Федеральное государственное автономное образовательное учреждение
высшего образования «Московский физико-технический институт
(национальный исследовательский университет)»

Физтех-школа биологической и медицинской физики
Кафедра системной и синтетической биологии

Направление подготовки: 03.03.01 Прикладные математика и физика
Направленность (профиль) подготовки: Системная и синтетическая биология

Исследование роли транспортных потоков в развитии эпидемии на основе компьютерной агентной модели

(бакалаврская работа)

Студент:
Клочков Константин Александрович

Научный руководитель:
Манолов Александр Иванович,
к.б.н

Москва 2025

Аннотация

В данной работе мы исследовали роль транспортных потоков в распространении эпидемии на основе агентной модели Covasim.

Была реализована возможность моделирования пассажиропотоков (обмена агентами между населенными пунктами), добавлена опция введения ограничений на передвижение, а также изучено влияние уровня заразности инфекционного агента и конфигурации транспортной сети на развитие эпидемического процесса.

В ходе исследования были изучены характер зависимости смещения пика эпидемиологической кривой от интенсивности транспортного потока, чувствительность различных эпидемических метрик к изменению входящих и исходящих потоков, а также возможность идентификации населенного пункта, в котором началась эпидемия.

Кроме того, продемонстрирована значимость своевременного ограничения транспортных потоков для снижения нагрузки на систему здравоохранения — на примере распространения вируса SARS-CoV-2.

Список сокращений

`Covasim` COVID-19 Agent-based Simulator (агентный симулятор COVID-19). 4, 12, 13, 16, 17, 24

ABM Agent-based model (агентная модель). 8, 10–12

COVID-19 Coronavirus Disease 2019 (коронавирусная инфекция 2019). 12

GIL Global Interpreter Lock (глобальная блокировка интерпретатора). 17

SARS-CoV-2 Severe acute respiratory syndrome-related coronavirus 2 (коронавирус 2, связанный с тяжелым острым респираторным синдромом). 4, 24, 37, 39, 44

SEIR susceptible → exposed → infectious → recovered (восприимчивый → подверженный → инфицированный → восстановившийся). 10, 11

SIR susceptible → infectious → recovered (восприимчивый → инфицированный → восстановившийся). 5, 6, 10

SIS susceptible → infectious → susceptible (восприимчивый → инфицированный → восприимчивый). 5–7, 9

Содержание

1 Введение

С увеличением вычислительных мощностей математические модели стали важнейшими инструментами анализа распространения инфекционных заболеваний и поиска способов борьбы с ними [1—3]. Агентные модели при наличии достаточных объемов демографических данных способны описать развитие эпидемии на уровне отдельных агентов (людей), что позволяет рассматривать сложные социальные структуры, динамические изменения в них, иммунологическую историю индивидов и многие другие особенности, недоступные детерминистическим компартментным моделям, основанным на дифференциальных уравнениях [4].

Целью данной работы является исследование влияния транспортных потоков и ограничений, накладываемых на них, на ход протекания эпидемии в модельных системах на основе высокопроизводительной агентной модели.

Исследования в рамках данной работы проводились на модели, разработанной нами ранее на основе COVID-19 Agent-based Simulator (агентный симулятор COVID-19) (`Covasim`). `Covasim` — агентная модель с открытым исходным кодом, содержащая реалистичную систему передачи инфекции в разных социальных слоях и протекания болезни, зависящую от возраста агента [5].

Мы внедрили в `Covasim` модель транспортных потоков: вычисления, соответствующие различным населенным пунктам, ведутся параллельно. На каждом шаге симуляции в каждом городе выбираются случайные агенты-туристы, количество которых согласуется с величинами транспортных потоков. Агент-турист удаляется из сети контактов своего населенного пункта и внедряется в город прибытия, вступая в фиксированное число случайных контактов. По завершении периода поездки производится обратная операция. Перемещениям не подвержены агенты в тяжелом или критическом состоянии: если они оказываются в списке туристов, то не покидают свой город. Если же они оказались в таком состоянии по завершении периода поездки, они возвращаются в родной город только при выздоровлении.

Данная работа демонстрирует важность транспортных потоков и своевременного их ограничения в развитии эпидемии. В дальнейшем планируется исследовать распространение Severe acute respiratory syndrome-related coronavirus 2 (коронавирус 2, связанный с тяжелым острым респираторным синдромом) (SARS-CoV-2) и других менее трансмиссивных заболеваний в случае туристических поездок по Российской Федерации.

2 Обзор литературы

2.1 Виды эпидемиологических моделей

Математические модели стали важными инструментами анализа распространения инфекционных заболеваний и способов борьбы с ними. В процессе создания модели уточняются предположения и параметры. Само же моделирование позволяет оценить критические значения параметров модели, приводящие к развитию эпидемии: базовый индекс репродукции инфекции R_0, эффективный R_{eff}, среднее число контактов между людьми. Математические модели и компьютерное моделирование являются полезными экспериментальными инструментами для проверки теорий, оценки количественных гипотез, ответов на конкретные вопросы, определения чувствительности к изменениям значений параметров и оценки ключевых параметров по данным. Понимание особенностей передачи инфекционных заболеваний в сообществах, регионах и странах может привести к улучшению подходов к снижению распространения этих заболеваний. Математические модели используются при сравнении, планировании, реализации, оценке и оптимизации различных методов программ по выявлению, профилактике, лечению и контролю. Моделирование в эпидемиологии может способствовать разработке и анализу эпидемиологических обследований, предлагать важнейшие данные которые необходимо собрать, выявить тенденции, сделать общие прогнозы и оценить неопределенность прогнозов [1—3].

2.1.1 Детерминистические компартментные модели

Подавляющее большинство эпидемиологических моделей основано на компартментализации людей или иных моделируемых сущностей в зависимости от их состояния [6—9]. Базовые модели описывают лишь два компартмента: восприимчивых и инфицированных. В таком случае происходит пренебрежение многими деталями развития эпидемии, тем не менее такие модели все еще активно применяются. Такая susceptible → infectious → susceptible (восприимчивый → инфицированный → восприимчивый) (SIS) модель описывается системой дифференциальных уравнений

$$\begin{cases} \dfrac{dS}{dt} = gI - \lambda S, \\ \dfrac{dI}{dt} = \lambda S - gI, \end{cases} \tag{1}$$

другой из основных используемых является susceptible → infectious → recovered (восприимчивый → инфицированный → восстановившийся) (SIR) модель, в которой вводится третий компартмент восстановившихся

$$\begin{cases} \dfrac{dS}{dt} = bN - \lambda S - dS, \\ \dfrac{dI}{dt} = \lambda S - gI - dI, \\ \dfrac{dR}{dt} = gI - dR. \end{cases} \tag{2}$$

Переменные S, I, R обозначают количества восприимчивых, инфицированных, восстановившихся людей соответственно; N — размер популяции; b, d, g — коэффициенты рождаемости, смертности и восстановления (доля рождающихся, умирающих,

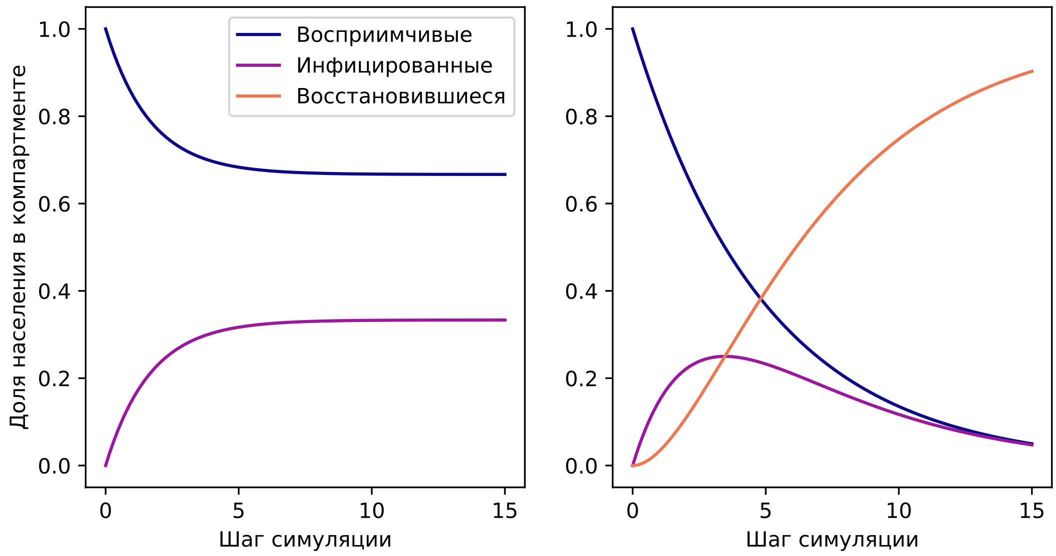

Рис. 1: Численные решения задачи Коши детерминистических компартментных моделей с начальными условиями $S(0) = 1$, $I(0) = 0$, $R(0) = 0$ и параметрами $b = d = 0$, $g = 0.4$, $\lambda = 0.2$, $N = 1$: слева для SIS модели, справа для SIR.

восстанавливающихся людей в единицу времени); λ — сила инфекции (доля восприимчивых людей, инфицируемых в единицу времени).

SIR модели применяются для описания инфекционных заболеваний, формирующих очень стойкий или пожизненный иммунитет, например, кори, коклюша [6; 7; 9—11]. SIS модели используют преимущественно при описании распространения заболеваний, передающихся половым путем, реинфицирование которыми вполне возможно, например, хламидиоза, гонореи [6; 12; 13].

Также активно применяются модифицированные модели с уточненной компартментализацией для более сложного протекания болезни [6; 14; 15] и с более проработанной структурой популяции [12; 16; 17].

2.1.2 Детерминистические модели среднего поля

Компартментные модели строятся в предположении, что каждый человек взаимодействует с каждым. Однако в большой популяции у каждого индивида есть свой ограниченный круг общения. Именно им определяется путь распространения инфекции, причем число контактов очень сильно меняется от человека к человеку [18]. Для более реалистичного описания таких социальных взаимодействий распространение эпидемии моделируют на графе контактов (или сети контактов). В нем каждому человеку соответствует узел, а каждой социальной связи — ребро графа [19].

В связи с тем, что прямой анализ стохастического поведения эпидемии на графе является сложным, его описывают детерминистическим приближением среднего поля. Анализируют некоторые средние характеристики моделей [19]. Другими словами, в таких моделях для упрощения математического описания вводятся различные приближения: попарные модели рассматривают состояния пар узлов, соединенных ребром [20; 21]; message passing (обмен сообщениями) (MP) модели рассматривают каждый узел и все возможные пути передачи инфекции в этот узел независимо [22]; модели средней степени вершины графа учитывают звездчатые структуры — узлы с их соседями с учетом их состояния [23]; edge-based compartmental models (компартментные модели на ребрах графа) (EBCM) оценивают вероятность случайного узла остаться восприимчивым [24]. Все эти модели можно построить на одном и том же графе контактов, исходя из разных предположений. При этом часть из них оказывается эквивалентна друг другу [19].

Эти модели лучше описывают социальное устройство популяции, однако они рассматривают марковский процесс. Они исходят из предположения, что состояние графа на следующем шаге определяется лишь его состоянием на текущем шаге и никак не зависит от предыдущих шагов. Эти предположения часто оказываются неверны, так как инфицирование или выздоровление происходит не случайно в любой момент времени, а после конкретного инкубационного периода или периода болезни соответственно [6; 21; 25].

2.1.3 Моделирование при помощи марковских цепей

Результатами применения марковских цепей также являются временные зависимости некоторых глобальных величин [26; 27]. Однако, в отличие от детерминистических моделей, стохастический подход допускает временные флуктуации как на глобальном уровне, так и на локальном — флуктуации на уровне агентов в зависимости от их контактов [28].

Основой таких моделей, как и с моделями среднего поля, часто является граф взаимодействий, описываемый матрицей контактов [29]. В отличие от компартментных детерминистических моделей, матрица контактов может содержать в себе крайне сложные социальные структуры. Это позволяет наблюдать более широкий спектр эпидемиологических явлений [30]. Распространение инфекции задается матрицей перехода $\hat{T}$, матричные элементы $T_{\mu\nu}$ которой равны вероятностям перехода из состояния графа ν (конфигурации всех узлов) в μ [31]. Эта разница между детерминистическими компартментными моделями и моделями на марковских цепях приводит к различию между временными зависимостями эпидемиологических статистик. К примеру, оказывается, что, хоть SIS модель демонстрирует схожее с аналогичной стохастической поведение при различных соотношениях между силой инфекции и коэффициентом восстановления, на ранних этапах развития эпидемии, когда среднее число инфицированных сильно меньше размера популяции, ключевым фактором являются именно флуктуации числа инфицированных. И именно такому режиму соответствуют санитарные меры по сдерживанию эпидемий на ранних стадиях [28].

2.1.4 Агентные модели

Агентные модели (agent-based models) (ABMs) применяются при моделировании распространения инфекционных заболеваний уже более 40 лет [32; 33], однако из-за недостатка данных и вычислительных мощностей раскрывать свой потенциал они начали совсем недавно [34].

Устройство Agent-based model (агентная модель) (ABM) позволяет напрямую описывать сложные социальные и физические системы. Хоть для воспроизведения социальной структуры конкретной страны или города может потребоваться очень много данных, ABM может отразить неоднородность социальных взаимодействий и динамические изменения в структуре сети [4].

Такой подход очень полезен для реализации учета иммунологической истории индивидов, вариации климата между регионами, социальных структур, что позволяет проводить более тонкую калибровку. Также ABM позволяют рассматривать распространение инфекции в маленьких группах людей [4].

Типичное построение ABM требует, во-первых, воссоздания искусственной популяции, соответствующей демографическим данным в регионе, и, во-вторых, создания сети публичных объектов, таких как школ, рабочих мест, где агенты могут взаимодействовать друг с другом. Главной сложностью при построении достаточно подробной ABM является тот факт, что требуемая информация часто скрыта при агрегировании федеральными службами государственной статистики. В связи с этим возникает необходимость в разработке методов, позволяющих оценить исходные данные из агрегатов, или моделей, приближающих искомые данные [4].

2.2 Подходы к изучению транспортных потоков в эпидемиологии

Транспорт играет важнейшую роль в жизни людей и в то же время активно способствует распространению эпидемий. В связи с этим введение ограничений на транспортные потоки является важной противоэпидемической мерой [35], а понимание влияния этих мер позволит человечеству эффективнее бороться с распространением инфекционных заболеваний.

В имеющихся исследованиях по этой теме транспорт описывается преимущественно детерминистическими компартментными моделями, марковскими моделями на графах и агентными моделями. Статистические (регрессионные и авторегрессионные модели на пространственных данных) же для этих целей практически не применяются в связи с необходимостью иметь огромное множество данных о мобильности людей, в отсутствие которых качество результатов работы таких моделей крайне ограничено. Клеточными автоматами симулировать перемещение индивидов также затруднительно, так что их для этих целей тоже применяют редко [35].

2.2.1 Детерминистические компартментные модели с транспортными потоками

Одним из преимуществ детерминистических компартментных моделей является тот факт, что для их калибровки требуются малые объемы данных. По этой причине было предпринято множество попыток адаптирования существующих моделей под описание системы нескольких городов. Хоть эти модели и предполагают, что города обладают одинаковыми демографическими параметрами, что перемещения происходят мгновенно, некоторые игнорируют возможность инфицирования при перемещении, из них можно получить важные выводы [35].

Дж. Хайман и др. (2003) [36] для описания распространения гриппа в США в присутствие авиатранспорта каждый исследуемый город описали моделью типа susceptible $\rightarrow$ infectious $\rightarrow$ recovered $\rightarrow$ partially immune (восприимчивый $\rightarrow$ инфицированный $\rightarrow$ восстановившийся $\rightarrow$ частично иммунный) (SIRP) — добавили компартмент, содержащий частично имунных агентов, в который попадает часть восстановившихся. Перемещения людей ввели, определив постоянную матрицу миграции M, содержащую долю перемещающихся людей в единицу времени, внеся в производные $\frac{d}{dt}(S_k, I_k, R_k, P_k)$ слагаемое $\sum_{j=1}^{n}\left[m_{jk}\frac{(S_j,I_j,R_j,P_j)}{N_j} - m_{kj}\frac{(S_k,I_k,R_k,P_k)}{N_k}\right]$. В результате авторы выявили, что наиболее эффективно замедляет эпидемию сокращение продолжительности инфекционной стадии, сокращение числа контактов r, уменьшение вероятности передачи β. Этот подход полностью игнорирует возможность инфицирования при перемещении, что некорректно, особенно при рассмотрении быстро распространяющихся заболеваний как грипп или атипичная пневмония, ведь во время поездок люди в течение долгого времени находятся в очень близком контакте друг с другом [35].

Ванг и Жао (2004) [37] на совокупности двух SIS моделей на двух примерах проиллюстрировали важность перемещения агентов в развитии эпидемии. Первый показал: если в отсутствие транспортных потоков базовый индекс репродукции одного города $R_{01} \geqslant 5/3$, а второго $R_{02} = 1/3$, то транспортные потоки распространяют эпидемию на оба города. Если же $1 < R_{01} < 5/3$ и $R_{02} = 1/3$, транспортные потоки ослабляют распространение эпидемии. Второй пример проиллюстрировал: если $R_{01} = 0.75$ и $R_{02} = 0.25$, то наличие транспортных потоков обеспечивает распространение эпидемии в такой системе.

Такеучи и др. (2006) [38] предложили аналогичный подход для системы из двух SIS моделей и ввели слагаемое, учитывающее инфицирование агентов в процессе перемещения между населенными пунктами. Авторы считали частоту заболевания туристов равной $\gamma(\alpha S_j)(\alpha I_j)/(\alpha S_j + \alpha I_j)$, где γ — трансмиссивность инфекции в процессе перемещения, α — доля перемещающихся агентов в единицу времени. Анализ этой модели привел авторов к выводу, что инфицирования, связанные с транспортными потоками, приводят к повышению как абсолютного, так и относительного числа больных. Это подтверждает необходимость вводить транспортные ограничения сразу при обнаружении вспышки инфекции [38].

Лиу и Такеучи (2006) [39] рассматривали возможность тестирования инфицированных агентов и помещения их в карантинный компартмент в системе двух susceptible → infectious → quarantined → susceptible (восприимчивый → инфицированный → изолированный → восприимчивый) (SIQS) моделей, учитывая туризм аналогично авторам упомянутым выше. Они пришли к выводу, что входной скрининг с последующим карантином инфицированных крайне полезен. Он может привести к завершению эпидемии, даже если инфекция эндемична в каждом из рассматриваемых городов [39].

Ван и Цуй (2007) [40] исследовали возможности контроля распространения инфекционных заболеваний, изучая модель susceptible → exposed → infectious → susceptible (восприимчивый → подверженный → инфицированный → восприимчивый) (SEIS). Оказалось, что даже если запретить инфекционным агентам перемещаться, инфицированные в инкубационной фазе все равно приносят болезнь в города. Помимо этого они подтвердили результаты, полученные ранее Вангом и Жао [37] и Такеучи и др. [38], о том, что наличие транспортных потоков может качественно изменить ход эпидемии.

2.2.2 Агентные модели с транспортными потоками

В ABM каждому агенту могут быть присвоены свои атрибуты. При этом каждый принимает индивидуальные решения, что позволяет отразить детали межличностных взаимодействий. В связи с развитием систем общественного транспорта им пользуется все больше людей, что значительно усложняет сеть контактов. Поэтому моделирование транспортных систем очень важно в эпидемиологических ABM.

Симоэс (2006) [41] для описания эпидемии паротита в Португалии 1996 года разработала ABM, включающую в себя состояния susceptible → exposed → infectious → recovered (восприимчивый → подверженный → инфицированный → восстановившийся) (SEIR) и транспортную модель. Она базировалась на делении Португалии на регионы, между которыми вводились 4 типа перемещений, связанные с различными типами социальных активностей: движения в пределах квартала; региона; между соседними регионами и между удаленными регионами. Перемещение агента на каждом шаге по времени определялось взвешенной суммой всевозможных перемещений между регионами с весами, равными вероятностям этих перемещений. Такой подход учитывает пространственное распределение популяции и перемещения агентов, что позволило получить результаты моделирования эпидемии паротита близкие к реальным.

Раковски и др. (2010) [4] при построении ABM распространения гриппа в Польше использовали простые правила перемещений для имитации поездок людей и позволили в модели инфекции передаваться непосредственно в транспорте. Авторы выбрали 38 крупнейших городов Польши и на их основе построили граф транспортной сети. В нем веса ребер были пропорциональны расстоянию между городами и обратно пропорциональны размерам их популяций. С помощью алгоритма Дейкстры находился кратчайший маршрут между любыми двумя городами, а список промежуточных городов на этом пути заносился в матрицу. На каждом шаге симуляции выбирались случайные агенты, отправляемые в другой город; каждому агенту назначался город прибытия (вероятность выбора города пропорциональна его населению); определялись города, в которых производились пересадки; формировались группы путешественников размером до 10 человек. Эта модель, хоть и не учитывает точные маршруты, позволяет оценить их исключительно из карты распределения плотности населения, что сильно снижает требования к объемам необходимых для моделирования данных.

Фриас-Мартинез и др. (2011) [42] первыми использовали реальные данные сотовых операторов в ABM. Ранее для решения задачи моделирования транспортных потоков применялись результаты опросов, которые не дают требуемого временного и пространственного разрешения. Авторы статьи оценили покрытия вышек операторов сотовой связи и на основе имеющихся данных создали модель перемещения пользователей, которая оценивала положение людей в каждый момент времени, и модель социальных связей, которая определяла наиболее близкие взаимоотношения. Помимо этого каждый агент подчинялся модели заболевания SEIR. Так авторы выяснили, что ограничение на передвижение людей в 2009 году в Мексике позволило снизить пиковое число зараженных H1N1 гриппом на 10% и отложить момент наступления этого пика на 2 дня.

Крукс и др. (2014) [43] моделировали вспышки холеры в лагере для беженцев Дадааб в Кении с помощью ABM. Распорядок дня беженцев учитывался в поведении агентов. Конкретные их действия могли выполняться в специализированных для этого местах: школе, религиозном центре, магазине и так далее. Для этого карта Дадааба была оцифрована и поделена на участки по видам возможной там деятельности. Также каждый агент отличался личными характеристиками (пол, возраст), социальными связями (размер семьи, круг общения), наличием или отсутствием симптомов при инфицировании, целями и приоритетами. На каждом шаге днем агент принимает решение, остаться там, где он находится, или переместиться, чтобы закрыть свои потребности (в воде, еде, образовании). Ночью возвращается домой. Само распространение холеры определяется SEIR моделью, где инфицированные могут быть как симптоматическими, так и асимптоматическими. Само заражение происходит из источников воды, ставших заразными после того, как туда попали фекалии инфицированных. Благодаря разработанной модели авторы смогли рассмотреть два сценария развития холеры: радиально от источника загрязненной воды и через дождевые стоки, причем их результаты качественно совпали с реальными данными, например, для лагеря Дагахлей. Потенциально такая модель может стать частью системы раннего предупреждения вспышек холеры.

2.3 Агентная модель Covasim

В этой работе в дальнейшем речь будет идти о модели `Covasim` и ее модернизации. `Covasim` — ABM с открытым кодом, учитывающая информацию о возрастной структуре и размере популяции, реалистичную систему передачи инфекции в разных

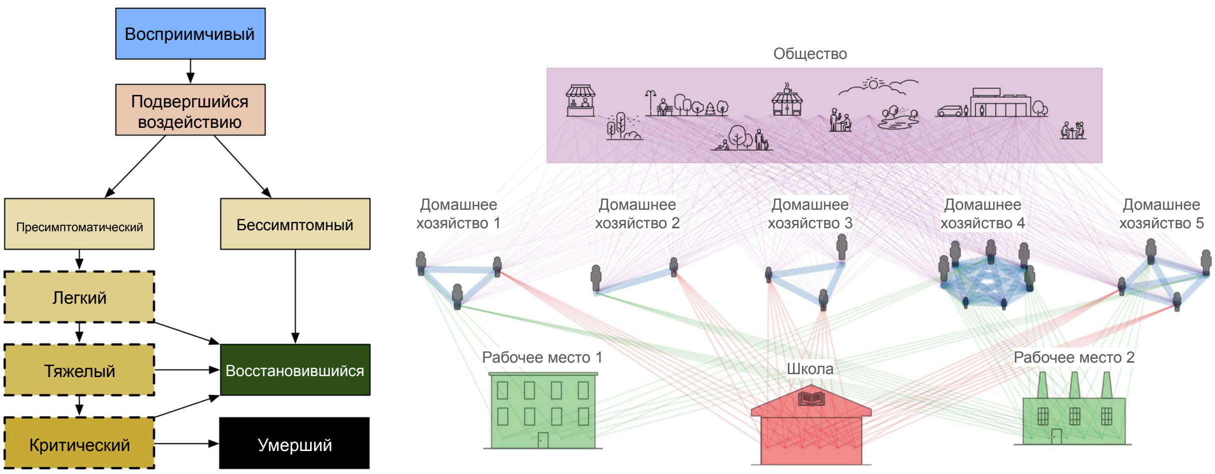

Рис. 2: Структура модели развития инфекции (слева) и иллюстрация сети контактов в Covasim. Источник: [5].

социальных слоях (домашние хозяйства, школы, рабочие места, случайные контакты и др.), течение болезни, зависящее от возраста, иммунитет, определяемый уровнем антител. Также в данной модели реализовано множество различных противоэпидемических мер: социальное дистанцирование, ношение масок, вакцинирование, тестирование и другие.

Каждый запуск Covasim состоит из нескольких этапов. Сначала создается объект симуляции и загружаются все необходимые параметры. Затем в соответствии с данными о распределении возрастов создаются агенты и объединяются в сеть контактов согласно выбранному способу генерации синтетической популяции. Затем в цикле на каждом шаге по времени производится масштабирование популяции (если необходимо для повышения производительности); обновляются состояния агентов, в том числе связанные с развитием инфекции; инфицируются случайные агенты; применяются противоэпидемические меры; вычисляются вероятности дальнейшего инфицирования по сети контактов; вычисляются результирующие метрики.

В Covasim реализована довольно сложная модель протекания болезни, в которой каждый агент может пребывать в одном из 9 состояний: восприимчивый, подвергшийся воздействию (инфицированный, но пока не инфекционный), инфекционный (разделяется по тяжести симптомов на пресимптоматический, бессимптомный, легкий, тяжелый, критический), восстановившийся и умерший. Длительности нахождения в каждой фазе заболевания определяются для каждого агента отдельно из логнормального распределения с параметрами, взятыми из исследований Coronavirus Disease 2019 (коронавирусная инфекция 2019) (COVID-19) [44—53]. Вероятности приобретения симптомов, развития тяжелой, критической фазы или смерти различны для агентов в зависимости от их возраста, и эти данные также взяты из исследований COVID-19 [51; 53—56]. В зависимости от слоя, которому принадлежит связь, вероятность передачи инфекции домножается на 0.05 для домохозяйства, на 0.01 для школ и рабочих мест, на 0.005 для случайных контактов, что согласуется с литературными данными [57; 58].

В зависимости от объемов имеющихся демографических данных для запуска Covasim можно использовать 3 способа генерации синтетической популяции: случайные сети, алгоритм SynthPops и гибридные сети.

В отсутствие данных применяются случайные сети: всем агентам из Пуассоновского распределения определяются числа контактов, после чего в соответствии с этими значениями формируются произвольные связи между людьми.

SynthPops применяется при наличии информации о вероятностях взаимодействий между возрастными группами в домохозяйствах, школах, на рабочих местах, на улице; распределении численностей людей в школах; возрастной структуре школьников; отношении числа преподавателей к числу школьников; распределении численностей рабочих мест и числах трудоустроенных людей разных возрастов; распределении размеров домохозяйств, возрастов и полов членов семей.

При формировании связей в домохозяйствах SynthPops сначала выбирает размеры семей из их распределения, после чего за каждой семьей закрепляет главу семейства. Остальные члены домохозяйств набираются в соответствии с вероятностями контактов в семьях и распределением возрастов в них.

Контакты между учениками в школах формируются аналогично, собирая агентов по семьям (дети из одних домохозяйств в одних школах), после чего выбираются агенты-работники учебного заведения и в зависимости от размера школы либо создаются контакты между всеми школьниками и работниками, либо каждому работнику определяется число связей с учениками, которые затем формируются.

Связи на рабочих местах формируются аналогично из агентов случайных домохозяйств.

Для создания гибридной сети необходимы данные о распределении возрастов и размеров семей, однако они уже загружены в Covasim для множества стран. Гибридный подход не учитывает распределение возрастов в домохозяйствах, все дети (все агенты от 6 до 22 лет) объединяются в школы, взрослые (агенты возрастом между 22 и 65) в рабочие места с предопределенными для них числами контактов в этих слоях из распределения Пуассона с различными средними значениями. Это позволяет воссоздать некоторую популяционную гетерогенность в случае недостатка демографических данных.

2.4 Дисперсионный анализ чувствительности математических моделей

В данной работе анализ чувствительности модели проводился с помощью дисперсионных методов в связи с ее стохастичностью и их информативностью. Недостатком являются большие вычислительные затраты. Такие методы никак не зависят от природы модели и позволяют оценить влияние изменений любых комбинаций параметров [59].

Пусть задана квадратично интегрируемая функция f на Ω^k, единичном k-мерном гиперкубе,

$$\Omega^k = (X | 0 \le x_i \le 1; i = 1, \ldots, k),$$

Илья Меерович Соболь предложил раскладывать эту функцию в сумму по возрастающим размерностям:

$$f = f_0 + \sum_i f_i + \sum_i \sum_{j>i} f_{ij} + \ldots + f_{12\ldots k},$$

где каждое слагаемое также является квадратично интегрируемым на области определения и является функцией лишь параметров, указанных в его индексе, то есть $f_i = f_i(X_i)$, $f_{ij} = f_{ij}(X_i, X_j)$ и так далее. Соболь показал, что слагаемые суммы можно последовательно вычислить из результатов расчетов модели Y как

$$\begin{aligned}
&f_0 = E(Y),\\
&f_i = E(Y|X_i) - E(Y),\\
&f_{ij} = E(Y|X_i, X_j) - f_i - f_j - E(Y).
\end{aligned}$$

Оказывается, что дисперсия условного матожидания может быть использована как показатель чувствительности. Дисперсии слагаемых в приведенном выше разложении и есть искомые меры. К примеру, $V(f_i(X_i))$ есть $V[E(Y|X_i)]$, поэтому индекс чувствительности первого порядка вводится как

$$S_i = \frac{V[E(Y|X_i)]}{V(Y)}.$$

Индекс первого порядка представляет собой основной вклад варьирования каждого параметра в дисперсию модели и был описан разными учеными как мера важности [60—64].

Тогда же Соболь предложил эквивалентное определение [65], основанное на корреляции между результатами модели Y и условным матожиданием $E(Y|X_i)$

$$S_i = \mathrm{Corr}(Y, E(Y|X_i)).$$

С учетом того, что $V_i = V(f_i(X_i)) = V[E(Y|X_i)]$ выражение

$$V_{ij} = V(f_{ij}(X_i, X_j)) = V(E(Y|X_i, X_j)) - V(E(Y|X_i)) - V(E(Y|X_j))$$

позволяет оценить эффект совместного варьирования (X_i, X_j) на результат Y, который называют индексом чувствительности второго порядка S_{ij} [66]. Аналогичные формулы могут быть выписаны для индексов высших порядков, основываясь на analysis of variance high-dimensional model representation (анализ дисперсии по многомерному представлению) (ANOVA-HDMR) разложении

$$V(Y) = \sum_i V_i + \sum_i \sum_{j>i} V_{ij} + \ldots + V_{123\ldots k}.$$

Деля обе части равенства на $V(Y)$, получаем

$$\sum_i S_i + \sum_i \sum_{j>i} S_{ij} + \ldots + S_{123\ldots k} = 1.$$

Также возможно оценить вклад варьирования параметра и всех связанных с ним взаимодействий вычислением полного индекса Соболя

$$S_{T_i} = \frac{E[V(Y|\boldsymbol{X}_{\sim i})]}{V(Y)} = 1 - \frac{V[E(Y|\boldsymbol{X}_{\sim i})]}{V(Y)},$$

где $\boldsymbol{X}_{\sim i}$ — множество всех параметров, кроме X_i.

2.5 Гравитационные модели транспортных потоков

В отсутствие достаточных объемов данных о перемещениях людей величины транспортных потоков оценивают с помощью гравитационных моделей. Такую модель исследовали Бушар и Паерс (1965) [67]. На основе данных опросов разных лет о перемещениях в Вашингтоне авторы проверили точность гравитационной модели при прогнозировании распределений городских поездок. В качестве метрики качества предсказания авторы использовали среднеквадратичное отклонение оцененных чисел поездок между парами зон Вашингтона от фактических из результатов опросов, деленное на среднее число поездок. Модель представляется уравнением

$$T_{i-j} = \frac{P_i A_j F_{(t_{i-j})} K_{(i-j)}}{\sum\limits_{x=1}^{n} A_x F_{(t_{i-x})} K_{(i-x)}},$$

где

P_i — количество поездок из зоны i,

A_j — количество поездок в зону j,

t_{i-j} — мера пространственного разделения (сумма минимального времени перемещения из зоны i в j и терминального времени),

$F_{(t_{i-j})}$ — эмпирический временной фактор, показывающий убывание транспортных потоков с возрастанием времени поездки,

$K_{(i-j)}$ — калибровочный коэффициент.

Авторы показали, что при правильной калибровке такая модель показывает высокую точность (нормированное среднеквадратичное отклонение чисел поездок менее 15% для крупных потоков), а факторы $F_{(t_{i-j})}$ стабильны во времени. Однако для достижения такой точности необходима стратификация поездок по их цели.

3 Материалы и методы

3.1 Используемые программные пакеты

Таблица 1: Используемые модули Python.

Модуль	Назначение
numpy [68]	Обработка многомерных массивов
pandas [69]	Обработка табличных баз данных
matplotlib [70]	Визуализация данных
seaborn [71]	Визуализация данных (тепловые карты, ящики с усами)
scipy [72]	Проведение статистических тестов, линейной аппроксимации, вычисление коэффициента корреляции Пирсона
SALib [73; 74]	Вычисление последовательности Соболя и индексов Соболя

3.2 Реализация транспортных потоков в исследуемой модели

Все основные алгоритмы в цикле симуляции Covasim, такие как вычисление восприимчивости и трансмиссивности агентов и определение новых зараженных агентов, реализованы с помощью высокооптимизированных операций над 32-битными массивами Numba. Для достижения большей эффективности агенты представлены не отдельными объектами, а в виде срезов набора массивов состояний.

На каждом шаге симуляции в каждом городе выбираются случайные агенты-туристы, количество которых согласуется с величинами транспортных потоков. Агент-турист удаляется из сети контактов своего населенного пункта и внедряется в город прибытия. Там он вступает в фиксированное число случайных контактов. По завершении периода поездки производится обратная операция. Перемещениям не подвержены агенты в тяжелом или критическом состоянии. Если они оказываются в списке туристов, то не покидают свой город. Если же они оказались в таком состоянии по завершении периода поездки, они возвращаются в родной город только при выздоровлении.

Для реализации транспортных потоков мы расширили набор массивов состояний 4 атрибутами агентов, что изображено на Рис. 3:

1. trueId хранит идентификаторы, которые агенты имели, находясь в родном городе, чтобы восстановить их при возвращении агентов;

2. inCity хранит булевые значения нахождения агентов в городе;

3. restInAnotherCityDays хранит число оставшихся дней нахождения агентов в другом городе;

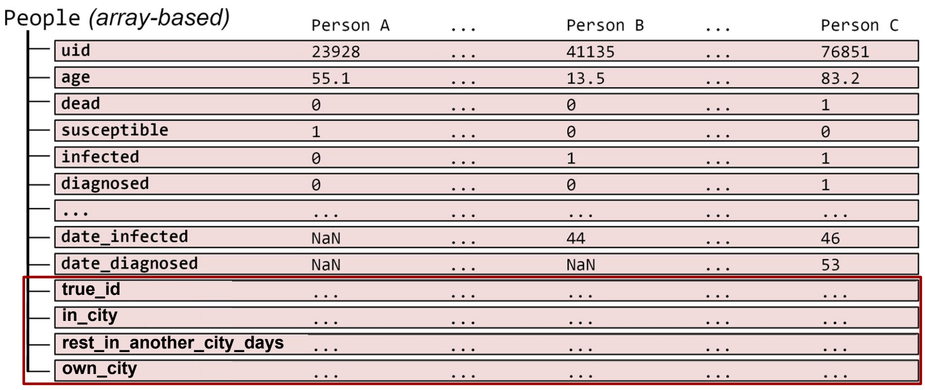

Рис. 3: Матричный подход представления агентов в `Covasim` с добавленными нами атрибутами. Источник: [5].

4. `ownCity` хранит идентификаторы родного города агентов.

Возможность помещения агентов-туристов в города обеспечивается выделением в них агентов-«пустышек», не участвующих во взаимодействиях, так как отмечены как отсутствующие в городе (`inCity` = `False`). К синтетической популяции они подключены посредством слоя контактов `touristLayer` с основными параметрами из `randomLayer`. По приезде в город агенты-туристы вместе со всеми своими свойствами замещают в городе прибытия агентов-«пустышек» и включаются в социальные взаимодействия. Такой подход позволяет не изменять размеры массивов во время вычислений.

Для произведения параллельных вычислений используется модуль `multiprocessing` языка `Python`, с помощью которого создаются процессы, каждый из которых параллельно симулирует происходящее в одном городе, периодически синхронизируясь и обмениваясь информацией о перемещающихся агентов. Именно `multiprocessing`, а не `multithreading` применяется нами, так как создает отдельные процессы со своими Global Interpreter Lock (глобальная блокировка интерпретатора) (GIL). Это позволяет эффективно работать на нескольких ядрах центрального процессора, а значит, и масштабировать вычисления при работе на кластерах. `Multithreading` же используют один процесс, в рамках которого выделяются потоки, из-за чего имеют один GIL на все потоки и истинный параллелизм оказывается невозможен — такой подход обычно применяют, если задачи не являются чисто вычислительными и потоки много времени проводят в режиме ожидания внешних операций.

В текущей версии модели у туристических потоков можно настроить следующие параметры:

1. `adjacencyMatrix`: матрица контактов городов, показывающая величины транспортных потоков из одного города в другой в долях населения города отправления;

2. `timeRelax`: средняя длительность нахождения агента в другом городе, по умолчанию 7;

3. `interventionData`: информация об ограничениях на транспортные потоки — словарь, содержащий информацию о том, какой конкретно транспортный поток ограничивается, в какой день вступают в силу эти ограничения, какова их длительность и во сколько раз уменьшаются потоки из-за введения ограничений;

4. `contactCount`: среднее число контактов у агента-туриста с местными жителями, по умолчанию 40;

5. `beta`: множитель вероятности передачи инфекции в туристическом слое, по умолчанию 0.3.

При запуске симуляции в режиме нескольких городов схема вычислений:

1. Инициализация;

2. Барьер синхронизации процессов;

3. Цикл симуляции:

   (a) Обмен агентами-туристами между процессами (функции `AddBack` и `AddTourists`);

   (b) Основной шаг симуляции;

   (c) Барьер синхронизации процессов;

   (d) Определение отправляемых агентов-туристов (функции `ExtractBack` и `ExtractTourists`);

   (e) Барьер синхронизации процессов.

4. Обработка и визуализация результатов.

3.2.1 Инициализация

На этапе инициализации для каждого города определяется выделяемое число мест под туристов как доля численности населения города равная

$$\max\left(0.1, \sum_j \frac{2 \cdot \texttt{timeRelax}}{\texttt{maxMultCoef} \cdot \texttt{adjacencyMatrix}[i,j]}\right),$$

другими словами, под туристов выделяется число мест больше или равное удвоенному математическому ожиданию их количества при значении `maxMultCoef` по умолчанию равном единице.

Затем эти агенты добавляются в города как отсутствующие ($\texttt{inCity} = \texttt{False}$) и внедряются в социальные взаимодействия с помощью `touristLayer` в соответствии с заданными параметрами.

После инициализации модель входит в главный цикл симуляции.

3.2.2 Функция внедрения перемещаемых агентов в город назначения

Данная функция передает агентам в городе назначения под индексами `inds` все атрибуты агентов-туристов `rest`, кроме идентификатора `uid`, после чего обозначает их присутствие в городе ($people.inCity[inds] \leftarrow True$). Этот алгоритм приведен в Псевдокоде 1.

Algorithm 1 Функция внедрения перемещаемых агентов в город назначения

function UPDATEPEOPLEBYREST($people, rest, inds$)
  for *each* $arrayMember$ **in** $people.arrayMembers$ **do**
    if $arrayMember \neq uid$ **then**
      $people[arrayMember][inds] \leftarrow rest[arrayMember]$
    end if
  end for
  $people.inCity[inds] \leftarrow True$
end function

3.2.3 Функция возвращения агентов в родной город

Данная функция возвращает агентов `backPeople` в соответствии с их родным идентификатором `backPeople.trueUid` с помощью функции `UpdatePeopleByRest`.

3.2.4 Функция размещения агентов-туристов в городе назначения

Данная функция определяет доступные под размещение туристов идентификаторы в городе назначения, выбирает из них необходимое количество соответствующее числу приезжающих туристов, если им хватает мест, или все оставшиеся места в противном случае. Затем на места по выбранным идентификаторам размещаются агенты с помощью функции `UpdatePeopleByRest`. Этот алгоритм приведен в Псевдокоде 2.

Algorithm 2 Функция размещения агентов-туристов в городе назначения

function ADDTOURISTS($people, tourists$)
  $allTouristsInds \leftarrow people.uid[(people.uid \geqslant people.popSize) * (\textbf{not}\ people.inCity)]$
  $freeTouristsInds \leftarrow allTouristsInds[: tourists.popSize]$
  $updatePeopleByRest(people.tourists, freeTouristsInds)$
end function

3.2.5 Функция выбора уезжающих агентов и их извлечения из города

Данная функция в соответствии с долей ежедневно уезжающих из города туристов `outflowRatio` задает их количество из распределения Пуассона. Затем это число идентификаторов агентов выбирается среди местных жителей (`people.uid` < `people.popSize`), находящихся в городе (`people.inCity == True`). Этот список разбивается на списки по городам назначения из известных соотношений транспортных потоков `outflowRatioToCitiesPercent`. Для каждого города назначения `cityInd` из списка туристов удаляются тяжелые, критические, мертвые. Оставшимся назначается длительность пребывания в другом городе `people.restInAnotherCityDays` либо тождественно равная 1 для всех при моделировании потоков рабочих, либо из распределения Пуассона без 0 в противном случае. Затем все отъезжающие агенты из списка `touristPeople` извлекаются из города отправления и сохраняются в виде списка пар (`cityInd`, `touristPeople`). Этот алгоритм приведен в Псевдокоде 3.

Algorithm 3 Функция выбора уезжающих агентов и их извлечения из города

```
function ExtractTourists(people)
    allPeopleLeftCityCount ← Poisson(people.popSize * outflowRatio)
    inCityRestrictionInds ← people.uid[people.inCity * (people.uid < people.popSize)]
    allPeopleLeftCityInds ← random selection of allPeopleLeftCityCount indices from 0 to len(inCityRestrictionInds)
    peopleLeftCityToCityInds ← split(allPeopleLeftCityInds, outflowRatioToCitiesPercent)
    listTouristPeople ← empty list
    for cityInd ← 0 to citiesCount − 1 do
        if cityInd = ownInd then
            continue
        end if
        touristPeopleIndsInCity ← peopleLeftCityToCityInds[cityInd]
        touristPeopleInds ← inCityRestrictionInds[touristPeopleIndsInCity]
        touristPeopleIndsFiltered ← select indices from touristPeopleInds where not (people.severe or people.critical or people.dead)
        touristPeople ← select people with indices touristPeopleIndsFiltered
        if lambda == 0 then
            restInAnotherCityDays ← Ones(touristPeople.popSize)
        else
            restInAnotherCityDays ← ZeroTruncatedPoisson(lambda, touristPeople.popSize)
        end if
        append(cityInd, touristPeople) to listTouristPeople
    end for
    remove people with indices allPeopleLeftCityInds from people
    return listTouristPeople
end function
```

3.2.6 Функция выбора возвращаемых агентов и их извлечения из города

Данная функция определяет индексы агентов, чей родной город соответствует индексу в цикле (`people.ownCity == cityInd`), время поездки которых кончилось (`people.restInAnotherCityDays == 0`) и которые на текущий шаг симуляции не являются тяжелыми, критическими или мертвыми. После этого агенты с этими индексами `backPeople` извлекаются из города их нахождения и добавляются в список возвращаемых как пара (`cityInd`, `backPeople`). Этот алгоритм приведен в Псевдокоде 4.

Algorithm 4 Функция выбора возвращаемых агентов и их извлечения из города

function ExtractBack(people)
  listBackPeople ← *empty list*
  for *cityInd* ← 0 **to** *citiesCount* − 1 **do**
    if *cityInd* = *ownInd* **then**
      continue
    end if
    shouldBackCityInds ← *find people who should return to their own city* **and** *who is* **not** (*people.severe* **or** *people.critical* **or** *people.dead*)
    backPeople ← *select people with indices shouldBackCityInds*
    remove people with indices shouldBackCityInds from people
    append(*cityInd*, *backPeople*) *to listBackPeople*
  end for
  return *listBackPeople*
end function

3.3 Вычисление индексов Соболя

Вычисление индексов Соболя производилось с помощью модуля `SALib` [73; 74] по следующему алгоритму [59]:

- Создавалась матрица случайных значений параметров k размера $(N, 2k)$, определяющая матрицы A и B. Параметры выбирались, используя последовательности псевдослучайных чисел для более равномерного заполнения пространства параметров [75; 76] (этот факт продемонстрирован на Рис. 4)

$$A = \begin{bmatrix} x_1^{(1)} & x_2^{(1)} & \dots & x_i^{(1)} & \dots & x_k^{(1)} \\ x_1^{(2)} & x_2^{(2)} & \dots & x_i^{(2)} & \dots & x_k^{(2)} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ x_1^{(N-1)} & x_2^{(N-1)} & \dots & x_i^{(N-1)} & \dots & x_k^{(N-1)} \\ x_1^{(N)} & x_2^{(N)} & \dots & x_i^{(N)} & \dots & x_k^{(N)} \end{bmatrix},$$

$$B = \begin{bmatrix} x_{k+1}^{(1)} & x_{k+2}^{(1)} & \dots & x_{k+i}^{(1)} & \dots & x_{2k}^{(1)} \\ x_{k+1}^{(2)} & x_{k+2}^{(2)} & \dots & x_{k+i}^{(2)} & \dots & x_{2k}^{(2)} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ x_{k+1}^{(N-1)} & x_{k+2}^{(N-1)} & \dots & x_{k+i}^{(N-1)} & \dots & x_{2k}^{(N-1)} \\ x_{k+1}^{(N)} & x_{k+2}^{(N)} & \dots & x_{k+i}^{(N)} & \dots & x_{2k}^{(N)} \end{bmatrix}.$$

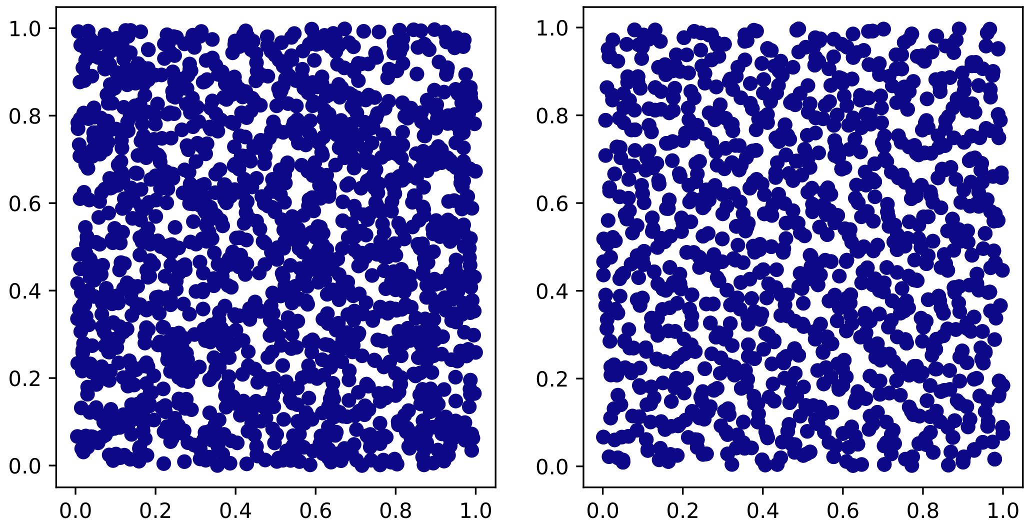

Рис. 4: Наборы точек из двумерного пространства, взятые из равномерного случайного распределения (слева) и последовательности псевдослучайных чисел Соболя (справа).

- Создавалась матрица C_i так, что ее столбцы с индексами, кроме i, были равны соответствующим столбцам матрицы B, а столбец с индексом i был равен столбцу матрицы A с тем же индексом:

$$C_i = \begin{bmatrix} x_{k+1}^{(1)} & x_{k+2}^{(1)} & \dots & x_i^{(1)} & \dots & x_{2k}^{(1)} \\ x_{k+1}^{(2)} & x_{k+2}^{(2)} & \dots & x_i^{(2)} & \dots & x_{2k}^{(2)} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ x_{k+1}^{(N-1)} & x_{k+2}^{(N-1)} & \dots & x_i^{(N-1)} & \dots & x_{2k}^{(N-1)} \\ x_{k+1}^{(N)} & x_{k+2}^{(N)} & \dots & x_i^{(N)} & \dots & x_{2k}^{(N)} \end{bmatrix}.$$

- Вычислялся вывод модели при всех наборах параметров из матриц A, B, C_i, то есть получались три вектора длины N:

$$y_A = f(A) \qquad y_B = f(B) \qquad y_{C_i} = f(C_i).$$

- Вычислялся индекс Соболя первого порядка

$$S_i = \frac{V[E(Y|X_i)]}{V(Y)} = \frac{y_A \cdot y_{C_i} - f_0^2}{y_A \cdot y_A - f_0^2} = \frac{(1/N)\sum_{j=1}^{N} y_A^{(j)} y_{C_i}^{(j)} - f_0^2}{(1/N)\sum_{j=1}^{N} (y_A^{(j)})^2 - f_0^2},$$

где

$$f_0^2 = \left(\frac{1}{N} \sum_{j=1}^{N} y_A^{(j)} \right)^2.$$

Аналогично полный индекс Соболя

$$S_{T_i} = 1 - \frac{V[E(Y|\boldsymbol{X}_{\sim i})]}{V(Y)} = 1 - \frac{y_B \cdot y_{C_i} - f_0^2}{y_A \cdot y_A - f_0^2} = 1 - \frac{(1/N)\sum_{j=1}^{N} y_B^{(j)} y_{C_i}^{(j)} - f_0^2}{(1/N)\sum_{j=1}^{N} (y_A^{(j)})^2 - f_0^2}.$$

Преимуществом такого подхода является тот факт, что для вычисления индексов по k параметрам требуется только $N(k+2)$ запусков симуляции в отличие от полного перебора N^2 точек пространства параметров. Метод позволяет вычислить только полные индексы и индексы первого порядка, но они являются самыми информативными, поэтому мы ограничились ими в данной работе [59].

3.4 Доступ к исходным материалам

Материалы проекта, в том числе исходные коды для проведения экспериментов и обработки их результатов, доступны в репозитории GitHub: https://github.com/KonstantinKlochkovv/bachelor-thesis.

4 Результаты

4.1 Модель связи двух городов с одинаковой численностью агентов

Данный эксперимент был поставлен на модели из двух идентичных городов с населением по 100 тысяч человек, транспортные потоки между которыми были равны. Схематичное представление транспортной модели представлено на Рис. 5. При различных величинах этих потоков (в долях населения в день) и трансмиссивности инфекции (в долях трансмиссивности уханьского варианта SARS-CoV-2) были запущены по 150 симуляций эпидемии при начале в одном из городов. Среднее время пребывания туриста в городе назначения — 7 дней, число контактов — 40, множитель трансмиссивности для туристов — 0.3 (как у случайных контактов в `Covasim`). Перед дальнейшей обработкой полученных эпидемиологических кривых были удалены выбросы — симуляции, в которых эпидемия не началась вовсе и в которых она не перешла во второй город.

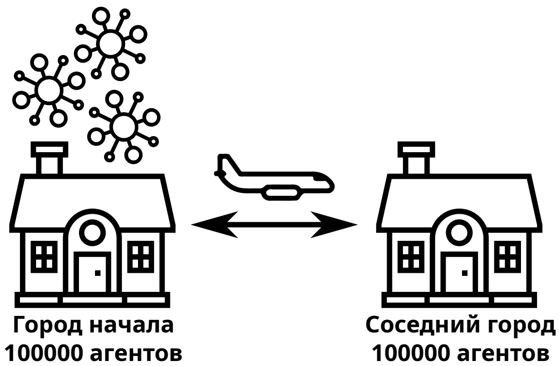

Рис. 5: Схематичное представление простейшей исследуемой модели транспортных потоков.

В результате оказалось, что сдвиг эпидемиологической кривой эпидемии во втором городе относительно первого оказался монотонно убывающей функцией трансмиссивности и транспортного потока, как и разброс результатов, что продемонстрировано на Рис. 6 – 9.

На Рис. 8 приведены тепловые карты среднего сдвига дня пика эпидемии (слева) и t-статистики при сравнении дней пика в рассматриваемых городах (справа). Наблюдается, что в отличие от среднего сдвига дня пика t-статистика монотонно возрастает с увеличением трансмиссивности и убывает с увеличением транспортного потока. Также следует отметить, что при пороговом уровне значимости 0.05 критическое значение t-статистики составляет 1.96, то есть при транспортных потоках $\leqslant 0.003$ между двумя городами наблюдается статистически значимая разница между временами пиков инфицирований.

На Рис. 9 изображены зависимости среднего смещения (слева) и t-статистики смещения дня пика (справа) от величины транспортного потока при различных трансмиссивностях инфекции. Оказалось, что средний сдвиг дня пика является линейной функцией логарифма величины транспортного потока, причем угловой коэффициент ее наклона тем меньше по модулю, чем больше трансмиссивность (коэффициент детерминации $R^2 \in [0.983, 0.995]$). t-статистика же является немонотонной выпуклой вверх функцией транспортного потока с максимумом при величине потока $< 10^{-4}$.

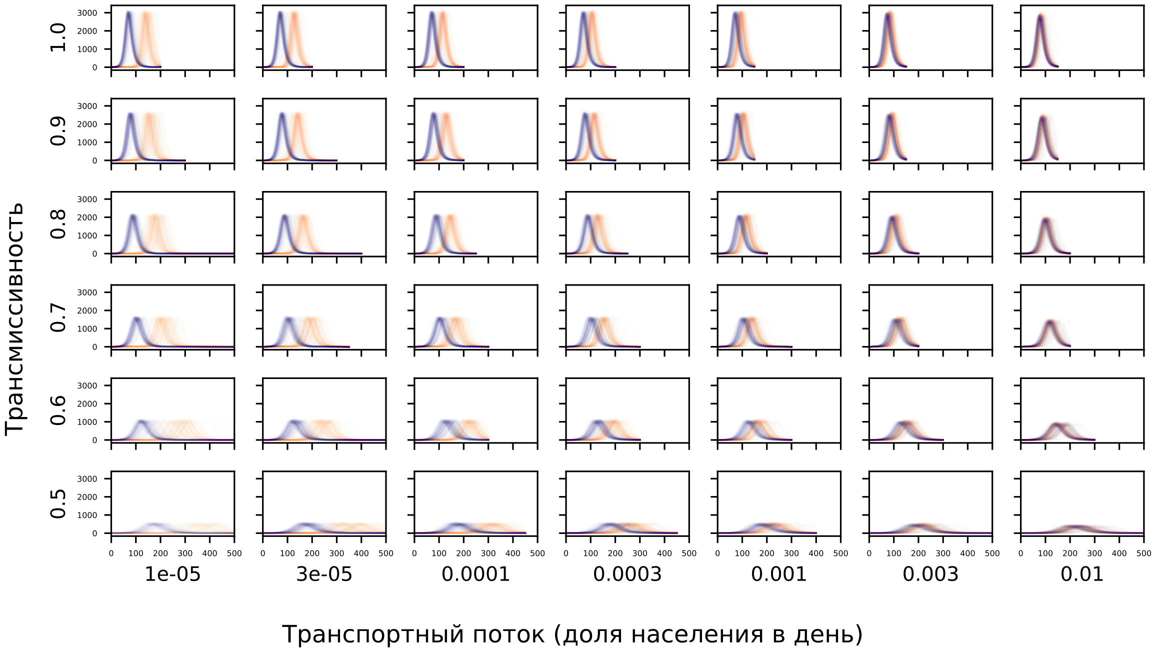

Рис. 6: Эпидемиологические кривые по результатам моделирования двух городов при разных трансмиссивностях инфекции и пассажиропотоках (каждый график соответствует 150 экспериментам).

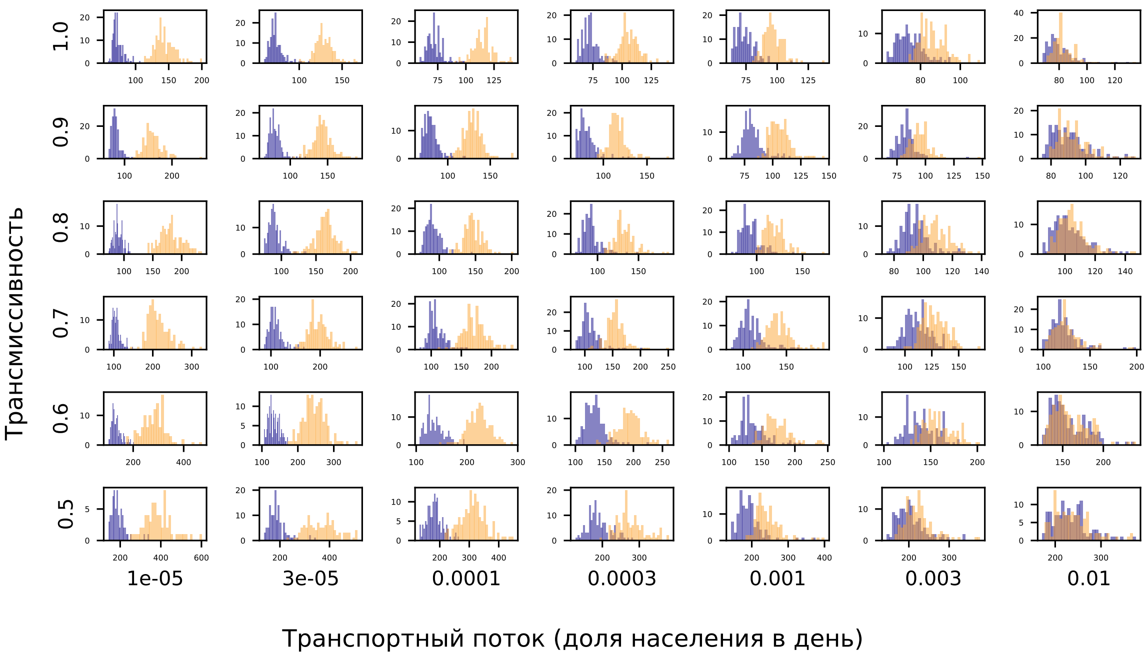

Рис. 7: Гистограммы распределения дня пика инфицирований для двух городов (синяя соответствует городу начала эпидемии, оранжевая — второму городу).

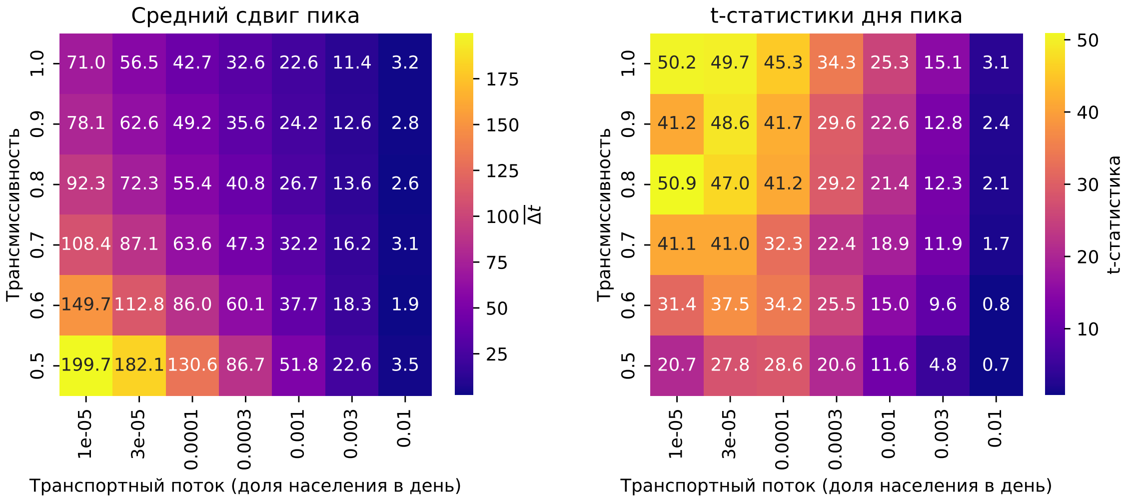

Рис. 8: Тепловые карты среднего сдвига дня пика (слева) и t-статистики сдвига дня пика (справа).

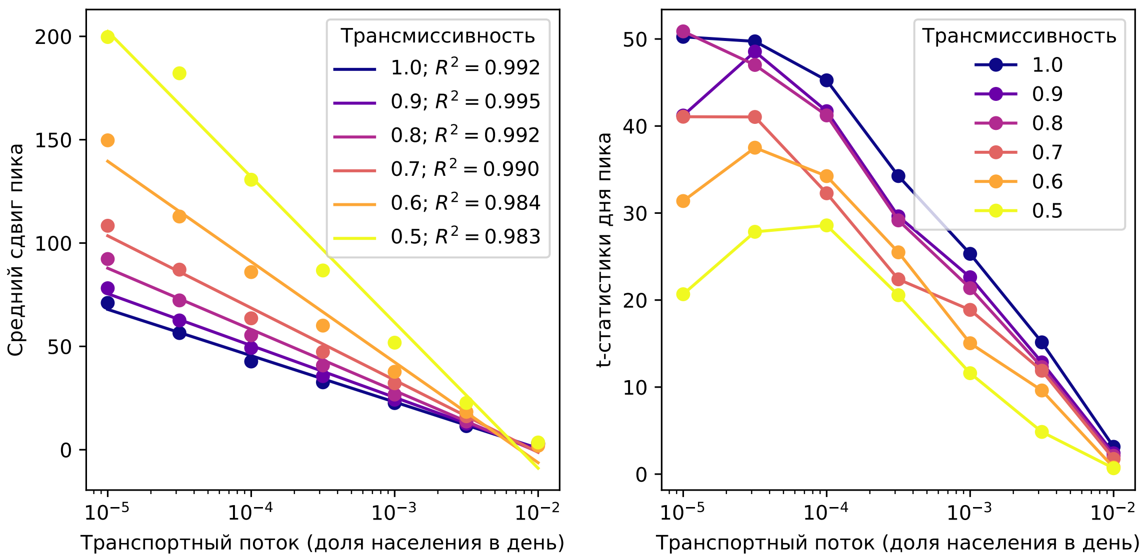

Рис. 9: Зависимости среднего сдвига дня пика (слева) и t-статистики сдвига дня пика (справа) от величины транспортного потока.

4.2 Индексы Соболя транспортных потоков модели двух городов

Данный эксперимент ставился на модели двух городов с суммарным числом агентов 200 тысяч при различных размерах популяций. Проводилось вычисление индексов Соболя первого порядка потока из города начала и в город начала в отношении различных метрик эпидемии (кумулятивное число инфицирований, максимальное число инфицирований, день пика инфицирований) в исследуемых городах по 4096 точкам двумерной последовательности Соболя.

Минимальным исследуемым размером популяции города начала было 10 тысяч агентов. Схематичный вид модели транспортных потоков представлен на Рис. 10.

Наблюдается, что в такой системе исследуемые метрики практически не зависят от величины пассажиропотока из города начала ($S < 0.05$). При этом кумулятивное число инфицирований в городе начала возрастает с увеличением величины транспортного потока в него и индекс Соболя $S = 0.97 \pm 0.08$. Максимальное же число инфицирований обладает таким же характером монотонности, а $S = 0.53 \pm 0.08$. Эти зависимости представлены на Рис. 11. Помимо этого кумулятивное число инфицирований в соседнем городе убывает с увеличением потока из него ($S = 0.71 \pm 0.07$), а день пика инфицирований в нем есть нелинейная убывающая функция потока из него ($S = 0.69 \pm 0.08$). При этом метрики соседнего города также не зависят от потока из города начала ($S < 0.1$). Максимальное число инфицирований соседнего города не зависит ни от одного из рассматриваемых потоков ($S < 0.1$). Эти зависимости приведены на Рис. 12.

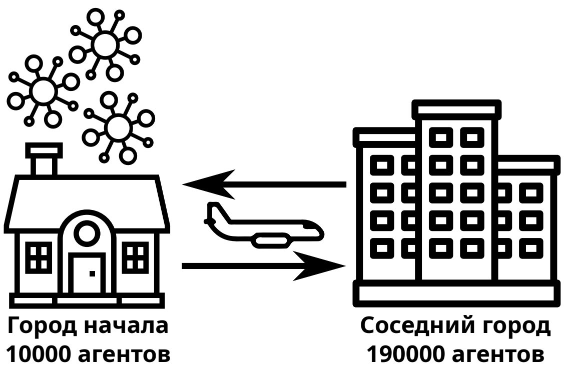

Рис. 10: Схематичное представление модели транспортных потоков при минимальном размере города начала в эксперименте по нахождению индексов Соболя транспортных потоков.

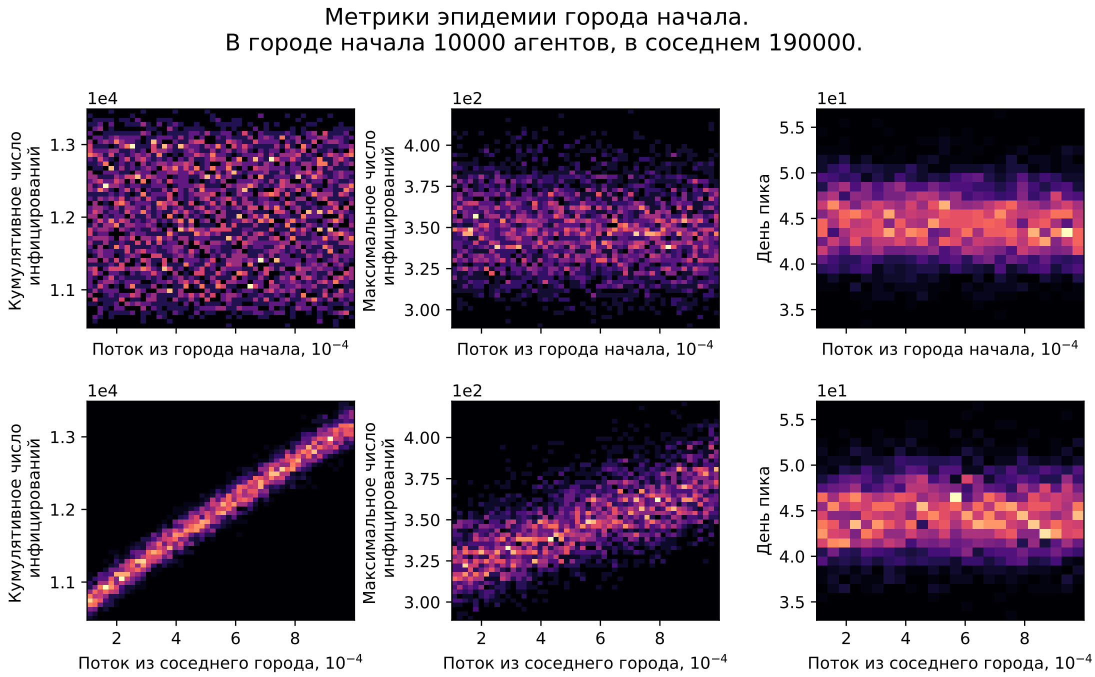

Рис. 11: Тепловые карты зависимости различных метрик эпидемии города начала от транспортных потоков модели при минимальном рассматриваемом размере города начала.

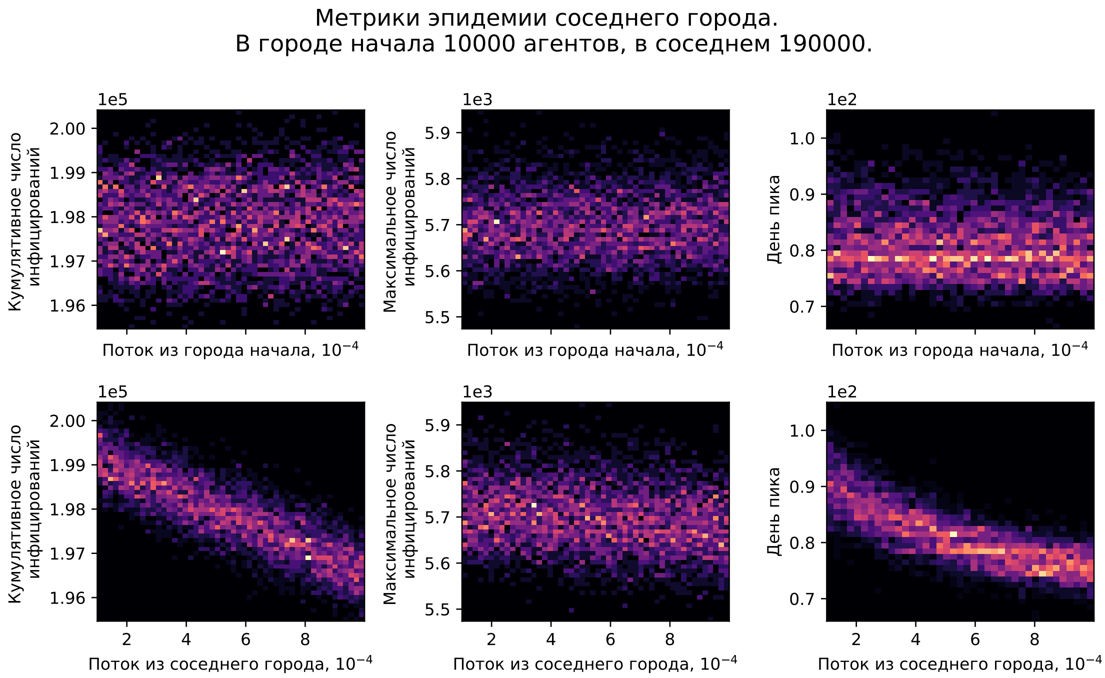

Рис. 12: Тепловые карты зависимости различных метрик эпидемии соседнего города от транспортных потоков модели при минимальном рассматриваемом размере города начала.

Максимальный исследуемый размер популяции города начала — 190 тысяч агентов, при этом в соседнем городе — 10 тысяч. Модель транспортных потоков в данном случае приведена на Рис. 13.

В отличие от предыдущей модели в данном случае метрики практически не зависят от потоков из соседнего города ($S < 0.05$). Теперь кумулятивное число инфицирований и максимальное число инфицирований города начала убывают с ростом потока из него (S соответственно 0.45 ± 0.09, 0.14 ± 0.10). Кумулятивное и максимальное число инфицирований соседнего города возрастает с увеличением потока из города начала, а день наступления пика убывает (S соответственно 0.88 ± 0.08, 0.66 ± 0.07, 0.58 ± 0.08). Эти результаты приведены на Рис. 14, 15.

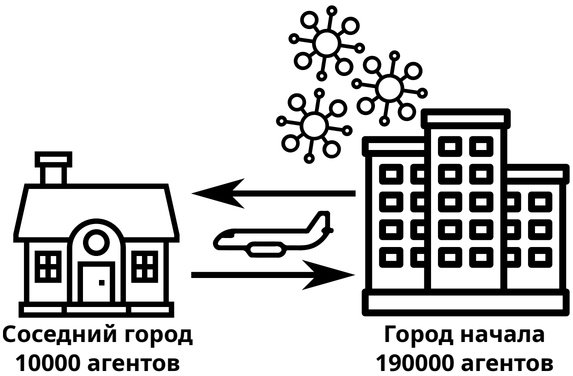

Рис. 13: Схематичное представление модели транспортных потоков при максимальном размере города начала в эксперименте по нахождению индексов Соболя транспортных потоков.

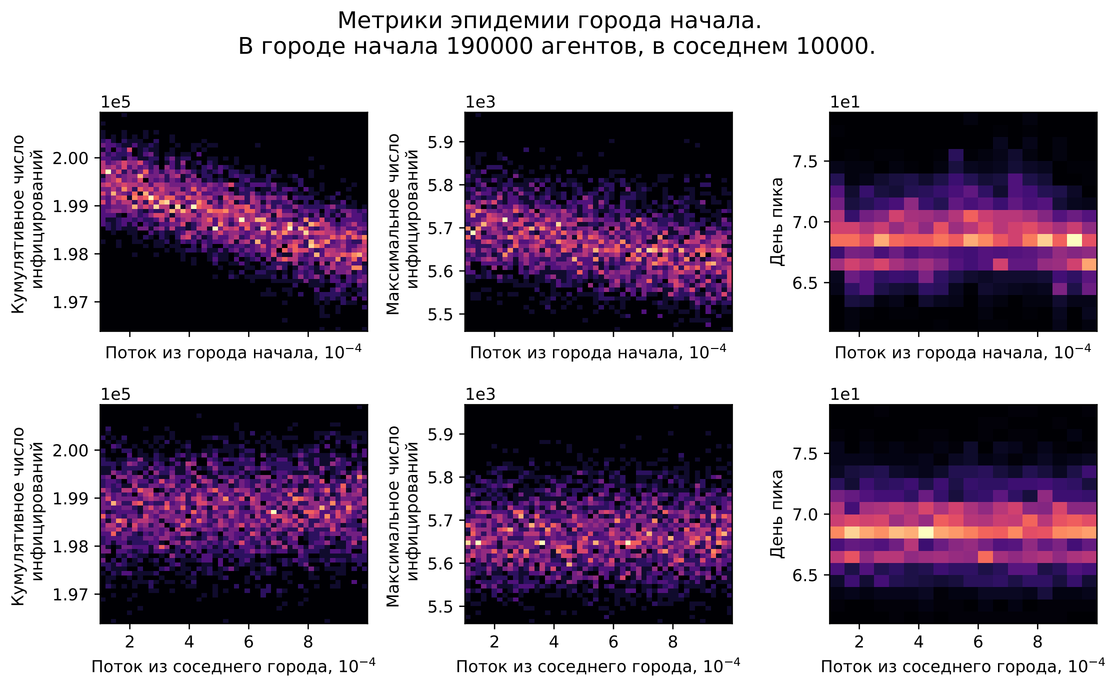

Рис. 14: Тепловые карты зависимости различных метрик эпидемии города начала от транспортных потоков модели при максимальном рассматриваемом размере города начала.

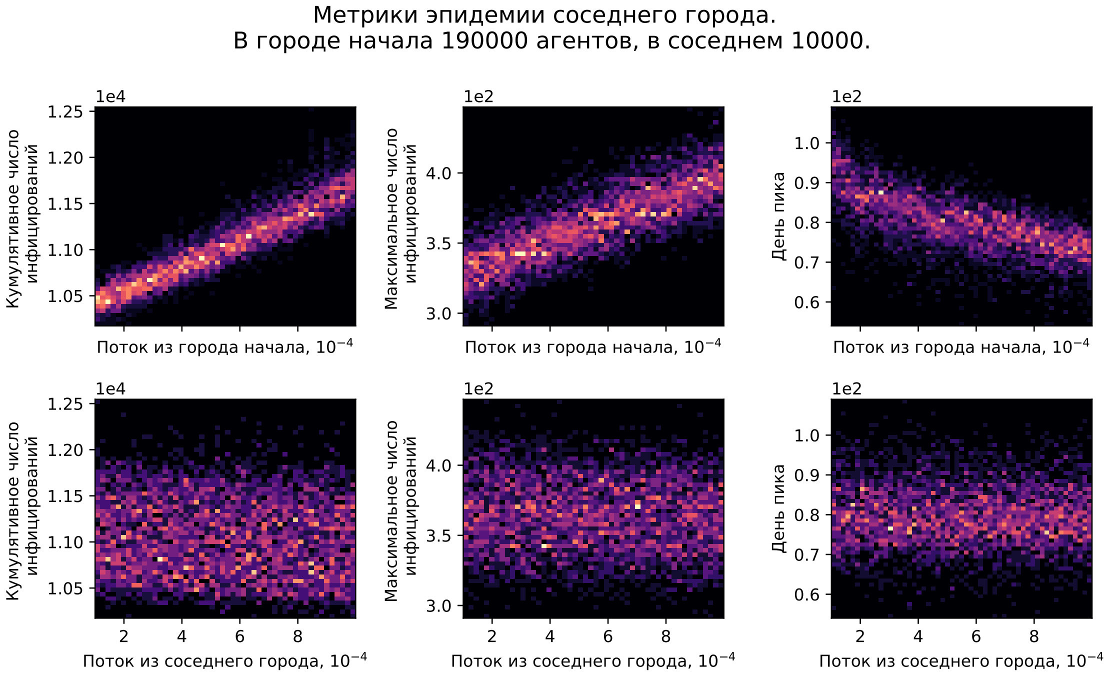

Рис. 15: Тепловые карты зависимости различных метрик эпидемии соседнего города от транспортных потоков модели при максимальном рассматриваемом размере города начала.

Индексы Соболя исследуемых метрик при промежуточных соотношениях размеров популяций городов приведены на Рис. 16. Наблюдается, что с увеличением размера популяции города начала чувствительность кумулятивного числа инфицирований для обеих городов к потоку из города начала растет, а в город начала — убывает.

При этом максимальное число инфицирований практически не зависит от потока из соответствующего города ($S < 0.25$). Также для соседнего города чувствительность максимального числа инфицирований от потока из города начала монотонно возрастает с увеличением размера города начала, а чувствительность для города начала к потоку в него резко убывает так, что при размере популяции города начала большей 50 тысяч агентов практически нулевая ($S < 0.2$).

Чувствительность дня пика инфицирований в соседнем городе к потоку из города начала убывает, а из соседнего города возрастает с увеличением размера города начала. День пика в городе начала практически не зависит от потока в него ($S < 0.2$).

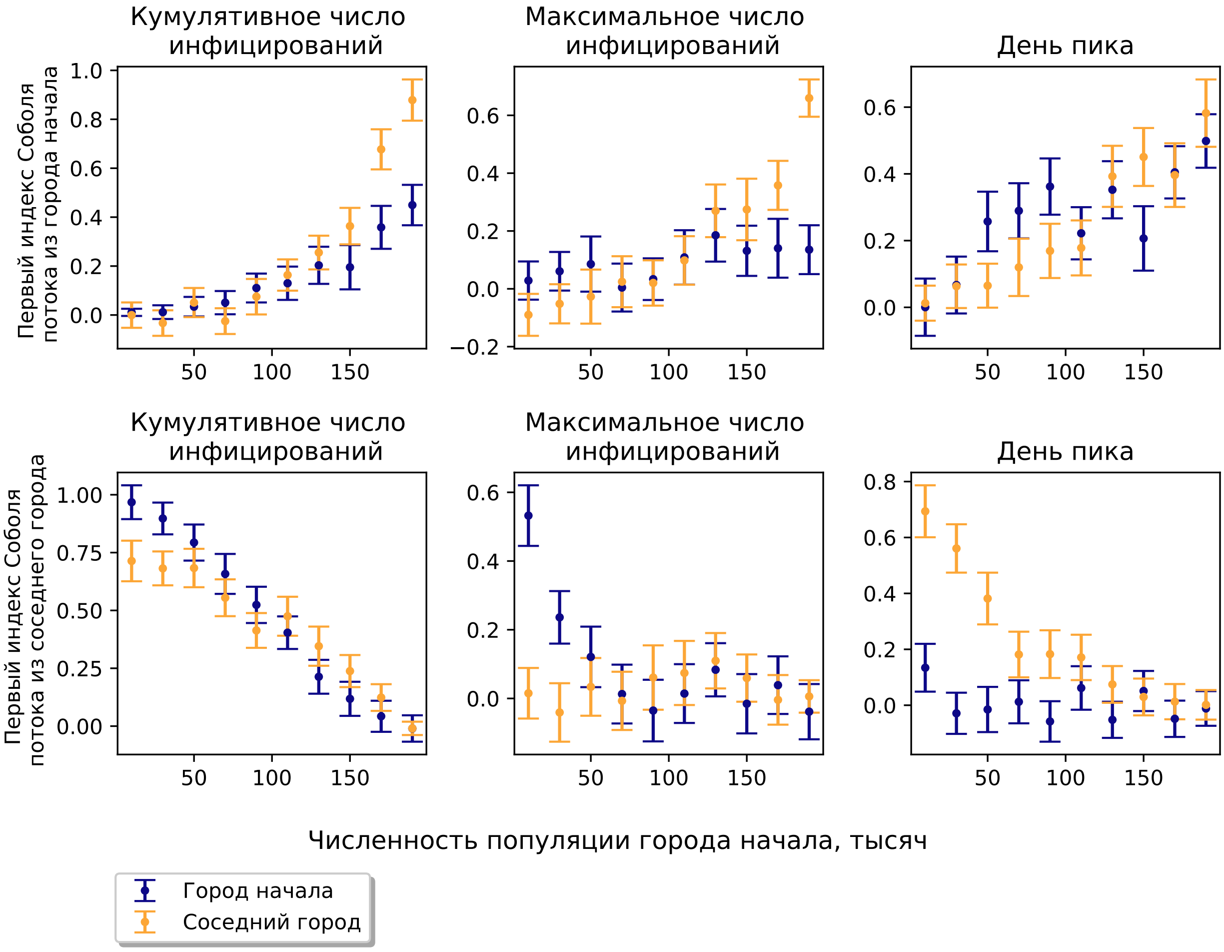

Рис. 16: Индексы Соболя транспортных потоков по отношению к различным метрикам эпидемии в городе начала и соседнем городе.

4.3 Модель трудовой миграции в системе пяти городов («Хаб и сателлиты»)

В данном эксперименте была построена модель пассажиропотоков хаб-сателлиты. Хаб — большой центральный город, вокруг которого сформированы более мелкие города-сателлиты. Отношение количеств агентов в хабе и сателлитах соответствует отношению численностей населения Москвы и Московской области, суммарное число агентов в 20 раз меньше суммарного числа людей для сокращения длительности расчетов. Область разделена на 4 города. Транспортный поток из хаба в сателлиты и из сателлитов в хаб соответствует данным операторов сотовой связи, взятым из [77]. Остальные потоки оценены из гравитационной модели, предполагающей пропорциональность потоков численностям популяций и обратную их пропорциональность квадрату расстояния между городами (частный случай модели, рассмотренной в разделе «Обзор литературы»). Для простоты хаб предполагался располагающимся в центре квадрата с вершинами в сателлитах (то есть расстояние между соседними сателлитами в $\sqrt{2}$ раз больше расстояния хаб-сателлиты, а между противоположными — в 2 раза). Время нахождения агента в городе назначения было равно 1 дню, число контактов — 40, множитель трансмиссивности слоя туристов — 0.6, что соответствует взаимодействиям людей при перемещениях в соседние города на свои рабочие места. Модель транспортных потоков в данной системе приведена на Рис. 17.

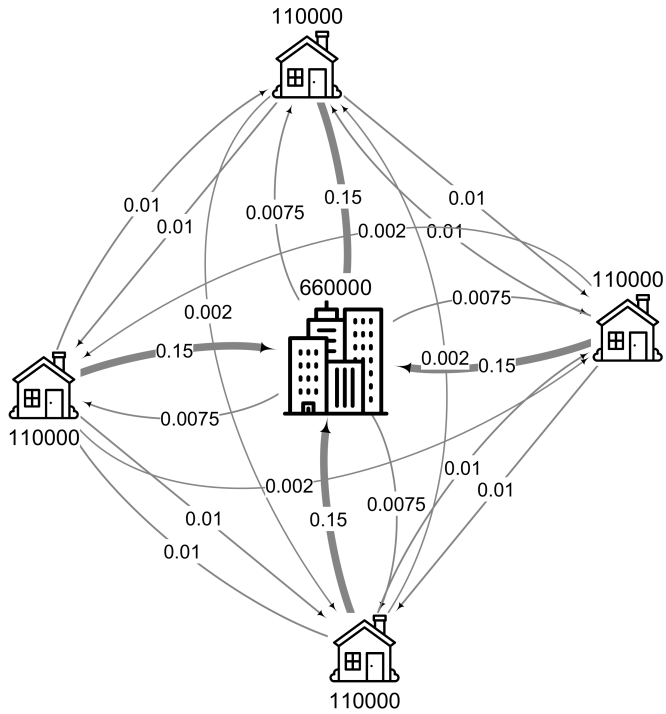

Рис. 17: Представление рассматриваемой модели транспортных потоков в системе хаб-сателлиты в виде графа (подписи у вершин — количество агентов в популяции населенного пункта, у ребер — доля населения города, ежедневно отправляющаяся в однодневную поездку в соответствии с положением ребра).

Запускались по 30 симуляций при начале эпидемии в сателлите и в хабе, после чего на различных шагах симуляции вводились ограничения: уменьшение в 10 или в 100 раз всех транспортных потоков системы или уменьшение в 10 или 100 раз потоков на дорогах, непосредственно связанных с городом начала. Также для сравнения были запущены по 30 симуляций в отсутствие ограничений. По результатам вычислений определялись исследуемые метрики: максимальное число ежедневных инфицирований, критических случаев, смертей, а также день пика инфицирований.

На Рис. 19 приведены зависимости исследуемых метрик системы от дня введения эпидемиологических мер и их вида при начале эпидемии в хабе. Серой полосой отмечены метрики и их среднеквадратичные отклонения по 30 симуляциям без введения ограничений. Доверительный интервал отмечен звездой, если разница наблюдаемой метрики в присутствие ограничений и в отсутствие является статистически значимой при тесте Стьюдента при пороговом уровне значимости 0.05 (с поправкой на множественные сравнения). Для иллюстрации положения дня введения мер на Рис. 18 приведены эпидемиологические кривые системы при начале эпидемии в хабе (слева) и сателлите (справа).

Эксперименты показали, что при начале эпидемии в хабе введение ограничений вплоть до дня пика эпидемии позволяет достигнуть статистически значимого снижения пиковых чисел инфицированных, критических случаев, смертей. При введении ограничений с 40 по 70 день (в конце экспоненциальной фазы) эпидемии разница между уменьшением потоков в 10 и 100 раз между всеми городами или только между хабом и сателлитами незначительна, также эффект от введения этих мер практически не меняется в конце экспоненциальной фазы, это проиллюстрировано на Рис. 19. Даже при введении мер в это время удается достичь снижения пиковых чисел зараженных на 10% ($p = 10^{-49}$), критических случаев на 20% ($p = 10^{-36}$), смертей на 20% ($p = 10^{-25}$). Разница между ограничением потоков в 10 и 100 раз наблюдается на ранних этапах эпидемии — там ограничения пассажиропотоков в 100 раз позволяют уменьшить пиковые значения еще практически в два раза в сравнении с введением в конце экспоненциальной фазы.

На Рис. 20 приведены аналогичные зависимости при начале эпидемии в сателлите. Ключевой разницей является тот факт, что ограничения только на потоки, связанные с этим сателлитом, намного менее эффективны, чем ограничения на все потоки, хотя снижение целевых метрик и является статистически значимым.

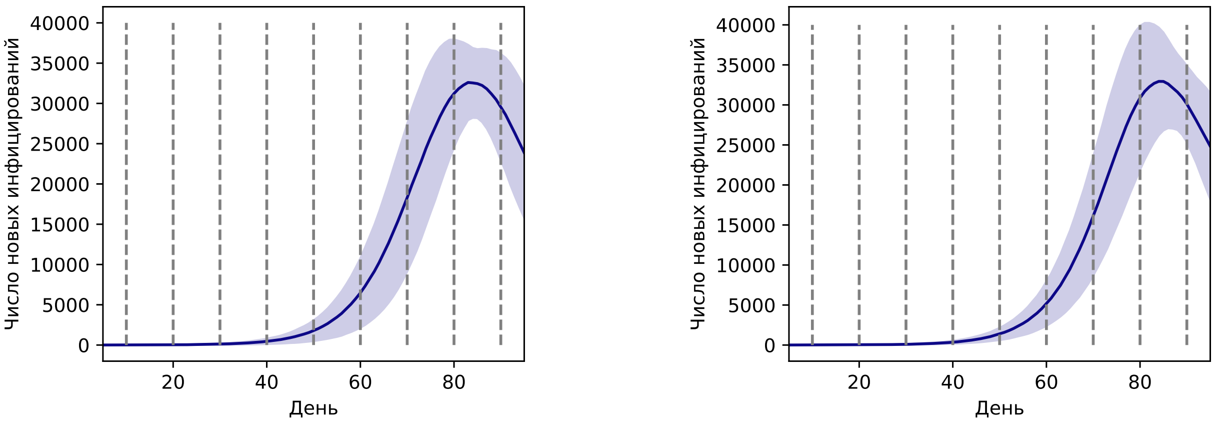

Рис. 18: Суммарное число новых ежедневных инфицирований в системе при моделировании в отсутствие ограничительных мер при начале эпидемии в хабе (слева) и сателлите (справа) (заполнение цветом соответствует среднеквадратичному отклонению по результатам 30 запусков симуляции).

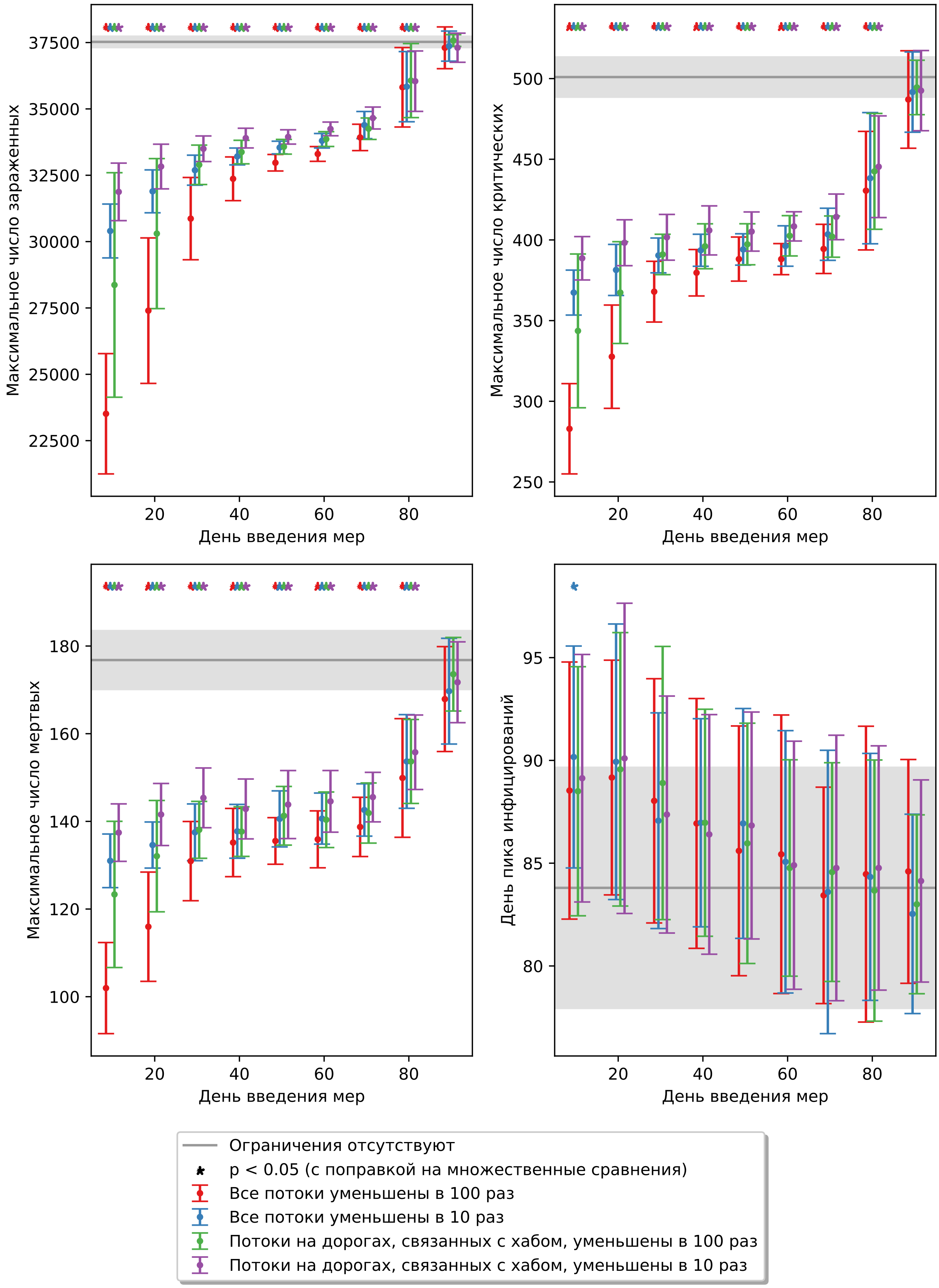

Рис. 19: Зависимости исследуемых метрик системы от дня введения ограничений на транспортные потоки и величины этих ограничений при начале эпидемии в хабе.

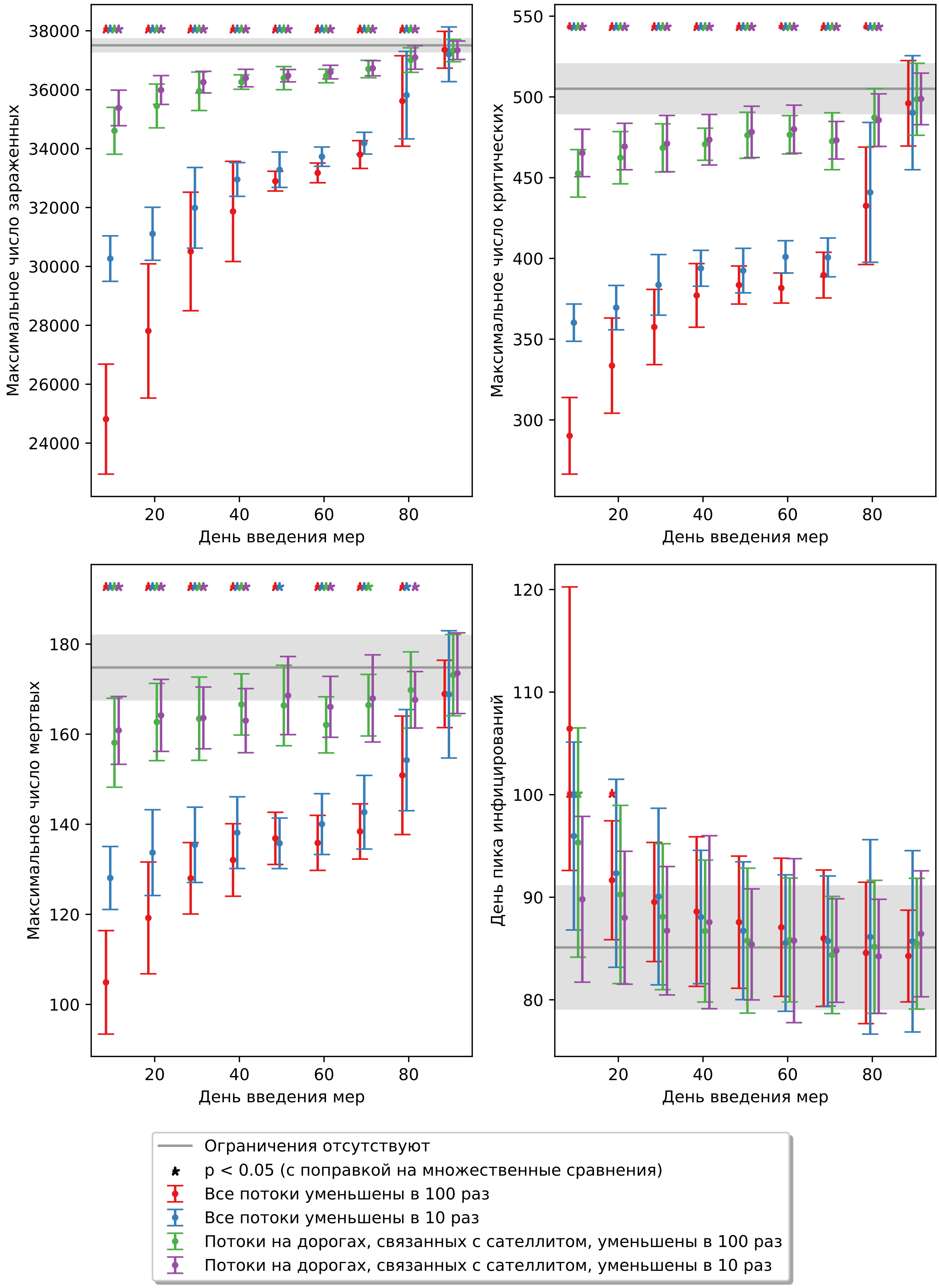

Рис. 20: Зависимости исследуемых метрик системы от дня введения ограничений на транспортные потоки и величины этих ограничений при начале эпидемии в сателлите.

На Рис. 21 изображены те же метрики, но для отдельных городов при начале эпидемии в сателлите и в хабе в отсутствие ограничений. Звездой отмечена статистически значимая разница метрик города при начале в хабе и в сателлите. Оказалось, что в подобной системе детектировать место начала эпидемии по рассматриваемым метрикам не представляется возможным. Это также иллюстрирует Рис. 22.

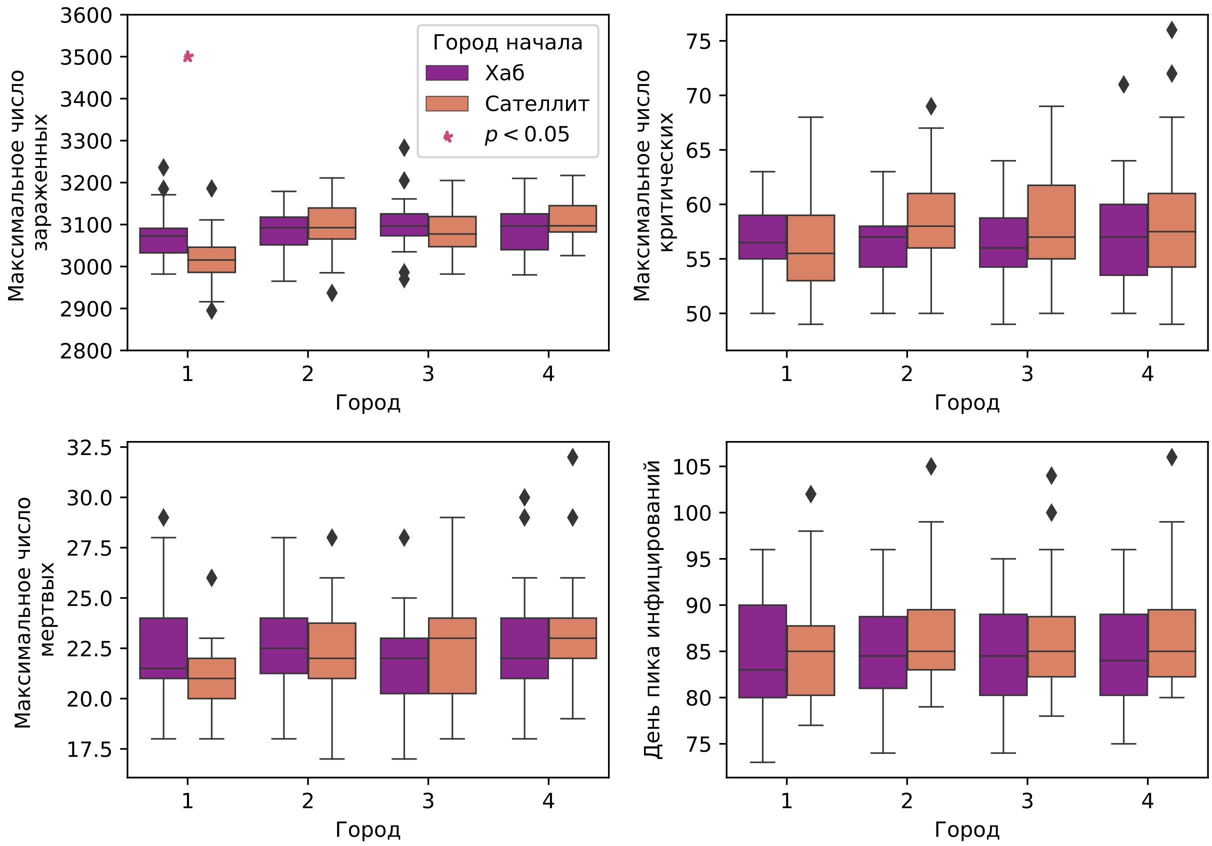

Рис. 21: Исследуемые метрики в различных сателлитах в зависимости от места начала эпидемии (ограничительные меры отсутствуют).

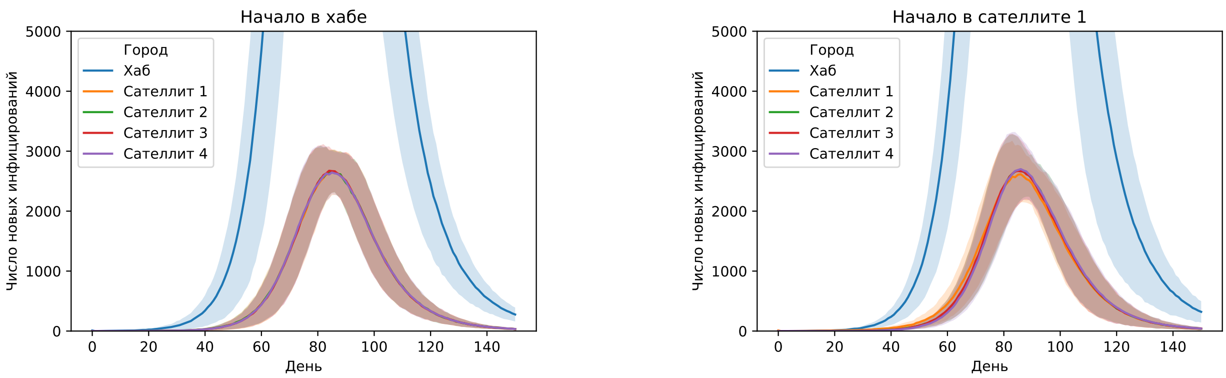

Рис. 22: Суммарное число новых ежедневных инфицирований в отдельных городах в системе хаб-сателлиты при моделировании в отсутствие ограничительных мер при начале эпидемии в хабе (слева) и в сателлите (справа).

4.4 Детекция города начада эпидемии в модели «хаб-сателлиты»

Эксперименты проводились над моделью транспортных потоков хаб-сателлиты, описанной выше, при различных значениях трансмиссивности инфекции и множителя на все величины транспортных потоков. Эпидемиологические кривые для различных сателлитов усреднялись, после чего вычислялось среднее абсолютное отклонение (MAE) этих кривых от средней, нормированное на максимум кривой. Распределение сравниваемых метрик проверялось на нормальность тестом Шапиро-Уилка. В связи с несоответствием распределения метрики нормальному для сравнения результатов симуляции при начале эпидемии в хабе и в сателлите проводился тест Манна-Уитни.

Оказалось, что при начале в хабе, вне зависимости от скорости распространения инфекции в системе, эпидемиологические кривые сателлитов являются более синхронными при всех рассматриваемых множителях транспортных потоков. Это продемонстрировано на Рис. 23.

На Рис. 24 приведены тепловые карты средней разницы нормированных значений MAE эпидемиологических кривых при симулировании начала эпидемии в сателлите и в хабе (слева) и p-значение при сравнении их распределений тестом Манна-Уитни (справа). Среднее значение — монотонно убывающая функция множителя транспортного потока, p-значение — немонотонная функция с максимумом вблизи множителя $3 \cdot 10^{-2}$, что может быть связано с увеличением разброса результатов в этой области, являющимся проявлением дискретности модели на малых величинах транспортных потоков.

При пороговом уровне значимости $p < 0.05$ оказывается, что при трансмиссивности инфекции, соответствующей уханьскому варианту SARS-CoV-2, случаи начала эпидемии в хабе и сателлите удается уверенно различить, то есть предложенная нами в этом эксперименте метрика является более чувствительной для нашей задачи, чем стандартные, рассматриваемые в предыдущем эксперименте. Также при любых других рассматриваемых трансмиссивностях при множителе транспортного потока меньше единицы детекция города начала оказывается возможной. Эти результаты приведены на Рис. 24.

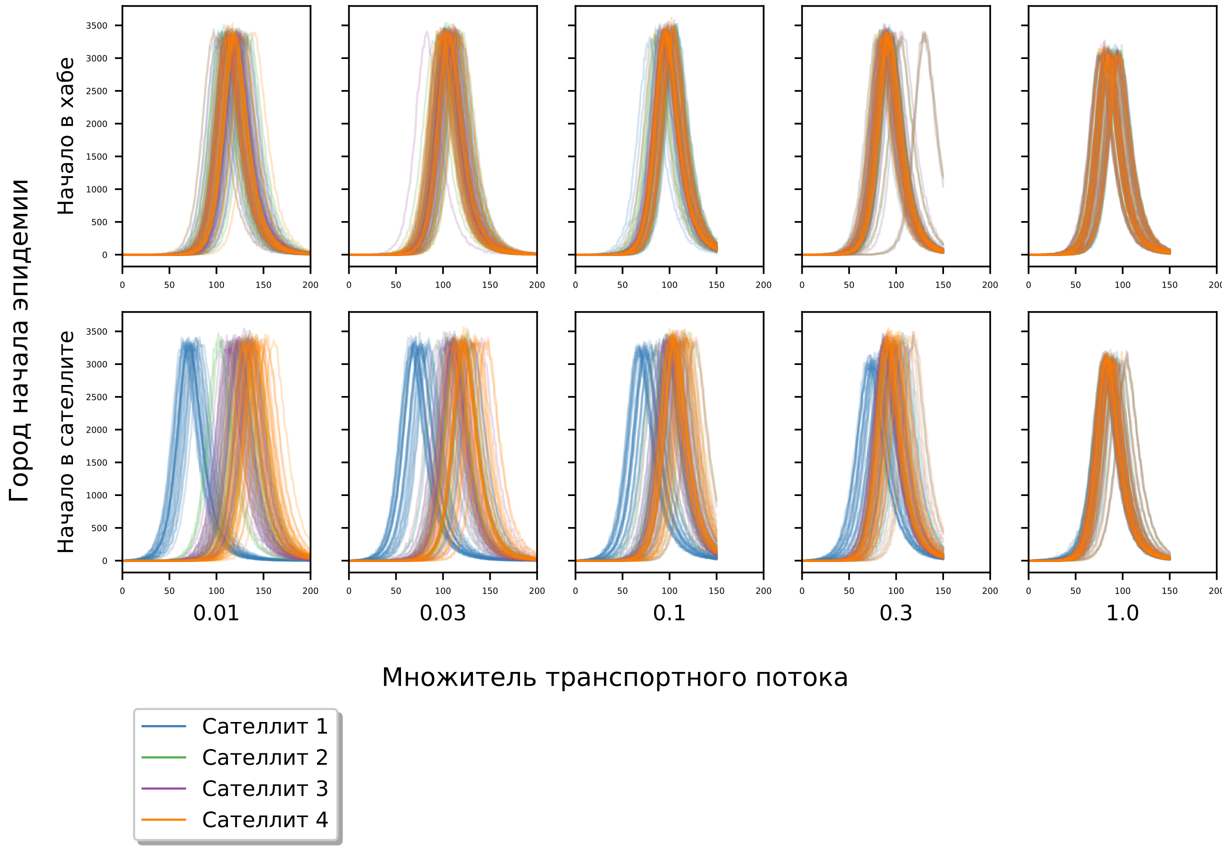

Рис. 23: Эпидемиологические кривые различных сателлитов системы хаб-сателлиты при различных величинах множителя транспортного потока при начале эпидемии в хабе (сверху) и в сателлите 1 (снизу).

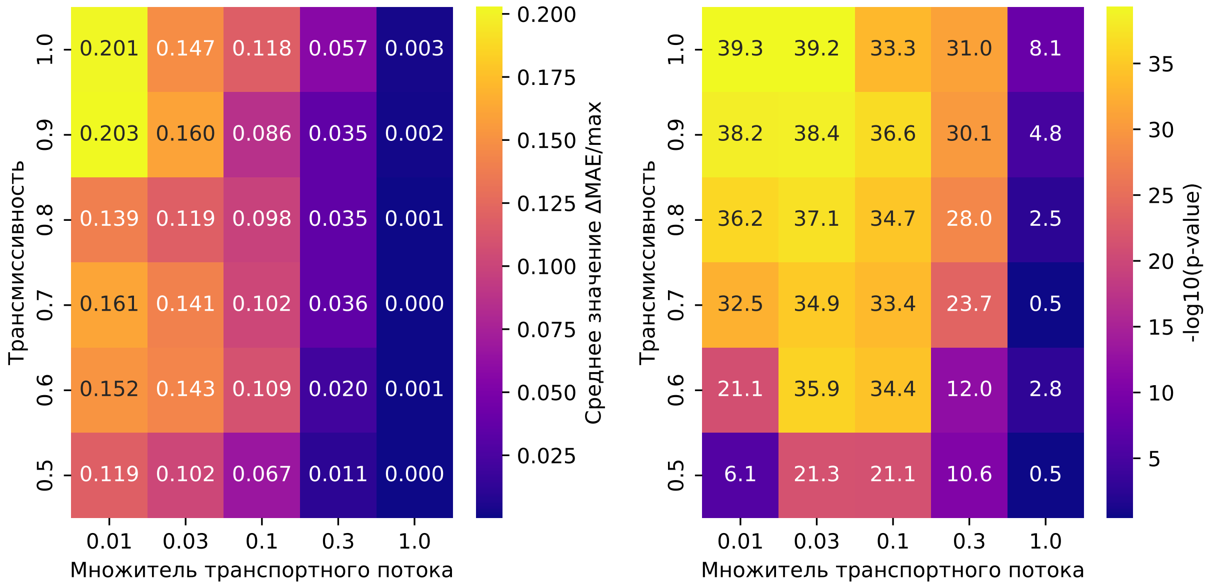

Рис. 24: Тепловые карты средней разницы нормированных MAE эпидемиологических кривых (слева) и p-значения теста Манна-Уитни (справа).

4.5 Модель внутрироссийских авиаперелетов

В данном эксперименте на основе географических данных о территориальных границах субъектов Российской Федерации и переписи населения 2021 года была создана транспортная модель, на которой были проведены эксперименты и анализ зависимости времени между пиками эпидемиологических кривых от логарифма величины транспортного потока.

Симуляция проводилась с числом агентов в 10 раз меньшем населения Российской Федерации для сокращения времени вычислений. Величины транспортных потоков рассчитывались из гравитационной модели, согласно которой поток между субъектами пропорционален произведению численностей их населений и обратно пропорционален расстоянию между ними к степени 1.5. В качестве расстояния использовалось расстояние между геометрическими центрами субъектов, а не какими-либо конкретными городами. Калибровочная константа модели оценивалась на основе данных о годовом числе внутрироссийских пассажирских авиаперелетов и предположения о том, что в рассматриваемой модели все транспортные потоки определяются авиаперелетами. Были произведены вычисления для распространения эпидемии SARS-CoV-2 при ее начале в Приморском крае и Москве.

На Рис. 25 приведены результаты симуляции. Цветом на карте показано отношение числа новых инфицирований на данном шаге симуляции к максимальному ее значению в ходе симуляции в рассматриваемом регионе. При начале эпидемии в Москве наблюдается волнообразное распространение инфекции из эпицентра. При начале в Приморском крае после завершения эпидемии в нем инфекционный агент появляется в большинстве субъектов практически одновременно.

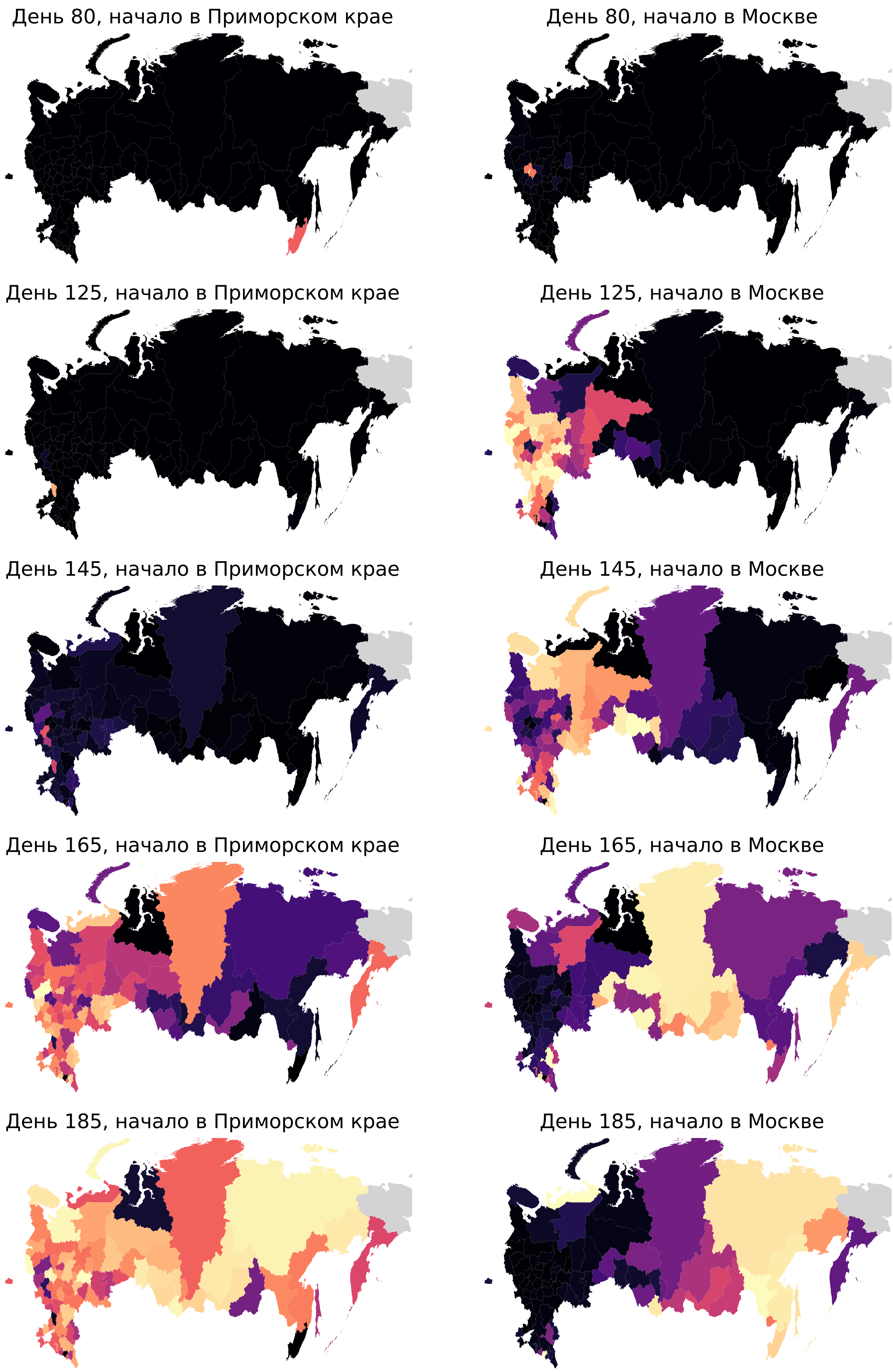

Рис. 25: Ход развития эпидемии в модели транспортных потоков Российской Федерации при симуляции начала в Приморском крае (слева) и Москве (справа). Цветом обозначено отношение числа новых инфицирований на данном шаге симуляции к максимальному в рассматриваемом субъекте.

5 Выводы

В результате исследований в рамках данной работы были построены 3 модели транспортных потоков: модель системы двух городов, модель хаб-сателлиты, модель внутрироссийских авиаперелетов. После их исследования были получены следующие результаты.

Модель двух городов с одинаковой численностью агентов:

- Сдвиг эпидемиологической кривой эпидемии во втором городе относительно города начала есть монотонно убывающая функция трансмиссивности и транспортного потока, как и разброс результатов, и является линейной функцией логарифма величины транспортного потока;
- t-статистика при сравнении средних дней наступления пика инфицирований монотонно возрастает с увеличением трансмиссивности, убывает с увеличением транспортного потока и является немонотонной выпуклой вверх функцией транспортного потока с максимумом при величине потока $< 10^{-4}$.

Модель хаб-сателлиты:

- Введение ограничений на транспортные потоки вплоть до дня пика эпидемии позволяет достигнуть статистически значимого снижения пиковых чисел инфицированных, критических случаев, смертей;
- При начале эпидемии в сателлите ограничения только на потоки дорог, непосредственно связанных с этим сателлитом, значительно менее эффективны, чем ограничения на все потоки.

Список литературы

1. *Hethcote H. W.* The mathematics of infectious diseases // SIAM review. — 2000. — Т. 42, № 4. — С. 599—653.
2. *Hethcote H. W.* Three basic epidemiological models // Applied mathematical ecology. — Springer, 1989. — С. 119—144.
3. *Hethcote H.*, *Van Ark J.* Transmission and AIDS in the United States, Lect // Notes Biomath. — 1992. — Т. 95.
4. Influenza epidemic spread simulation for Poland—a large scale, individual based model study / F. Rakowski [и др.] // Physica A: Statistical Mechanics and its Applications. — 2010. — Т. 389, № 16. — С. 3149—3165.
5. Covasim: an agent-based model of COVID-19 dynamics and interventions / C. C. Kerr [и др.] // PLOS Computational Biology. — 2021. — Т. 17, № 7. — e1009149.
6. *Keeling M. J.*, *Eames K. T.* Networks and epidemic models // Journal of The Royal Society Interface. — 2005. — Т. 2, № 4. — С. 295—307.
7. *Kermack W. O.*, *McKendrick A. G.* A contribution to the mathematical theory of epidemics // Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character. — 1927. — Т. 115, № 772. — С. 700—721.
8. *Bailey N. T.* The mathematical theory of epidemics // (No Title). — 1957.
9. *Anderson R. M.*, *Robert M.* May. Infectious diseases of humans: dynamics and control. — 1992.
10. *Grenfell B.* Chance and chaos in measles dynamics // Journal of the Royal Statistical Society: Series B (Methodological). — 1992. — Т. 54, № 2. — С. 383—398.
11. *Rohani P.*, *Earn D. J.*, *Grenfell B. T.* Impact of immunisation on pertussis transmission in England and Wales // The Lancet. — 2000. — Т. 355, № 9200. — С. 285—286.
12. *Hethcote H.*, *Yorke J.* Springer Lecture Notes in Biomathematics // Lecture Notes in Biomathematics. — 1984.
13. *Garnett G. P.*, *Anderson R. M.* Sexually transmitted diseases and sexual behavior: insights from mathematical models // Journal of Infectious Diseases. — 1996. — Т. 174, Supplement_2. — S150—S161.
14. *Anderson R. M.* The epidemiology of HIV infection: variable incubation plus infectious periods and heterogeneity in sexual activity // Journal of the Royal Statistical Society: Series A (Statistics in Society). — 1988. — Т. 151, № 1. — С. 66—93.
15. *Grenfell B. T.*, *Bjørnstad O. N.*, *Kappey J.* Travelling waves and spatial hierarchies in measles epidemics // Nature. — 2001. — Т. 414, № 6865. — С. 716—723.

16. *Ghani A. C.*, *Swinton J.*, *Garnett G. P.* The role of sexual partnership networks in the epidemiology of gonorrhea // Sexually transmitted diseases. — 1997. — Т. 24, № 1. — С. 45—56.

17. *Keeling M. J.* Modelling the persistence of measles // Trends in microbiology. — 1997. — Т. 5, № 12. — С. 513—518.

18. *Pastor-Satorras R.*, *Vespignani A.* Epidemic spreading in scale-free networks // Physical review letters. — 2001. — Т. 86, № 14. — С. 3200.

19. Mean-field models for non-Markovian epidemics on networks / N. Sherborne [и др.] // Journal of mathematical biology. — 2018. — Т. 76. — С. 755—778.

20. *Keeling M. J.* The effects of local spatial structure on epidemiological invasions // Proceedings of the Royal Society of London. Series B: Biological Sciences. — 1999. — Т. 266, № 1421. — С. 859—867.

21. *House T.*, *Keeling M. J.* Insights from unifying modern approximations to infections on networks // Journal of The Royal Society Interface. — 2011. — Т. 8, № 54. — С. 67—73.

22. *Karrer B.*, *Newman M. E.* Message passing approach for general epidemic models // Physical Review E—Statistical, Nonlinear, and Soft Matter Physics. — 2010. — Т. 82, № 1. — С. 016101.

23. Effective degree network disease models / J. Lindquist [и др.] // Journal of mathematical biology. — 2011. — Т. 62. — С. 143—164.

24. *Miller J. C.*, *Slim A. C.*, *Volz E. M.* Edge-based compartmental modelling for infectious disease spread // Journal of the Royal Society Interface. — 2012. — Т. 9, № 70. — С. 890—906.

25. *Volz E.* SIR dynamics in random networks with heterogeneous connectivity // Journal of mathematical biology. — 2008. — Т. 56. — С. 293—310.

26. *Aiéllo O. E.*, *Silva M. A. da.* New approach to dynamical Monte Carlo methods: application to an epidemic model // Physica A: Statistical Mechanics and its Applications. — 2003. — Т. 327, № 3/4. — С. 525—534.

27. *Haas V. J.*, *Caliri A.*, *Da Silva M.* Temporal duration and event size distribution at the epidemic threshold // Journal of Biological Physics. — 1999. — Т. 25. — С. 309—324.

28. Efficient method for comprehensive computation of agent-level epidemic dissemination in networks / G. M. Nakamura [и др.] // Scientific reports. — 2017. — Т. 7, № 1. — С. 40885.

29. *Albert R.*, *Barabási A.-L.* Statistical mechanics of complex networks // Reviews of modern physics. — 2002. — Т. 74, № 1. — С. 47.

30. Epidemic processes in complex networks / R. Pastor-Satorras [и др.] // Reviews of modern physics. — 2015. — Т. 87, № 3. — С. 925—979.

31. *Van Kampen N. G.* Stochastic processes in physics and chemistry. Т. 1. — Elsevier, 1992.

32. Herd immunity: basic concept and relevance to public health immunization practices / J. P. Fox [и др.] // American journal of epidemiology. — 1971. — Т. 94, № 3. — С. 179—189.

33. An influmza simulation model for immunization studies / L. R. Elveback [и др.] // American journal of epidemiology. — 1976. — Т. 103, № 2. — С. 152—165.

34. *Koopman J.* Controlling smallpox // Science. — 2002. — Т. 298, № 5597. — С. 1342—1344.

35. *Li J.*, *Xiang T.*, *He L.* Modeling epidemic spread in transportation networks: A review // Journal of Traffic and Transportation Engineering (English Edition). — 2021. — Т. 8, № 2. — С. 139—152.

36. *Hyman J. M.*, *LaForce T.* Modeling the spread of influenza among cities // Bioterrorism: Mathematical modeling applications in homeland security. — SIAM, 2003. — С. 211—236.

37. *Wang W.*, *Zhao X.-Q.* An epidemic model in a patchy environment // Mathematical biosciences. — 2004. — Т. 190, № 1. — С. 97—112.

38. Spreading disease with transport-related infection / Y. Takeuchi, Y. Saito [и др.] // Journal of theoretical biology. — 2006. — Т. 239, № 3. — С. 376—390.

39. *Liu X.*, *Takeuchi Y.* Spread of disease with transport-related infection and entry screening // Journal of Theoretical Biology. — 2006. — Т. 242, № 2. — С. 517—528.

40. An SEIS epidemic model with transport-related infection / H. Wan [и др.] // Journal of theoretical biology. — 2007. — Т. 247, № 3. — С. 507—524.

41. *Simoes J.* Modelling a mumps outbreak through spatially explicit agents // Potentials of Complexity Science for Business, Governments, and the Media. — 2006.

42. *Frias-Martinez E.*, *Williamson G.*, *Frias-Martinez V.* An agent-based model of epidemic spread using human mobility and social network information // 2011 IEEE third international conference on privacy, security, risk and trust and 2011 IEEE third international conference on social computing. — IEEE. 2011. — С. 57—64.

43. *Crooks A. T.*, *Hailegiorgis A. B.* An agent-based modeling approach applied to the spread of cholera // Environmental Modelling & Software. — 2014. — Т. 62. — С. 164—177.

44. The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application / S. A. Lauer [и др.] // Annals of internal medicine. — 2020. — Т. 172, № 9. — С. 577—582.

45. Serial interval of COVID-19 among publicly reported confirmed cases / Z. Du [и др.] // Emerging infectious diseases. — 2020. — Т. 26, № 6. — С. 1341.

46. *Nishiura H.*, *Linton N. M.*, *Akhmetzhanov A. R.* Serial interval of novel coronavirus (COVID-19) infections // International journal of infectious diseases. — 2020. — Т. 93. — С. 284—286.

47. Investigation of three clusters of COVID-19 in Singapore: implications for surveillance and response measures / R. Pung [и др.] // The Lancet. — 2020. — Т. 395, № 10229. — С. 1039—1046.

48. Incubation period and other epidemiological characteristics of 2019 novel coronavirus infections with right truncation: a statistical analysis of publicly available case data / N. M. Linton [и др.] // Journal of clinical medicine. — 2020. — Т. 9, № 2. — С. 538.

49. Temporal dynamics in viral shedding and transmissibility of COVID-19 / X. He [и др.] // Nature medicine. — 2020. — Т. 26, № 5. — С. 672—675.

50. Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus–infected pneumonia in Wuhan, China / D. Wang [и др.] // jama. — 2020. — Т. 323, № 11. — С. 1061—1069.

51. Clinical progression of patients with COVID-19 in Shanghai, China / J. Chen [и др.] // Journal of infection. — 2020. — Т. 80, № 5. — e1—e6.

52. Estimates of the severity of coronavirus disease 2019: a model-based analysis / R. Verity [и др.] // The Lancet infectious diseases. — 2020. — Т. 20, № 6. — С. 669—677.

53. Virological assessment of hospitalized patients with COVID-2019 / R. Wölfel [и др.] // Nature. — 2020. — Т. 581, № 7809. — С. 465—469.

54. Age-specific mortality and immunity patterns of SARS-CoV-2 / M. O'Driscoll [и др.] // Nature. — 2021. — Т. 590, № 7844. — С. 140—145.

55. Report 34: COVID-19 Infection Fatality Ratio: Estimates from Seroprevalence / L. W. Baguelin [и др.] // See paper. — 2020. — Т. 201029.

56. Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand. Imperial College COVID-19 Response Team / N. M. Ferguson [и др.] // Imperial College COVID-19 Response Team. — 2020. — Т. 20, № 10.25561. — С. 77482.

57. Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China / J. Zhang [и др.] // Science. — 2020. — Т. 368, № 6498. — С. 1481—1486.

58. *Lader D.*, *Short S.*, *Gershuny J.* The time use survey, 2005 // Office for National Statistics, London. — 2006.

59. Global sensitivity analysis: the primer / A. Saltelli [и др.]. — John Wiley & Sons, 2008.

60. *Hora S.*, *Iman R.* A comparison of maximum/bounding and Bayesian/Monte Carlo for fault tree uncertainty analysis // Report SAND85-2839, Sandia Laboratories. — 1986.

61. *Ishigami T., Homma T.* An importance quantification technique in uncertainty analysis for computer models // [1990] Proceedings. First international symposium on uncertainty modeling and analysis. — IEEE. 1990. — С. 398—403.

62. *Iman R. L., Hora S. C.* A robust measure of uncertainty importance for use in fault tree system analysis // Risk analysis. — 1990. — Т. 10, № 3. — С. 401—406.

63. *Saltelli A., Andres T., Homma T.* Sensitivity analysis of model output: An investigation of new techniques // Computational statistics & data analysis. — 1993. — Т. 15, № 2. — С. 211—238.

64. *Homma T., Saltelli A.* Importance measures in global sensitivity analysis of nonlinear models // Reliability Engineering & System Safety. — 1996. — Т. 52, № 1. — С. 1—17.

65. *Sobol' I. M.* On the "freezing" of inessential variables // Vestnik Moskovskogo Universiteta. Seriya 1. Matematika. Mekhanika. — 1996. — № 6. — С. 92—94.

66. Statistics for experimenters. Т. 664 / G. E. Box, W. H. Hunter, S. Hunter [и др.]. — John Wiley, sons New York, 1978.

67. *Bouchard R. J., Pyers C. E.* Use of gravity model for describing urban travel // Highway Research Record. — 1965. — Т. 88, № 196. — С. 5.

68. Array programming with NumPy / C. R. Harris [и др.] // Nature. — 2020. — Сент. — Т. 585, № 7825. — С. 357—362. — DOI: 10.1038/s41586-020-2649-2. — URL: https://doi.org/10.1038/s41586-020-2649-2.

69. *team T. pandas development.* pandas-dev/pandas: Pandas. — Вер. latest. — 02.2020. — DOI: 10.5281/zenodo.3509134. — URL: https://doi.org/10.5281/zenodo.3509134.

70. *Hunter J. D.* Matplotlib: A 2D graphics environment // Computing In Science & Engineering. — 2007. — Т. 9, № 3. — С. 90—95.

71. mwaskom/seaborn: v0.8.1 (September 2017) / M. Waskom [и др.]. — Вер. v0.8.1. — 09.2017. — DOI: 10.5281/zenodo.883859. — URL: https://doi.org/10.5281/zenodo.883859.

72. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python / P. Virtanen [и др.] // Nature Methods. — 2020. — Т. 17. — С. 261—272. — DOI: 10.1038/s41592-019-0686-2.

73. *Iwanaga T., Usher W., Herman J.* Toward SALib 2.0: Advancing the accessibility and interpretability of global sensitivity analyses // Socio-Environmental Systems Modelling. — 2022. — Май. — Т. 4. — С. 18155. — DOI: 10.18174/sesmo.18155. — URL: https://sesmo.org/article/view/18155.

74. *Herman J., Usher W.* SALib: An open-source Python library for Sensitivity Analysis // The Journal of Open Source Software. — 2017. — Янв. — Т. 2, № 9. — DOI: 10.21105/joss.00097. — URL: https://doi.org/10.21105/joss.00097.

75. *Sobol I. M.* The distribution of points in a cube and the approximate evaluation of integrals // USSR Computational mathematics and mathematical physics. — 1967. — T. 7. — C. 86—112.

76. *Soboĺ I.* Uniformly distributed sequences with additional uniformity properties // USSR Computational mathematics and mathematical physics. — 1976. — T. 16, № 5. — C. 1332—1337.

77. *Makhrova A. G.*, *Kirillov P. L.*, *Bochkarev A. N.* Work commuting of the population in the Moscow agglomeration: estimating commuting flows using mobile operator data // Regional Research of Russia. — 2017. — T. 7. — C. 36—44.