

Санкт-Петербургский государственный университет

*Денисова Юлия Сергеевна*

**Выпускная квалификационная работа**  
***Цепочки распространения эпидемиологических***  
***процессов в разных странах***

Уровень образования:

Направление: 01.04.02 Прикладная математика и информатика  
Образовательная программа: ВМ.5505.2021 Математическое и  
информационное обеспечение экономической деятельности

Научный руководитель  
Кандидат физ.-мат. наук,  
доцент кафедры математической  
теории экономических решений  
Раевская Анастасия Павловна

Рецензент:  
Инженер, Отдел развития и  
проектирования информационных систем,  
Управление развития информационных  
систем в сфере городской инфраструктуры  
и безопасного города  
Фурсов Дмитрий Викторович

Санкт-Петербург  
2023

# Содержание

Введение.....	3
Постановка задачи.....	6
1. Основные направления исследования моделирования заболеваний.....	8
1.1. Обзор литературы о влиянии различных заболеваний на экономику стран .....	8
1.2. Обзор исследований по математическому моделированию заболеваний и эпидемий .....	21
2. Методологическая специфика моделирования заболеваний.....	32
2.1. Описание основных источников данных.....	32
2.2. Описание модель SIR.....	33
2.3. Описание модели CIR.....	36
3. Анализ математических моделей заболевания COVID-19 .....	39
3.1. Цепочки распространения пандемии COVID-19 .....	39
3.2. Построение прогноза с помощью SIR-модели.....	44
3.3. Построение прогноза с помощью модели CIR.....	47
Заключение .....	68
Список литературы .....	70
Приложение .....	78

## Введение

Инфекционные заболевания являются основной причиной смерти во всем мире, и в прошлом они убили гораздо больше людей, чем все войны. Например, эпидемия «испанки» 1918–1919 годов унесла более 50 000 000 жизней во всем мире, а ежегодные сезонные эпидемии гриппа уносят до 35 000 жизней во всем мире.

Пандемии и эпидемии вызывают как краткосрочные, так и долгосрочные экономические последствия для стран по всему миру. Усилия по сдерживанию пандемии включают введение карантина, подготовку медицинских учреждений, изоляцию зараженных и отслеживание контактов с привлечением ресурсов общественного здравоохранения, человеческих ресурсов и затрат на реализацию. Сюда также входят расходы системы здравоохранения на закупку расходных материалов, таких как антибиотики, медикаменты и средства индивидуальной защиты.

Эпидемии также могут привести к сокращению налоговых поступлений и увеличению расходов, что вызывает финансовые проблемы, особенно в странах с уровнем дохода ниже среднего. Экономические потрясения часто связаны с нехваткой рабочей силы из-за болезней, роста смертности и поведения, вызванного страхом. Помимо нехватки рабочей силы, нарушения работы транспорта, закрытие рабочих мест, ограничения торговли и поездок, а также закрытые сухопутные границы являются причинами экономического спада из-за пандемии.

Изучение данных об инфекционных заболеваниях началось с работы Джона Граунта (англ. John Graunt), описанной в его книге 1662 года «Естественные и политические наблюдения над списками умерших». Эти списки представляли собой еженедельные записи о количестве и причинах смерти в лондонских приходах. Он проанализировал различные причины

смерти и дал метод оценки сравнительных рисков смерти от различных болезней.

Первую математическую модель для исследования эффективности профилактических прививок против оспы создал швейцарский математик и физик Даниил Бернулли в 1760 году. Идея продолжила свое развитие в работах английского ученого Уильяма Фарра (англ. William Farr), который смоделировал показатели смертности населения Уэльса от эпидемии оспы в 1837–1839 годах. На основании данной математической модели были выявлены статистические закономерности распространения эпидемии, что позволило составить прогноз.

В начале XX века Уильям Хеймер (англ. William Hamer) и Рональд Росс (англ. Ronald Ross) применили закон действующих масс для объяснения эпидемического поведения. В 1906 г. Хеймер предположил, что распространение инфекции должно зависеть от числа восприимчивых особей и числа заразных особей. Рональд Росс был удостоен второй Нобелевской премии по медицине в 1902 году за демонстрацию динамики передачи малярии между комарами и людьми. Труды Хеймера и Росса, наряду с работами Андерсона Маккендрика (англ. Anderson G. McKendrick) и Уильяма Кермака (англ. William O. Kermack), послужили теоретической основой для дальнейших исследований в области математического моделирования эпидемий.

Работа Кермака и Маккендрика, опубликованная в 1927 году, оказала большое влияние на основу моделирования. Их модель SIR до сих пор используется для моделирования эпидемий инфекционных заболеваний. На основе этой модели стало развиваться исследование и прогнозирование эпидемиологических процессов.

Математическое представление и анализ инфекционных заболеваний занимают центральное место в эпидемиологии с момента ее зарождения как дисциплины. Математический анализ и модели способны объяснить ранее

полученные труднообъяснимые наблюдения и играют важную роль в стратегиях общественного здравоохранения во многих странах.

Математические модели используют, чтобы связать биологический процесс передачи и возникающую динамику инфекции на уровне популяции. В простейшем случае эпидемию можно описать, перечислив, кто кого заразил и когда.

Целью данной работы является построение цепочек распространения эпидемиологических процессов, связанных с COVID-19, проведение анализа и прогнозирования протекания заболевания в странах мира.

Исходя из поставленной цели, были определены следующие задачи работы:

1. Систематизация литературные источники по теме исследования;
2. Изучение имеющиеся модели исследования инфекционных заболеваний;
3. Сбор и обработка данных об основных показателях эпидемии заболевания COVID-19;
4. Построение цепочек распространения заболевания;
5. Разработать программное обеспечение для анализа и прогнозирования SIR-модели;
6. Разработать программное обеспечение для анализа и прогнозирования балансовой модели.

## Постановка задачи

Для достижения поставленной цели были собраны сведения об эпидемии COVID-19, собранные ресурсным центром Джона Хопкинса. Данные включают в себя количество подтвержденных случаев заболевания с момента начала эпидемии по текущий момент, количество умерших от этого заболевания, а также количество выздоровевших. Для исследования были взяты данные с 1 февраля 2020 года по 1 января 2021 года. Имеющиеся страны разбиваем на пять групп по континентам, а именно на Европу, Азию, Африку, Южную и Северную Америки. Каждая группа рассматривается отдельно.

Сначала находим процентный прирост, то есть изменение количества случаев заболевания во времени, выраженное в процентах. Для этого из текущего значения вычитаем предыдущее, эту разность делим на текущее значение и умножаем на сто процентов.

Затем по каждой стране в группе эмпирическим методом определяем день, в который процентный прирост случаев заболевания достиг примерно 20%, при этом после данного дня прирост, имеет тенденцию к снижению. Кратковременные всплески игнорируются. Назовем такой день контрольной точкой. Когда они определены, найдем страну, у которой эта точка достигается позднее, чем у остальных в группе. Такое государство является страной-последователем. Страны, имеющие контрольную точку примерно на 5–7 дней раньше будем считать странами-предшественниками. Таких государств может быть несколько. Затем уже для этих стран находим предшественников. Таким образом формируется цепочка, в которой отображена последовательность передачи заболевания на континенте.

Полученные цепочки будут использовать для моделирования количества выбывших и инфицированных в одной стране по данным стран-предшественников. Далее определяем даты, в которых процентный прирост

достигает уровня в двадцать, десять, пять и два процента. Будем считать, что промежуток между датами достижения десяти и пяти процентов для базовой страны относится к промежутку между временем перехода с двадцати до десяти процентов в этой же стране, как периоды у страны-последователя.

И наконец, сравним прогнозирование пандемии COVID-19 с помощью простой SIR-модели и балансовой модели CIR. Для построения прогноза был написан код на языке программирования Python, что позволило быстро обработать большой объем данных и упростить построение множества прогнозов.

# 1. Основные направления исследования моделирования заболеваний

## 1.1. Обзор литературы о влиянии различных заболеваний на экономику стран

Здоровье населения влияет на экономический рост при помощи множества различных механизмов. Существует ряд исследований, которые доказывают, что здоровье влияет многие аспекты экономической жизни людей. Хорошее здоровье граждан страны означает здоровых работников, а значит, и более высокую производительность труда, и более высокие доходы. Кроме того, это означает здоровых детей и растущую продолжительность жизни, что ведет к накоплению человеческого капитала. Увеличение ожидаемой продолжительности жизни также приводит к увеличению сбережений и инвестиций.

Инфекционные заболевания, наоборот, во многом ставят под угрозу экономический рост. Некоторые смертельные заболевания, такие как СПИД, могут привести к значительному сокращению предложения рабочей силы. Другие заболевания, такие как грипп, распространяются через социальное взаимодействие, что заставляет людей избегать такое общение и, следовательно, снижает как потребление, так и предложение труда. Другие инфекционные заболевания, такие как малярия в странах Африки, влияют на снижение человеческого капитала в регионе в результате высоких показателей заболеваемости и смертности при достаточно низких перспективах выздоровления. Также существуют болезни, которым подвержены в большей степени дети. Из-за подобных заболеваний может пострадать долгосрочный экономический рост.

Многие исследователи затрагивали вопрос влияния болезней на экономику и социальную сферу [6, 14, 16, 28], в том числе на человеческий капитал [10, 27], предлагая различные механизмы для выявления причинно-



следственной связи между экономическим ростом и инфекционными заболеваниями.

В исследовании [43] использовал данные 88 стран за период с 1999 по 2018 год. Чтобы проверить взаимосвязь между экономическим ростом и инфекционными заболеваниями, использовалась эмпирическая модель

$$GDP_{it} = \beta_0 + \beta_1 diseases_{it} + \beta_2 X_{it} + \varepsilon_{it},$$

где  $GDP_{it}$  — это годовой темп роста ВВП на душу населения с индексом  $i$ , обозначающим страну;  $diseases_{it}$  обозначает меру здоровья;  $X_{it}$  представляет собой матрицу  $k$  контрольных переменных:  $Pop$  относится к рабочей силе, которая определяется приростом населения (годовой %);  $inf$  относится к инфляции, которая аппроксимируется дефлятором ВВП (годовой %);  $GOVEXP$  относится к расходам на конечное потребление органов государственного управления (% от ВВП);  $INVST$  относится к внутренним инвестициям, измеряемым валовым накоплением основного капитала (% от ВВП),  $\varepsilon_{it}$  является погрешностью.

В исследовании автор использовал три показателя здоровья: количество смертей от малярии; количество умерших от кишечных инфекционных заболеваний и количество смертей от болезни, вызванной вирусом иммунодефицита человека (ВИЧ). Для того, чтобы доказать существование влияния переменных друг на друга, в длительном периоде был использован тест Педрони.

Также автор оценил модели FMOLS (Fully Modified Ordinary Least Squares) и DOLS (Dynamic Ordinary Least Squares), чтобы найти долгосрочную связь между всеми переменными в моделях и использовал модель VECM (Vector Error Correction Model), чтобы изучить причинно-следственную связь между экономическим ростом и болезнями и между открытостью торговли и инфекционными заболеваниями.

В результате в статье было доказано, что инфекционные заболевания ухудшают экономический рост в долгосрочной перспективе. Все три заболевания (малярия, ВИЧ, кишечные инфекционные заболевания) влияют на снижение ВВП на душу населения во всех странах, но более интенсивно в развивающихся странах, чем в развитых. Это объясняется более развитой системой здравоохранения во втором случае. Кроме того, было выявлено, что в долгосрочной перспективе государственные расходы оказывают негативное влияние на экономический рост во всех характеристиках в развитых и развивающихся странах. Автор связывает это с непродуктивным эффектом государственных инвестиционных расходов.

Высокие темпы роста населения в долгосрочной перспективе по результатам исследования снизили темпы экономического развития. Быстрый рост населения увеличивает потребности населения, что повышает потребительские расходы, в то время как сбережения увеличиваются в значительно меньшей степени. Это увеличивает потребительские расходы. Таким образом, норма сбережений и накопление капитала не сильно увеличиваются. Следовательно, экономический рост ухудшается в долгосрочной перспективе. Высокий прирост населения ухудшает экономический рост в большей степени в развивающихся странах, чем в развитых.

Кроме того, в рассматриваемой статье [43] изучено влияние болезней на внешнюю торговлю, использованием следующей модели

$$OPEN_{it} = \beta_0 + \beta_1 diseases_{it} + \beta_2 Z_{it} + \mu_{it}$$

где  $OPEN_{it}$  относится к внешней торговле, которая измеряется суммой экспорта и импорта по отношению к ВВП (%);  $Z_{it}$  — матрица  $k$  контрольных переменных:  $Pop$  — рабочая сила, которая выражается приростом населения (годовой %);  $FD$  — финансовое развитие;  $FDI$  — прямые иностранные

инвестиции; *GOVEXP* — расходы на конечное потребление органов государственного управления (% от ВВП);  $\mu_{it}$  — погрешность.

Исходя из результатов, полученных при помощи данной модели, автор приходит к ряду выводов. Во-первых, различные заболевания (малярия, ВИЧ, кишечные инфекционные заболевания) оказывают сильное негативное влияние на внешнюю торговлю всех стран в долгосрочном периоде, но более интенсивно в развитых странах, чем в развивающихся. Это можно объяснить тем, что экспорт развитых стран более важен для развивающихся стран. В период эпидемий экспорт чаще всего сокращается.

Для переменной прироста населения коэффициент отрицателен и значим в большинстве моделей. Хотя увеличение количества населения положительно влияет на торговые потоки в краткосрочной перспективе, потому что оно может увеличить количество рабочей силы, в долгосрочной перспективе рост населения снижает доход на душу населения, что делает людей беднее, поэтому производство и экспорт сокращаются. Кроме того, снижение дохода на душу населения также имеет тенденцию к снижению спроса на импорт.

Увеличение государственных расходов оказывает негативное влияние на внешнюю торговлю в развивающихся и развитых странах, поскольку ведет к увеличению импорта, что может привести к ухудшению торгового баланса.

Автор с помощью модели VEC определил, что в краткосрочной перспективе причинно-следственная связь между экономическим ростом и различными показателями инфекционных заболеваний является однонаправленной как в развивающихся, так и в развитых странах. То есть болезни влияют на экономический рост.

В долгосрочной перспективе экономический рост, болезни, население, инвестиции, государственные расходы и инфляция также тесно связаны между собой. Политика государства по улучшению здравоохранения,

управлению государственными расходами, увеличению инвестиций в человеческий капитал могут привести к экономической устойчивости в долгосрочной перспективе, как в развитых странах, так и в развивающихся странах.

Влияние инфекционных заболеваний на экономический рост может являться положительным, поскольку неинфицированные страны оказывают поддержку в виде технологий здравоохранения, передачи знаний, международных займов, чтобы инфицированное государство могло справиться с болезнью. Однако COVID-19 за короткий срок захватил все государства, поэтому помощь не могла быть оказана никому. Уязвимые места в экономических, социальных и других сферах способствовали распространению эпидемии.

В исследовании [38] использовалась традиционная модель роста Солоу — Свона, которая включает в себя труд, капитал, технологии и влияние этих параметров на объем производства.

$$Y_{it} = \beta_0 + \beta_1 L_{it} + \beta_2 K_{it} + \beta_3 (L \times T)_{it} + \beta_4 (K \times T)_{it} + \varepsilon_{it}$$

где  $Y$  — объем производства,  $L$  — количество рабочей силы,  $K$  — капитал,  $T$  — технология,  $i$  и  $t$  показывают разрез и период времени, а  $\varepsilon$  — погрешность.

Данное уравнение было модифицировано и дополнено набором параметров, чтобы показать влияние пандемией коронавируса,

$$EG_{it} = \beta_0 + \beta_1 COMD_{it} + \beta_2 HE_{it} + \beta_3 PD_{it} + \beta_4 PI_{it} + \beta_5 COVID-DUM_{it} + \varepsilon_{it}$$

где  $EG$  показывает экономический рост,  $COMD$  — инфекционные заболевания,  $HE$  — расходы на здравоохранение,  $PD$  — плотность населения,  $PI$  — уровень бедности,  $COVID-DUM$  показывает модель COVID,  $i$  и  $t$  показывают 76 стран и период времени с 2020 по 2019 год.

Уровень бедности в модели используется, поскольку бедные слои населения более подвержены инфекционным заболеваниям из-за низкого или отсутствующего дохода, постоянной безработицы и плохого медицинского обслуживания. Показатель COVID-DUM используется для оценки масштабов и интенсивности коронавируса, который в значительной степени увеличивается из-за высоких социальных контактов между членами населения, поскольку этот вирус легко распространяется при тесном контакте в сообществе, например, при рукопожатии, чихании, кашле. Поэтому, чем выше плотность населения на заданной территории, тем выше вероятность заражения вирусом. COVID-DUM присваиваются значения 1 и 0. Данные о смертности, вызванной инфекционными заболеваниями, и расходах на здравоохранение на душу населения, позволяют свести к минимуму вероятность систематической ошибки в данной модели. Данные для исследования были взяты из базы данных Всемирного банка и POVCAL Net.

Результаты исследований показывают, что расходы на здравоохранение чрезвычайно увеличиваются в случае с COVID-19. Множество стран не имеют доступа к базовым медицинским учреждениям, к простейшим средствам личной гигиены: мылу и воде для мытья рук, а также имеют ограниченный доступ к аппаратам искусственной вентиляции легких, маскам, инструментам для анализов и защитному оборудованию. Бедные страны в значительной степени пострадали от пандемии коронавируса, поскольку у них уже меньше ресурсов, нехватка продовольствия, низкий уровень здравоохранения, безработица, и многие другие факторы, которые негативно сказываются на здоровье населения и на экономическом росте страны.

Для изучения связи между заданными параметрами в исследовании использовалась матрица учета инноваций, основанная на двух функциях: функции импульсного отклика (IRF) и анализе дисперсионной декомпозиции (VDA). Оценки показывают, что различные инфекционные заболевания и расходы на здравоохранение, вероятно, увеличат экономический рост стран в

течение определенного периода времени, тогда как плотность населения и COVID-19 в основном усугубят экономические проблемы в виде снижения объема производства в течение следующих 10 лет. В группе с низким доходом происходит снижение экономического роста до 2023 года, однако после 2023 года экономический рост в этих странах начинает увеличиваться из-за роста неравенства доходов между странами. На рис. 1 представлено лишь несколько примеров графиков, полученных автором. Более подробно с результатами можно ознакомиться в статье [38].

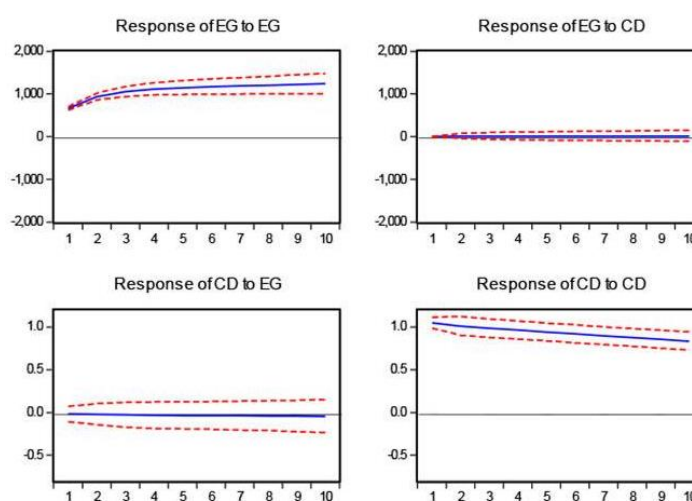


Рисунок 1. IRF оценки [38]

Оценки VDA показывают, что плотность населения будет оказывать большее влияние на экономический рост страны, как и расходы на здравоохранение, COVID-19 и инфекционные заболевания, в то время как уровень бедности окажет наименьшее влияние на экономический рост страны за определенный период времени. На рис. 2 также представлено лишь несколько примеров графиков, полученных авторами.

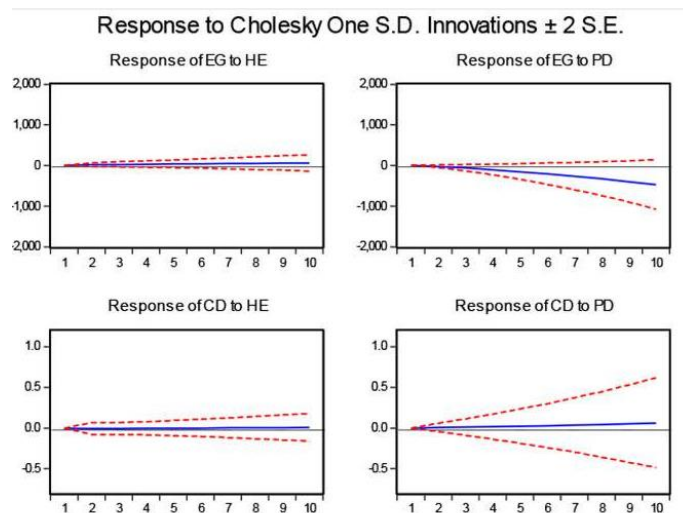


Рисунок 2. Оценки VDA [38]

Авторы статьи пришли к выводу, что инфекционные заболевания, в том числе COVID-19, в значительной степени усугубляют экономические проблемы из-за повышенного спроса на систему здравоохранения, усугубляемого уровнем бедности и социальной компактностью. Чтобы снизить смертность, необходимо тратить деньги на тестирование на коронавирус, скрининг, лечение и сдерживание эпидемии. Эти действия могут снизить интенсивность и страх перед пандемией коронавируса, что поможет контролировать вспышку заболевания.

Группа с низким доходом в значительной степени пострадает от этой пандемии из-за низкой осведомленности, проблем со здравоохранением, безработицы, неграмотности, отсутствия программ социальной защиты. Таким образом, существует большая потребность в поддержке более бедных стран посредством программ денежных переводов. Предоставление основных продуктов питания, базовых медицинских услуг, санитарно-гигиенических условий, а также надлежащего консультирования и руководства сведет к минимуму риск распространения инфекционных заболеваний.

В статье [52] рассматриваются тенденции капиталистической экономики во время пандемии. В исследовании авторы провели качественный и количественный анализ для оценки экономических последствий для

скоординированной рыночной экономики (СМЕ) и либеральной рыночной экономики (ЛМЕ).

Для количественного анализа используется вычисляемое общее равновесие (CGE) для моделирования экономических последствий вспышки пандемии. Модель CGE предполагает использование фактических экономических данных для оценки изменений в экономике как реакции на изменение внешних факторов, таких как технологические изменения, изменение политики и другие. Она работает на основе предположения, что страны не имеют барьеров и постоянно участвуют в торговле с идеальными конкурентными рынками и однородными технологиями.

Авторы обозначают запас капитала на рынках совершенной конкуренции в определенный период как

$$K_r^{t+1} = K_r^t + I_r^t - D_r^t$$

где  $K_r^{t+1}$  представляет собой запас капитала, накопленный в экономике  $r$  в заданном году.  $K_r^t$  — доступный капитал в регионе  $r$  в год  $t$ ,  $I_r^t$  — инвестиции в новый капитал в регионе  $r$  и году  $t$  и  $D_r^t$  — амортизация капитала региона  $r$  в году  $t$ .

Предложения труда на рынке труда в экономике обозначается следующим уравнением:

$$\frac{LS_r^t}{Pop_r^t} = (RW_r^t)^\beta A_r^t$$

где  $LS_r^t$  можно описать как предложение труда в экономике в период  $t$  в регионе  $r$ .  $Pop_r^t$  - доля населения в году  $t$  в регионе  $r$ .  $(RW_r^t)^\beta A_r^t$  обозначает реальную заработную плату после уплаты налогов для вида труда  $l$  в году  $t$  и регионе  $r$  с  $A_r^t$  и  $\beta$  — положительными константами, определяющими предложение труда в экономике.



Авторы рассматривают характер и степень воздействия на использование экономического капитала и предложение труда, оценивая экономические последствия при помощи переменных роста, таких как ВВП, инвестиции, ставки заработной платы. Данные взяты в периоды распространения следующих болезней: атипичной пневмонии (2002–2004 гг.), свиного гриппа H1N1 (2009–2010 гг.), а также текущей пандемии. Экономические данные были взяты для разных типов: для СМЕ (Япония, Швеция, Германия) и для LME (Великобритания, Австралия, Канада и Новая Зеландия), чтобы изучить разницу в экономическом росте до и после пандемии.

На рисунке 3 показано влияние тяжелого острого респираторного синдрома (SARS), который представляет собой вирусное респираторное заболевание, на торговлю в Китае. Болезнь возникла в китайской провинции Гуандун в 2002 г. и была заразна как вирус простуды. Также он был передан в такие страны, как Южная Африка, Гонконг, Канада, Австралия, Бразилия, Испания и США, и был локализован к июлю 2003 года. Было инфицировано около 10 000 человек, из которых 10% умерли.

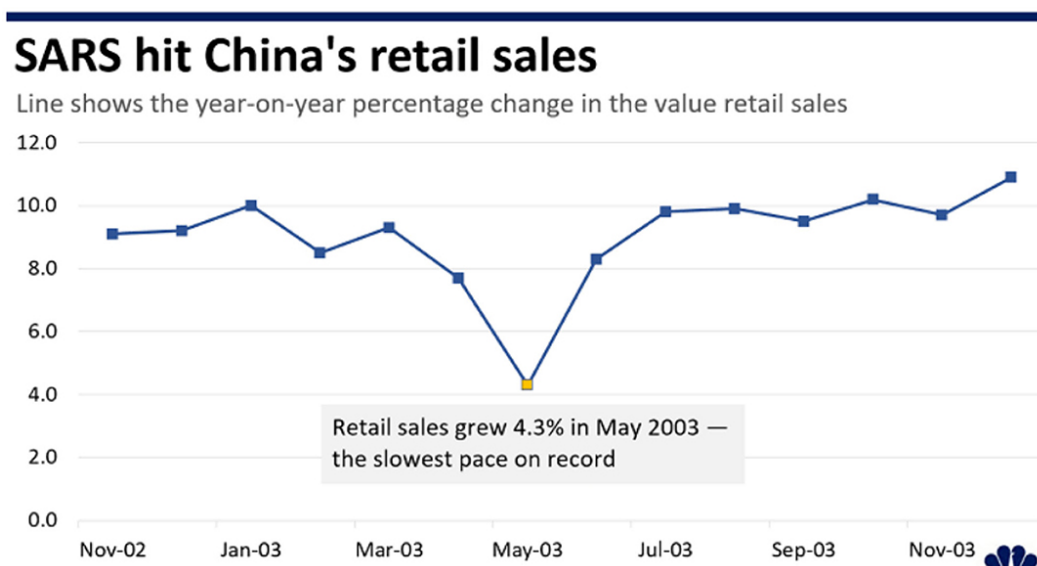


Рисунок 3. Влияние SARS на розничные продажи Китая [52]

SARS оказал большое влияние на инвестиции, розничную торговлю и туризм в Китае и Гонконге, сделав их наиболее пострадавшими районами. В этих регионах было значительное число погибших, а также большие краткосрочные экономические потери.

Экономические последствия атипичной пневмонии, вызванной SARS, с точки зрения расходов на здравоохранение и демографических последствий невелики по сравнению с экономическими последствиями таких эпидемий, как ВИЧ/СПИД или малярия. Однако болезнь быстро распространилась по странам, повлияв на здоровье жителей, а экономика была разрушена из-за проблем в торговых и финансовых связях между странами.

Исследование авторов показывает, что в случае с LME уровень безработицы и уровень инфляции негативно повлияли на экономический рост. Средняя продолжительность рабочего дня и государственные инвестиции положительно повлияли на уровень экономического роста. Модель, построенная в СМЕ (Японии, Германии и Швеции) для оценки воздействия вспышки атипичной пневмонии, показывает схожие результаты.

Далее авторы исследуют влияние свиного гриппа H1N1, который был вызван штаммом вируса гриппа, обычно встречающимся у свиней и имеющим симптомы, похожими на грипп. После его распространения болезни страны пережили огромный экономический спад, что повлекло за собой серьезный обвал на фондовом рынке. Произошел всплеск спроса на госпитали и медицинские услуги. Был временный значительный рост отпусков по болезни и закрытия школ, что потребовало ухода родителей из рабочей сферы. Была высокая смертность, что привело к соответствующему сокращению рабочей силы.

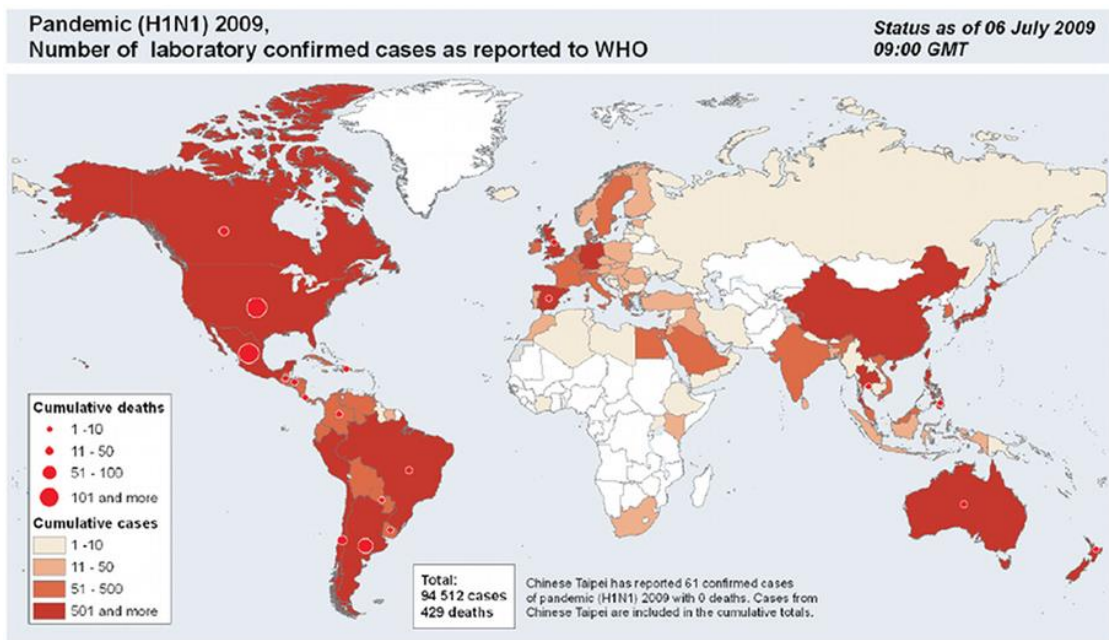


Рисунок 4. График, показывающий влияние пандемии H1N1 на весь мир [52]

Авторы исследуют влияние вспышки H1N1 на экономику Великобритании, Австралии, Канады и Новой Зеландии. В случае с LME уровень безработицы негативно повлиял на экономический рост. Напротив, средняя продолжительность рабочей силы и государственные инвестиции положительно повлияли на уровень экономического роста.

В случае СМЕ независимые переменные, включая уровень безработицы, отработанное время в неделю, уровень инфляции и инвестиции правительства, положительно повлияли на экономический рост во время вспышки вируса H1N1.

Covid-19 повлиял на общество гораздо больше, чем на здоровье пострадавших. Воздействие пандемии серьезно и варьируется от страны к стране. Были краткосрочные проблемы, такие как остановка туризма и развлечений, и долгосрочные последствия, такие как сбои в торговле и инвестициях. Заболевание имеет обширные последствия для здравоохранения, экономики и социальной сферы.

Из-за карантина и риска распространения болезни производство товаров первой необходимости замедлилось. Цепочка поставок продуктов была

нарушена, и национальные, и международные предприятия несут убытки. Миллионы рабочих потеряли работу из-за закрытия предприятий. На ВВП многих стран также повлияло нарушение производства в отраслях.

В случае с LME уровень безработицы негативно повлиял на экономический рост. Средняя продолжительность рабочей силы, уровень инфляции и государственные инвестиции положительно повлияли на экономический рост. В случае влияния распространения коронавируса на экономику СМЕ, каждая макроэкономическая переменная, включая безработицу, среднее количество рабочих часов, уровень инфляции и государственные инвестиции, положительно влияли на экономический рост.

Таким образом, авторы пришли к выводу, что существуют значительные различия в том, как вспышки заболеваний влияют на экономику. Экономика LME, включающая следующие страны: Великобританию, Канаду и Австралию), показала, что уровень безработицы и уровень инфляции негативно повлияли на экономический рост. Государственные инвестиции положительно повлияли на экономический рост, как и ожидалось. С другой стороны, в странах СМЕ все переменные положительно связаны с экономическим ростом экономики. Это свидетельство можно объяснить внутренней природой скоординированной рыночной экономики, в которой проблемы, стоящие перед бизнесом, решаются государственными институтами.

Анализ показал, что экономика LME была более восприимчива к последствиям вспышек заболеваний, чем экономика СМЕ, где участие правительства смягчило болезнь. влияние. Эти данные позволяют сделать вывод о том, что усиление участия государственных органов может смягчить воздействие вспышек заболеваний.

Подводя итоги, можно подчеркнуть, что тему влияний различных заболеваний на экономику как отдельных стран, так и всего мира в своих

статьях освещали множество авторов. Это показывает важность данной проблемы и ее актуальность. Для того, чтобы государства могли подготовить экономику страны в связи с болезнью следует проводить прогнозирование процесса болезни.

## 1.2. Обзор исследований по математическому моделированию заболеваний и эпидемий

Распространение эпидемии среди населения в значительной степени поддается математическому моделированию. Важным преимуществом математического моделирования является то, что оно требует прозрачности и точности в отношении предположений, что позволяет проверить понимание эпидемиологии заболевания путем сравнения результатов моделирования с наблюдаемыми закономерностями. Модели также могут помочь в принятии решений путем составления прогнозов относительно важных вопросов, таких как вызванные вмешательством изменения в распространении болезней. Существует множество исследований с разнообразными моделями заболеваний [3, 4, 17, 20, 24, 32].

Двумя наиболее популярными подходами к моделированию заболеваний являются моделирование распространения болезни посредством стохастического моделирования на основе агентов или использование детерминированных моделей. Стохастические агентные модели представляют конкретных людей или группы людей как агентов. Действия каждого агента регулируются набором правил, которые сами по себе могут быть функциями характеристик каждого агента. Через взаимодействие между агентами могут распространяться инфекционные заболевания. Нередко для моделирования заболевания используют временные ряды [33]. В последних исследованиях часто также используют методы машинного обучения [5, 9].

Модель SIR является одной из самых основных моделей для исследования эпидемического процесса. В этом сценарии каждый человек

может находиться в одном из трех эпидемиологических состояний в любой момент времени: восприимчивый, инфицированный и заразный или выбывший, что может означать выздоровевший или мертвый. Данная модель имеет ряд модификаций. Более подробно данная модель раскрыта во второй главе. Однако существуют исследования, которые модифицируют данную модель [2, 11, 35, 37, 36].

Стохастическая модель SIR, работающая поверх сложных сетей охватывает более реалистичную ситуацию. Простейший сценарий может быть представлен случайным графом как модель, предложенная Эрдёшем и Реньи (ER), где сеть строится, начиная с набора  $N$  узлов, и все пары узлов имеют одинаковую вероятность соединения. В результате получается однородный граф, в котором вершины имеют число соседей, называемое  $k$  степенью, мало отличающееся от средней степени  $\langle k \rangle$ . Распределение связности для этого графа может быть представлено распределением Пуассона.

Определим производящую функцию

$$F_0(z) = \sum_{k=0}^{\infty} p_k z^k.$$

Пока  $\sum_{k=0}^{\infty} p_k = 1$ , этот степенной ряд сходится для  $0 \leq z \leq 1$ , и его можно дифференцировать почленно. Таким образом

$$p_k = \frac{F_0^{(k)}(0)}{k!}, k = 0, 1, 2, \dots$$

Рассматривается вспышка заболевания, начинающаяся с единичного инфицированного индивидуума («нулевого пациента»), передающего инфекцию каждому индивидууму, с которым этот индивидуум связан, то есть по каждому ребру графа из вершины, соответствующей этому индивидууму. Другими словами, вспышка заболевания начинается, когда один возбудитель передает инфекцию всем людям, с которыми он контактирует [40].

Интересным подходом в математическом моделировании является вероятностную модель заражения (PIM), которая сочетает в себе неоднородность стохастических моделей с вычислительной эффективностью и детерминированным характером моделей. Авторы исследования [47] выделяют следующую идею своей модели: необходимо вычислить для каждой вершины контактной сети вероятности четырех состояний SEIR, связанных с этой вершиной. В модели SEIR рассматривается период, называемый Exposed (E), в течение которого человек инфицирован, но еще не передает вирус. Эта стадия соответствует промежуточному периоду между восприимчивым и инфицированным. Вероятности для каждого состояния и каждой вершины суммируются за промежутки времени, соответствующие латентному и инфекционному периодам данного заболевания. Это позволяет получить вероятностные значения различных состояний с течением времени на индивидуальном уровне, а также предоставляет ожидаемые значения размеров подгрупп SEIR, соответствующих каждому состоянию.

В данной вероятностной модели заражения контакт происходит между соседними вершинами. Таким образом, все соседи определенной вершины имеют вероятность установить контакт. Вероятность того, что вершина  $v$  вступит в контакт с вершиной  $u$  в результате единственной встречи

$$\psi(u, v) = \frac{w(u, v)}{\sum_{x \in N(v)} w(u, x)}$$

$w(u, v)$  — вес ребра  $(u, v)$ , а  $N(v)$  — множество соседей вершины  $v$ . Вероятность контакта из вершины  $u$  в вершину  $v$  будет отличаться от вероятности контакта из вершины  $v$  в вершину  $u$ , в зависимости от количества соседей каждой вершины и весов смежных ребер.

Каждый раз, когда  $v$  контактирует с инфицированным агентом  $u$ , существует вероятность передачи  $T(u, v)$ . Вероятность того, что вершина  $v$

будет заражена  $u$  в день  $t$  в результате единичного контакта, совершенного  $u$ , затем определяется как

$$\delta_t(u, v) = \psi(u, v) \cdot I_t(u) \cdot T(u, v).$$

Пусть  $\Omega_t(u)$  обозначает количество контактов, которые  $u$  совершает в день  $t$ . Вероятность того, что  $v$  не заразится в результате одного контакта в день  $t$ , равна  $1 - \delta_t(u, v)$ . Принимая во внимание всех соседей  $v$ , вероятность того, что  $v$  не заражен ни одним из соседей

$$\prod_{u \in n(v)} (1 - \delta_t(u, v))^{\Omega_t(u)}.$$

Вероятность того, что вершина находится в восприимчивом состоянии, то есть  $v$  не заражена ни одним из соседей с нулевого дня до текущего дня  $t$ :

$$S_t(v) = \prod_{n=0}^t \prod_{u \in n(v)} (1 - \delta_t(u, v))^{\Omega_t(u)}.$$

Вероятность состояния exposed равна вероятности пребывания в восприимчивом состоянии в день  $\max(0, t - \delta_v)$  минус текущая вероятность восприимчивого состояния в день  $t$ :

$$E_t(v) = S_{\max(0, t - \delta_v)}(v) - S_t(v).$$

Любой восприимчивый агент, который был инфицирован  $\delta_v + \gamma_v$  (инкубационный период + инфекционный период) дней назад, будет находиться в инфекционном состоянии:

$$I_t(v) = S_{\max(0, t - (\delta_v + \gamma_v))}(v) - S_{\max(0, t - \delta_v)}(v),$$

$$R_t(v) = 1 - S_{\max(0, t - (\delta_v + \gamma_v))}(v).$$



В результате исследования авторы сравнили модель PIM и стохастическую модель Монте-Карло для заболеваний. Обе модели дали схожие результаты и аналогичные значения для общего числа инфицированных людей во время вспышки. Однако модель PIM требует значительно меньшие вычислительные затраты.

Другим примером модификации SIR модели является знаменитая модель SIRS, в которой человек имеет только временный иммунитет и может снова стать восприимчивым к болезни через определенное время. Недавно была предложена модель для изучения распространения пандемии COVID-19 на основе 10 компартментов [39]. По данным этой работы популяция подразделяется на: восприимчивых (S), контактировавших (E), бессимптомных инфекционных (A), симптоматических инфекционных (I), подлежащих госпитализации в отделение интенсивной терапии (предварительно госпитализирован в ОИТ, PH), летальный прогноз (предварительно умерший, PD), поступивший в отделение интенсивной терапии, который выздоровеет (HR) или умрет (HD), выздоровевший (R) и умерший (D).

Недавно были проведены исследования с использованием статистических инструментов для изучения пространственно-временного распространения болезни COVID-19. В другой работе использовался корреляционный анализ для количественной оценки связи между количеством пассажиров из Уханя и количеством инфицированных людей в десяти близлежащих городах [45]. Нередко применяются агентные модели [21, 22].

Кроме того, при математическом моделировании заболеваний нужно принимать во внимание особенности конкретных болезней [25, 34]. Некоторые заболевания передаются от человека к человеку опосредованно, через переносчиков — живых организмов, которые могут передавать инфекционные заболевания. Такими организмами являются кровососущими насекомыми, которые поглощают болезнетворные микроорганизмы во время приема крови от инфицированного (человека) хозяина, а затем вводят их

новому хозяину во время последующего приема крови. Таким способом передаются такие заболевания, как малярия, вирус Зика, лихорадка Рифт-Валли, желтая лихорадка, японский энцефалит, лимфатический филяриатоз и лихорадка Западного Нила. Кроме того, клещи распространяют болезнь Лайма и туляремию, клопы — болезнь Лайма и болезнь Шагаса, мухи — онхоцеркоз, москиты — лейшманиоз, блохи — чуму, передаваемую от крыс к человеку, а также некоторые пресноводные улитки являются переносчиками некоторых болезней [40].

Болезни, передающиеся переносчиками, требуют моделей, включающих как переносчиков, так и хозяев. Для большинства болезней, передаваемых переносчиками, переносчиками являются насекомые с гораздо более короткой продолжительностью жизни, чем у хозяев, которыми могут быть люди, как в случае малярии, или животные, как в случае лихорадки Западного Нила.

Существуют болезни человека, передающиеся гетеросексуальным путем, которые также имеют свои особенности. Их также можно рассматривать как болезни, передаваемые переносчиками, потому что мужчин и женщин следует рассматривать как отдельные популяции, а болезнь передается от одной популяции к другой.

Еще одна интересная модификация SIR модели представлена в статье [49]. Авторы пытаются предложить новую модель эпидемии, которая обеспечивает прямое моделирование, основанное на дискретном дифференциальном уравнении с задержкой, которое включает период изоляции (госпитализацию). Первая модель предполагает, что все инфицированные случаи приводят к появлению симптомов, а затем к изоляции. Во второй модели добавляются и бессимптомные случаи (такие больные не изолируются).

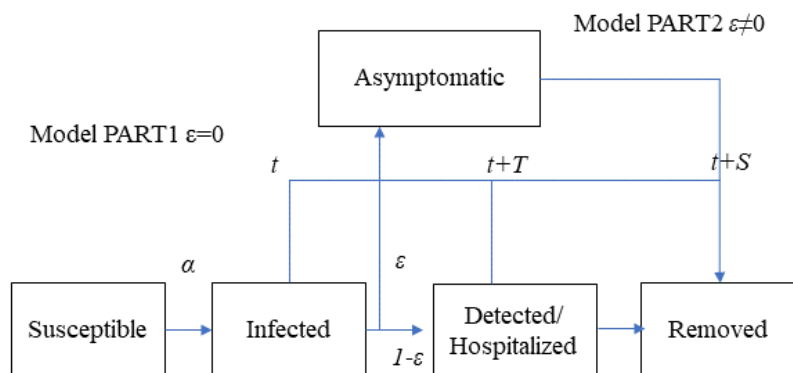


Рисунок 5. Основная концепция модели [49].

На рисунке 5 авторы представили основную концепцию эпидемиологической модели. Как только восприимчивые люди заражаются в момент времени  $t$  со скоростью передачи  $\alpha$ , они сразу становятся заразными. Бессимптомные случаи остаются заразными до тех пор, пока человек не выйдет из модели (выздоровление/смерть) в момент времени  $t + S$ , в то время как случаи с симптомами остаются заразными до госпитализации (изоляция) в момент времени  $t + T$ . Параметры  $S$  и  $T$  должны определяться эмпирическим методом, основанным на данных наблюдений. Часть  $\varepsilon$  случаев обозначают как бессимптомные, а остальные  $1-\varepsilon$  — как симптоматические. Принимая суммарные случаи  $x(t)$  как первичную зависимую переменную, тогда количество инфекционных случаев в целом ( $Q$ ) можно рассчитать, как сумму бессимптомной части  $\varepsilon(x(t) - x(t - S))$  и симптоматической части  $(1 - \varepsilon)(x(t) - x(t - T))$ . Восприимчивая часть популяции ( $M$ ) может быть выражена как  $\frac{1-x(t)}{M}$ . Также существует предположение, что частота случаев пропорциональна произведению  $Q$  на восприимчивую часть, то есть  $\frac{dx(t)}{dt} = \alpha Q \left( \frac{1-x(t)}{M} \right)$ . Тогда  $x(t)$  определяется следующим дифференциальным уравнением с запаздыванием.

Для  $\varepsilon = 1$  модель сводится к обычной SIR модели. А при  $0 < \varepsilon < 1$  эпидемическая модель описывается уравнением:

$$\frac{dx}{dt} = \alpha \{ x - (1 - \varepsilon)x(t - T)u(t - T) - \varepsilon x(t - S)u(t - S) \} \left( 1 - \frac{x}{M} \right),$$

где  $u(t)$  — функция Хевисайда (единичная ступенька):

$$u(t) = \begin{cases} 0, & t < 0, \\ 1, & t \geq 0. \end{cases}$$

Модель для  $\varepsilon = 0$  лечит только клинические симптомы и не принимает во внимание бессимптомные случаи. Следовательно, уравнение принимает следующий вид:

$$\frac{dx}{dt} = \alpha \{x - x(t - T)u(t - T)\} \left(1 - \frac{x}{M}\right).$$

Эффективное репродуктивное число  $R_t$  относится к количеству заболевших на один случай заражения в популяции до их удаления. Автор данной статьи рассчитывают  $R_t$  как отношение ежедневных новых случаев к среднему количеству инфицированных в день в течение промежутка времени  $[t - T, t]$  как:

$$R_t = \frac{\frac{dx}{dt}}{\left(\frac{Q}{T}\right)},$$

или

$$R_t = \alpha T \left(1 - \frac{x}{M}\right).$$

Проведя исследования, авторы пришли к следующим выводам. Существующая эпидемиологическая модель ATLM достаточно проста с точки зрения математики и включает небольшое количество подгоночных параметров. Однако существуют и более сложные подходы, состоящие из нескольких уравнений с переменными, зависящими от количества сегментов сообщества в модели. Авторам удалось с помощью ATLM смоделировать тенденцию данных первой волны эпидемии COVID-19 в Токио.

Также по результатам исследования был сделан вывод, о необходимости стратегий общественного контроля над эпидемией (то есть вмешательства общественного здравоохранения в нужном масштабе и в нужное время).

Одним из них является оценка режима самоизоляции по сокращению контактов между людьми на 80%. Основываясь на предсказании инфекционных случаев в целом с помощью модели, объявление должно было

быть сделано на один месяц раньше. Если это так, умеренное сокращение контактов на 43% было бы достаточно эффективным, чтобы уменьшить как стационарное лечение, так и влияние на социальную и экономическую деятельность. И наконец, авторы предположили, что, судя по поведению инфекционных случаев в целом, чрезвычайное положение следовало отложить на одну неделю, когда расчетные инфекционные случаи в целом опустились бы ниже единицы, после чего эпидемия прекратилась.

В другой статье [41] модель строится с использованием программного обеспечения RAMAS® Metapop v. 6.0, которое представляет собой вычислительный инструмент для построения популяционных решений, но также и эпидемиологические подходы, основанных на вероятностном моделировании методом Монте-Карло. Каждое состояние SIR подразделяется на возрастные классы (молодые и пожилые) с различной вероятностью летального исхода и заражения, а также эффект зависимости от плотности (DD) для учета снижения уровня инфицирования по мере уменьшения плотности восприимчивой популяции. То есть первичная структура модели соответствует моделированию SIR, однако существуют вторичные структуры, которые интегрированы в новую модель.

Модель описывает популяцию, состоящую из шести групп шести на континентах. Структура состояния каждой популяции:

[ $s = 1$ ] — восприимчивые молодые,

[ $s = 2$ ] — инфицированные молодые,

[ $s = 3$ ] — выздоровевшие молодые,

[ $s = 4$ ] — восприимчивые пожилые,

[ $s = 5$ ] — инфицированные пожилые и

[ $s = 6$ ] — пожилые люди, которые выздоровели,

где  $s$  — состояние. К пожилым людям относятся те, кому 60 и больше лет [41].

Эта модель обладала как преимуществами перед другими моделями, так и недостатками. Например, большим достоинством такого подхода по сравнению с детерминированным анализом заключается в том, что результаты

показывают не только то, что может произойти, но и насколько вероятен каждый исход. Однако проблему с данными данная модель не смягчает. Существует значительная неопределенность и отсутствие ясности в отношении того, как собирались данные о количестве подтвержденных случаев в каждой стране. Несмотря на то, что результаты не удовлетворили авторов, данный подход наравне с другими вносит значительный вклад в моделирование эпидемий.

Также существует модель, основанная на идеи пространственного распространения болезней [31], а именно времени пребывания. То есть описывает ситуации ситуациях, в которых жители проводят часть своего времени в других участках. Например, распространение инфекционного заболевания из одной деревни в другую через людей, посещающих другие места. Другая интерпретация заключается в предположении, что люди проводят часть своего времени в среде, которая с большей вероятностью допускает передачу болезни.

Тогда рассматриваются два участка с общей численностью постоянной популяции  $N_1$  и  $N_2$  соответственно, причем каждая популяция будет разделена на восприимчивых, заразных и выбывших людей. Общая численность населения, проживающего на каждом участке, постоянна.

Жители участка  $i$  проводят долю  $p_{ij}$  своего времени на участке  $j$ , причем

$$p_{11} + p_{12} = 1, \quad p_{21} + p_{22} = 1.$$

Частота контактов в  $i$  равна  $\beta_i$ .

Каждый из восприимчивых к  $p_{11}S_1$  из группы 1, присутствующих на участке 1, может быть заражен инфекциями из группы 1 и из группы 2, присутствующими на участке 1. Аналогично, каждый из восприимчивых к  $p_{12}S_1$ , присутствующих на участке 2, может быть заражен инфекциями из группы 1 и из группы 2, присутствующими на участке 2. Плотность инфицированных агентов на участке 1 в момент времени  $t$ , которые могут заразить только агентов, находящихся в данный момент на участке 1 в момент

времени  $t$ , то есть эффективная доля инфицированных на участке 1 определяется как

$$p_{11} \frac{I_1(t)}{N_1} + p_{21} \frac{I_2(t)}{N_2}.$$

Таким образом, частота новых заражений агентов участка 1 на участке 1 составляет

$$\beta_1 p_{11} S_1 \left( p_{11} \frac{I_1(t)}{N_1} + p_{21} \frac{I_2(t)}{N_2} \right).$$

Тогда модель SIR задается следующим образом

$$S'_1 = -\beta_1 p_{11} S_1 \left[ p_{11} \frac{I_1}{N_1} + p_{21} \frac{I_2}{N_2} \right] - \beta_2 p_{12} S_1 \left[ p_{12} \frac{I_1}{N_1} + p_{22} \frac{I_2}{N_2} \right],$$

$$I'_1 = \beta_1 p_{11} S_1 \left[ p_{11} \frac{I_1}{N_1} + p_{21} \frac{I_2}{N_2} \right] + \beta_2 p_{12} S_1 \left[ p_{12} \frac{I_1}{N_1} + p_{22} \frac{I_2}{N_2} \right] - \gamma I_1.$$

В этой структуре фактическое перемещение между участками явно не описывается, что делает анализ менее сложным.

Таким образом, можно сказать, что математическое моделирование достаточно распространенная тема исследований, что говорит об актуальности данной проблемы. Одной из целей эпидемиологии является определение причины болезни, ее течение и разработка способов борьбы. Математическое моделирование помогает исследовать заболевания, а также делать прогнозы. На данный момент не одна из моделей не является идеальной, каждая имеет определенные преимущества и недостатки, и подходит для определенных ситуаций. А это значит, что данная тема еще долгое время будет востребованной.

## 2. Методологическая специфика моделирования заболеваний

### 2.1. Описание основных источников данных

Выбор метода зависит от многих факторов — контекста прогноза, актуальности и доступности исторических данных, желаемой степени точности, периода времени для прогнозирования и других факторов. Однако основные шаги, используемые для построения моделей и прогнозирования, одинаковы для всех методов. Одним из этапов этих процессов является получение данных. Чаще всего используются два вида информации: статистические данные и накопленный опыт и знания.

Для исследования, а именно для построения цепочек распространения и построения прогноза, были взяты статистические данные с официальных сайтов ресурсного центра Джона Хопкинса [54]. Ресурсный центр по коронавирусу Джона Хопкинса (CRC) является постоянно обновляемым источником данных о COVID-19. Здесь собираются и анализируются доступные данные о случаях заражения, смертях, тестах, госпитализациях и вакцинах. Данные формируются из официальных источников стран, таких как сайты министерств здравоохранения и других. Однако с 10 марта 2023 года Ресурсный центр Джона Хопкинса прекратил сбор и представление глобальных данных о COVID-19.

С этого ресурса были взяты данные о количестве подтверждённых случаях заболевания, выздоровления, а также о количестве смертельных исходов по всем странам мира (197 признанные страны) с 22 января 2020 года по 22 ноября 2022 года. Из них были выбраны и проанализированы страны наиболее интересные для исследования и имеющие подходящие статистические данные.

Кроме того, для получения данных по России был использован интернет-ресурс для информирования населения по вопросам коронавируса



Стоп Коронавирус [53]. Данный сайт является основным всероссийским источником официальной информации по всем вопросам эпидемии.

И наконец, информация о плотности населения была взята из отчета «World Population Prospects» за 2019 год. «Демографические перспективы мира» за 2019 год является двадцать шестым изданием Организации Объединенных Наций демографических оценок и прогнозов. В нем представлены оценки численности населения с 1950 года по настоящее время по 235 странам или районам.

Таким образом, были собраны необходимые для прогнозирования данные. Источники являются официальными и достоверными базами данных.

## 2.2. Описание модель SIR

Как уже говорилось ранее наиболее распространенной эпидемиологической моделью является модель SIR, разработанная Россом, Хеймером и другими исследователями в начале двадцатого века, состоящая из системы трех связанных нелинейных обыкновенных дифференциальных уравнений.

Математические модели все чаще используются для выяснения передачи нескольких заболеваний. Эти модели могут быть довольно простыми, но их изучение имеет решающее значение для получения важных знаний об основных аспектах распространения инфекционных заболеваний. Модель SIR является отправной точкой для понимания того, что происходит на коротком промежутке времени. Далее модель модифицируется и начинает дополняться дополнительными деталями.

Математические модели представляют собой упрощенное представление о том, как заболевание распространяется в популяции с течением времени. В модели SIR люди делятся на три типа:

- Восприимчивые (*Susceptible*): это агенты, которые являются восприимчивыми к инфекции, то есть те, кто может заболеть; также сюда могут входить пассивно невосприимчивые, как только они теряют иммунитет.
- Инфицированные (*Infected*): это агенты, у которых высокая вероятность передачи инфекции другим восприимчивым агентам;
- Выздоровевший или выбывший (*Recovered/Resistant*): включает агентов, которые были инфицированы и выздоровели, либо умерли.

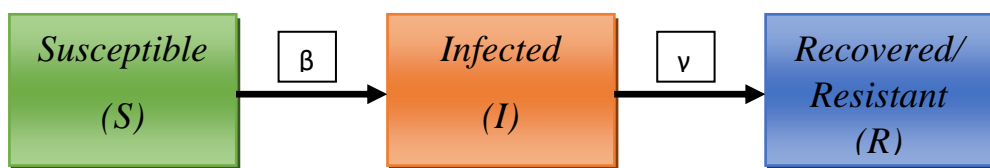


Рисунок 6. Схематичное представление модели SIR

В данной модели заболеваний агенты после выздоровления приобретают пожизненный иммунитет. Общая численность населения считается постоянной и равна  $N = S + I + R$ . Разделив популяцию на группы, представим набор уравнений, определяющих, как эти групп изменяются с течением времени. Для начала рассмотрим модель SIR, исключая рождение и смерть:

$$\frac{dS}{dt} = -\beta SI,$$

$$\frac{dI}{dt} = \beta SI - \gamma I,$$

$$\frac{dR}{dt} = \gamma I,$$

при соблюдении начальных условий  $S(0) > 0, I(0) \geq 0, R(0) \geq 0$ .

Кроме того, существует коэффициент заражаемости на душу населения  $\beta$ , скорость выздоровления —  $\gamma$ . Также часто используется пороговое

значение  $R_0 = \frac{\beta}{\gamma}$ , называемое базовым репродуктивным числом.  $R_0$  определяется как среднее число вторичных случаев инфицирования, при попадании одного возбудителя в полностью восприимчивую популяцию.  $\sigma$  — среднее число контактов, достаточных для передачи инфекции, восприимчивому агенту от инфицированного в течение инфекционного периода. Обратим внимание, что  $R_0 \geq \sigma$ .

Если значение  $R_0 < 1$  болезнь не охватит всю популяцию, распространение болезни закончится в течение определенного периода времени, который зависит от того, насколько мал  $R_0$ . Когда  $R_0 > 1$  инфекция может распространиться среди всего населения. Как правило, чем больше значение  $R_0$ , тем более серьезной и, возможно, широко распространенной будет эпидемия.

Модель SIR с демографией включает естественную продолжительность жизни агента, равную  $\frac{1}{\mu}$  лет. Коэффициент естественной смертности отдельных лиц определяется через  $\mu$ . Таким образом, модель SIR, включающая рождения и смерти, описывается следующим образом

$$\frac{dS}{dt} = \mu - \beta SI - \mu S,$$

$$\frac{dI}{dt} = \beta SI - \gamma I - \mu I,$$

$$\frac{dR}{dt} = \gamma I - \mu R.$$

с начальными условиями  $S(0) > 0$ ,  $I(0) \geq 0$ ,  $R(0) \geq 0$ . Для данной модели базовое репродуктивное число  $R_0 = \frac{\beta}{\gamma + \mu}$ .

Если  $R_0 < 1$ , каждый инфицированный агент производит, в среднем, заражает менее одного нового агента, а значит заболевание не сможет

охватить популяцию и затухнет. Если  $R_0 > 1$ , то инфекция проникнет в восприимчивую популяцию.

Если предположить, что  $S \approx N$ , то система дифференциальных уравнений примет следующий вид

$$\begin{aligned}\frac{dI(t)}{dt} &= (\beta - \gamma)I(t), \\ \frac{dR(t)}{dt} &= \gamma I(t),\end{aligned}$$

и её решение будет иметь вид

$$\begin{aligned}I(t) &= I(t_0)e^{(\beta-\gamma)(t-t_0)}, \\ R(t) &= \frac{\gamma I(t_0)}{\beta - \gamma} (e^{(\beta-\gamma)(t-t_0)} - 1) + R(t_0).\end{aligned}$$

Таким образом, была рассмотрена самая распространенная математическая модель заболевания — модель SIR. Поскольку в исследовании, отсутствуют данные о количестве восприимчивых, в следующей главе будет использоваться система уравнений, описанная в конце параграфа.

### 2.3. Описание модели CIR

В модели CIR в место параметра  $S$  — восприимчивые используется новый параметр  $C$  (Confirmed) — общее число подтвержденных случаев инфицирования с момента начала эпидемии до определенного нами дня. Параметр  $\Delta(t)$  обозначает количество новых случаев заболевания в день  $t$ .

$$C(t) = C(t - 1) + \Delta(t),$$

или

$$C(t) = \alpha(t)C(t - 1),$$

где

$$\begin{aligned}\alpha(t) &= 1 + \frac{r(t)}{100}, \\ r(t) &= 100 \frac{\Delta(t)}{C(t - 1)}.\end{aligned}$$

Формула приобретает следующий окончательный вид

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

Параметр  $r(t)$  показывает отношение в процентах абсолютного прироста общего количества выявленных случаев инфицирования до дня  $t$  включительно к общему количеству выявленных случаев инфицирования до предыдущего дня включительно.

Таким образом модель CIR описывается следующей системой дискретных уравнений.

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

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

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

Одним из способов прогнозирования динамики показателя  $C(t)$  является подход, основанный на методе прецедентов. В методе используется эвристический выбор будущих значений процентного прироста общего количества выявленных случаев инфицирования в одной стране на основе информации о развитии заболевания в других странах [12, 13].

Для начала выбранный промежуток времени эпидемии разбивается на  $M$  промежутков  $(T_{m-1}, T_m]$ ,  $0 \leq T_{m-1} \leq T_m \leq T$ ,  $m = 1, 2, \dots, M$ . Границами промежутков являются моменты достижения процентным приростом определенных значений. Например, были выбраны моменты достижения 20%, 10%, 5%, 2%. Количество дней в промежутке заданном промежутке равно  $T_m - T_{m-1}$ . Изменения прогнозируемых значений общего количества, инфицированных внутри определенного промежутка, описывается как

$$\tilde{C}(T_{m-1} + k) = \left(1 + \frac{r(T_{m-1} + k)}{100}\right) \tilde{C}(T_{m-1} + k - 1),$$

$$k = 1, 2, \dots, (T_m - T_{m-1}).$$

Данная CIR модель используется как основа для дальнейшего исследования. Она позволяет на основе данных одних стран строить прогноз для других. Кроме того, для данной модели не нужно находить дополнительные параметры, как например, коэффициент заражаемости и скорость выздоровления для модели SIR. Поэтому эта модель более проста в использовании.

### 3. Анализ математических моделей заболевания COVID-19

#### 3.1. Цепочки распространения пандемии COVID-19

В данном параграфе представлены цепочки распространения эпидемии COVID-19 в пяти группах: в Европе, Африке, Северной и Южной Америке, и наконец, в Азии. В каждой группе была найдена одна страна, которая позже остальных достигла примерно 20% стабильного (после этого дня нет сильных скачков) прироста случаев заболевших. Эта страна является страной-последователем. Далее были найдены страны, чей уровень в 20% прироста был достигнут раньше, примерно на 4–7 дней. Таких стран может быть несколько. Далее уже для новых государств определяем страны-предшественники.

На основе полученных данных были построены следующие цепочки распространения. Так, например, для страны-последователя Российской Федерации странами-предшественниками являются Испания, Италия, Франция, Латвия и остальные страны, которые расположены на схеме левее самой России. Следовательно, можно построить следующую цепочку стран: Италия — Франция — Бельгия — Россия. Для Австрии цепочка может иметь вид: Испания — Латвия — Австрия или Испания — Швейцария — Австрия.

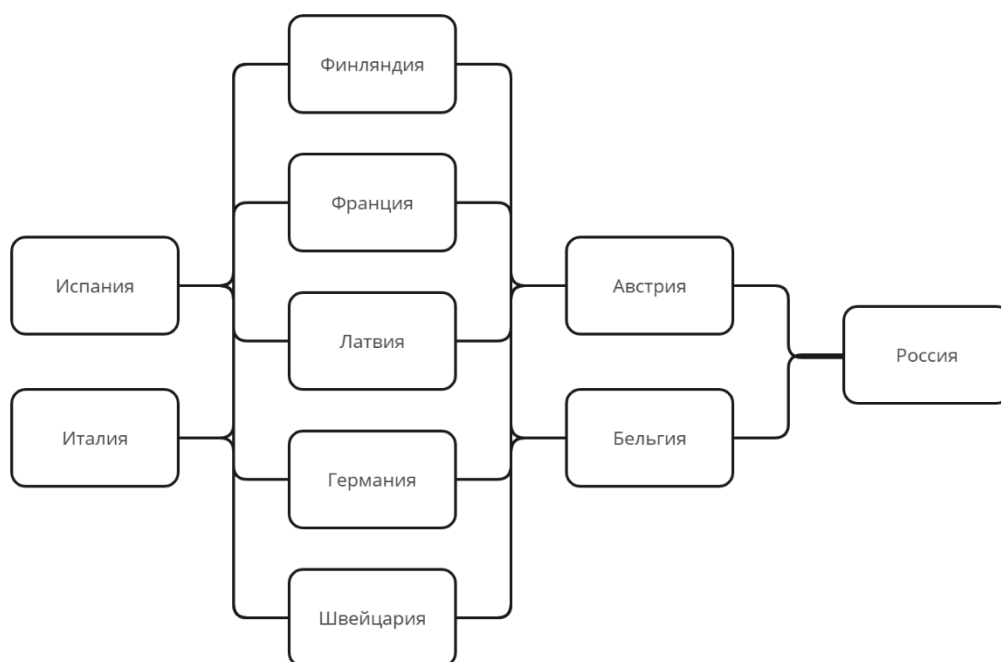


Схема 1. Страны Европы в цепочке распространения эпидемии

Таблица 1. Характеристики стран Европы в цепочке распространения эпидемии

<b>Номер страны</b>	<b>Название страны</b>	<b>Количество дней перехода</b>	<b>Дата начала отсчета</b>	<b>Плотность населения, чел./км<sup>2</sup></b>
1	Россия	70	27 марта	8,9
2	Австрия	17	21 марта	108,7
2	Бельгия	36	21 марта	381,1
3	Финляндия	32	18 марта	18,2
3	Франция	30	18 марта	118,9
3	Латвия	28	18 марта	30,7
3	Германия	29	16 марта	239,6
3	Швейцария	27	16 марта	217,4
4	Испания	31	13 марта	93,7
4	Италия	35	11 марта	205,9



Для стран Азии цепочка получилась более последовательная. Например, получаем последовательность: Сингапур — Япония — Южная Корея — Иран — Грузия — Азербайджан — Израиль — Афганистан — Киргизия — Бангладеш.



Схема 2. Страны Азии в цепочке распространения эпидемии

Таблица 2. Характеристики стран Азии в цепочке распространения эпидемии

Номер страны	Название страны	Количество дней перехода	Дата начала отсчета	Плотность населения, чел./км <sup>2</sup>
1	Бангладеш	82	15 апреля	1 252,6
2	Киргизия	20	8 апреля	33,5
3	Афганистан	78	2 апреля	58,3
4	Израиль	27	26 марта	393,7
5	Азербайджан	34	21 марта	121,6
5	Индонезия	67	21 марта	149,4
6	Грузия	51	14 марта	57,5
7	Иран	42	7 марта	50,9
8	Южная Корея	12	1 марта	526,8
9	Япония	72	15 февраля	348,0
10	Сингапур	98	9 февраля	8 291,9

Также для некоторых группы из анализа были исключены некоторые страны. Это было сделано поскольку некоторые государства имеют неподходящие данные, либо много данных отсутствует. Из-за получились следующие цепочки, представленные ниже.

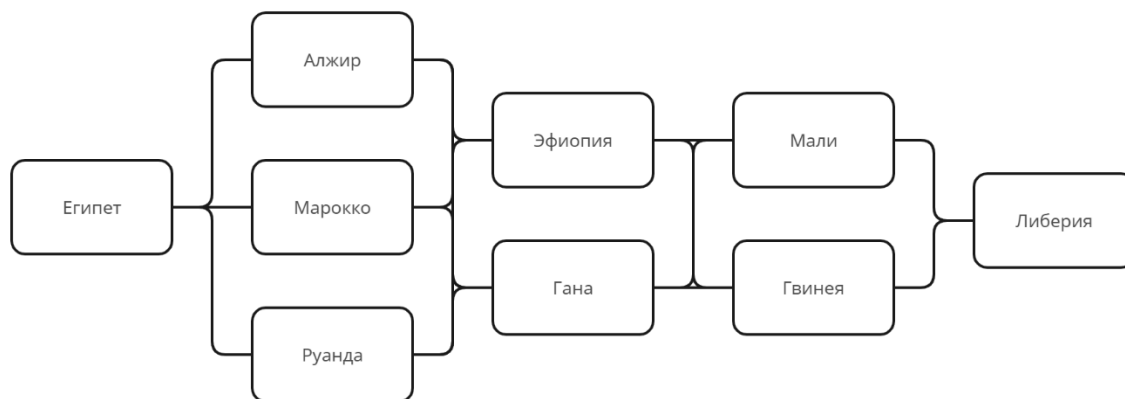


Схема 3. Страны Африки в цепочке распространения эпидемии

Таблица 3. Характеристики стран Африки в цепочке распространения эпидемии

Номер страны	Название страны	Количество дней перехода	Дата начала отсчета	Плотность населения, чел./км <sup>2</sup>
1	Либерия	30	17 апреля	51,3
2	Мали	55	12 апреля	16,1
2	Гвинея	55	9 апреля	52,0
3	Эфиопия	20	3 апреля	112,1
3	Гана	102	1 апреля	133,7
4	Алжир	63	26 марта	18,1
4	Марокко	49	26 марта	81,7
4	Руанда	21	26 марта	511,8
5	Египет	100	19 марта	100,8



Схема 4. Страны Южной Америки в цепочке распространения эпидемии

Таблица 4. Характеристики стран Южной Америки в цепочке распространения эпидемии

Номер страны	Название страны	Количество дней перехода	Дата начала отсчета	Плотность населения, чел./км <sup>2</sup>
1	Боливия	94	28 марта	10,6
2	Чили	91	24 марта	25,5
2	Бразилия	99	23 марта	25,3
3	Парагвай	68	17 марта	17,7



Схема 5. Страны Северной Америки в цепочке распространения эпидемии

Таблица 5. Характеристики стран Северной Америки в цепочке распространения эпидемии

Номер страны	Название страны	Количество дней перехода	Дата начала отсчета	Плотность населения, чел./км <sup>2</sup>
1	США	43	28 марта	33,8
2	Куба	42	24 марта	106,5
2	Доминиканская Республика	53	23 марта	215,7
3	Мексика	100	21 марта	65,6
3	Коста-Рика	29	18 марта	98,9
4	Канада	60	9 марта	4,1

Таким образом, в данном параграфе страны в каждой группе были распределены в хронологическом порядке в зависимости от времени прохождения ими двадцатипроцентного уровня прироста.

### 3.2. Построение прогноза с помощью SIR-модели

Модель SIR является одной из самых основных моделей для описания временной динамики инфекционного заболевания в популяции. Опираясь на нее множество авторов строит более сложные модели, которые описывают действительную ситуацию.

Базовая модель SIR достаточно упрощена, что может обеспечить прогнозы лучше, чем более реалистичная модель с большим количеством параметров, когда данных недостаточно или качество данных низкое. Проверим, насколько данная модель подходит для прогнозирования пандемии COVID-19.

Построим прогноз с помощью SIR модели. Для этого в работе использовался язык программирования Python. На оси ординат отмечены дни,

а на оси ординат — количество человек. На графике можно увидеть реальные данные по количеству заболевших и количеству выбывших, а их также прогнозные значения. Знаком «\*» на графиках обозначен день, с которого начинается прогноз. Для расчета параметров  $\beta$  и  $\gamma$ , а также построения прогноза мы использовали формулы

$$I(t) = I(t_0)e^{(\beta-\gamma)(t-t_0)},$$

$$R(t) = \frac{\gamma I(t_0)}{\beta - \gamma} (e^{(\beta-\gamma)(t-t_0)} - 1) + R(t_0).$$

Рассмотрим, как пример, Бангладеш. С помощью метода наименьших квадратов, были получены показатели  $\beta \approx 0.2419$  и  $\gamma \approx 0.0165$ . Исходя из полученных параметров был построен следующий прогноз.

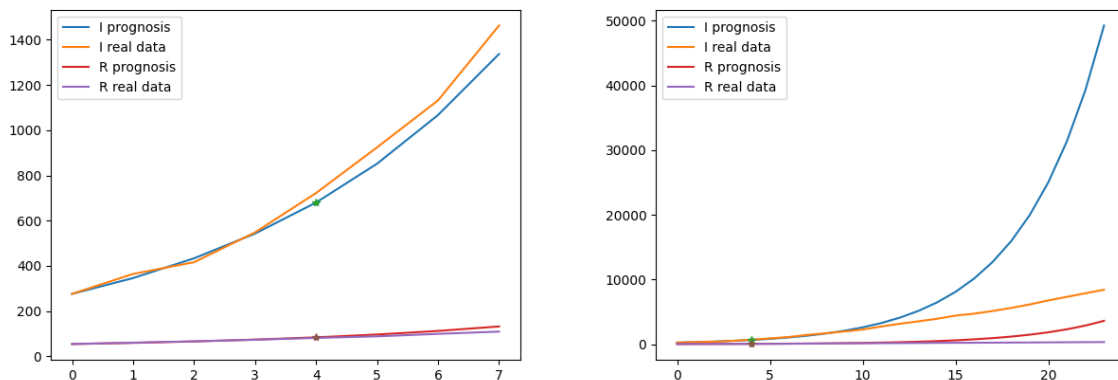


Рисунок 7. График значений количества инфицированных и выбывших для страны Бангладеш. Слева — короткий период прогнозирования. Справа — длительный срок прогнозирования

Как можно увидеть из рисунка 7 данные на непродолжительный период времени можно спрогнозировать достаточно точно. Для количества инфицированных прогнозные значения несколько ниже реальных, а для количества выбывших наоборот — прогноз немного выше имеющихся значений. Если попытаться построить прогноз на более длительный срок, он становится не точным, поскольку график растет экспоненциально. Это происходит поскольку параметры  $\beta$  и  $\gamma$  были найдены и зафиксированы в начале эпидемии, когда еще не предприняты меры защиты от болезни и активно не используются методы лечения.

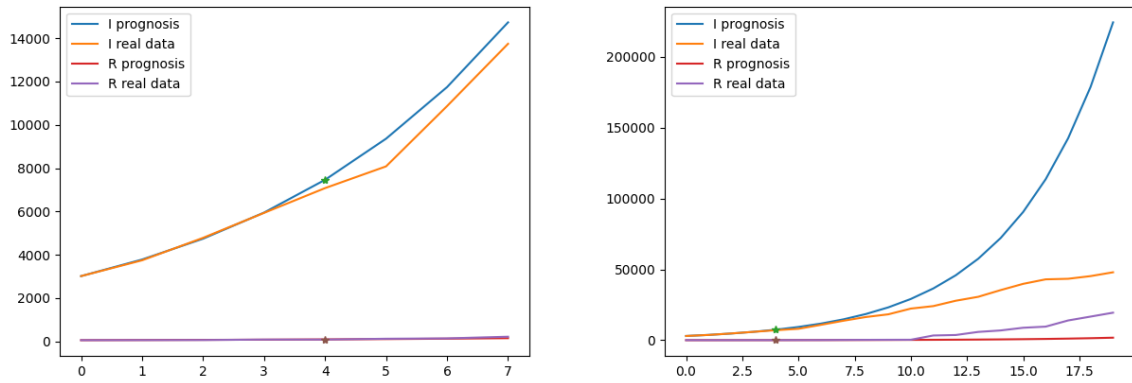


Рисунок 8. График значений количества инфицированных и выбывших для Германии. Слева — короткий период прогнозирования. Справа — длительный срок прогнозирования

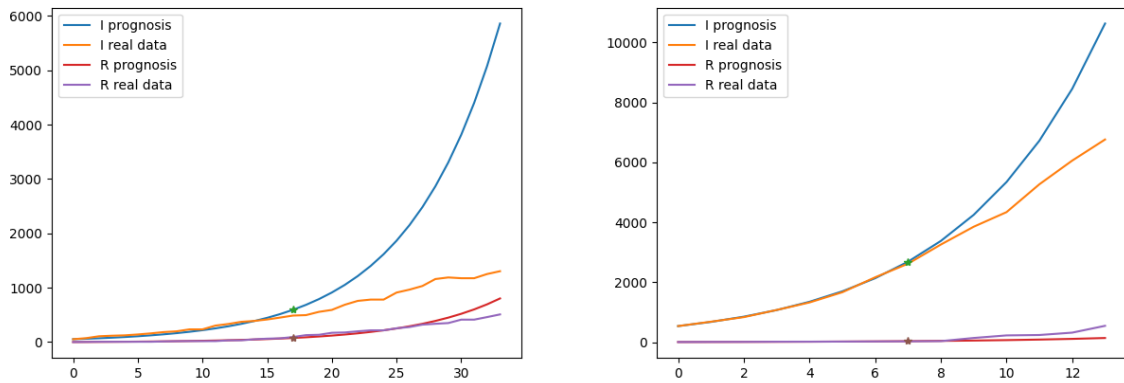


Рисунок 9. График значений количества инфицированных и выбывших. Слева — прогноз для Гвинеи. Справа — прогноз для Коста-Рики

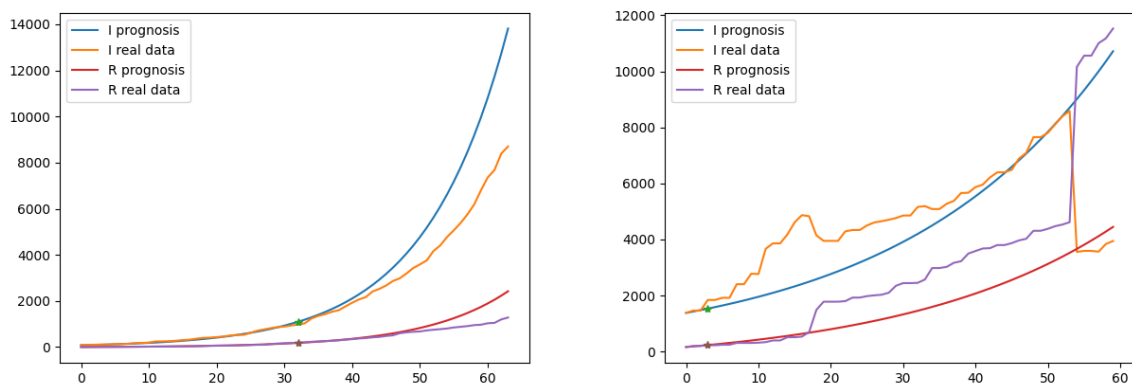


Рисунок 10. График значений количества инфицированных и выбывших. Слева — прогноз для Боливии. Справа — прогноз для Ганы

Для других стран получаем схожую картину. Как видно из графиков, представленных на рисунках 8–10, количество выбывших практически для всех стран прогнозируются лучше, чем количество заболевших. Приведем несколько примеров, чтобы это проиллюстрировать.

Таблица 6. Значения  $\beta$  и  $\gamma$  для представленных стран

Страны	Параметр $\beta$	Параметр $\gamma$
Германия	0.228648	0.001797
Гвинея	0.162932	0.019761
Коста-Рика	0.096667	0.049431
Гана	0.050616	0.015926
Боливия	0.096178	0.014404

Модель SIR лучше подходит для анализа краткосрочной эпидемии в локальных районах. Модель широко используется в основном в развивающихся странах, нуждающихся в борьбе с различными инфекционными заболеваниями, где население невелико. Для анализа распространения заболевания в контексте COVID-19 SIR модель не подходит, поскольку распространение инфекции преобладает в развитых странах с большим, тесно контактирующим между собой населением. К тому же пандемия продлилась несколько лет, что достаточно долго.

### 3.3. Построение прогноза с помощью модели CIR

В данном параграфе рассмотрим результаты прогнозирования с использованием дискретной CIR-модели распространения эпидемии, основанной на значении процентного прироста. Сначала построим графики, на которых будет показан процентный прирост количества заболевших с 10 марта до 28 июля 2020 года. Исключением являются графики для Азии, поскольку в странах этой группы эпидемия началась раньше. Также на

графике знаком «\*» отмечены дни, в которые значение прироста последний раз превышает 20%, 10%, 5%, 2%. На оси ординат расположены дни, а на оси абсцисс — процентный прирост заболевших. Ниже приведены несколько примеров (рисунки 11–15).

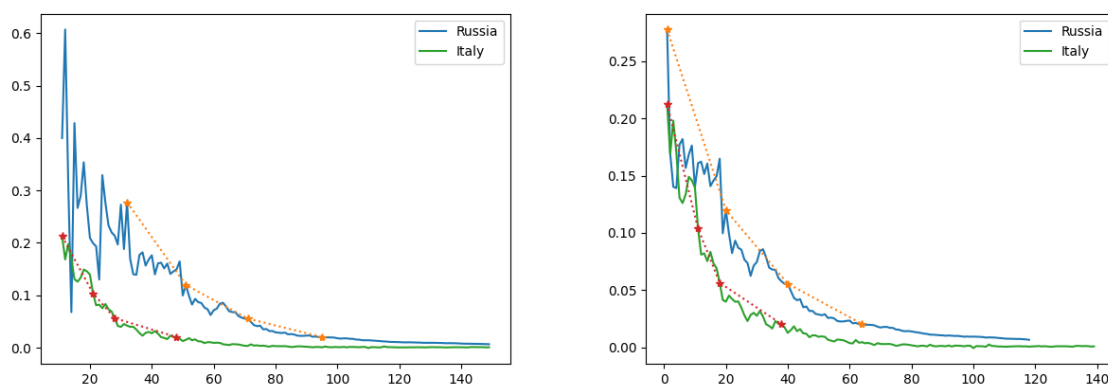


Рисунок 11. Зависимость значения процентного прироста числа заболевших в России и Италии от даты. Слева изображен график с началом отсчета с 10 марта 2020 года. Справа — началом отсчета является день, в который процентный прирост последний раз превышает 20%

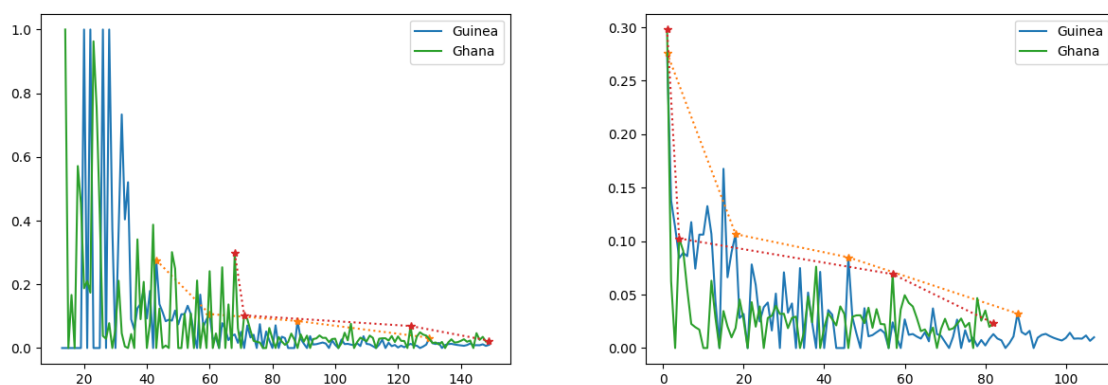


Рисунок 12. Зависимость значения процентного прироста числа заболевших в Гвинее и Гане от даты. Слева изображен график с началом отсчета с 10 марта 2020 года. Справа — началом отсчета является день, в который процентный прирост последний раз превышает 20%



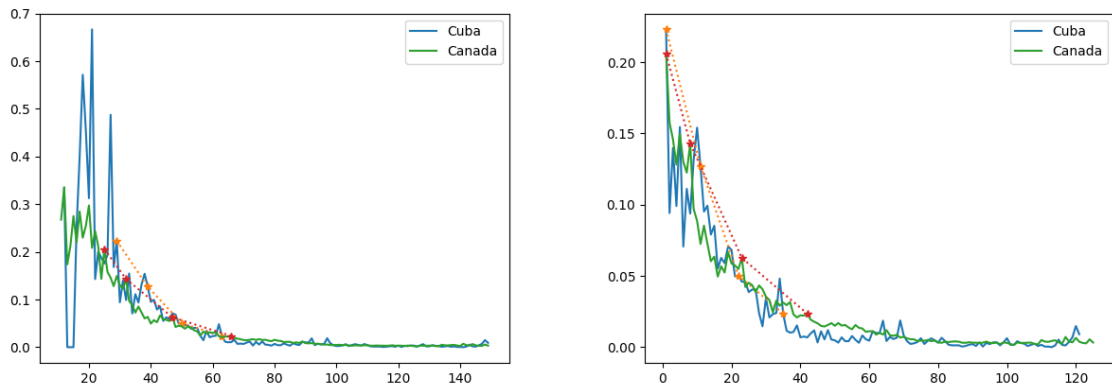


Рисунок 13. Зависимость значения процентного прироста числа заболевших на Кубе и в Канаде от даты. Слева изображен график с началом отсчета с 10 марта 2020 года. Справа — началом отсчета является день, в который процентный прирост последний раз превышает 20%

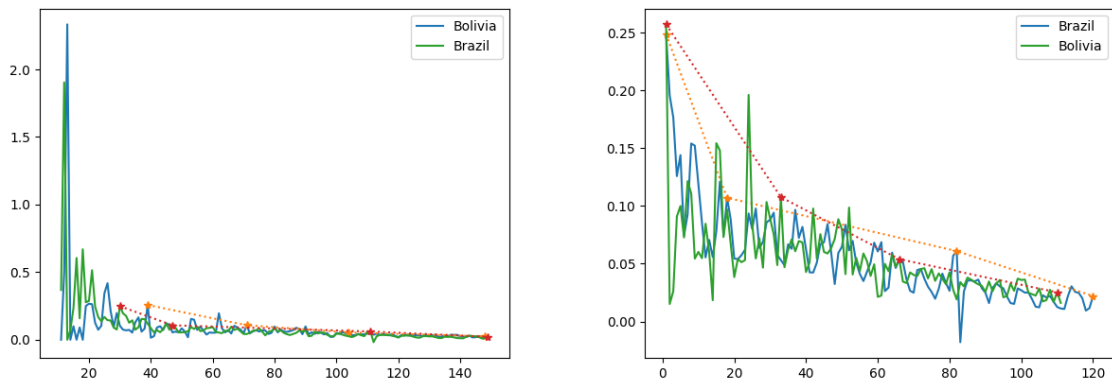


Рисунок 14. Зависимость значения процентного прироста числа заболевших в Боливии и Бразилии от даты. Слева изображен график с началом отсчета с 10 марта 2020 года. Справа — началом отсчета является день, в который процентный прирост последний раз превышает 20%

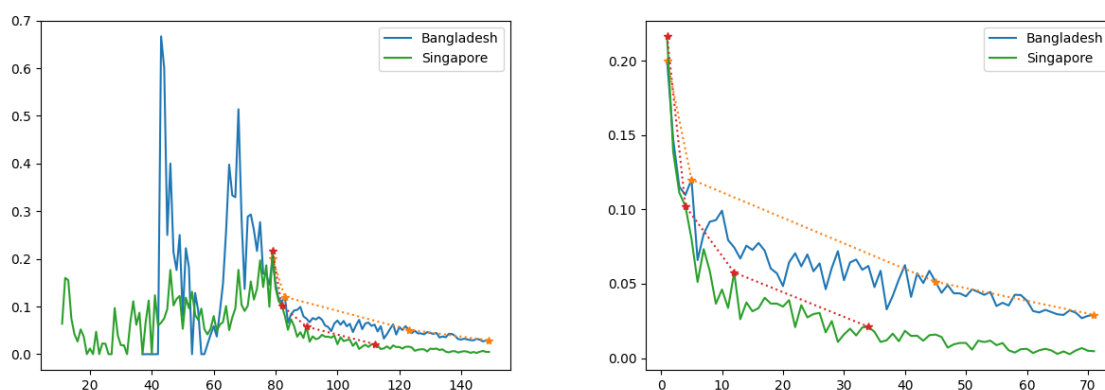


Рисунок 15. Зависимость значения процентного прироста числа заболевших в Бангладеш и Сингапуре от даты. Слева изображен график с началом отсчета с 10 февраля 2020 года. Справа — началом отсчета является день, в который процентный прирост последний раз превышает 20%

В представленной ниже таблице обозначены дни достижения выбранных уровней процентного прироста и период времени между ними для стран, изображенных на графиках.

Таблица 6. Дни достижения разными странами двадцати-, десяти-, пяти- и двухпроцентного прироста количества заболевших.

Страна	20%	10%	5%	2%	20%–10%	10%–5%	5%–2%
Россия	31	50	70	94	19	20	24
Италия	10	20	27	47	10	7	20
Гвинея	46	119	119	147	73	0	28
Гана	24	119	127	148	95	8	21
Куба	28	38	49	62	10	11	13
Канада	24	31	46	65	7	15	19
Боливия	38	70	103	147	32	33	44
Бразилия	29	46	110	148	17	64	38
Бангладеш	78	82	122	148	4	40	26
Сингапур	78	81	89	111	3	8	22

Далее построим прогноз для страны  $X$  на интервале  $[t_{10}, t_5]$ . Пусть данные для всех стран на  $[t_{20}, t_{10}]$  у нас есть. Время  $p_5[X]$  перехода от  $t_{10}$  к  $t_5$  для страны  $X$  так же относится ко времени перехода  $p_{10}[X]$  от  $t_{20}$  к  $t_{10}$ , как соответствующие сроки у страны  $Y$ . То же самое можно сказать и времени  $p_2[X]$ .

Также были выбраны несколько стратегий для определения времени переходов, а именно:

1. Первая стратегия — найдем дни, в которые прирост последний раз превышает выбранный уровень процентного прироста и за время  $p$  обозначаем промежуток времени между этими датами;

2. Вторая стратегия — находим дни, в которые прирост последний раз, а также дни, к которым первый раз прирост достигает данных значений, временем  $p$  будем считать промежутки между серединами этих значений;

3. Третья стратегия — находим дни, в которые прирост последний раз превышает 20%, 10%, 5%, 2% у двух стран-предшественников, за время  $p$  принимаем среднее значение времени переходов этих двух стран.

Рассмотрим и сравним стратегии на примере различных стран-предшественников и стран-последователей в каждой группе. Для начала была взята страна, для которой выполняется прогноз — Россия и страны-предшественники Италия и Финляндия. На оси ординат расположены дни, а на оси абсцисс — количество человек.

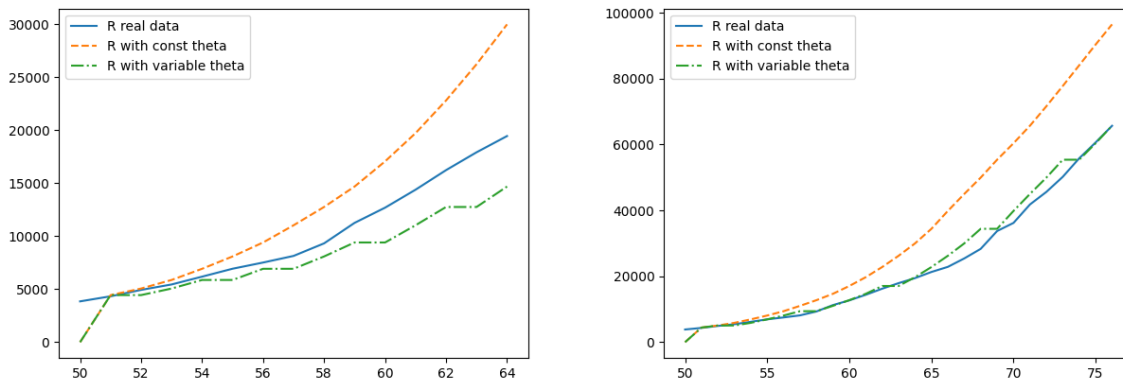


Рисунок 16. Прогноз количества выбывших в России (первая стратегия). Слева изображен график с базовой страной Италией. Справа —Финляндией

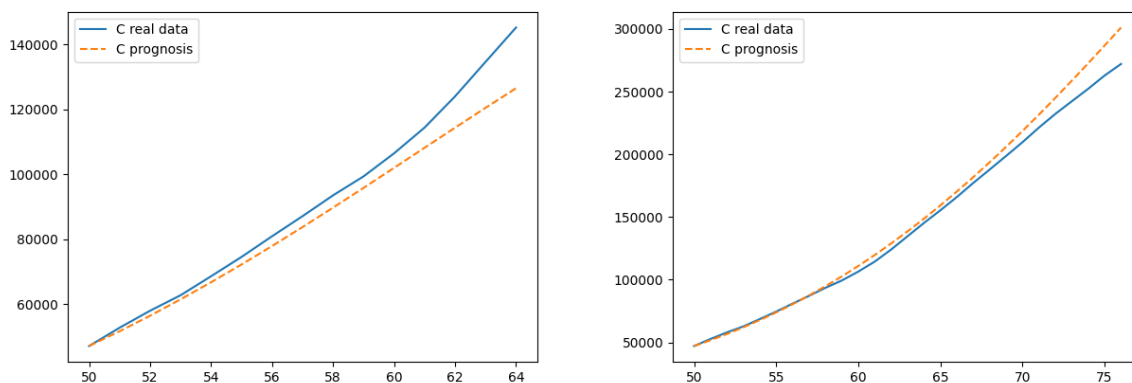


Рисунок 17. Прогноз количества инфицированных в России (первая стратегия). Слева изображен график с базовой страной Италией. Справа —Финляндией

Как видно из графиков, прогноз с использованием первой стратегии по Финляндии получился более точным, чем Италии. По расчетам на основе данных Италии и Франции, день последнего достижения пятипроцентного уровня прироста получился 64 и 76 соответственно. В действительности для России это был 70 день.

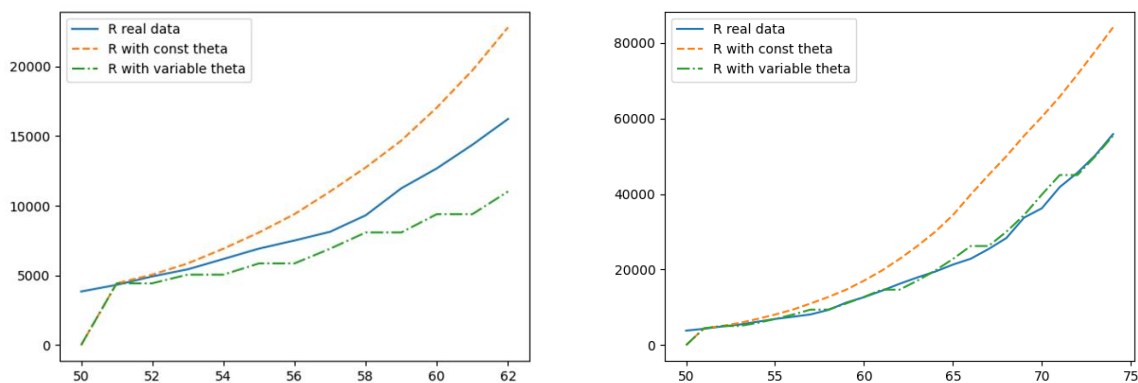


Рисунок 18. Прогноз количества выбывших в России (вторая стратегия). Слева изображен график с базовой страной Италией. Справа — Финляндией

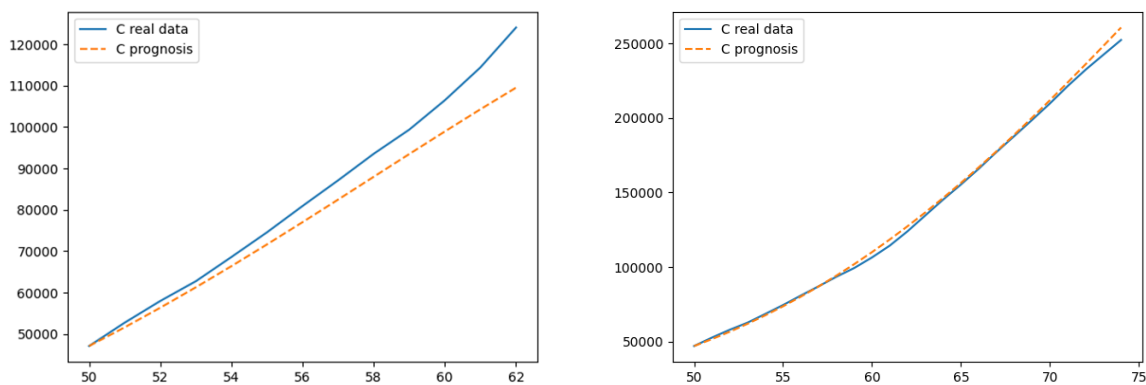


Рисунок 19. Прогноз количества инфицированных в России (вторая стратегия). Слева изображен график с базовой страной Италией. Справа — Финляндией

Прогноз с использованием второй стратегии по Финляндии также получился более точным, чем Италии. Кроме того, вторая стратегия показала себя лучше первой по Финляндии, а вот Италии хуже. По расчетам на основе данных Италии и Франции, день последнего достижения пятипроцентного уровня прироста получился 62 и 74 соответственно. В действительности для России это был 70 день.

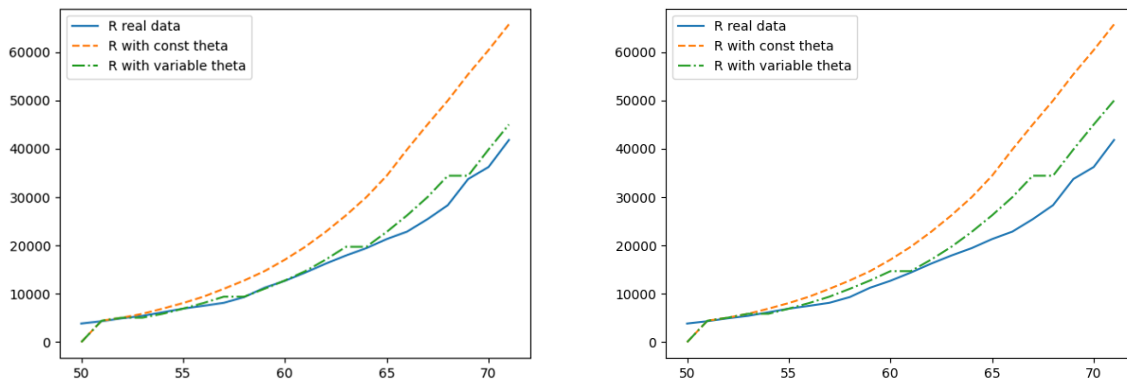


Рисунок 20. Прогноз количества выбывших в России (третья стратегия). Слева изображен график с базовыми странами Италией и Испанией. Справа — Финляндия и Латвия

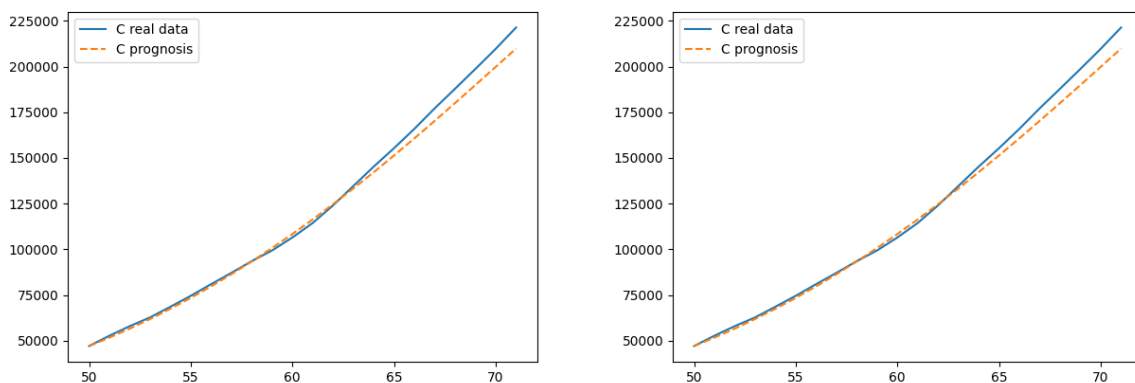


Рисунок 21. Прогноз количества инфицированных в России (третья стратегия). Слева изображен график с базовыми странами Италией и Испанией. Справа — Финляндия и Латвия

Третья стратегия предполагает использование сразу двух базовых стран. Прогноз по Финляндии и Латвии получился хуже, чем по Италии и Испании. По расчетам на основе данных этих пар стран, день последнего достижения пятипроцентного уровня прироста получился 71 и 71 соответственно, что очень близко к действительному значению.

Рассмотри следующую группу — страны Азии. В этом примере возьмем две базовые страны — Афганистан и Иран и страну-последователя, а именно Бангладеш.

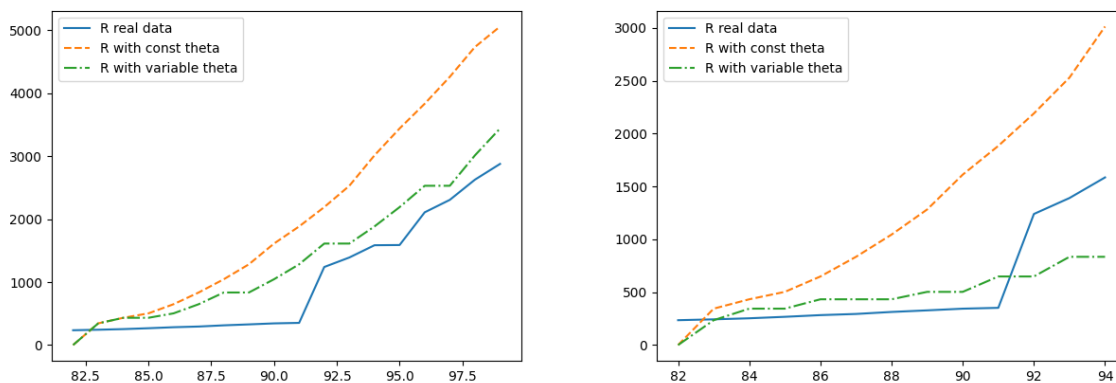


Рисунок 22. Прогноз количества выбывших в Бангладеш (первая стратегия). Слева изображен график по Афганистану. Справа — по Ирану

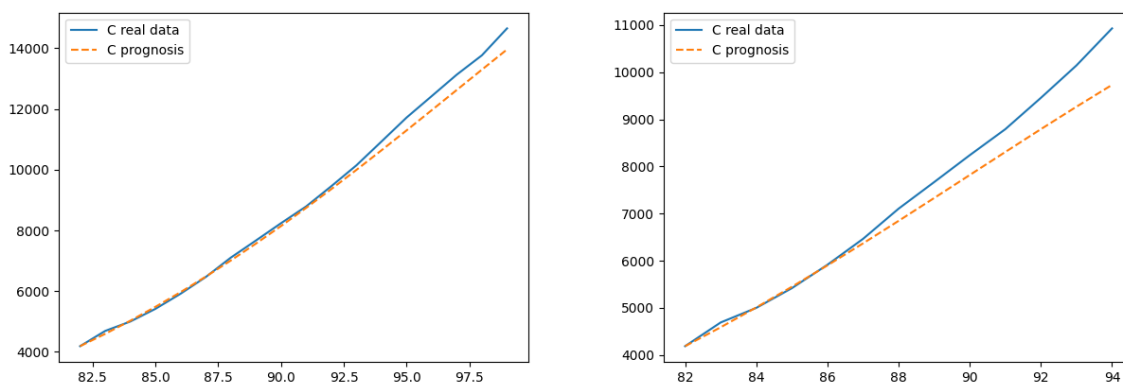


Рисунок 23. Прогноз количества инфицированных в Бангладеш (первая стратегия). Слева изображен график по Афганистану. Справа — по Ирану

Как видно из графиков, прогноз с использованием первой стратегии по данным Афганистана получился немного лучше, чем по данным Ирана. По расчетам на основе данных Афганистана и Ирана, день последнего достижения пятипроцентного уровня прироста получился 99 и 94 соответственно. В действительности для Бангладеш это был 122 день, что довольно сильно отличается от полученных значений.

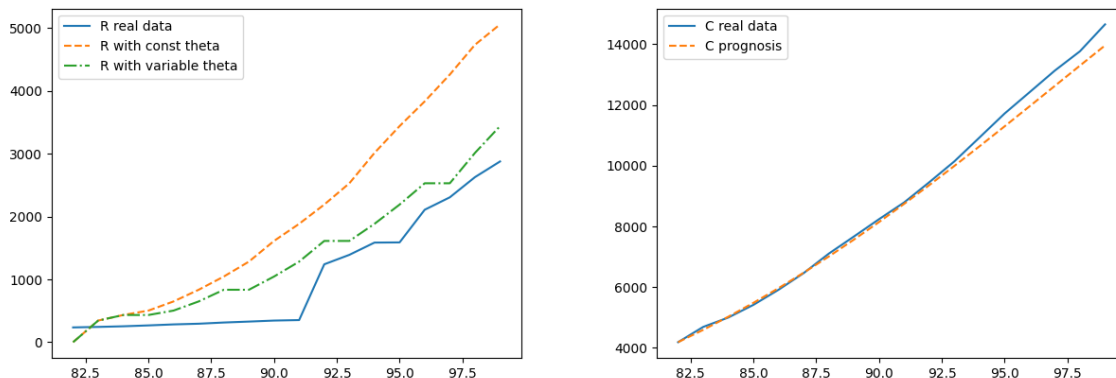


Рисунок 24. Слева изображен прогноз количества выбывших в Бангладеш по Ирану. Справа — прогноз количества инфицированных в Бангладеш по Ирану (вторая стратегия)

Вторая стратегия для данных по Афганистану показала идентичные результаты. Для Ирана же прогноз стал лучше. День последнего достижения пятипроцентного уровня прироста уже стал 98, что лучше, чем предыдущее значение 94.

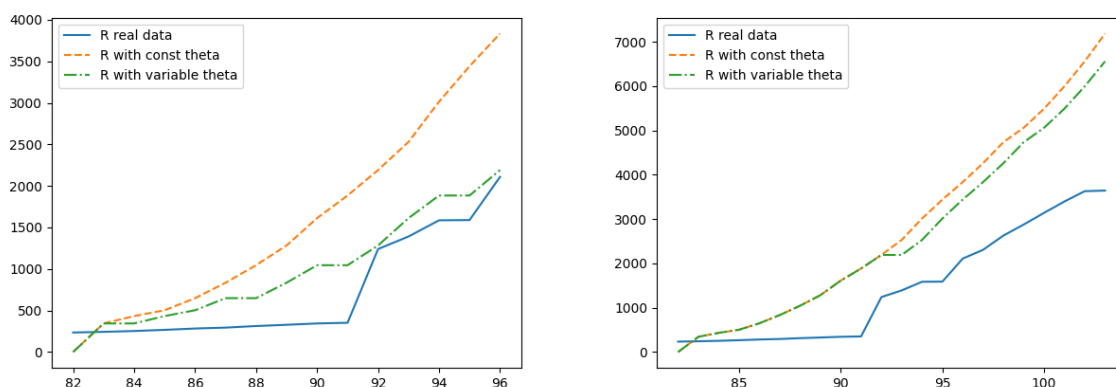


Рисунок 25. Прогноз количества выбывших в Бангладеш (третья стратегия). Слева изображен график по данным Афганистана и Ирана. Справа — по данным Ирана и Азербайджана

Для третьей стратегия были взяты данные следующим парам базовых стран Афганистан и Иран, Иран и Азербайджан. Для первой пары прогноз стал чуть лучше. Для второй пары прогноз выбывших стал хуже, однако прогноз инфицированных стал точнее всех предыдущих. День последнего достижения пятипроцентного уровня прироста 96 и 103 соответственно.



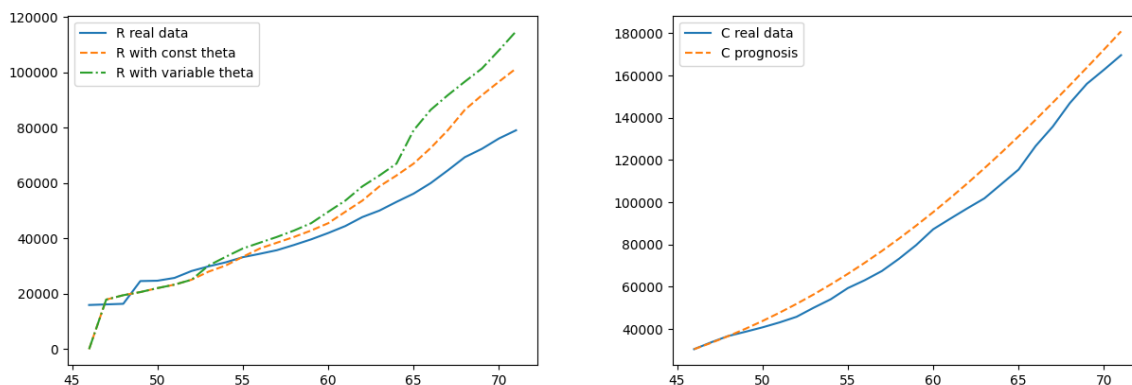


Рисунок 26. Прогноз для Бразилии (первая стратегия). Слева изображен прогноз количества выбывших по данным Парагвая. Справа — количества инфицированных по Парагваю

Первая и вторая стратегия показали идентичный прогноз для Бразилии по данным Парагвая. Днем последнего достижения пятипроцентного уровня прироста по расчетам является 71, а в действительности 110. Это довольно большой разрыв, однако количество инфицированных удалось спрогнозировать достаточно хорошо.

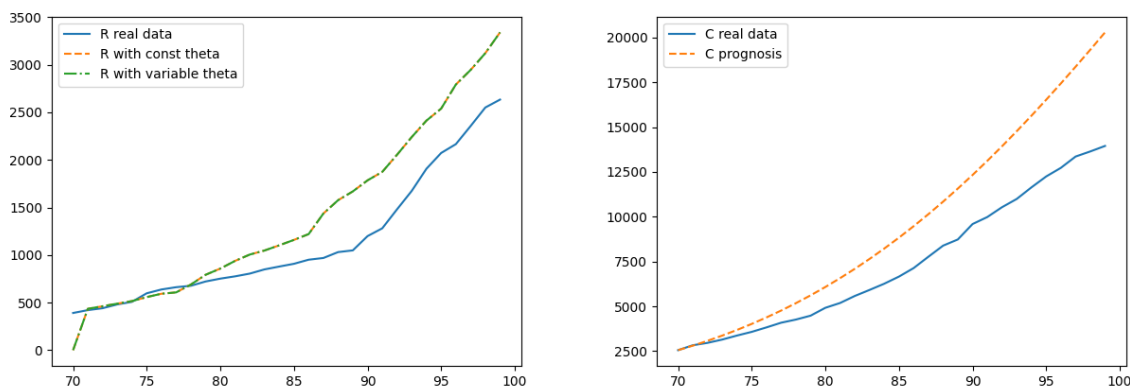


Рисунок 27. Прогноз для Боливии (третья стратегия). Слева изображен прогноз количества выбывших по данным Чили и Бразилии. Справа — количества инфицированных по Чили и Бразилии

Для третьей стратегии базовыми страна были взяты Чили и Бразилия, а страной-последователем стала Боливия. Прогноз получилась также недостаточно точны по сравнению с предыдущими группа. Возможно, это связано с теми данными, которые эти страны предоставляют. Днем последнего

достижения пятипроцентного уровня прироста по расчетам является 99 день, в действительности 103.

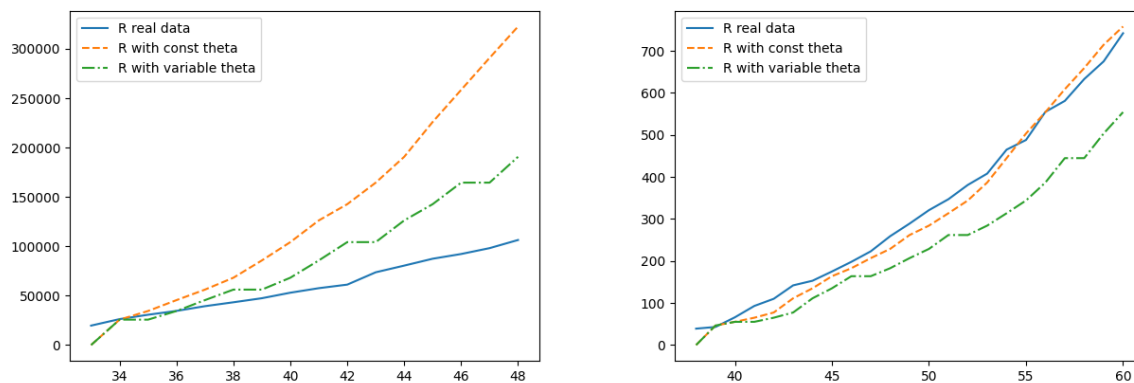


Рисунок 28. Прогноз выбывших по данным базовой страны — Канады (первая стратегия). Слева изображен прогноз для США. Справа — для Кубы

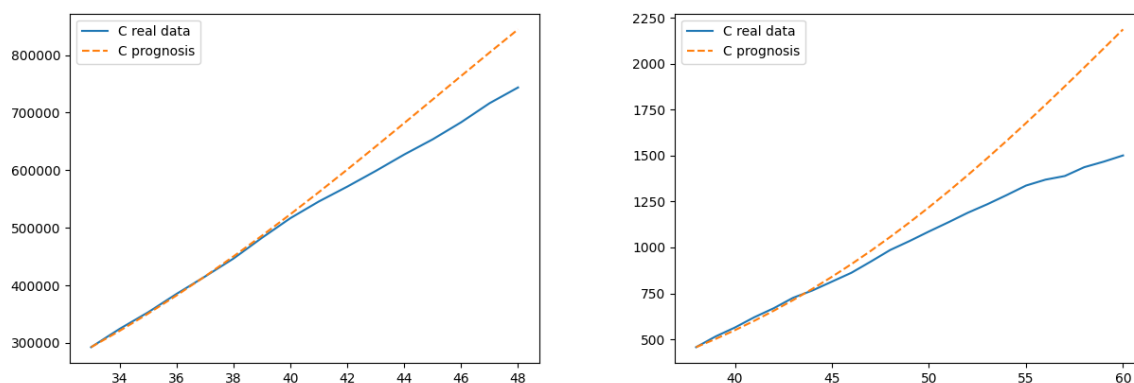


Рисунок 29. Прогноз инфицированных по данным базовой страны — Канады (первая стратегия). Слева изображен прогноз для США. Справа — для Кубы

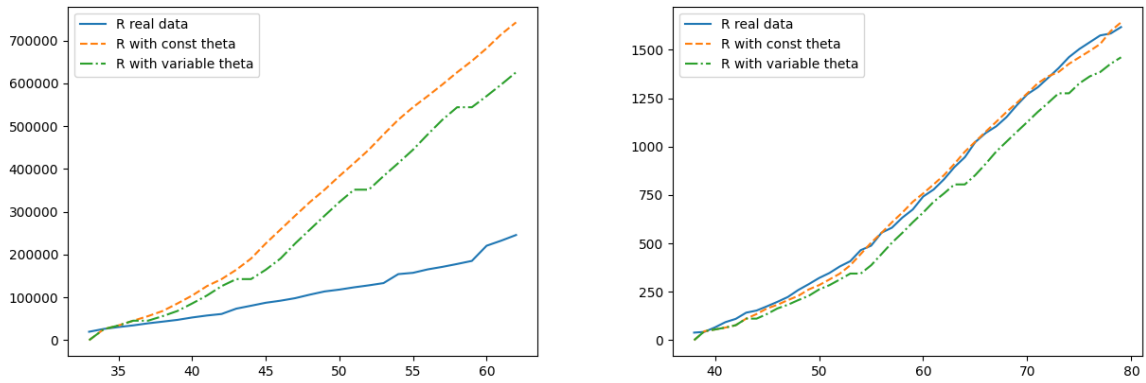


Рисунок 30. Прогноз выбывших по данным базовой страны — Канады (вторая стратегия). Слева изображен прогноз для США. Справа — для Кубы

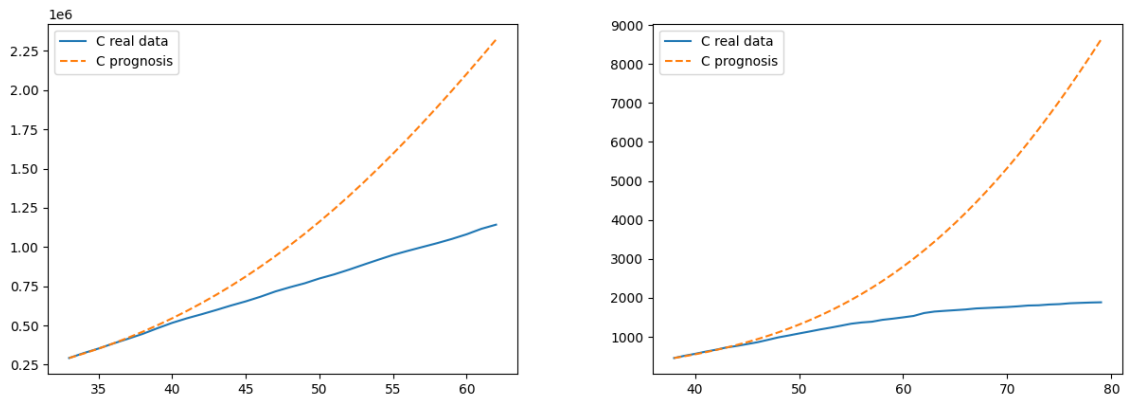


Рисунок 31. Прогноз инфицированных по данным базовой страны — Канады (вторая стратегия). Слева изображен прогноз для США. Справа — для Кубы

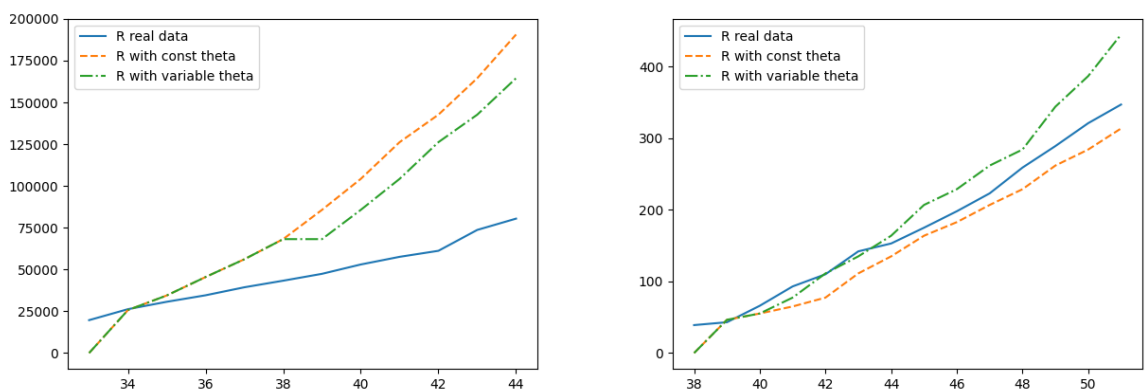


Рисунок 32. Прогноз выбывших (третья стратегия). Слева изображен прогноз для США на основе данных Кубы и Канады. Справа — для Кубы на основе данных Мексики и Канады

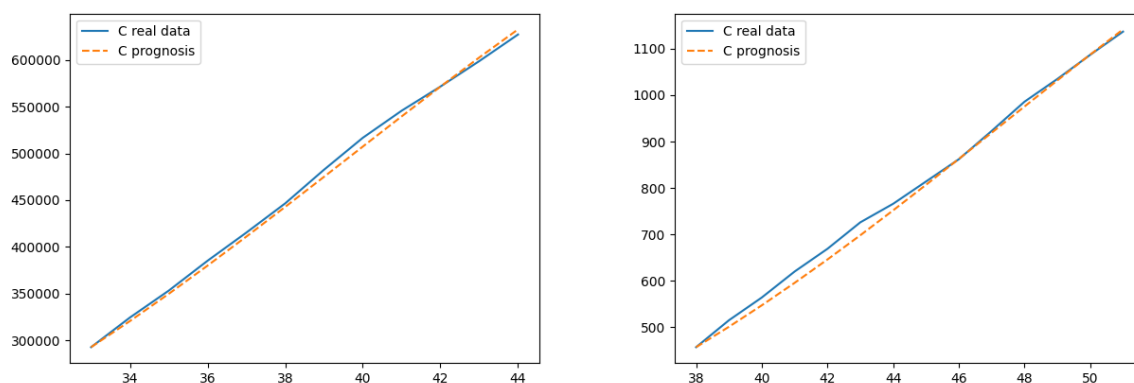


Рисунок 33. Прогноз инфицированных (третья стратегия). Слева изображен прогноз для США на основе данных Кубы и Канады. Справа — для Кубы на основе данных Мексики и Канады

Как можно видеть из представленных выше графиков, наилучшим образом количество инфицированных было спрогнозировано с помощью третьей стратегии как для Кубы, так и для США. Количество выбывших здесь также было получено также достаточно хорошо, по сравнению с остальными стратегиями. Исключение является лишь прогноз для Кубы по второй стратегии. День последнего достижения пятипроцентного уровня прироста ближе всего также получился благодаря использованию двух базовых стран. Результаты расчетов показаны в таблице.

Таблица 7. День последнего достижения пятипроцентного уровня прироста

Страна/стратегия	США		Куба	
	Прогнозное значение	Реальные данные	Прогнозное значение	Реальные данные
Первая стратегия	48	40	60	49
Вторая стратегия	62		79	
Третья стратегия	44		51	

Таким же образом построим прогноз для страны  $X$  на интервале  $[t_5, t_2]$ . Пусть теперь у нас имеются данные для всех стран на  $[t_{10}, t_5]$ . Время  $p_2[X]$  перехода от  $t_5$  к  $t_2$  для страны  $X$  так же относится ко времени перехода  $p_5[X]$  от  $t_{10}$  к  $t_5$ , как соответствующие сроки у страны  $Y$ .

Приведем в пример снова страны из Европы. Страной-последователем также будет являться Россия. Базовыми странами выберем Германию и Швейцарию, хотя Финляндии, которая была взята ранее имеет также хороший прогноз для нового промежутка.

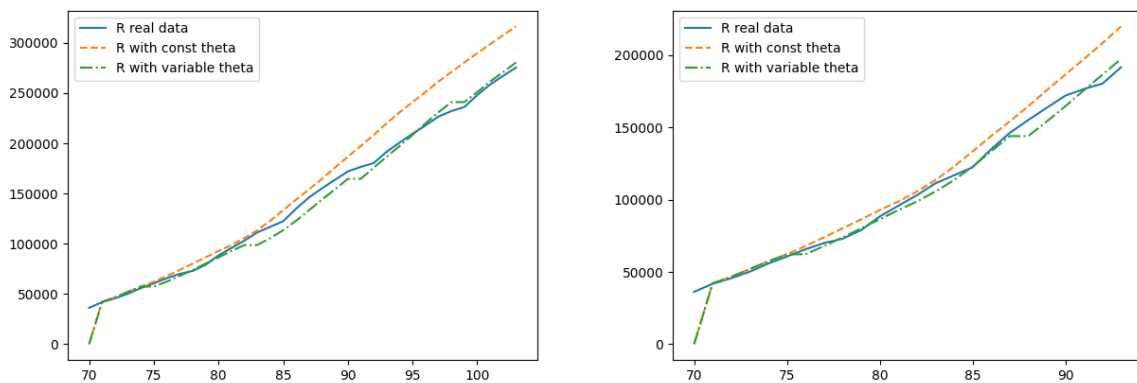


Рисунок 34. Прогноз количества выбывших в России (первая стратегия). Слева изображен график с базовой страной Германией. Справа — Швейцарией

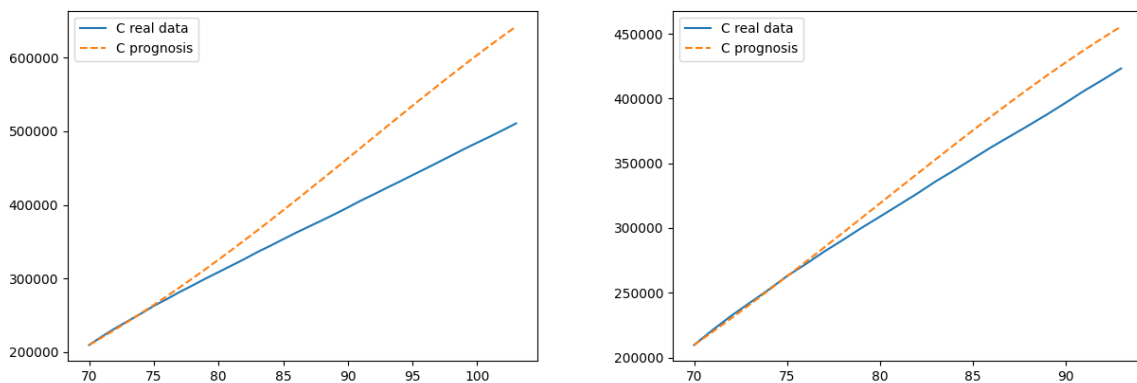


Рисунок 35. Прогноз количества инфицированных в России (первая стратегия). Слева изображен график с базовой страной Германией. Справа — Швейцарией

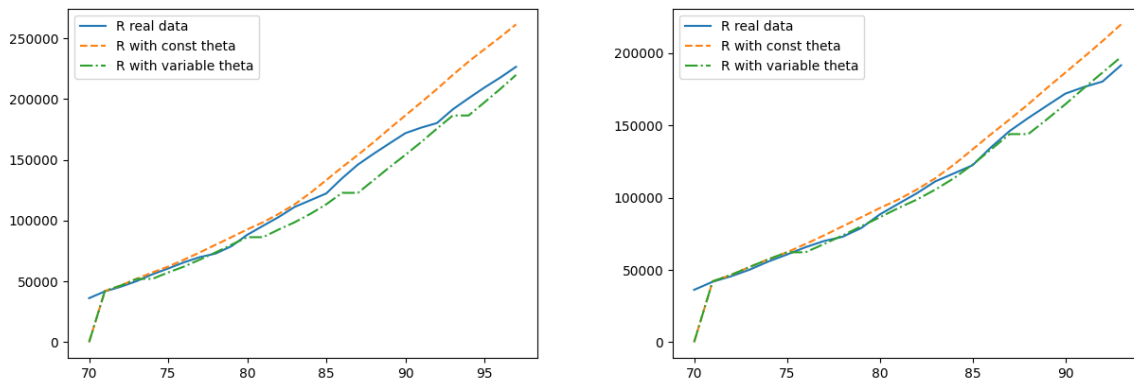


Рисунок 36. Прогноз количества выбывших в России (вторая стратегия). Слева изображен график с базовой страной Германией. Справа — Швейцарией

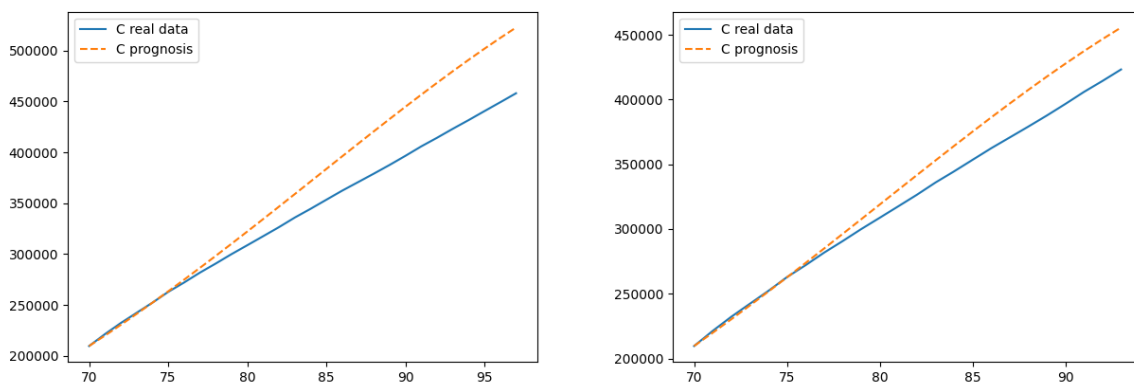


Рисунок 37. Прогноз количества инфицированных в России (вторая стратегия). Слева изображен график с базовой страной Германией. Справа — Швейцарией

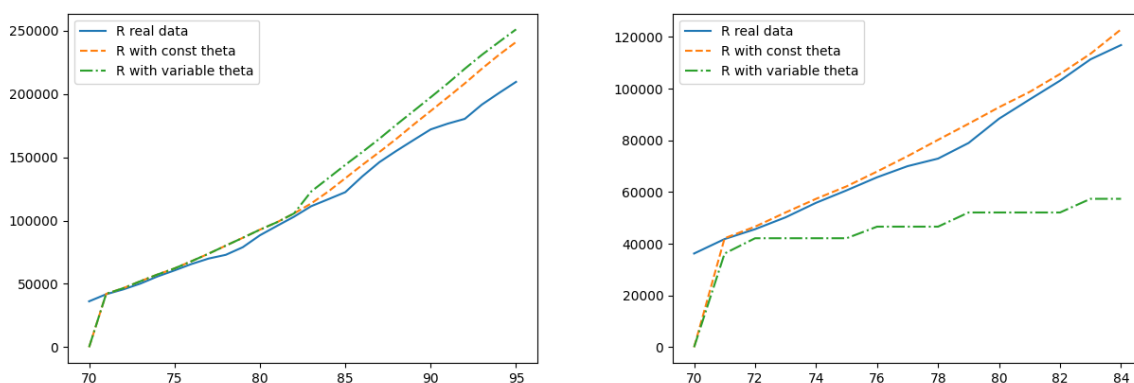


Рисунок 38. Прогноз количества выбывших в России (третья стратегия). Слева изображен график с базовыми странами Германией и Италией. Справа — Швейцарией и Испанией.

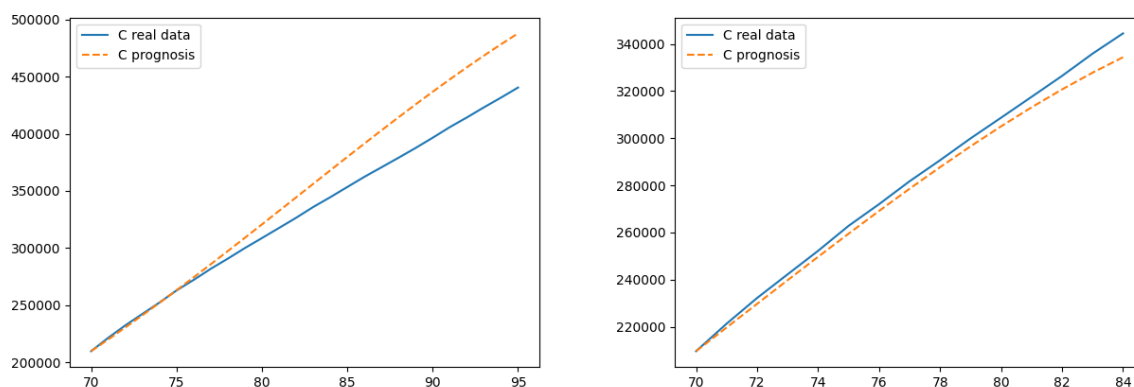


Рисунок 39. Прогноз количества инфицированных в России (третья стратегия). Слева изображен график с базовыми странами Германией и Италией. Справа — Швейцарией и Испанией

На графиках видно, что для первых двух стратегий на промежутке  $[t_5, t_2]$  количество выбывших прогнозировалось достаточно, хотя по данным Швейцарии немного лучше, чем по данным Германии. Использование первого и последнего дня достижения уровня процентного прироста даже улучшило прогноз по данным Германии. А во включение в модель дополнительной базовой страны ухудшило прогноз для Швейцарии. Причем пара Швейцария и Испания показывает наилучший результат среди всех пар стран Европы. Добавление же данных Италии к данным Германии, наоборот позволило улучшить прогноз и приблизить значение дня последнего достижения двухпроцентного уровня прироста.

Таблица 8. День последнего достижения двухпроцентного уровня прироста в России по данным Германии и Швейцарии

Страна/стратегия	Германия		Швейцария	
	Прогнозное значение	Реальные данные	Прогнозное значение	Реальные данные
Первая стратегия	103	94	93	94
Вторая стратегия	97		93	
Третья стратегия	95		84	

Самыми сложными для прогнозирования стали страны Африки. Возможно, это связано с низким уровнем здравоохранения в стране, в том числе сбором и публикацией информации об инфицированных, умерших и выбывших. Рассмотрим сначала промежуток  $[t_{10}, t_5]$ .

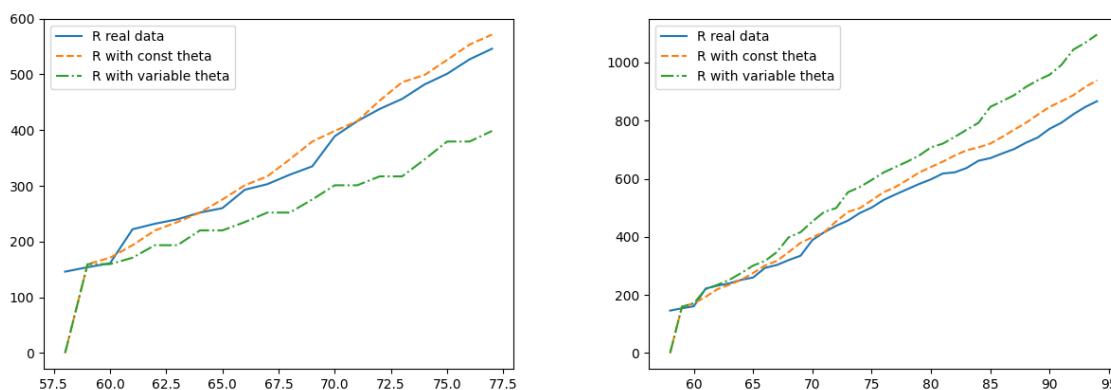


Рисунок 40. Прогноз количества выбывших в Мали (первая стратегия). Слева изображен график с базовой страной Гвинея. Справа — с базовой страной Марокко

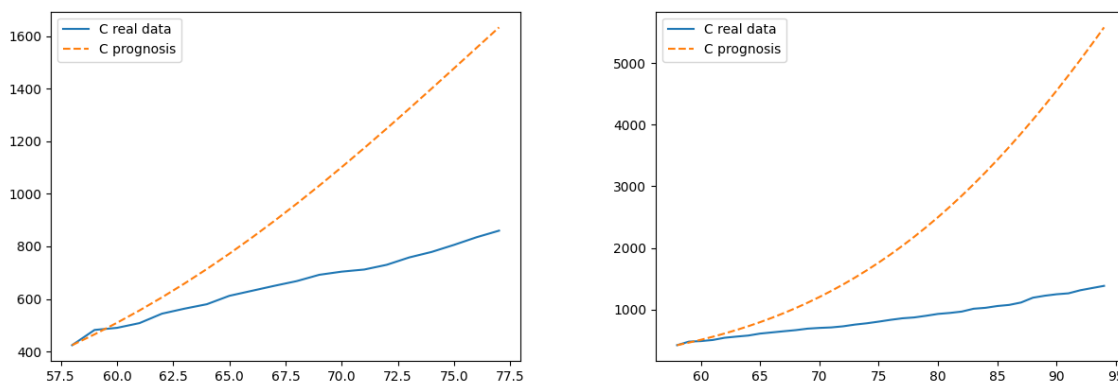


Рисунок 41. Прогноз количества инфицированных в Мали (первая стратегия). Слева изображен график с базовой страной Гвинея. Справа — с базовой страной Марокко

Первая и вторая стратегия дает одинаковый прогноз. Как видно из графиков, изменение стратегии с первой или второй на третью улучшает прогноз для выбывших в обоих случаях. А вот количество инфицированных во всех стратегиях прогнозируется недостаточно точно.



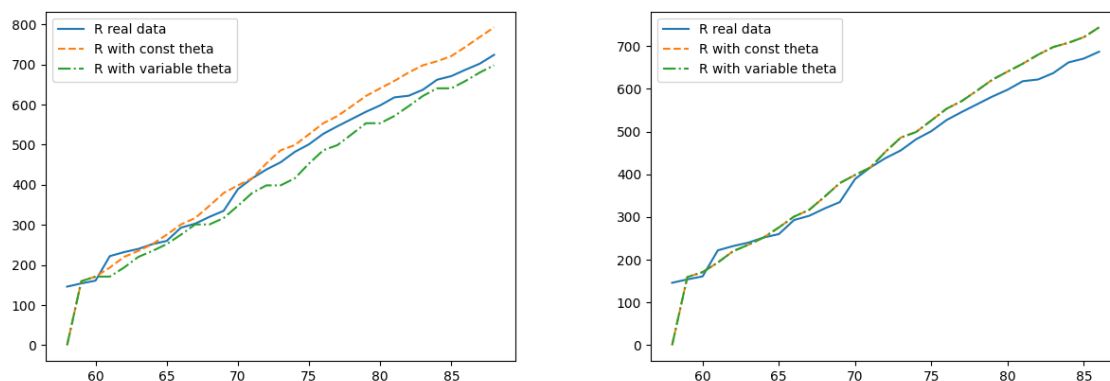


Рисунок 42. Прогноз количества выбывших в Мали (первая стратегия). Слева изображен график с базовыми странами Гвинея и Алжир. Справа — с базовыми странами Марокко и Гвинея

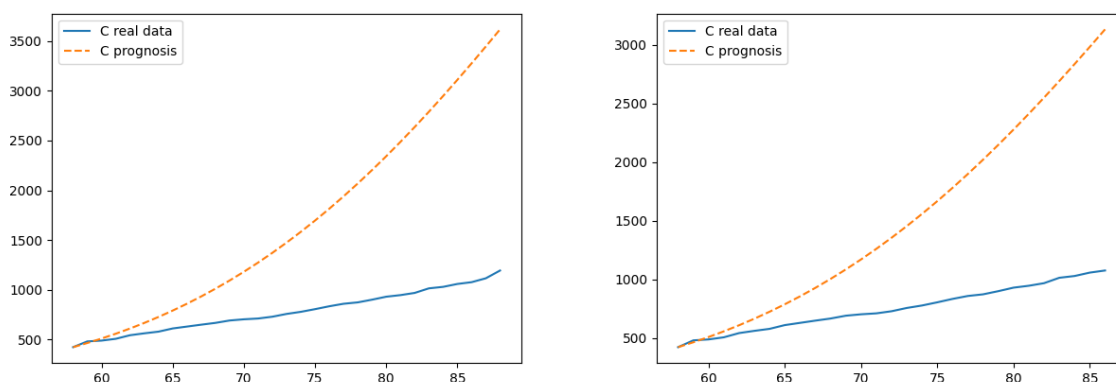


Рисунок 43. Прогноз количества инфицированных в Мали (первая стратегия). Слева изображен график с базовыми странами Гвинея и Алжир. Справа — с базовыми странами Марокко и Гвинея

Таблица 9. День последнего достижения пятипроцентного уровня прироста в Мали по данным Гвинеи и Марокко

Страна/стратегия	Гвинея		Марокко	
	Прогнозное значение	Реальные данные	Прогнозное значение	Реальные данные
Первая стратегия	77	100	94	100
Вторая стратегия	77		94	
Третья стратегия	88		86	

Похожая ситуация происходит и на промежутке  $[t_5, t_2]$ . Количество выбывших, как и в предыдущем случае, можно получить более точно, используя третью стратегию, но на прогноз количества инфицированных это не влияет. Кроме того, применяя данные двух стран был получен день последнего достижения двухпроцентного уровня прироста, который совпадает с реальными данными.

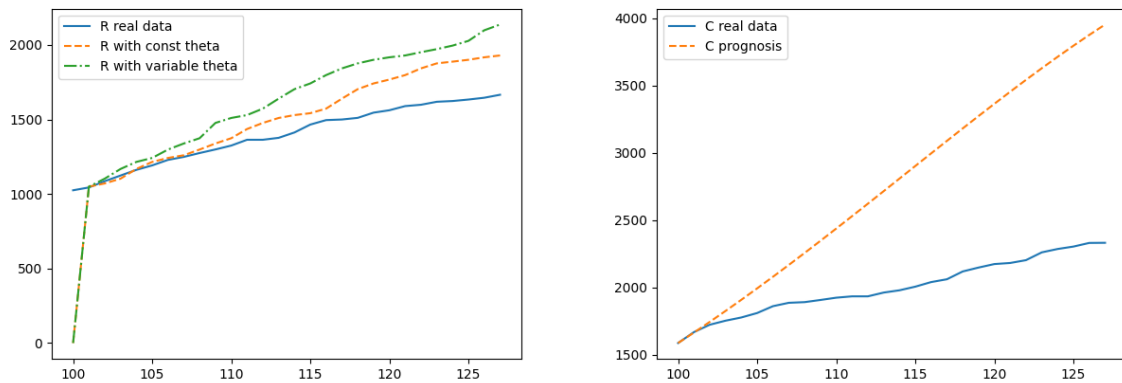


Рисунок 44. Прогноз для Мали с базовой страной Гана (первая стратегия). Слева изображен прогноз выбывших. Справа — инфицированных

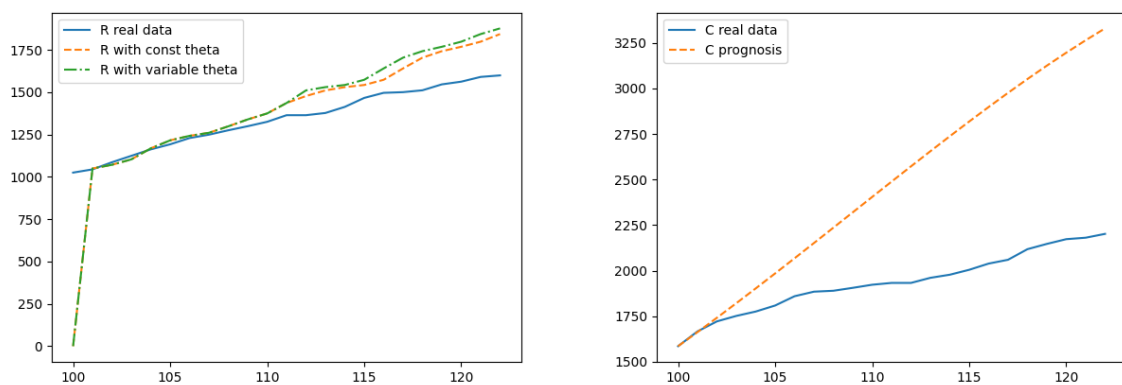


Рисунок 45. Прогноз для Мали с базовыми странами Гана и Эфиопия (третья стратегия). Слева изображен прогноз выбывших. Справа — инфицированных

Таким образом, в данном параграфе с помощью SIR-модели были спрогнозированы количество выбывших и инфицированных для различных стран. Благодаря данной модели были получены более точные прогнозы на длительный промежуток времени по сравнению с рассмотренной ранее моделью SIR. Кроме того, если в модели использовать несколько стран-

предшественников, в подавляющем числе случаев показатели принимают значения, которые ближе к реальным. Это может быть связано со одинаковыми мерами, которые страны принимают для того, чтобы справиться с болезнью. Также на схожую тенденцию протекания эпидемии может влиять менталитет и отношение населения к принятым ограничениям (халатное или наоборот строгое соблюдение правил), с погодные условия, которые преобладают в определенном регионе и другие объединяющие факторы.

## Заключение

Тема математического моделирования инфекционных заболеваний достаточно популярна в научной среде. Это объясняется в том числе тем, что болезни влияют на все сферы жизни человека. Эпидемии и даже сезонные заболевания, такие как грипп, воздействуют на экономику как страны, так и всего мира в целом. Некоторые исследования показывают, что дети, рожденные во время эпидемий, чаще болеют в течении всей жизни и имеют высокий риск приобретения серьезных медицинских проблем. Кроме того, слабеет человеческий капитал, поскольку взрослые люди либо болеют сами, либо выходят на больничный, чтобы присматривать за детьми. Торговля из-за ограничений также страдает, как и другие финансовые сферы. На борьбу с болезнями государства вынуждены тратить огромные денежные ресурсы. Именно поэтому важно понимать, как будет протекать инфекционный процесс, чтобы правительство могло подбирать эффективную политику, способную минимизировать ущерб для экономики и других сфер жизни человека.

В работе, были рассмотрены различные методы математического моделирования заболеваний. Самой распространенной является модель SIR. Основным преимуществом данной модели является ее простота, поскольку требует небольшое количество данных. Однако, это преимущество может стать и ее недостатком. Именно поэтому стали появляться различные модификации, такие как модель SIER. Также часто используются модели с использованием временных рядов или же методов машинного обучения. Еще одним видом являются сетевые модели заболеваний, которые позволяют оценить в том числе влияние мер по борьбе с инфекциями.

На данном исследовании были использованы две модели, а именно SIR и CIR. Для прогноза были взяты данные 40 стран с разных континентов о количестве заболевших, умерших и выбывших со времени начала пандемии

COVID-19 по начало 2021 года. Более того для обеих моделей был написан программный код. На основе проведенного исследования был сделан вывод, что для прогнозирования данного заболевания больше подходит модель CIR. Она позволяет делать прогноз для одной страны на основе данных другой страны. Также лучше использовать несколько базовых государств. Однако страны должны иметь схожие меры борьбы с болезнью, схожее отношение населения к данным ограничениям и другими факторами. Кроме того, данная модель позволяет получить хороший прогноз на более длительный срок, чем модель SIR. Именно поэтому данному способу было отдано предпочтение в моделировании пандемии COVID-19.

## Список литературы

1. Барулина, М. А. Основы математического моделирования и обработки данных в медицине: учебно-методическое пособие / М. А. Барулина. — Самара: 2022. — 63 с. — Текст: электронный // Лань: электронно-библиотечная система. — URL: <https://e.lanbook.com/book/326510> (дата обращения: 15.05.2023). — Режим доступа: для авториз. пользователей.
2. Баталин Р.М. Оптимальное управление в моделях эпидемий трансмиссивных заболеваний с SEI-SEIR системами / Р.М. Баталин, В.А. Терлецкий // Известия Иркутского государственного университета. Серия: Математика. — 2015. — №14 — С.18-30
3. Белоусова Е.П. Построение модели оценки уровня заболеваемости в период эпидемии / Е.П. Белоусова, А.А. Данилов // Регион: системы, экономика, управление. — 2021. — №1 (52). — С. 295-206.
4. Бруттан Ю.В. Исследование подходов к моделированию вирусных инфекций / Ю.В. Бруттан, Е.С. Туманова // Вестник псковского государственного университета. Серия: технические науки. — 2020. — №11. — С.38-48.
5. Вотякова Л.Р., Кадочникова О.А., Прогнозирование вспышек опасных заболеваний методами машинного обучения / Л.Р. Вотякова, О.А. Кадочникова // Заметки ученого. — 2021. — вып.2. — С.35-40
6. Гареева И.А. Социальные изменения в условиях пандемии COVID-19 / И.А. Гареева // Власть и управление на Востоке России. — 2021. — № 4 (97). — С. 99-109.
7. Герасименко П. В. Моделирование количества случаев заболевания коронавирусом COVID-19 в Санкт-Петербурге в период 2020-2022 гг. // Здоровье мегаполиса — 2022. — Т. 3 № 3. — С. 30-38
8. Головинский П. А. Математическое моделирование распространения вирусов с длинной инкубационной фазой в тесном мире. / П.А.

- Головинский // Вестник ВГУ. Серия: Системный анализ и информационные технологии. — 2020. — №2. — С. 5-14.
9. Горлач, Б. А. Математическое моделирование. Построение моделей и численная реализация / Б. А. Горлач, В. Г. Шахов. — 5-е изд., стер. — Санкт-Петербург: Лань, 2023. — 292 с. — ISBN 978-5-507-46275-9. — Текст: электронный. — Режим доступа: <https://e.lanbook.com/book/305219>, для авториз. пользователей. (дата обращения: 15.05.2023)
10. Горчакова Н.С. Роль экономики здоровья в формировании человеческого потенциала / Н.С. Горчакова // Известия Саратовского университета. Новая серия. — 2020. — Т. 20 №2 — С.134-139.
11. Жумартова Б.О. Применение SIR модели в моделировании эпидемий / Б.О. Жумартова, Р.С. Ысмагул // Международный журнал гуманитарных и естественных наук. — 2021. — vol. 12-2(63) — С.6-8.
12. Захаров В.В. Балансовая модель эпидемии COVID-19 на основе процентного прироста / В.В. Захаров, Ю.Е. Балыкина // Информатика и автоматизация. — 2021. — 20(5). — С.1034-1064. — Режим доступа: <https://doi.org/10.15622/20.5.2>, свободный (15.05.2023)
13. Захаров В.В. Прогнозирование динамики эпидемии коронавируса (COVID-19) на основе метода прецедентов. / В.В. Захаров, Ю.Е. Балыкина // Вестник Санкт-Петербургского университета. – Прикладная математика. Информатика. Процессы управления. — 2020. — т. 16. Вып. 3. — С. 249-259.
14. Зикунова И.В. Пандемийный шок: методологическое объяснение и новый общественный договор / И.В. Зикунова // Власть и управление на Востоке России. — 2021. — №3(96). — С. 45-51.
15. Иванов В. В. Математическое моделирование : учебное пособие / В. В. Иванов, О. В. Кузьмина. — 2-е, испр. и доп. — Йошкар-Ола: ПГТУ, 2021. — 116 с. — ISBN 978-5-8158-2246-7. — Текст: электронный // Лань: электронно-библиотечная система. — URL:

- <https://e.lanbook.com/book/237254> (дата обращения: 15.05.2023). — Режим доступа: для авториз. пользователей.
- 16.Иванова К. Д. Влияние эпидемий на исторические события / К.Д. Иванова, О.А. Горощенова // Молодежный вестник ИрГТУ. — 2022. — Т. 12 №4. — С. 809-815
- 17.Катаргин Н. В. Экономико-математическое моделирование / Н. В. Катаргин. — Санкт-Петербург: Лань, 2023. — 256 с. — ISBN 978-5-507-45667-3. — Текст: электронный // Лань: электронно-библиотечная система. — URL: <https://e.lanbook.com/book/279791> (дата обращения: 12.05.2023). — Режим доступа: для авториз. пользователей.
- 18.Кольцова Э. М. Математическое моделирование распространения эпидемии коронавируса COVID-19 в ряде европейских, азиатских стран, Израиле и России / Э.М. Кольцова, Е.С. Куркина, А.М. Васецкий // Проблемы экономики и юридической практики. — 2020. — Т. XVI. №2. — С. 154-165.
19. Кондратьев М.А. Методы прогнозирования и модели распространения заболеваний. / М.А. Кондратьев // Компьютерные исследования и моделирование. 2013. — т. 5. № 5 — С. 863-882.
- 20.Концевая Н. В. Об оценке эффективности карантинных мер и прогнозе сроков окончания эпидемии / Н.В. Концевая // Проблемы экономики и юридической практики. — 2020. — Т. 16. №3. — С. 69-75.
- 21.Лапшина С.Ю. Мультиагентное моделирование процессов распространения массовых эпидемий с использованием суперкомпьютеров / С.Ю. Лапшина // Программные продукты и системы. — 2018. — Т. 31 №3. — С. 640-644.
- 22.Макаров В.Л. Моделирование эпидемии COVID-19 - преимущества агент-ориентированного подхода / В.Л. Макаров, А.Р. Бахтизин, Е.Д. Сушко, А.Ф. Агеева// Экономические и социальные перемены: факты, тенденции, прогноз. — 2020. — Т.13 №4. — С.58-70.



23. Матвеев А.В. Математическое моделирование оценки эффективности мер против распространения эпидемии COVID-19. / А.В. Матвеев // Национальная безопасность и стратегическое планирование. — 2020. — № 1. — С. 23-39.
24. Мелик-Гусейнов Д.В. Регрессионные модели прогнозирования количества летальных исходов при новой коронавирусной инфекции / Д.В. Мелик-Гусейнов, Н.Н. Карякин, А.С. Благодирова, В.И. Клишко, А.П. Баврина, О.В. Другова, Н.В. Саперкин, О.В. Ковалишенина // Современные технологии в медицине. — 2020. — №2. — С.6-11
25. Ндиайе С.М. Эпидемическая модель малярии без вакцинации и при ее наличии. Ч.1. Модель малярии без вакцинации / С.М. Ндиайе, Е.М. Парилина // Вестник Санкт-Петербургского университета. Прикладная математика. Информатика. Процессы управления. — 2022. — Т. 18. Вып. 2. — С. 263-277.
26. Обеснюк В.Ф. Динамика локальной эпидемической вспышки COVID-19 через призму компартимент-моделирования / В.Ф. Обеснюк // Анализ риска здоровью. — 2020. — №2. — С. 83-89
27. Одинцова Е.В. Анализ феноменов здоровья и здорового образа жизни в формате социологии / Е.В. Одинцова // Известия Саратовского университета. Новая серия. Серия Социология. Политология. — 2020. — №3.
28. Панталева А.А. COVID-19 и проблемы занятости в странах БРИКС / А.А. Панталева, Е.Д. Кашкина, Т.В. Брянская // Вестник Пензенского государственного университета. — 2021. — № 1. — С. 46-49.
29. Романюха А. А. Математические модели в иммунологии и эпидемиологии инфекционных заболеваний: монография / А. А. Романюха. — 3-е изд. — Москва: Лаборатория знаний, 2020. — 296 с. — ISBN 978-5-00101-710-3. — Текст: электронный // Лань: электронно-библиотечная система. — URL: <https://e.lanbook.com/book/135547> (дата обращения: 15.05.2023). — Режим доступа: для авториз. пользователей.

30. Семахин А. М. Методы математического моделирования: учебное пособие / А. М. Семахин. — Курган: КГУ, 2022. — 160 с. — ISBN 978-5-4217-0607-6. — Текст: электронный // Лань: электронно-библиотечная система. — URL: <https://e.lanbook.com/book/300314> (дата обращения: 15.05.2023). — Режим доступа: для авториз. пользователей.
31. Соколовский В.Л. Пространственно-временное моделирование эпидемии COVID-19 / В.Л. Соколовский, Г.Б. Фурман, Д.А. Полянская, Е.Г. Фурман // Анализ риска здоровью. — 2021. — С.23-33. — Режим доступа: [10.21668/health.rlsk/2021.1.03](https://doi.org/10.21668/health.rlsk/2021.1.03), свободный (15.05.2023)
32. Степанов В.С. Зависимость уровня смертности в регионах от распространенности активных носителей SARS-CoV-2 и ресурсов организации здравоохранения / В.С. Степанов // Анализ риска здоровью. — 2020. — №4. — С.12-19
33. Тарасова С.А. Прогнозирование временного ряда инфекционной заболеваемости / С.А. Тарасова // Программные продукты и системы. — 2019. — Т. 32 №2.
34. Хорольская И.В. Оптимальные математические модели в изучении эпидемических процессов заболеваний гриппом // Междисциплинарные исследования: опыт прошлого, возможности настоящего, стратегии будущего. — 2021. — №3.
35. Шабунин А.В. SIRS-модель распространения инфекций с динамическим регулированием численности популяции: Исследование методом вероятностных клеточных автоматов / А.В. Шабунин // Известия вузов. ПНД. — 2019. — Т. 27, вып. 2. — С. 5-20. — Режим доступа: [10.18500/0869-6632-2019-27-2-5-20](https://doi.org/10.18500/0869-6632-2019-27-2-5-20), свободный (15.05.2023)
36. Шабунин А.В. Моделирование эпидемий решетками клеточных автоматов. SIRS модель с учетом воспроизводства и миграции / А.В. Шабунин // Известия Саратовского университета. Новая серия. Серия: Физика. — 2020. — Т. 20, вып. 4. — С. 278-287. — Режим доступа: [10.18500/1817-3020-2020-20-4-278-287](https://doi.org/10.18500/1817-3020-2020-20-4-278-287), свободный (15.05.2023)

37. Шабунин А.В. Гибридная SIRS-модель распространения инфекций / А.В. Шабунин // Известия высших учебных заведений. Прикладная нелинейная динамика. — 2022. — Т.30 №6. — С. 718-727.
38. Anser MK. Communicable Diseases (Including COVID-19)-Induced Global Depression: Caused by Inadequate Healthcare Expenditures, Population Density, and Mass Panic. / Anser MK, Yousaf Z, Khan MA, Sheikh AZ, Nassani AA, Abro MMQ, Zaman K. // Front Public Health. — 2020. — 18. Режим доступа: 10.3389/fpubh.2020.00398, свободный (15.05.2023)
39. Arenas A. Modeling the Spatiotemporal Epidemic Spreading of COVID-19 and the Impact of Mobility and Social Distancing Interventions / A. Arenas, W. Cota, J. Gómez-Gardeñes, S. Gómez, C. Granell, J.T. Matamalas, D. Soriano-Paños, B. Steinegger // Phys. Rev. — 2020. — vol 10, iss 4. — Режим доступа: <https://doi.org/10.1103/PhysRevX.10.041055>, свободный (15.05.2023)
40. Brauer F. Mathematical epidemiology: Past, present, and future / Brauer F. // Infect Dis Model. — 2017. — 2(2). — С.113-127. — Режим доступа: 10.1016/j.idm.2017.02.001, свободный (15.05.2023)
41. Duarte H.O. A probabilistic epidemiological model for infectious diseases: The case of COVID-19 at global-level / Duarte H.O., Siqueira P.G., Oliveira A.C.A., Moura M.D.C. // Risk Anal. — 2022. — 43(1). — С.183-201. — Режим доступа: doi: 10.1111/risa.13950, свободный (15.05.2023)
42. Gulbudak H. Sensitivity Analysis in an Immuno-Epidemiological Vector-Host Model. / Gulbudak H., Qu Z., Milner F. // Bull Math Biol. — 2022. — 84, 27. — Режим доступа: <https://doi.org/10.1007/s11538-021-00979-0>, свободный (15.05.2023)
43. Ismahene Y. Infectious Diseases, Trade, and Economic Growth: a Panel Analysis of Developed and Developing Countries. // J Knowl Econ. — 2022. — 13(3). Режим доступа: 10.1007/s13132-021-00811-z, свободный (15.05.2023)

44. Kermack W.O., McKendrick A.G. A contribution to the mathematical theory of epidemics. // *Proceedings of the Royal Society of London. Series A.* — 1927. — vol. 115. no. 772. — pp. 700–721.
45. Mata AS. Mathematical modeling applied to epidemics: an overview. / Mata AS, Dourado SMP // *São Paulo J. Math. Sci.* — 2021. — 15(2). Режим доступа: [10.1007/s40863-021-00268-7](https://doi.org/10.1007/s40863-021-00268-7), свободный (15.05.2023)
46. Mossong J. Social contacts and mixing patterns relevant to the spread of infectious diseases. / Mossong J, Hens N, Jit M, Beutels P, Auranen K, Mikolajczyk R, Massari M, Salmaso S, Tomba GS, Wallinga J, Heijne J, Sadkowska-Todys M, Rosinska M, Edmunds WJ // *PLoS Med.* — 2008. — 5(3). — Режим доступа: [10.1371/journal.pmed.0050074](https://doi.org/10.1371/journal.pmed.0050074), свободный (15.05.2023)
47. Qian W. A Probabilistic Infection Model for Efficient Trace-Prediction of Disease Outbreaks in Contact Networks / Qian W., Bhowmick S., O'Neill M., Ramisetty-Mikler S., Mikler A.R. // *Computational Science – ICCS 2020.* — 2020. — 12137:676–89. — Режим доступа: [10.1007/978-3-030-50371-0\\_50](https://doi.org/10.1007/978-3-030-50371-0_50), свободный (15.05.2023)
48. Sarkar K. An eco-epidemiological model with the impact of fear / Sarkar K., Khajanchi S. // *Chaos.* — 2022. — 32(8). — Режим доступа: [10.1063/5.0099584](https://doi.org/10.1063/5.0099584), свободный (15.05.2023)
49. Utamura M. An Epidemiological Model Considering Isolation to Predict COVID-19 Trends in Tokyo, Japan: Numerical Analysis / Utamura M, Koizumi M, Kirikami S // *JMIR Public Health Surveill* — 2020. — 6(4). — Режим доступа: [10.2196/23624](https://doi.org/10.2196/23624), свободный (15.05.2023)
50. Westerberg M. Simulation model of disease incidence driven by diagnostic activity / Westerberg M., Larsson R., Holmberg L., Stattin P., Garmo H. // *Stat Med.* — 2021. — 40(5). — С.1172-1188. — Режим доступа: [10.1002/sim.8833](https://doi.org/10.1002/sim.8833), свободный (15.05.2023)
51. Wong KLM. Social contact patterns during the COVID-19 pandemic in 21 European countries - evidence from a two-year study. / Wong KLM, Gimma

- A, Coletti P; CoMix Europe Working Group; Faes C, Beutels P, Hens N, Jaeger VK, Karch A, Johnson H, Edmunds W, Jarvis CI. // BMC Infect Dis. — 2023. —23(1). — Режим доступа: [10.1186/s12879-023-08214-y](https://doi.org/10.1186/s12879-023-08214-y), свободный (15.05.2023)
52. Yunfeng S. Effects of Pandemic Outbreak on Economies: Evidence From Business History Context / Yunfeng S., Haiwei L., Ren Z. // Front. Public Health. — 2021. — Vol. 9. — Режим доступа: <https://doi.org/10.3389/fpubh.2021.632043>, свободный (15.05.2023)
53. Стопкоронавирус.рф. Режим доступа: <https://xn--80aesfpebagmfblc0a.xn--p1ai/>, свободный (15.05.2023)
54. Johns Hopkins Coronavirus Resource Center. Режим доступа: <https://coronavirus.jhu.edu/data>, свободный (15.05.2023)

# Приложение

## Часть 1. Программный код для построения графиков процентного прироста

```
import numpy as np
import csv
import scipy.optimize as opt
from scipy.optimize import LinearConstraint as lincon
import matplotlib.pyplot as plt

with open('Название группы.csv', newline='') as f:
    data = list(csv.reader(f, delimiter=";"))

for X in [0,5]:
    Name = data[3*X + 1][0]
    C = np.array(data[3*X + 1][2:]).astype(np.float64)
    K0 = 10
    K = 150
    r = (C[K0+1:K] - C[K0:K-1])/C[K0:K-1]
    i20 = [i for i in range(len(r)) if r[i] >= 0.2][-1]
    i10 = [i for i in range(len(r)) if r[i] >= 0.1][-1]
    i5 = [i for i in range(len(r)) if r[i] >= 0.05][-1]
    i2 = [i for i in range(len(r)) if r[i] >= 0.02][-1]

    ind = np.array([i20, i10, i5, i2])
    inds = ind - i20
    print(Name, ": ", ind + K0)
    print("Average in [t20, t10]: ", np.sum(r[i20+1:i10+1]) / (i10 - i20))
    print("Average in [t10, t5]: ", np.sum(r[i10+1:i5+1]) / (i5 - i10))
    #plt.plot(range(K0+1,K), r, label=Name)
    #plt.plot(ind + K0 + 1, r[ind], marker='*', linestyle='dotted')
    plt.plot(range(1, K - K0 - i20), r[i20:], label=Name)
    plt.plot(inds + 1, r[ind], marker='*', linestyle='dotted')

plt.legend()
plt.show()
```

## Часть 2. Программный код для модели SIR

```
for X in [5]:
    Name = data[3*X + 1][0]
    print(Name)

    C = np.array(data[3*X + 1][2:]).astype(np.float64)
    R = np.array(data[3*X + 2][2:]).astype(np.float64) + np.array(data[3*X + 3][2:]).astype(np.float64)
```

```

I = C - R
# print(I)
K0 = 10
K = 150
r = (C[K0+1:K] - C[K0:K-1])/C[K0:K-1]
i20 = [i for i in range(len(r)) if r[i] >= 0.2][-1]
i10 = [i for i in range(len(r)) if r[i] >= 0.1][-1]
i5 = [i for i in range(len(r)) if r[i] >= 0.05][-1]
i2 = [i for i in range(len(r)) if r[i] >= 0.02][-1]

N = i10 - i20 # Количество дней для МНК
S = i20 # Первый день
ic = [0.05, 0.01]

def MSE(bg): # bg[0] = beta, bg[1] = gamma
    Ip = [I[i20] * np.exp((bg[0] - bg[1]) * k) for k in range(N)]
    Rp = R[i20] + [bg[1] / (bg[0] - bg[1]) * (Ip[k] - I[i20]) for k in
range(N)]
    E = [(I[i20+k] - Ip[k])**2 + (R[i20+k] - Rp[k])**2 for k in range(N)]

    return sum(E)

bounds = opt.Bounds([0, 0], [0.5, 0.2])
constr = lincon([[1, -1]], [0.001], [np.inf])
optim = opt.minimize(MSE, ic, bounds = bounds)
bg = optim.x
print(optim)

#print(bg)

ind = np.array([i20, i10, i5, i2])
print(ind)
NP = (i10 - i20) * 2
#NP = i5 - i20
NP = N+1
Ip = [I[i20] * np.exp((bg[0] - bg[1]) * k) for k in range(NP)]
plt.plot(range(NP), Ip)
plt.plot(range(NP), [I[i20+k] for k in range(NP)])
plt.plot(N, Ip[N], marker='*')
# plt.show()

Rp = [R[i20] + bg[1] / (bg[0] - bg[1]) * (Ip[k] - I[i20]) for k in range(NP)]
plt.plot(range(NP), Rp)
plt.plot(range(NP), [R[i20+k] for k in range(NP)])
plt.plot(N, Rp[N], marker='*')
plt.show()

```

Часть 3. Программный код для модели CIR (первая стратегия)

```

import numpy as np
import csv
import scipy.optimize as opt
import matplotlib.pyplot as plt

def FindTheta0(C, R, k = 0):
    T = next((i for i, e in enumerate(R) if e > C[k]), -1)
    t = next((j for j, e in enumerate(C) if e > R[T]), -1)
    return T, T - t

def FindTheta(C, R, T):
    t = next((j for j, e in enumerate(C) if e > R[T]), -1)
    return T - t

with open('Название группы.csv', newline='') as f:
    data = list(csv.reader(f, delimiter=";"))

Name = []
C = []
R = []
I = []
r = []
ind = []
for x in range(9):
    Name.append(data[3*x + 1][0])
    C.append(np.array(data[3*x + 1][2:]).astype(np.float64))
    R.append(np.array(data[3*x + 2][2:]).astype(np.float64) + np.array(data[3*x +
3][2:]).astype(np.float64))
    I.append(C[-1] - R[-1])
    K0 = 0
    K = 200
    r.append((C[-1][1:K] - C[-1][:K-1])/C[-1][:K-1])
    i20 = [i for i in range(len(r[-1])) if r[-1][i] >= 0.2][-1]
    i10 = [i for i in range(len(r[-1])) if r[-1][i] >= 0.1][-1]
    i5 = [i for i in range(len(r[-1])) if r[-1][i] >= 0.05][-1]
    i2 = [i for i in range(len(r[-1])) if r[-1][i] >= 0.02][-1]
    ind.append(np.array([i20, i10, i5, i2]))

X = 1
Y = 6
print("Base country: ", Name[Y], ", Work country: ", Name[X])

# Построим прогноз для страны X на интервал [t10, t5]
# Считаем, что данные на [t20, t10] у нас есть
# Согласно прогнозу время p5[X] перехода от t10 к t5 для страны X так же
относится
# ко времени перехода p10[X] от t20 к t10, как соответствующие сроки у страны Y
t20 = ind[X][0]
t10 = ind[X][1]

```



```

t5 = ind[X][2]
t20Y = ind[Y][0]
t10Y = ind[Y][1]
t5Y = ind[Y][2]
print("Checkpoints in ", Name[Y], ": ", ind[Y])
print("Checkpoints in ", Name[X], ": ", ind[X])

p5 = np.ceil((t10 - t20) * (t5Y - t10Y) / (t10Y - t20Y)).astype('int') # p10[X] *
p5[Y] / p10[Y]
#p5 = t5 - t10 # Истинный p5

# Прогноз по r на промежуток [t10, t5] --- линейное изменение на протяжении p5
дней от t10
rp = np.linspace(0.1, 0.05, p5 + 1)
print(t10, p5, t10 + p5)

# Прогноз развития C[X] на [t10, t5]
Cp = np.zeros_like(rp, dtype=np.float64)
Cp[0] = C[X][t10]
for i in range(1, len(Cp)):
    Cp[i] = Cp[i-1] * (1.0 + rp[i])

# Прогноз R строится на основе вычисления theta
theta10 = FindTheta(C[X], R[X], t10)
print("Theta 10 = ", theta10)

lambda10 = (R[X][t10] - C[X][t10-theta10-1]) / (C[X][t10 - theta10] - C[X][t10-
theta10-1])
print("Lambda 10 = ", lambda10)

# При постоянной theta
Rp = np.zeros_like(Cp)
for i in range(1, len(Rp)): # Нельзя залезать в прогнозируемый результат, что
может быть, если theta10 < p5
    Rp[i] = (1 - lambda10) * C[X][t10 + i - theta10 - 1] + lambda10 * C[X][t10 +
i - theta10]
Ip = Cp - Rp

# Сравним график прогноза и реальности
real = range(t10, t10 + p5 + 1)
plt.plot(real, R[X][real], label='R real data')
plt.plot(real, Rp, label='R with const theta', linestyle = '--')

# Попробуем так же линейно распространить theta из страны Y
theta10Y = FindTheta(C[Y], R[Y], t10Y)
theta5Y = FindTheta(C[Y], R[Y], t5Y)

print(t10Y, t5Y, theta10Y, theta5Y)

```

```

theta5p = np.round(theta5Y / theta10Y * theta10)
thetap = np.round(np.linspace(theta10, theta5p, p5+1)).astype('int')

print(thetap, FindTheta(C[X], R[X], t10+p5), FindTheta(C[X], R[X], ind[X][2]))

# При переменной theta
Rp1 = np.zeros_like(Cp)
for i in range(1, len(Rp)): # Нельзя залезать в прогнозируемый результат, что
    может быть, если theta10 < p5
    Rp1[i] = (1 - lambda10) * C[X][t10 + i - thetap[i] - 1] + lambda10 * C[X][t10
+ i - thetap[i]]
Ip1 = Cp - Rp1

# Сравним график прогноза и реальности
plt.plot(real, Rp1, label='R with variable theta', linestyle='-.')
plt.legend()
plt.show()

plt.plot(real, C[X][real], label='C real data')
plt.plot(real, Cp, label='C prognosis', linestyle = '--')
plt.legend()
plt.show()

```

#### Часть 4. Программный код для модели CIR (вторая стратегия)

```

import numpy as np
import csv
import scipy.optimize as opt
import matplotlib.pyplot as plt

def FindTheta0(C, R, k = 0):
    T = next((i for i, e in enumerate(R) if e > C[k]), -1)
    t = next((j for j, e in enumerate(C) if e > R[T]), -1)
    # print(T, t)
    return T, T - t

def FindTheta(C, R, T):
    t = next((j for j, e in enumerate(C) if e > R[T]), -1)
    return T - t

with open(Название группы.csv', newline='') as f:
    data = list(csv.reader(f, delimiter=";"))

Name = []
C = []
R = []
I = []
r = []

```

```

ind = []
for x in range(9):
    Name.append(data[3*x + 1][0])
    C.append(np.array(data[3*x + 1][2:]).astype(np.float64))
    R.append(np.array(data[3*x + 2][2:]).astype(np.float64) + np.array(data[3*x +
3][2:]).astype(np.float64))
    I.append(C[-1] - R[-1])
    K0 = 0
    K = 200
    r.append((C[-1][1:K] - C[-1][:K-1])/C[-1][:K-1])

    # Времена последних точек превышающих 20%, 10%, 5%, 2%
    i20 = [i for i in range(len(r[-1])) if r[-1][i] >= 0.2][-1]
    i10 = [i for i in range(len(r[-1])) if r[-1][i] >= 0.1][-1]
    i5 = [i for i in range(len(r[-1])) if r[-1][i] >= 0.05][-1]
    i2 = [i for i in range(len(r[-1])) if r[-1][i] >= 0.02][-1]
    # Времена первого пересечения 20%, 10%, 5%, 2%
    j20 = [i for i in range(len(r[-1])) if r[-1][i] < 0.2][0] - 1
    j10 = [i for i in range(len(r[-1])) if r[-1][i] < 0.1][0] - 1
    j5 = [i for i in range(len(r[-1])) if r[-1][i] < 0.05][0] - 1
    j2 = [i for i in range(len(r[-1])) if r[-1][i] < 0.02][0] - 1
    ind.append(np.array([[i20, j20], [i10, j10], [i5, j5], [i2, j2]]))
    # print(Name, ": ", ind + K0)
    #plt.plot(range(K0+1,K), r[x], label=Name[x])
    #plt.plot(ind[0] + K0 + 1, r[x][ind[0]], marker='*', linestyle='dotted')

X = 2
Y = 3
print("Base country: ", Name[Y], ", Work country: ", Name[X])

# Построим прогноз для страны X на интервал [t10, t5]
# Считаем, что данные на [t20, t10] у нас есть
# Согласно прогнозу время p5[X] перехода от t10 к t5 для страны X так же
относится
# ко времени перехода p10[X] от t20 к t10, как соответствующие сроки у страны Y
t20 = ind[X][0][0]
t10 = ind[X][1][0]
t5 = ind[X][2][0]
t20Y = (ind[Y][0][0] + ind[Y][0][1]) / 2
t10Y = (ind[Y][1][0] + ind[Y][1][1]) / 2
t5Y = (ind[Y][2][0] + ind[Y][2][1]) / 2
print("Checkpoints in ", Name[Y], ": ", ind[Y])
print("Checkpoints in ", Name[X], ": ", ind[X])

p5 = np.ceil((t10 - t20) * (t5Y - t10Y) / (t10Y - t20Y)).astype('int') # p10[X] *
p5[Y] / p10[Y]
#p5 = t5 - t10 # Истинный p5

```

```

# Прогноз по r на промежуток [t10, t5] --- линейное изменение на протяжении p5
дней от t10
rp = np.linspace(0.1, 0.05, p5 + 1)
#print(ind[X])
#print(ind[Y])
print(t10, p5, t10 + p5)

# Прогноз развития C[X] на [t10, t5]
Cp = np.zeros_like(rp, dtype=np.float64)
Cp[0] = C[X][t10]
for i in range(1, len(Cp)):
    Cp[i] = Cp[i-1] * (1.0 + rp[i])

# Прогноз R строится на основе вычисления theta
theta10 = FindTheta(C[X], R[X], t10)
#print(C[X][:t10])
#print(R[X][:t10])
print("Theta 10 = ", theta10)

lambda10 = (R[X][t10] - C[X][t10-theta10-1]) / (C[X][t10 - theta10] - C[X][t10-
theta10-1])
print("Lambda 10 = ", lambda10)

# При постоянной theta
Rp = np.zeros_like(Cp)
for i in range(1, len(Rp)): # Нельзя залезать в прогнозируемый результат, что
может быть, если theta10 < p5
    Rp[i] = (1 - lambda10) * C[X][t10 + i - theta10 - 1] + lambda10 * C[X][t10 +
i - theta10]
Ip = Cp - Rp

# Сравним график прогноза и реальности
real = range(t10, t10 + p5 + 1)
plt.plot(real, R[X][real], label='R real data')
plt.plot(real, Rp, label='R with const theta', linestyle = '--')

# Попробуем так же линейно распространить theta из страны Y
# theta10Y = FindTheta(C[Y], R[Y], t10Y.astype('int'))
# theta5Y = FindTheta(C[Y], R[Y], t5Y.astype('int'))
theta10Y = FindTheta(C[Y], R[Y], ind[Y][1][0])
theta5Y = FindTheta(C[Y], R[Y], ind[Y][2][0])

print(t10Y, t5Y, theta10Y, theta5Y)

theta5p = np.round(theta5Y / theta10Y * theta10)
thetap = np.round(np.linspace(theta10, theta5p, p5+1)).astype('int')

print(thetap, FindTheta(C[X], R[X], t10+p5), FindTheta(C[X], R[X], ind[X][2][0]))

```

```

# При переменной theta
Rp1 = np.zeros_like(Cp)
for i in range(1, len(Rp)): # Нельзя залезать в прогнозируемый результат, что
    может быть, если theta10 < p5
    Rp1[i] = (1 - lambda10) * C[X][t10 + i - thetap[i] - 1] + lambda10 * C[X][t10
+ i - thetap[i]]
Ip1 = Cp - Rp1

# Сравним график прогноза и реальности
plt.plot(real, Rp1, label='R with variable theta', linestyle='-.')
plt.legend()
plt.show()

plt.plot(real, C[X][real], label='C real data')
plt.plot(real, Cp, label='C prognosis', linestyle = '--')
plt.legend()
plt.show()

```

## Часть 5. Программный код для модели CIR (третья стратегия)

```

import numpy as np
import csv
import scipy.optimize as opt
import matplotlib.pyplot as plt

def FindTheta0(C, R, k = 0):
    T = next((i for i, e in enumerate(R) if e > C[k]), -1)
    t = next((j for j, e in enumerate(C) if e > R[T]), -1)
    # print(T, t)
    return T, T - t

def FindTheta(C, R, T):
    t = next((j for j, e in enumerate(C) if e > R[T]), -1)
    return T - t

with open(Название группы.csv', newline='') as f:
    data = list(csv.reader(f, delimiter=";"))

Name = []
C = []
R = []
I = []
r = []
ind = []
for x in range(9):
    Name.append(data[3*x + 1][0])
    C.append(np.array(data[3*x + 1][2:]).astype(np.float64))

```

```

    R.append(np.array(data[3*x + 2][2:]).astype(np.float64) + np.array(data[3*x +
3][2:]).astype(np.float64))
    I.append(C[-1] - R[-1])
    K0 = 0
    K = 200
    r.append((C[-1][1:K] - C[-1][:K-1])/C[-1][:K-1])
    i20 = [i for i in range(len(r[-1])) if r[-1][i] >= 0.2][-1]
    i10 = [i for i in range(len(r[-1])) if r[-1][i] >= 0.1][-1]
    i5 = [i for i in range(len(r[-1])) if r[-1][i] >= 0.05][-1]
    i2 = [i for i in range(len(r[-1])) if r[-1][i] >= 0.02][-1]
    ind.append(np.array([i20, i10, i5, i2]))
    # print(Name, ": ", ind + K0)
    # plt.plot(range(K0+1,K), r, label=Name)
    # plt.plot(ind + K0 + 1, r[ind], marker='*', linestyle='dotted')

X = 1
Y = 5
Z = 2
print("Base countries: ", Name[Y], " and ", Name[Z], ", Work country: ", Name[X])

# Построим прогноз для страны X на интервал [t10, t5]
# Считаем, что данные на [t20, t10] у нас есть
# Согласно прогнозу время p5[X] перехода от t10 к t5 для страны X так же
относится
# ко времени перехода p10[X] от t20 к t10, как соответствующие усреднённые сроки
у стран Y и Z
t20 = ind[X][0]
t10 = ind[X][1]
t5 = ind[X][2]
t20Y = (ind[Y][0] + ind[Z][0]) / 2
t10Y = (ind[Y][1] + ind[Z][1]) / 2
t5Y = (ind[Y][2] + ind[Z][2]) / 2
print("Checkpoints in ", Name[Y], ": ", ind[Y])
print("Checkpoints in ", Name[Z], ": ", ind[Z])
print("Checkpoints in ", Name[X], ": ", ind[X])

p5 = np.ceil((t10 - t20) * (t5Y - t10Y) / (t10Y - t20Y)).astype('int') # p10[X] *
p5[Y] / p10[Y]
#p5 = t5 - t10 # Истинный p5

# Прогноз по r на промежуток [t10, t5] --- линейное изменение на протяжении p5
дней от t10
rp = np.linspace(0.1, 0.05, p5 + 1)
#print(ind[X])
#print(ind[Y])
print(t10, p5, t10 + p5)

# Прогноз развития C[X] на [t10, t5]
Cp = np.zeros_like(rp, dtype=np.float64)
Cp[0] = C[X][t10]

```

```

for i in range(1, len(Cp)):
    Cp[i] = Cp[i-1] * (1.0 + rp[i])

#print(Cp)
#print(C[X][t10:t10+p5])

# Прогноз R строится на основе вычисления theta
theta10 = FindTheta(C[X], R[X], t10)
#print(C[X][:t10])
#print(R[X][:t10])
print("Theta 10 = ", theta10)

lambda10 = (R[X][t10] - C[X][t10-theta10-1]) / (C[X][t10 - theta10] - C[X][t10-
theta10-1])
print("Lambda 10 = ", lambda10)

# При постоянной theta
Rp = np.zeros_like(Cp)
for i in range(1, len(Rp)): # Нельзя залезать в прогнозируемый результат, что
    может быть, если theta10 < p5
    Rp[i] = (1 - lambda10) * C[X][t10 + i - theta10 - 1] + lambda10 * C[X][t10 +
    i - theta10]
Ip = Cp - Rp

# Сравним график прогноза и реальности
real = range(t10, t10 + p5 + 1)
plt.plot(real, R[X][real], label='R real data')
plt.plot(real, Rp, label='R with const theta', linestyle = '--')

# Попробуем так же линейно распространить theta из страны Y
theta10Y = FindTheta(C[Y], R[Y], ind[Y][1])
theta10Z = FindTheta(C[Z], R[Z], ind[Z][1])

theta5Y = FindTheta(C[Y], R[Y], ind[Y][2])
theta5Z = FindTheta(C[Z], R[Z], ind[Z][2])

print(t10Y, t5Y, theta10Y, theta5Y, theta10Z, theta5Z)

theta5p = np.round((theta5Y + theta5Z) / (theta10Y + theta10Z) * theta10)
thetap = np.round(np.linspace(theta10, theta5p, p5+1)).astype('int')

print(thetap, FindTheta(C[X], R[X], t10+p5), FindTheta(C[X], R[X], ind[X][2]))

# При переменной theta
Rp1 = np.zeros_like(Cp)
for i in range(1, len(Rp)): # Нельзя залезать в прогнозируемый результат, что
    может быть, если theta10 < p5
    Rp1[i] = (1 - lambda10) * C[X][t10 + i - thetap[i] - 1] + lambda10 * C[X][t10
    + i - thetap[i]]
Ip1 = Cp - Rp1

```

```
# Сравним график прогноза и реальности
plt.plot(real, Rp1, label='R with variable theta', linestyle='-.')
plt.legend()
plt.show()

plt.plot(real, C[X][real], label='C real data')
plt.plot(real, Cp, label='C prognosis', linestyle = '--')
plt.legend()
plt.show()
```