?

Математически нужно решить следующее нелинейное алгебраическое уравнение:

$$y_{model}(t) = p \cdot a \quad (18)$$

где  $p$  желаемая часть (например,  $p=98$ , если желаемый процент составляет 98%). Более простой подход к решению уравнения (18) состоит в представлении в таблицу Excel прогнозов модели, как функции времени и поиска в первый раз, когда явление превышает этот процент (так как известно асимптотическое значение  $a$ ).

Оценка суточной скорости изменения явления может быть выполнена либо дискретным способом, вычисляя разницу между двумя последовательными предсказаниями модели (например, в электронной таблице Excel), либо путем аналитического дифференцирования.

Первая производная экспоненциальной функции:

$$y' = a b \ln(10) 10^{bt} \quad (19)$$

Первая производная логистической функции:

$$y' = \frac{abc e^{ct}}{(b + e^{ct})^2} \quad (20)$$

Первая производная функции Гомпертца:

$$y' = abc e^{-b e^{-ct} - ct} \quad (21)$$

Уравнения (19)-(21) хороши, когда их рисуют на графике, и сравнивают с суточными колебаниями показателей как заражаемости, так и показателей смертности. Действительно, они помогают с дифференциальной точки зрения понять тенденцию явления, а максимальные и минимальные значения производной соответствуют точкам перегиба (т.е. максимальная скорость увеличения и уменьшения явления соответственно).

### 2.3 Идентификация модели

Модели, представленные выше могут быть использованы для моделирования в большинстве стран. Необходимо определить значения адаптивных параметров с помощью процедуры идентификации, которая основана на регрессии реальных данных. К счастью, в Excel реализован многомерный оптимизатор, который способен решить проблему минимизации уравнения (17). Основная цель этой надстройки - найти решение, то есть значения переменных решения в модели, которое удовлетворяет всем ограничениям и максимизирует или минимизирует значение целевой ячейки (если оно есть). В нашем случае, мы минимизируем сумму квадратов отклонений между моделью и реальными данными, в качестве изменяемых ячеек используются адаптивные параметры  $a$ ,  $b$ ,  $c$ .

Любая другая среда программирования также может решить эту проблему, если доступна многомерная процедура оптимизации.

### 2.4 Описание алгоритма

На практике, алгоритм решения данной задачи оптимизации выглядит так. В первый столбец таблицы вводим дни с начала эпидемии (временной параметр). Во второй, для удобства, дату, которая соответствует первому столбцу. В третий столбец помещаем кумулятивное число зараженных. Из него нетрудно вычислить количество зараженных за день (путем вычитания из кумулятивного числа зараженных за  $N+1$  день кумулятивного числа зараженных за  $N$  день). В следующий столбец помещаем формулу логистической модели (уравнение 13). Адаптивные параметры помещаем в отдельные ячейки соответственно. В качестве независимой переменной  $t$  указываем первый столбец. В следующий столбец помещаем квадрат разности между кумулятивным числом заболевших и результатом предсказания модели. Суммируем получившиеся значения в отдельной ячейке. Далее, с помощью надстройки «Поиск решения», минимизируем получившуюся сумму квадратов, изменяя ячейки, где находятся адаптивные параметры. Вычислим точку перегиба в отдельной ячейке при помощи уравнения (16). Для модели Гомпертца алгоритм такой же. Для удобства, сумму квадратов отклонений для логистической модели и для модели Гомпертца можно просуммировать в отдельной ячейке и в качестве целевой функции использовать ее.



	A	B	C	D	E	F	G
1	<b>Логистическая модель</b>			Математическое описание логистической модели		<b>Модель Гомперт</b>	
2	a	177515,971221		$y = \frac{a}{1+b \exp(-ct)} = \frac{a}{1+b e^{-ct}}$		a	
3	b	132,335156				b	
4	c	0,135607	Точка перегиба			c	
5					36,02570054		
6	Номер дня	Дата	Количество зараженных(кумулятивное)	Зараженные за день		Логистическая модель	Квадрат отклонений
7	1	22.фев	62			1523	2134659
8	2	23.фев	155	93		1742	2518795
9	3	24.фев	229	74		1992	3109001
10	4	25.фев	322	93		2278	3825384,639
11	5	26.фев	453	131		2604	4626044
12	6	27.фев	655	202		2976	5385384
13	7	28.фев	888	233		3400	6307756,825
14	8	29.фев	1128	240		3882	7586971
15	9	01.мар	1694	566		4432	7497857,852
16	10	02.мар	2036	342		5058	9129954,249
17	11	03.мар	2502	466		5768	10668291,97
18	12	04.мар	3089	587		6575	12151761,34

Рисунок 8. Скриншот документа Excel.

106	06.июн	234801			177503	
107	07.июн	234998			177504	
108	08.июн	235278			177506	
109	09.июн	235561			177507	
110	10.июн	235763			177508	
111	11.июн	236142			177509	
112	12.июн	236305			177510	
113	13.июн				177511	
					Логист.сумм	630 889 686,52

Рисунок 9. Скриншот документа Excel.

Параметры поиска решения

Оптимизировать целевую функцию:

До:  Максимум  Минимум  Значения:

Изменяя ячейки переменных:

В соответствии с ограничениями:

Сделать переменные без ограничений неотрицательными

Выберите метод решения:

Метод решения  
 Для гладких нелинейных задач используйте поиск решения нелинейных задач методом ОПГ, для линейных задач - поиск решения линейных задач симплекс-методом, а для негладких задач - эволюционный поиск решения.

Рисунок 10. Скриншот надстройки пакета Excel «Поиск решения».

### 3 Верификация алгоритмов

Далее будет представлено несколько графиков в соответствии с регионами. Первым графиком в каждом из регионов будет экспоненциальная модель, которая в той или иной степени точности демонстрирует динамику начала эпидемии, путем линейной аппроксимации в полулогарифмических координатах. После этого будет изображено еще по три рисунка, объясняющие зависимость точности моделей от количества входных данных, то есть с каждым последующим рисунком будет добавлено на две недели больше реальных данных. В случае России и Нижегородской области будет продемонстрирован прогноз на месяц вперед, то есть до середины июля.

#### 3.1 Верификация алгоритма (Италия)

Италия (60,5 млн. Жителей) является одной из самых страдающих наций в результате этой пандемии: почти 34000 официальных смертей в больницах (по состоянию на 5 июня 2020 года) и более 4000 коек для интенсивной терапии для лечения пациентов на пике чрезвычайной ситуации Covid-19 [21]. Аналогичным образом, Ломбардия (10 миллионов жителей) является самым многолюдным регионом Италии и наиболее пострадавшим и пораженным регионом Европы: более 13800 официальных смертей в больницах и почти 1400 коек в палатах интенсивной терапии [12].

До пандемии количество коек в отделениях интенсивной терапии для любых процедур в Ломбардии составляло около 700 [22], а в зимние месяцы эти койки обычно занимали 85-90% [12], Пандемия требовала удвоения этого числа в невероятном темпе, чтобы справиться с волной очень сложных пациентов, страдающих ДИП (то есть диффузной интерстициальной пневмонией). Этим пациентам требовалось все более интенсивное лечение, начиная от кислородных масок и заканчивая интубацией трахеи [8]. В некоторых больницах, таких как больница Лоди в Ломбардии до пандемии было 6 коек [2]. В течение нескольких недель количество кроватей стало 18, а в конечном итоге 27 на пике чрезвычайной ситуации (т.е. в 4,5 раза больше). Это потребовало огромных усилий с точки зрения нагрузки на врачей, медсестер и менеджеров, которые буквально преобразовали обычные палаты, операционные и палаты для восстановления, чтобы установить и эксплуатировать такое большое количество новых палат реанимации.

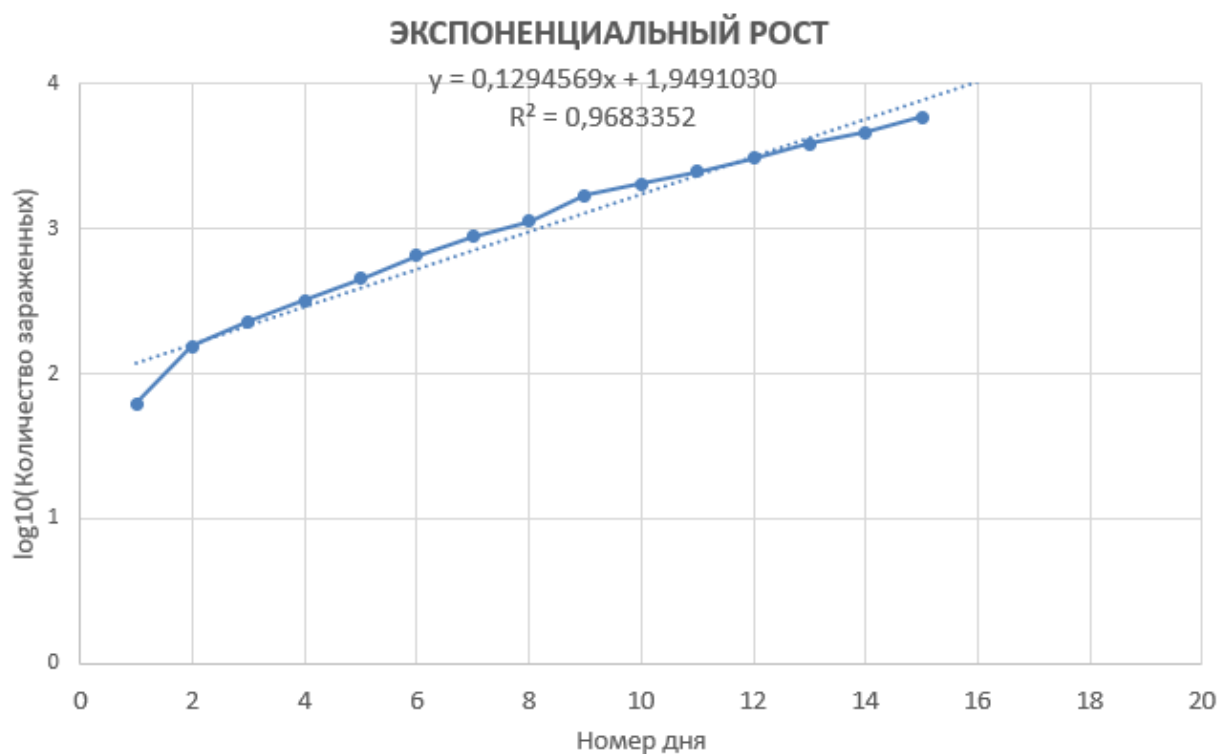


Рисунок 11. Экспоненциальный рост в начале эпидемии в Италии(первые 15 дней эпидемии).

На рисунке 11 показано, что количество заболевших в первые дни эпидемии растет экспоненциально и описываются уравнением (1), поскольку реальные данные по заболевшим лежат на аппроксимирующей прямой линии. Тут вертикальная ось является логарифмической, что означает, что ее значения представляют собой степени числа десять.

Введем в таблицу данные с момента начала эпидемии (22 февраля) до 18 апреля и проанализируем получившийся график.

Диаграмма 1- Италия

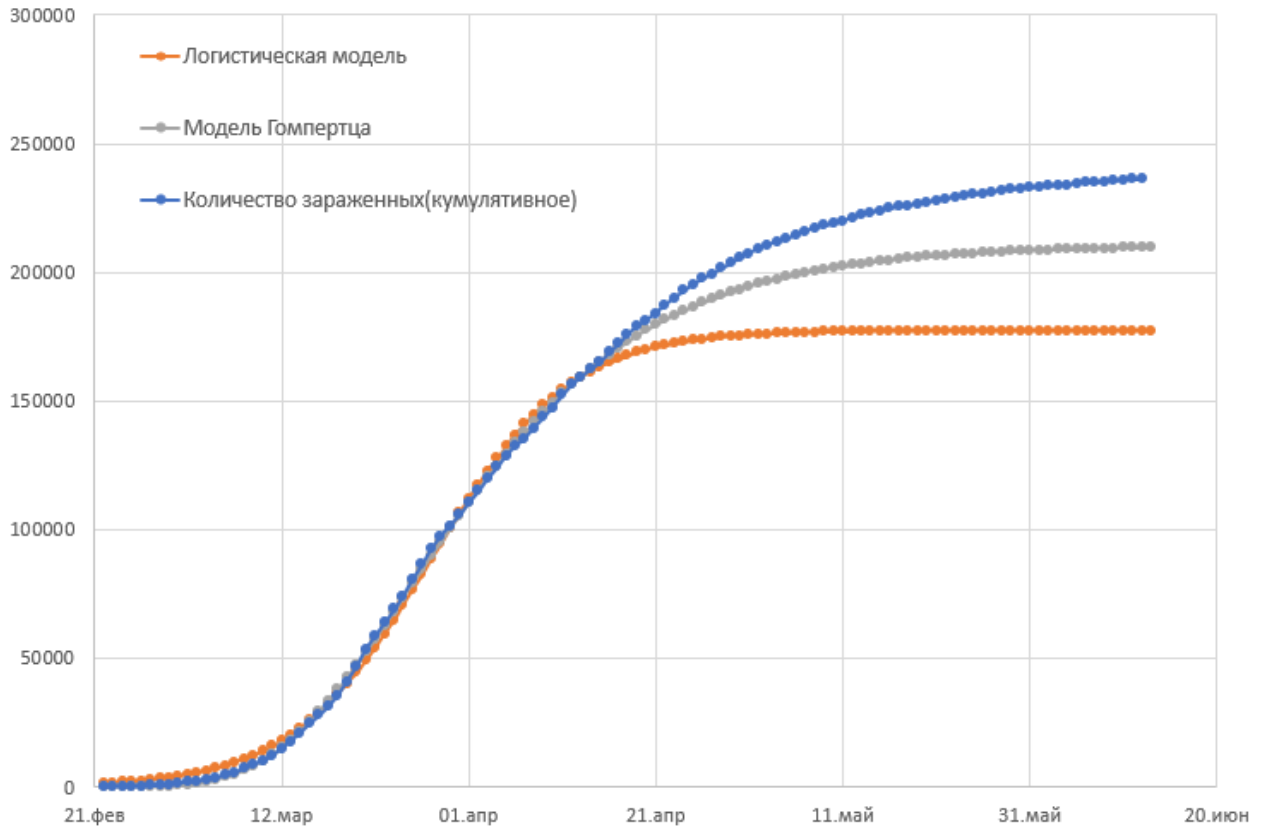


Рисунок 12. График заболеваемости в Италии, последние данные 18 апреля. Можно заметить, что модель Гомпертца даёт существенно более верный прогноз. Добавим в таблицу данные на две недели вперед.

Диаграмма 1- Италия

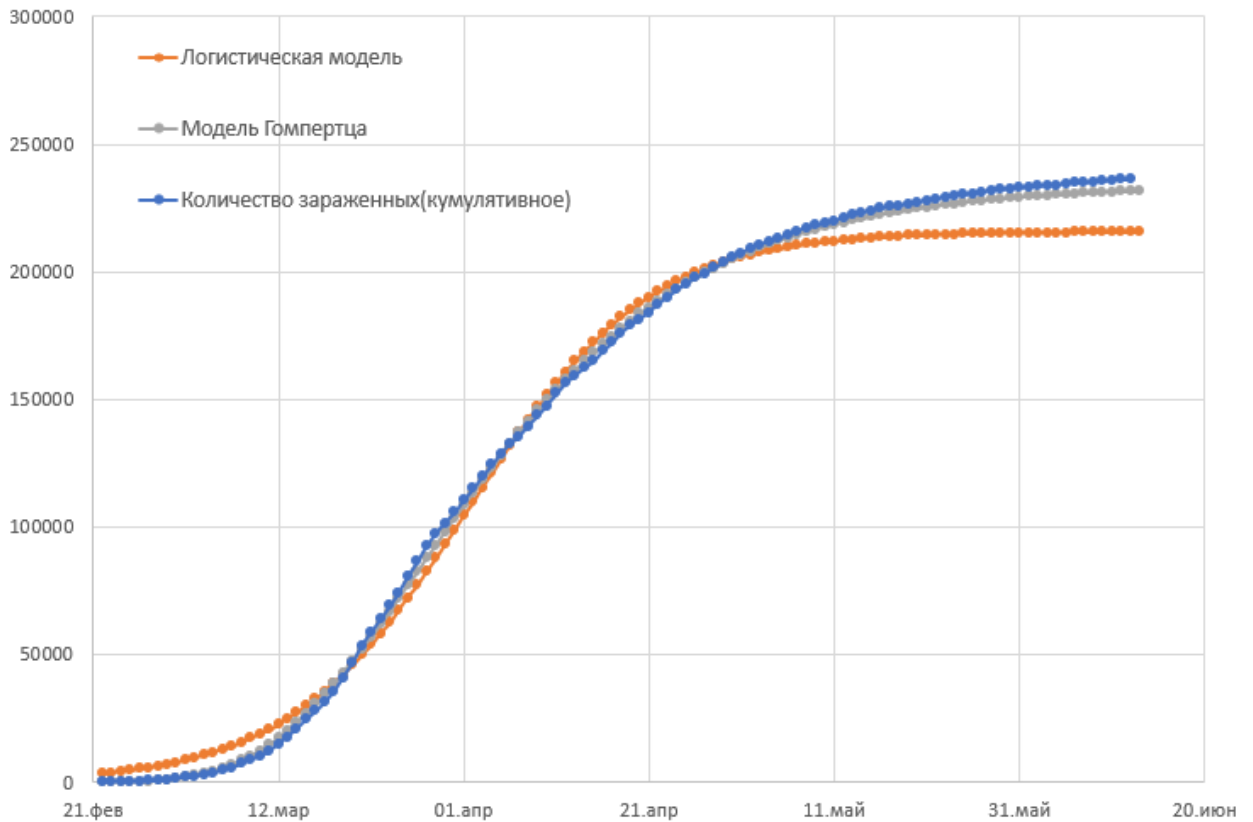


Рисунок 13. График заболеваемости в Италии. Последние данные 11 мая.

Для тех стран или регионов, где эпидемия уже почти побеждена, как видно, обе модели дают хороший прогноз, но модель Гомпертца ближе к реальным данным. Добавим данные ещё на две недели.

Диаграмма 1- Италия

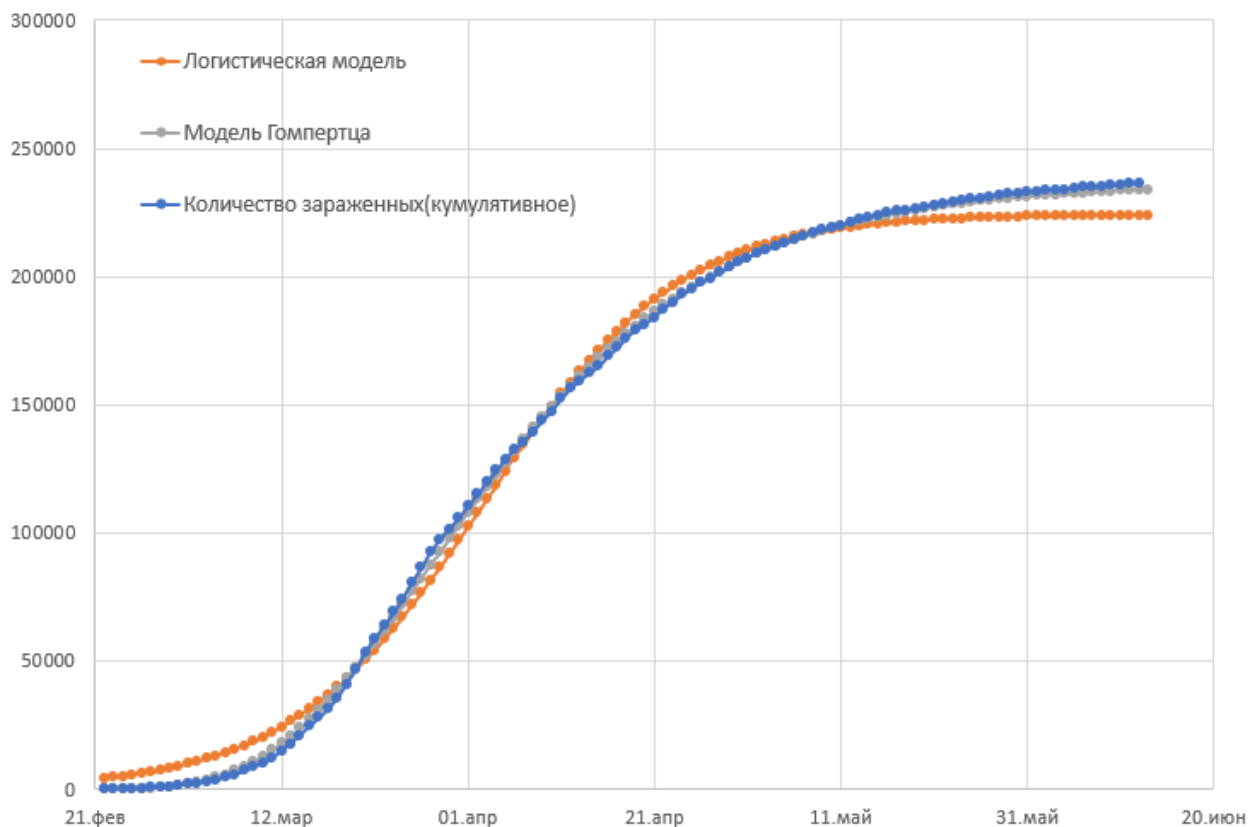


Рисунок 14. График заболеваемости в Италии. Последние данные 25 мая.

В Итальянском сценарии развития эпидемии погрешность обеих моделей незначительна. Но модель Гомпертца, все же была ближе. Она предсказывает что асимптотическое значение заболевших не привесит отметку в 236500 человек.

### 3.2 Верификация алгоритма (Россия)

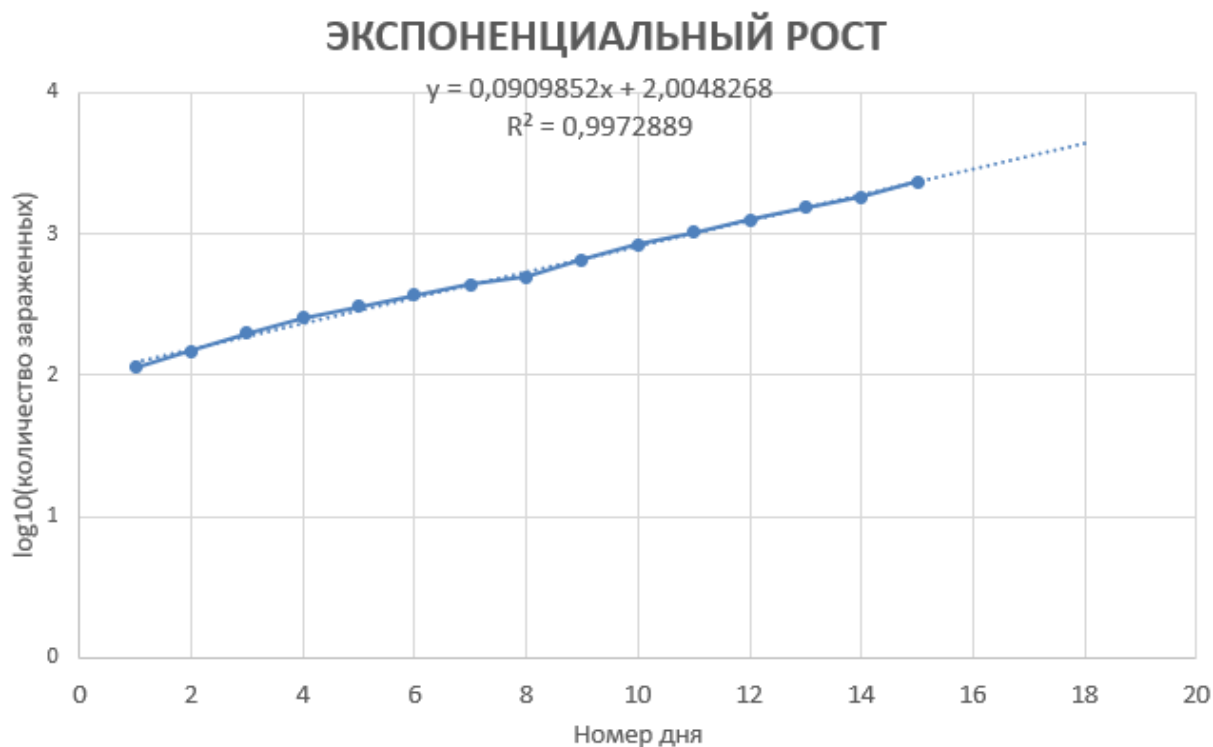


Рисунок 15. Экспоненциальный рост в начале эпидемии в России(первые 15 дней эпидемии).

Как видно на графике, кривая роста числа заболевших практически полностью соответствует росту по экспоненте. Это подтверждает высокий R- квадрат.



Диаграмма 1- Россия

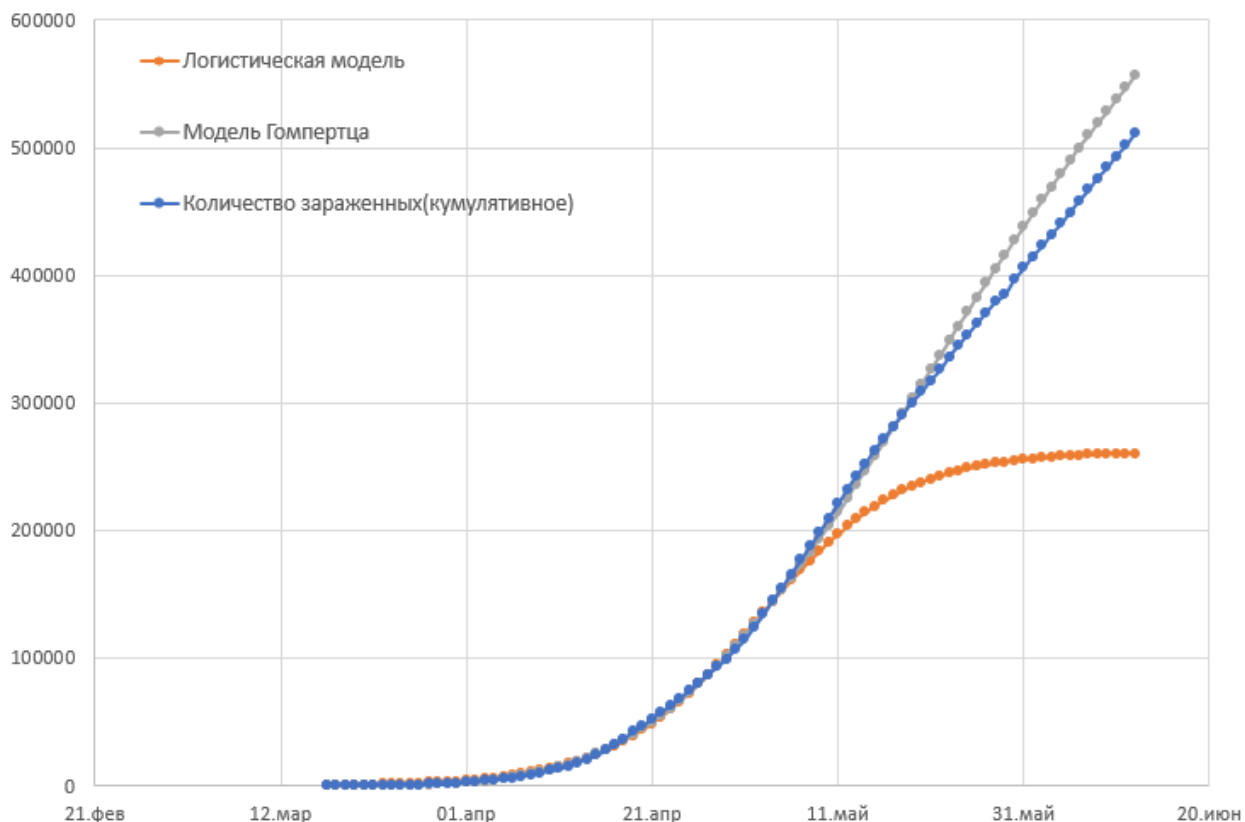


Рисунок 16. График заболеваемости в России. Последние данные 6 мая.

В условии недостатка данных, логистическая модель Ферхюльста даёт некорректный прогноз, но достаточно хорошо справляется модель Гомпертца, которая несколько завышает прогнозируемые показатели. Добавим в таблицу данные на две недели вперед.

Диаграмма 1- Россия

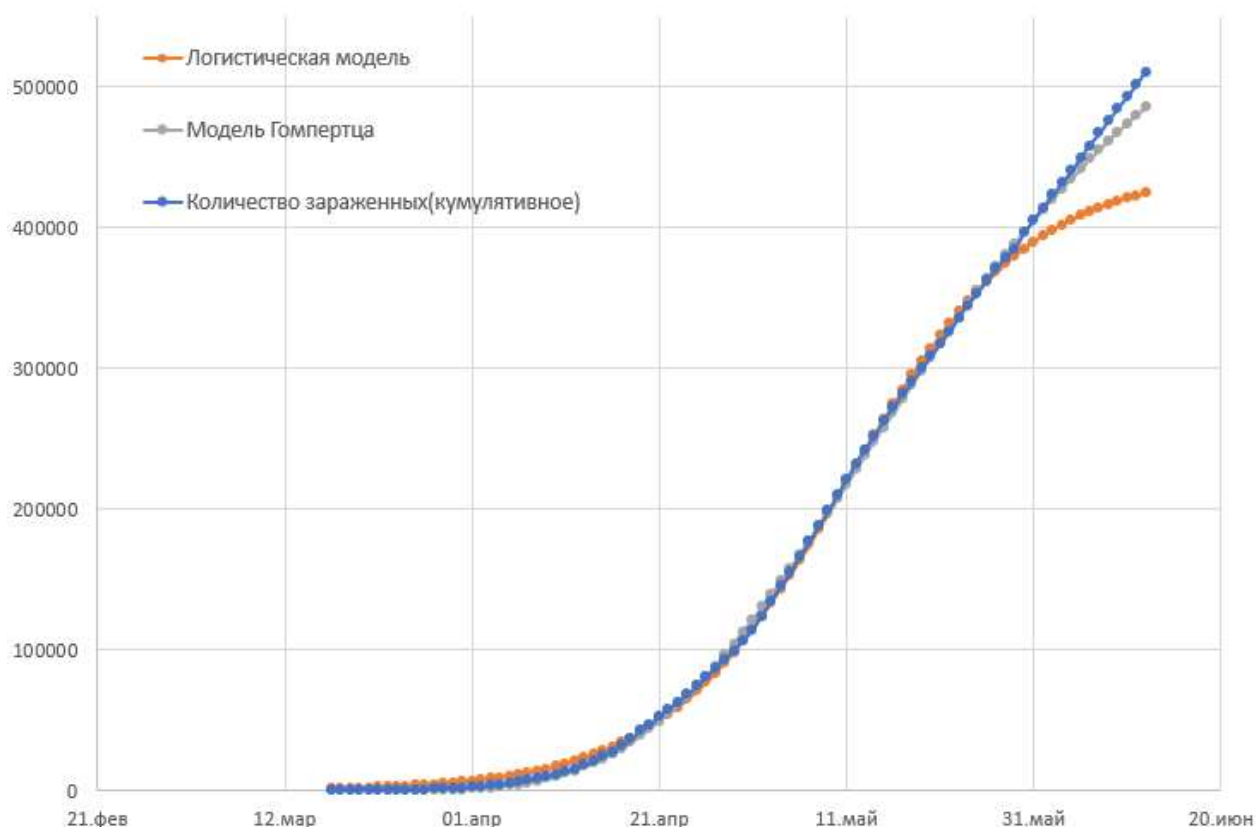


Рисунок 17. График заболеваемости в России. Последние данные 20 мая.

После того как добавили данные, погрешность моделирования стала меньше. Добавим данные ещё на две недели и сделаем небольшой прогноз.

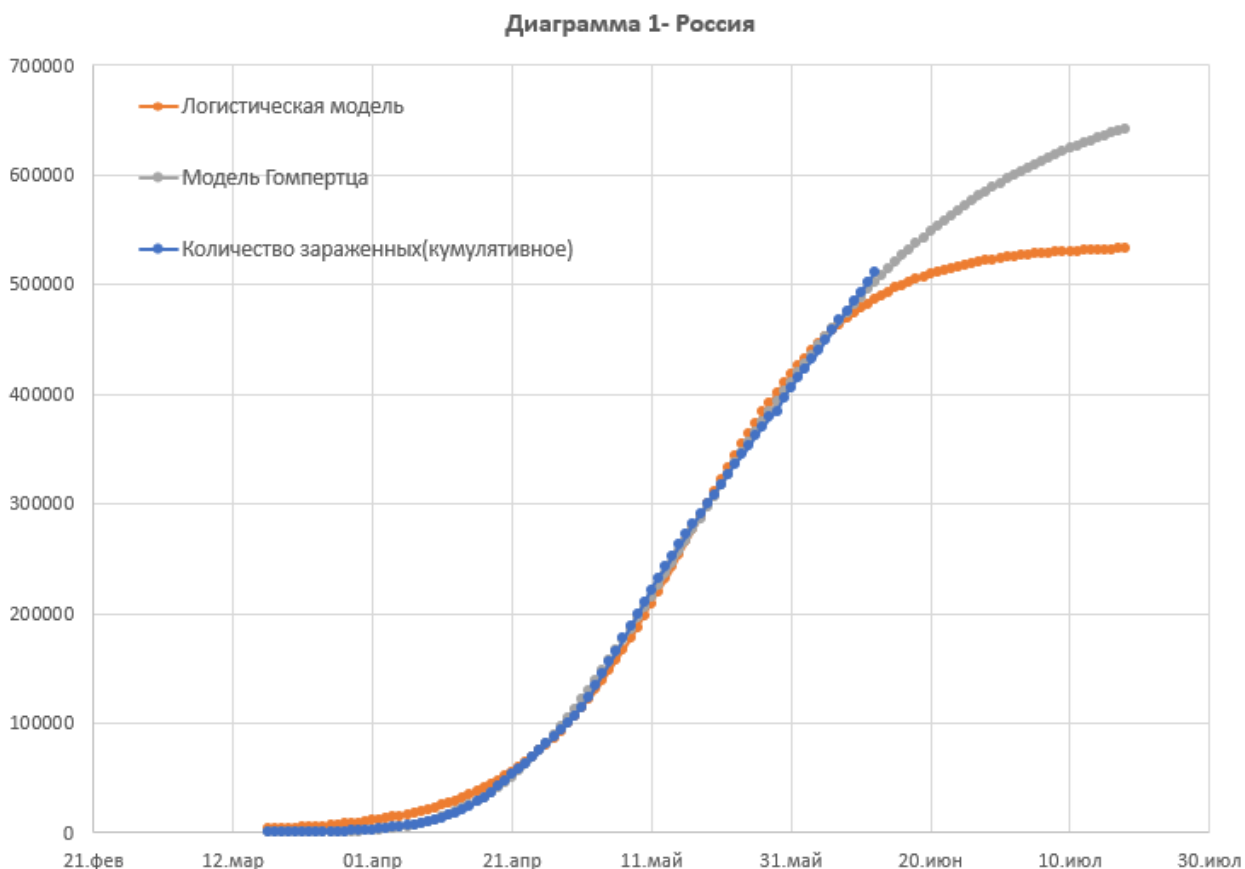


Рисунок 18. График заболеваемости в России. Последние данные 12 июня.

По сравнению с Италией, в России были введены не такие жесткие карантинные меры, из-за этого, кривая зараженных на графике более длинная и пологая, чем в Италии. А прогноз модели Гомпертца говорит нам, что к середине июля будет 650000 заболевших.

Изм.	Лист	№ докум.	Подп.	Дата

### 3.3 Верификация алгоритма (Нижегородская область)

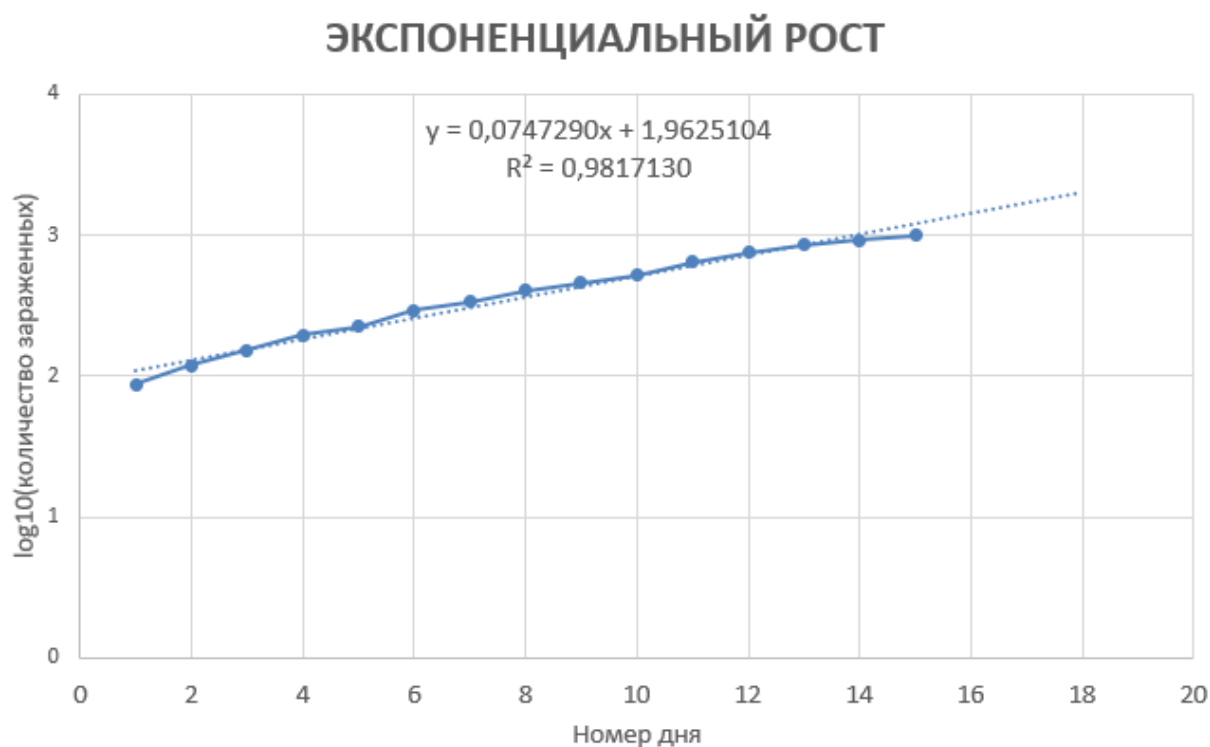


Рисунок 19. Экспоненциальный рост в начале эпидемии в Нижегородской области(первые 15 дней эпидемии).

Значения заболевших в начале эпидемии в Нижегородской области так же, как и в России в целом, хорошо соответствуют интерполяции.

Диаграмма 1- Нижегородская область

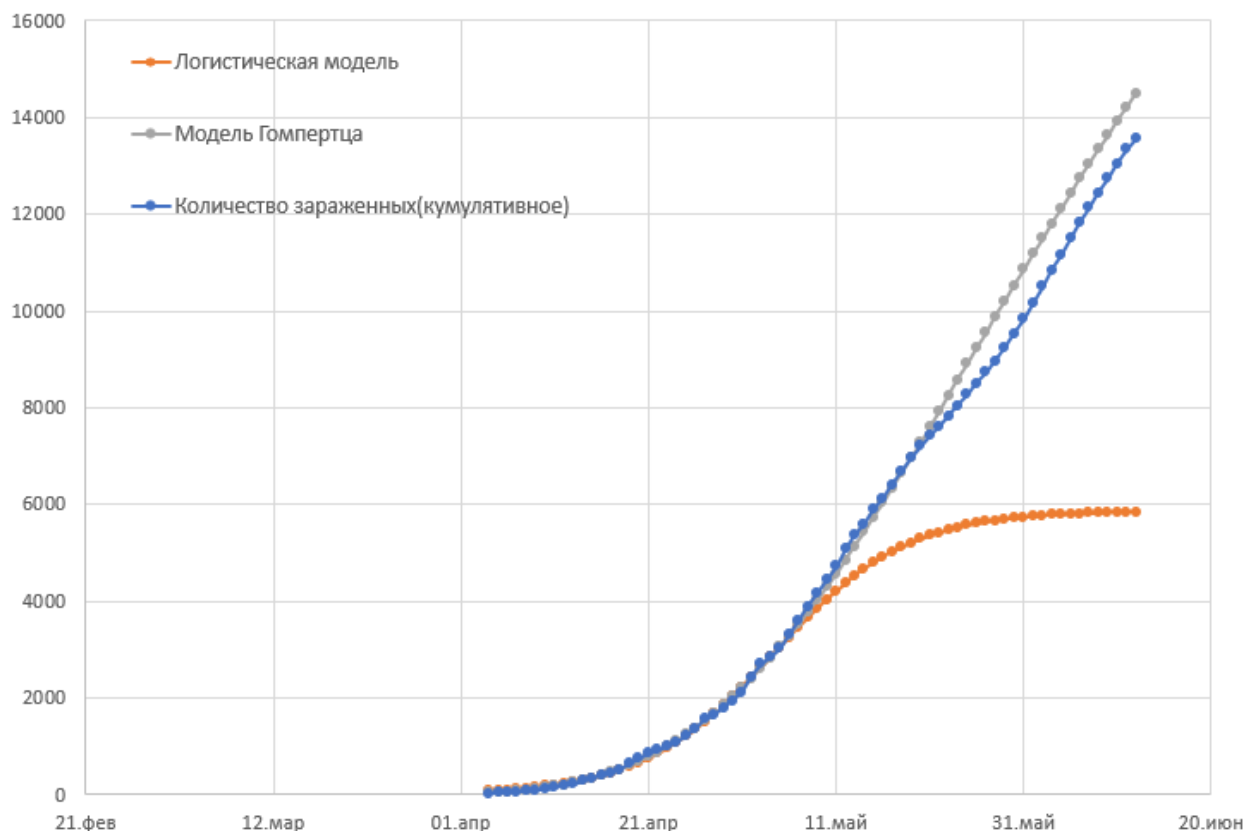


Рисунок 20. График заболеваемости в Нижегородской области. Последние данные 6 мая.

В условиях недостатка данных, как и в случае с Россией, модель Гомпертца чуть завышает показатели, а модель Ферхюльста дает нереалистичный прогноз. Добавим данные в таблицу с момента начала эпидемии до 6 мая.

Диаграмма 1- Нижегородская область

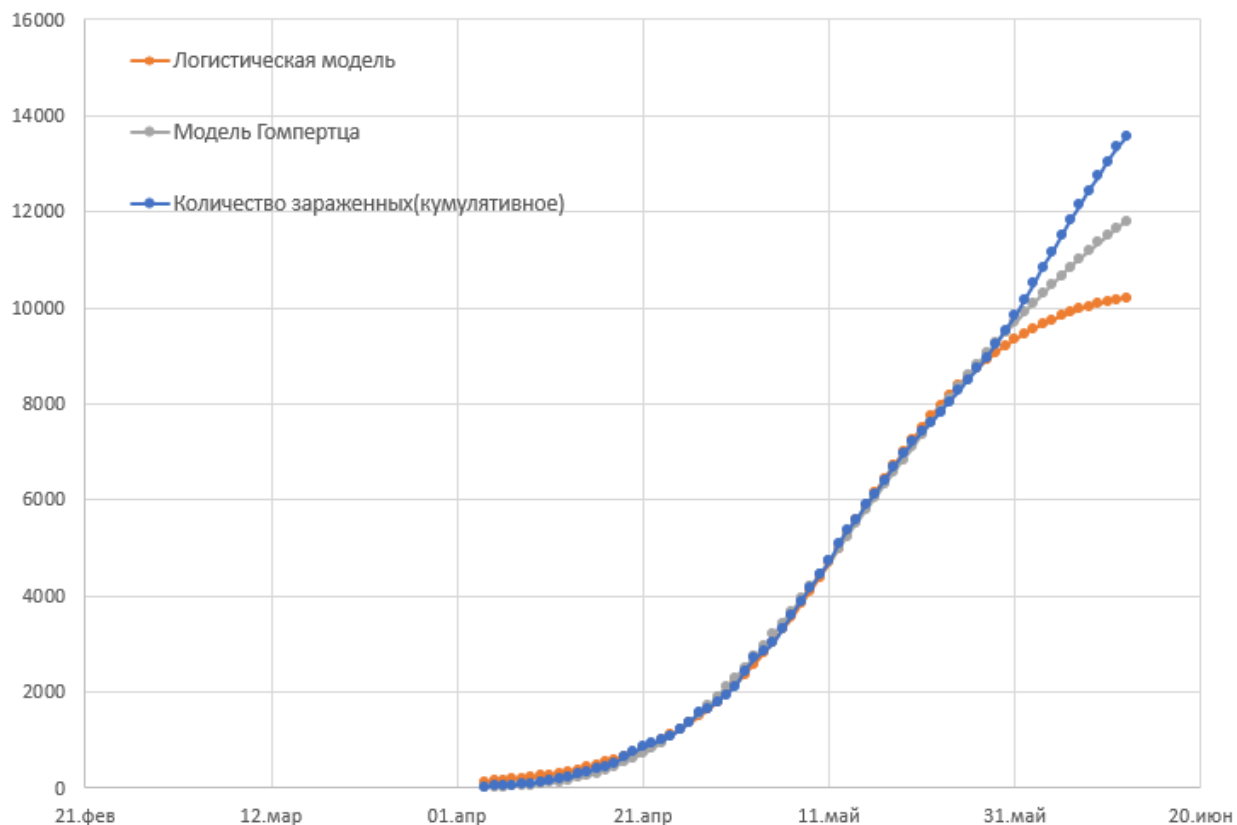


Рисунок 21. График заболеваемости в Нижегородской области. Последние данные 20 мая.

Добавим данные до 12 июня и сделаем прогноз на месяц.

На данном графике реальные данные обрываются как раз перед существенным скачком числа заболевших, из-за этого прогнозы обеих моделей были искажены.

Изм.	Лист	№ докум.	Подп.	Дата

Диаграмма 1- Нижегородская область

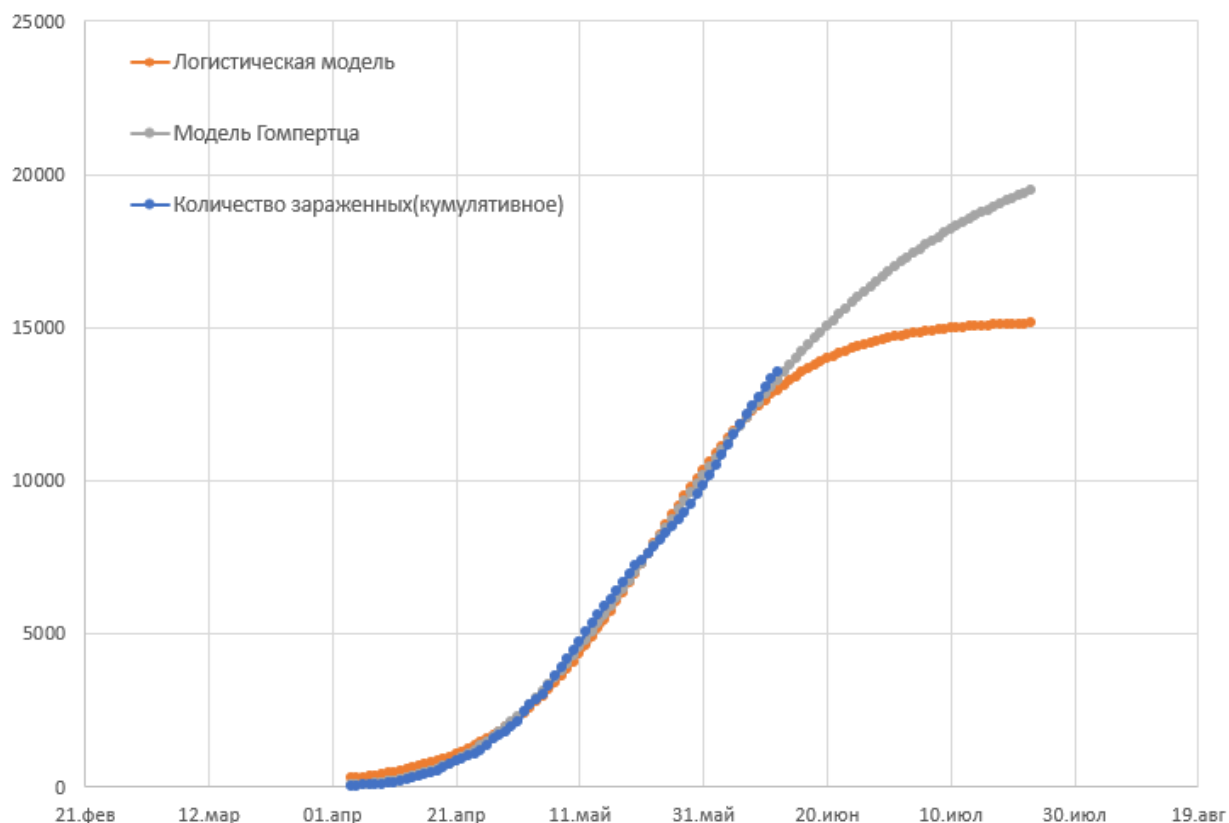


Рисунок 22. График заболеваемости в Нижегородской области. Последние данные 12 июня.

В Нижегородской области ситуация такая же, как и в целом по России. Тут модель Гомпертца тоже даёт существенно более точные результаты. Это связано с тем, что модель Гомпертца не симметрична, в связи с этим затухание эпидемии происходит не так быстро, как начало. Прогнозируемое значение количества зараженных в середине июля по ней – почти 19000 человек.

#### 4 Сравнение моделей

Как можно заметить, модель Гомпертца дает существенно более точный прогноз как на краткосрочные периоды, так и на длительные, особенно в местах, где распространение эпидемии протекает продолжительно. Обе модели хорошо справляются с описанием уже имеющихся данных, но в условиях длительного мониторинга, на экспоненциальном этапе, логистическая модель Ферхюльста отклоняется в большую сторону от реальных значений. Легко заметить, что это прогнозы, основанные на математической модели, которая зависит от наличия надежных данных. Чем больше реальных данных, тем лучше. Действительно, каждый день в национальных и региональных отчетах публикуются новые данные, и можно делать более надежные прогнозы. Кроме того, модель прогнозов основана на предположении, что будущие значения будут зарегистрированы и предоставлены в соответствии с теми же условиями с точки зрения мер социального дистанцирования и условий сбора данных.

					ВКР-ИГТУ-01.03.02-16-САИ-010-2020	Лист
Изм.	Лист	№ докум.	Подп.	Дата		40



## Заключение

В работе представлены и проанализированы несколько регрессионных моделей для прогнозирования двух наиболее важных переменных в пандемии с точки зрения принятия решений и планирования действий в чрезвычайных ситуациях с точки зрения наиболее важных, во время эпидемии, характеристик. Это число заболевших и число погибших. Эти модели могут быть применены к различным регионам и странам, поскольку феномен пандемии имеет одинаковые качественные характеристики. Два или три, адаптивных параметра позволяют количественно описать динамику заболеваемости и смертности разных регионов и стран. Действительно, каждый регион и каждая страна характеризуются различными особенностями, связанными с их территориальной природой, распределением населения (с точки зрения возраста, плотности, образа жизни, взаимодействия людей, семейных привычек) и политических решений. Математика, лежащая в основе предлагаемых моделей, довольно проста и может быть реализована в файле Excel, который может использоваться большинством лиц, принимающих решения, и врачей. Основываясь на реальных данных, которые ежедневно публикуются в открытых источниках для большинства стран/регионов, эти регрессионные модели можно настроить с точки зрения их адаптивных параметров и использовать для прогнозирования тенденций в заболеваемости и смертности на краткосрочных или долгосрочных периодах. Эти же модели также могут быть адаптированы для отслеживания других переменных, если надежность этих переменных достаточно высока. Две основные модели, предложенные в этой работе (то есть, логистическая и Гомпертца), могут использоваться для качественных и количественных целей в чрезвычайной ситуации Covid-19. Они играют две разные роли. Во-первых, они отслеживают реальные данные и позволяют дифференцировать модели, чтобы найти наиболее надежную, но также они позволяют понять, начинаются ли неожиданные тенденции. Они являются непрерывными и имеют аналитические производные любого порядка. Это означает, что можно наблюдать и отслеживать не только кумулятивные значения, но и суточные колебания, чтобы понять, происходят ли какие-либо изменения. Во-вторых, эти модели могут быть использованы для прогнозирования развития явления на краткосрочных или долгосрочных горизонтах. В первые трудные дни наиболее важны прогнозы на короткие промежутки времени, чтобы понять время с которым растет количество зараженных чтобы выделить / подготовить подходящие ресурсы, чтобы справиться с последующей волной пациентов в больницах.

					ВКР-НГТУ-01.03.02-16-САИ-010-2020	Лист
Изм.	Лист	№ докум.	Подп.	Дата		41

## Список литературы

1. Allen L. et al. A mathematical analysis and simulation of a localized measles epidemic //Applied Mathematics and Computation. – 1990. – Т. 39. – №. 1. – P. 61-77.
2. Anderson RM, May RM. Infectious disease of humans: Dynamics and control / Anderson RM, May RM.- Oxford University Press, Oxford-New York, 1991.- 768p.
3. Bernoulli D. Essai d'une nouvelle analyse de la mortalité causée par la petite vérole, et des avantages de l'inoculation pour la prévenir //Histoire de l'Acad., Roy. Sci.(Paris) avec Mem. – 1760. – P. 1-45.
4. Brauer F. Mathematical epidemiology is not an oxymoron //BMC Public Health. – 2009. – Т. 9. – №. 1. – P. 1-11.
5. Cowling B. J. et al. Impact assessment of non-pharmaceutical interventions against coronavirus disease 2019 and influenza in Hong Kong: an observational study //The Lancet Public Health. – 2020.
6. Cutuli A. COVID-19: the Lodi model (in Italian)// Esanum. - 2020.
7. Deguen S., Thomas G., Chau N. P. Estimation of the contact rate in a seasonal SEIR model: application to chickenpox incidence in France //Statistics in medicine. – 2000. – Т. 19. – №. 9. – P. 1207-1216.
8. Desai A. N., Aronoff D. M. Masks and coronavirus disease 2019 (covid-19) //Jama. – 2020.
9. Fraser C. et al. Pandemic potential of a strain of influenza A (H1N1): early findings //science. – 2009. – Т. 324. – №. 5934. – P. 1557-1561.
10. Golubev A. Exponentially modified Gaussian (EMG) relevance to distributions related to cell proliferation and differentiation //Journal of theoretical biology. – 2010. – Т. 262. – №. 2. – P. 257-266.
11. Grasselli G. et al. Baseline characteristics and outcomes of 1591 patients infected with SARS-CoV-2 admitted to ICUs of the Lombardy Region, Italy //Jama. – 2020. – Т. 323. – №. 16. – P. 1574-1581.
12. Grasselli G., Pesenti A., Cecconi M. Critical care utilization for the COVID-19 outbreak in Lombardy, Italy: early experience and forecast during an emergency response //Jama. – 2020. – Т. 323. – №. 16. – P. 1545-1546.
13. Guinard A. et al. Outbreak of influenza a (H1n1) v without travel history in a school in the Toulouse district, France. – 2009.
14. Gupte M. D. et al. Modelling epidemiology of leprosy //Indian journal of leprosy. – 2000. – Т. 72. – №. 3. – P. 305-316.

15. Gurav Y. K. et al. Pandemic influenza A (H1N1) 2009 outbreak in a residential school at Panchgani, Maharashtra, India //Indian J Med Res. – 2010. – Т. 132. – №. 1. – P. 67-71.
16. Hosmer Jr D. W., Lemeshow S., Sturdivant R. X. Applied logistic regression. – John Wiley & Sons, 2013. – Т. 398.
17. Hosmer Jr D. W., Lemeshow S., May S. Applied survival analysis: regression modeling of time-to-event data. – John Wiley & Sons, 2011. – Т. 618.
18. Kermack W. O., McKendrick A. G. Contributions to the mathematical theory of epidemics. II.—The problem of endemicity //Proceedings of the Royal Society of London. Series A, containing papers of a mathematical and physical character. – 1932. – Т. 138. – №. 834. – P. 55-83.
19. Kermack W. O., McKendrick A. G. Contributions to the mathematical theory of epidemics. III.—Further studies of the problem of endemicity //Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character. – 1933. – Т. 141. – №. 843. – P. 94-122.
20. Kermack M., McKendrick A. Contributions to the mathematical theory of epidemics. Part I //Proc. r. soc. a. – 1927. – Т. 115. – №. 5. – P. 700-721.
21. Livingston E., Bucher K. Coronavirus disease 2019 (COVID-19) in Italy //Jama. – 2020. – Т. 323. – №. 14. – P. 1335-1335.
22. Manca D. Analysis of the number growth of ICU patients with Covid-19 in Italy and Lombardy //European Society of Anaesthesiology. - 2020.
23. Manca D. Dynamics of ICU patients and deaths in Italy and Lombardy due to Covid-19// European Society of Anaesthesiology. - 2020.
24. Meima A. et al. SIMLEP: a simulation model for leprosy transmission and control //International Journal of Leprosy and Other Mycobacterial Diseases. – 1999. – Т. 67. – P. 215-236.
25. Murthy S., Gomersall C. D., Fowler R. A. Care for critically ill patients with COVID-19 //Jama. – 2020. – Т. 323. – №. 15. – P. 1499-1500.
26. Onder G., Rezza G., Brusaferro S. Case-fatality rate and characteristics of patients dying in relation to COVID-19 in Italy //Jama. – 2020. – Т. 323. – №. 18. – P. 1775-1776.
27. Panik M. J. Growth curve modeling: theory and applications. – John Wiley & Sons, 2014.
28. Panovska-Griffiths J. Can mathematical modelling solve the current Covid-19 crisis? //BMC Public Health. - 2020.

29. Poston J. T., Patel B. K., Davis A. M. Management of critically ill adults with COVID-19 //Jama. – 2020. – Т. 323. – №. 18. – P. 1839-1841.
30. Remuzzi A., Remuzzi G. COVID-19 and Italy: what next? //The Lancet. – 2020.
31. Ross R. The prevention of malaria. 2 edition //John Murray. - London, 1911. -311 p.
32. Rvachev L. A. Modelling experiment of a large-scale epidemic by means of a computer //Doklady Akademii Nauk. – Russian Academy of Sciences, 1968. – Т. 180. – №. 2. – P. 294-296.
33. Rvachev L. A., Longini Jr I. M. A mathematical model for the global spread of influenza //Mathematical biosciences. – 1985. – Т. 75. – №. 1. – P. 3-22.
34. Shil P. et al. Transmission dynamics of novel influenza A/H1N1 2009 outbreak in a residential school in India //Current Science. – 2011. – Т. 100. – №. 8. – P. 1177-1183.
35. Small M., Tse C. K. Clustering model for transmission of the SARS virus: application to epidemic control and risk assessment //Physica A: Statistical Mechanics and its Applications. – 2005. – Т. 351. – №. 2-4. – P. 499-511.
36. Small M., Tse C. K. Small world and scale free model of transmission of SARS //International Journal of Bifurcation and Chaos. – 2005. – Т. 15. – №. 05. – P. 1745-1755.
37. Smith A. et al. An outbreak of influenza A (H1N1) v in a boarding school in South East England, May-June 2009 //Eurosurveillance. – 2009. – Т. 14. – №. 27. – P. 19263.
38. Stone L., Olinky R., Huppert A. Seasonal dynamics of recurrent epidemics //Nature. – 2007. – Т. 446. – №. 7135. – P. 533-536.
39. Wang J. et al. Spatial dynamics of an epidemic of severe acute respiratory syndrome in an urban area //Bulletin of the World Health Organization. – 2006. – Т. 84. – P. 965-968.