



Федеральное государственное бюджетное учреждение науки Институт
вычислительной математики и математической геофизики Сибирского
отделения Российской академии наук (ИВМиМГ СО РАН)

*Алгоритм построения совмещенных
математических моделей эпидемиологии:
идентифицируемость и оптимизация*

*по материалам диссертации на соискание ученой степени доктора физико-математических наук по
специальности 1.2.2 – математическое моделирование, численные методы и комплексы программ*

Выступающий: к.ф.-м.н. О.И. Кривортько

Научный консультант: чл.-корр. РАН С.И. Кабанихин

Новосибирск, 2022

O себе

Руководитель коллектива, который рассчитывает сценарии распространения COVID-19 в Новосибирской области («Антивирус» СО РАН)

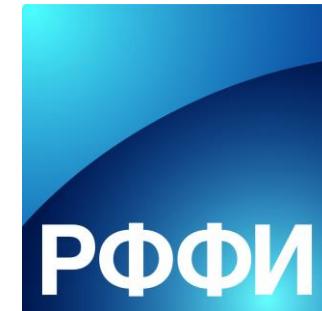
Руководитель рабочей группы научного совета ОМН РАН «Математическое моделирование распространения эпидемий с учетом социальных, экономических и экологических процессов» (Лебедев С.Н. (Снежинск), Кабанихин С.И. (Новосибирск), Шайдуров В.В. (Красноярск), Шананин А.А. (МФТИ)).

Принимает участие в расчетах сценариев распространения COVID-19 в Республике Казахстан (проект ИРН АР09260317), Великобритании (грант РФФИ № 21-51-10003) и США (сотрудничество с Институтом моделирования инфекционных заболеваний, Сиэтл).

Принимает участие в расчетах сценариев распространения ко-инфекции туберкулеза и ВИЧ в регионах Российской Федерации (сотрудничество с ФГБУ “НМИЦ ФПИ” Минздрава России).

НАУЧНАЯ ДЕЯТЕЛЬНОСТЬ 2018-2022 гг.:

- Статьи – 19 (Scopus & WoS: Q1 – 4, Q2 – 9)
- Тезисы конференций – 21 (6 пленарных докладов)
- Свидетельства о регистрации программы для ЭВМ – 7
- Участие в грантах: 5 (руководитель) + 4 (исполнитель)



ИНДЕКС ХИРША (Scopus): h=7

Оглавление

1. Результаты, выносимые на защиту
 - 1.1. Идентифицируемость агентных, дифференциальных и стохастических моделей эпидемиологии с учетом экономических и социальных процессов
 - 1.2. Комбинированные алгоритмы решения прямых и обратных задач
 - 1.3. Комплекс программ моделирования и построения сценариев распространения эпидемии COVID-19 в Новосибирской области
2. Обзор и анализ математических моделей эпидемиологии, участвующие в комбинированной модели
 - 2.1. Модели, основанные на анализе временных рядов
 - 2.2. Дифференциальные модели
 - 2.3. Агентно-ориентированные модели
3. Идентифицируемость математических моделей
 - 3.1. Дифференциальные модели (структурная и практическая идентифицируемость, анализ чувствительности)
 - 3.2. Чувствительность стохастических моделей
4. Алгоритмы решения обратных задач
 - 4.1. Локальная оптимизация – градиентные методы
 - 4.2. Глобальная оптимизация (эволюционные алгоритмы, тензорная оптимизация, байесовские подходы)
5. Комплекс программ
6. Численная реализация на примере моделирования распространения COVID-19 в Новосибирской области

Результаты, выносимые на защиту

1. Теоретические результаты

Впервые изучена идентифицируемость агентных, дифференциальных и стохастических моделей эпидемиологии с учетом экономических и социальных процессов. На этой основе построены новые градиентные методы решения обратных задач с использованием априорной информации.

2. Разработанные алгоритмы

Впервые разработаны комбинированные алгоритмы численного решения прямых и обратных задач, в которых агентные, дифференциальные и стохастические модели взаимозависимы, и включают в себя градиентные, тензорную оптимизацию и природоподобные алгоритмы (дифференциальной эволюции, генетический алгоритм, методы имитации отжига, роя частиц и древовидных оценок Парзена).

3. Комплекс программ

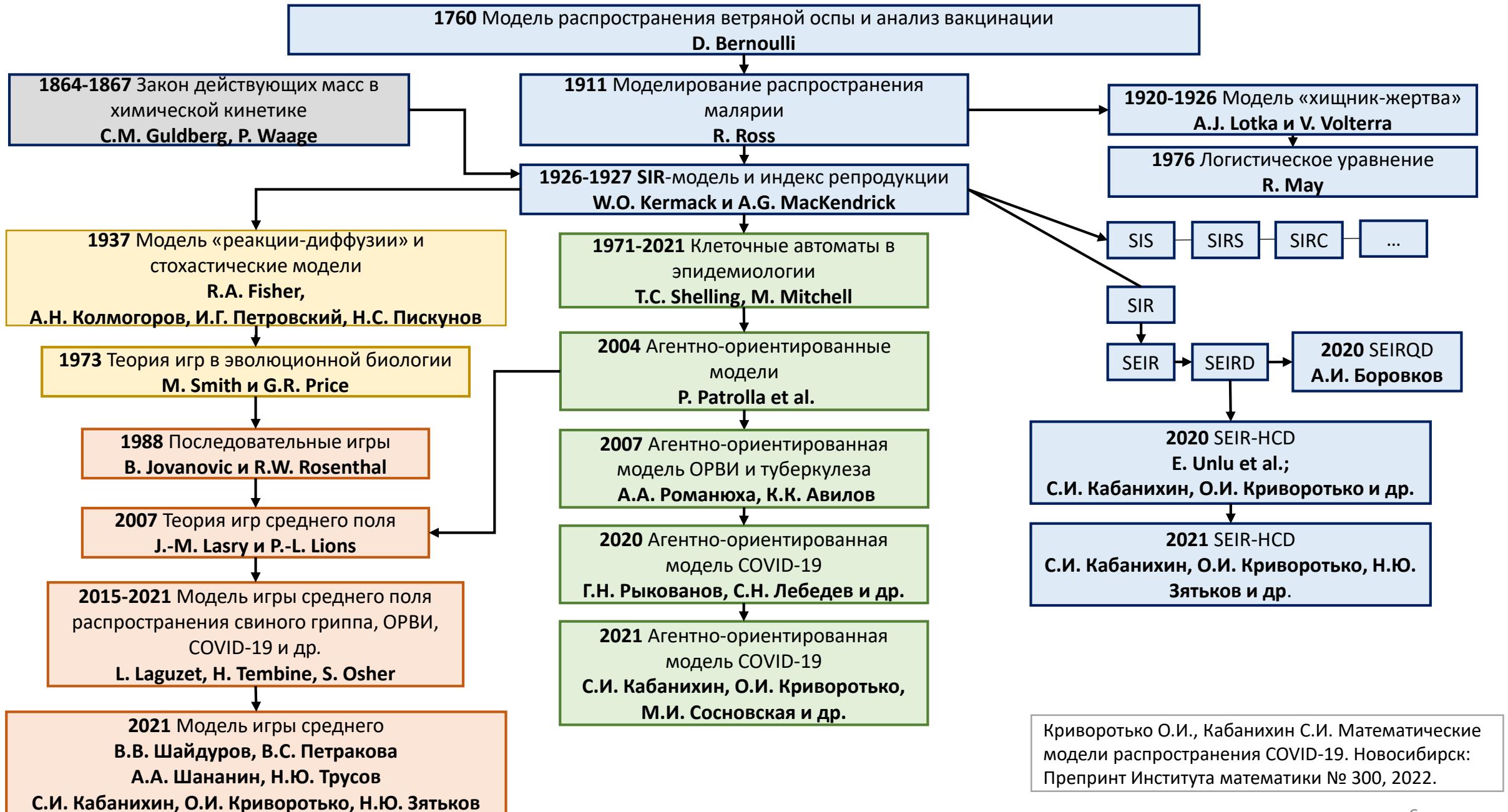
Впервые разработан комплекс программ моделирования и построения сценариев развития COVID-19 в Новосибирской области (<https://covid19-modeling.ru/>) с учетом эпидемической обстановки в прилегающих регионах, а также наиболее интенсивных транспортных потоков Новосибирск-Москва, Алтай, Казахстан, Красноярск, Омск.

4. Практическая значимость

Проведен анализ влияния ограничительных и фармацевтических мер на 45 дней на скорость распространения эпидемии для принятия стратегических решений.

Оглавление

1. Результаты, выносимые на защиту
 - 1.1. Идентифицируемость агентных, дифференциальных и стохастических моделей эпидемиологии с учетом экономических и социальных процессов
 - 1.2. Комбинированные алгоритмы решения прямых и обратных задач
 - 1.3. Комплекс программ моделирования и построения сценариев распространения эпидемии COVID-19 в Новосибирской области
2. Обзор и анализ математических моделей эпидемиологии, участвующие в комбинированной модели
 - 2.1. Модели, основанные на анализе временных рядов
 - 2.2. Дифференциальные модели
 - 2.3. Агентно-ориентированные модели
3. Идентифицируемость математических моделей
 - 3.1. Дифференциальные модели (структурная и практическая идентифицируемость, анализ чувствительности)
 - 3.2. Чувствительность стохастических моделей
4. Алгоритмы решения обратных задач
 - 4.1. Локальная оптимизация – градиентные методы
 - 4.2. Глобальная оптимизация (эволюционные алгоритмы, тензорная оптимизация, байесовские подходы)
5. Комплекс программ
6. Численная реализация на примере моделирования распространения COVID-19 в Новосибирской области



Математические модели эпидемиологии: два основных подхода

Камерные модели

Моделирование сверху-вниз

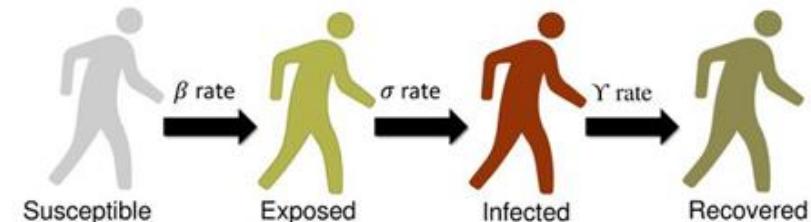
Детерминистские

Стохастические

Популяция разделена на группы со схожими признаками

SEIR Model

model the flows of people between four states: susceptible (S), exposed (E), infected (I), and recovered (R).

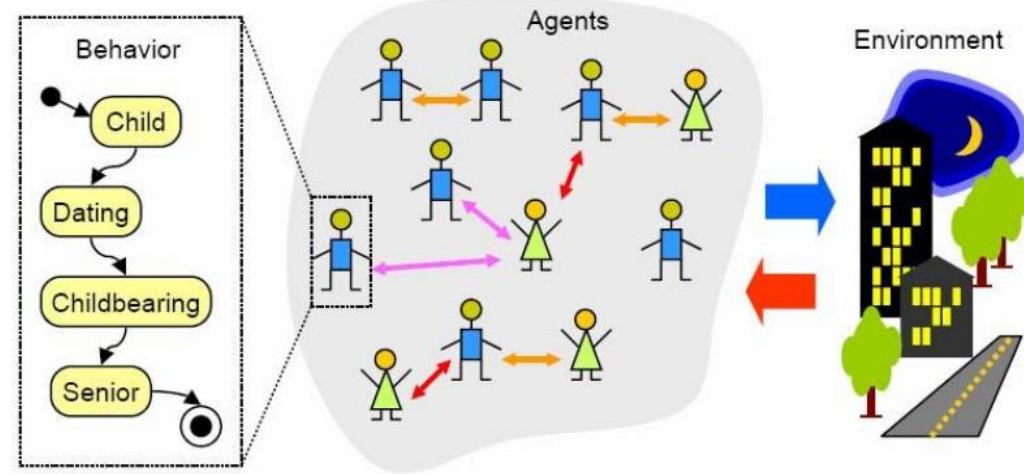


Закон действующих масс:

$$N = S(t) + E(t) + I(t) + R(t)$$
$$\frac{dN}{dt} = 0 \quad \forall t > 0.$$

Агентные модели

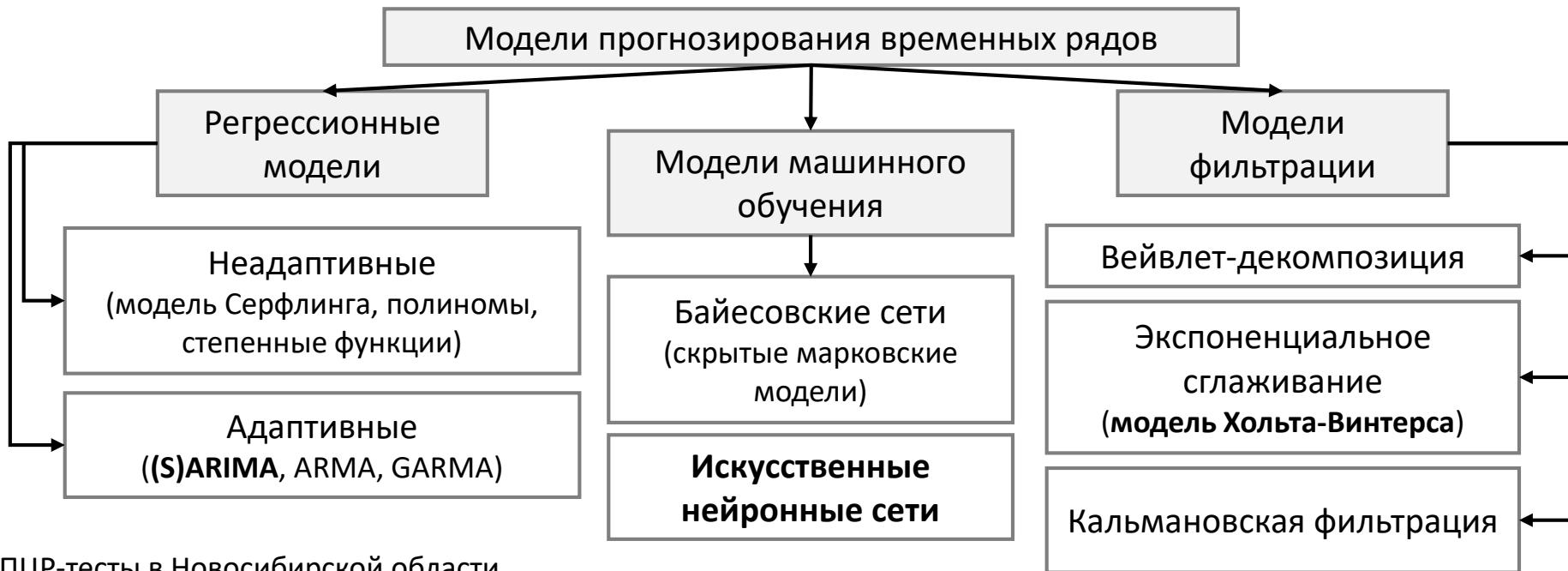
Моделирование снизу-вверх



Метод имитационного моделирования, исследующий поведение децентрализованных агентов и то, как такое поведение определяет поведение всей системы в целом. Агентное моделирование включает в себя клеточные автоматы, элементы теории игр, сложных систем, мультиагентных систем и эволюционного программирования, методы Монте-Карло.

Каждый агент обладает набором характеристик, а способ взаимодействия агентов из случайных графов определяется особенностями распространения заболевания.

2.1. Модели, основанные на анализе временных рядов



$T(t)$ - ПЦР-тесты в Новосибирской области
(ряд обладает "сезонностью" в 7 дней)



Сезонно-адаптивная

Линейная
регрессия

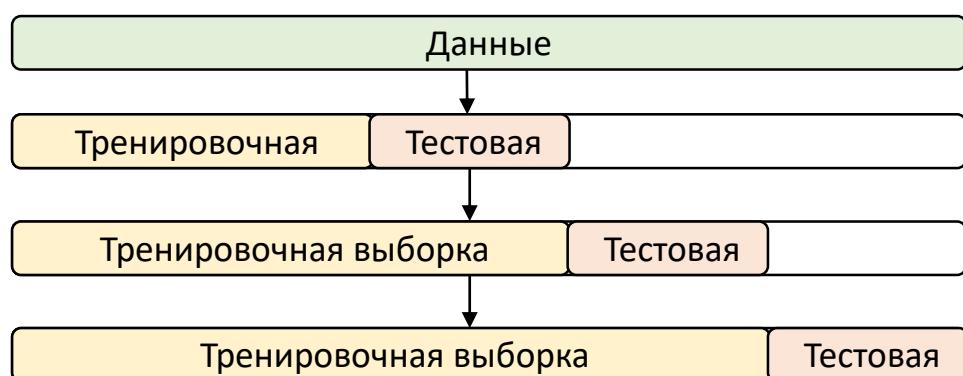
Тройное
экспоненциальное
сглаживание

$$\Delta^2 T(n) = c + a\Delta^2 T(n-1) + \sum_{j=1}^2 b_j \epsilon(n-j) + \epsilon(n).$$

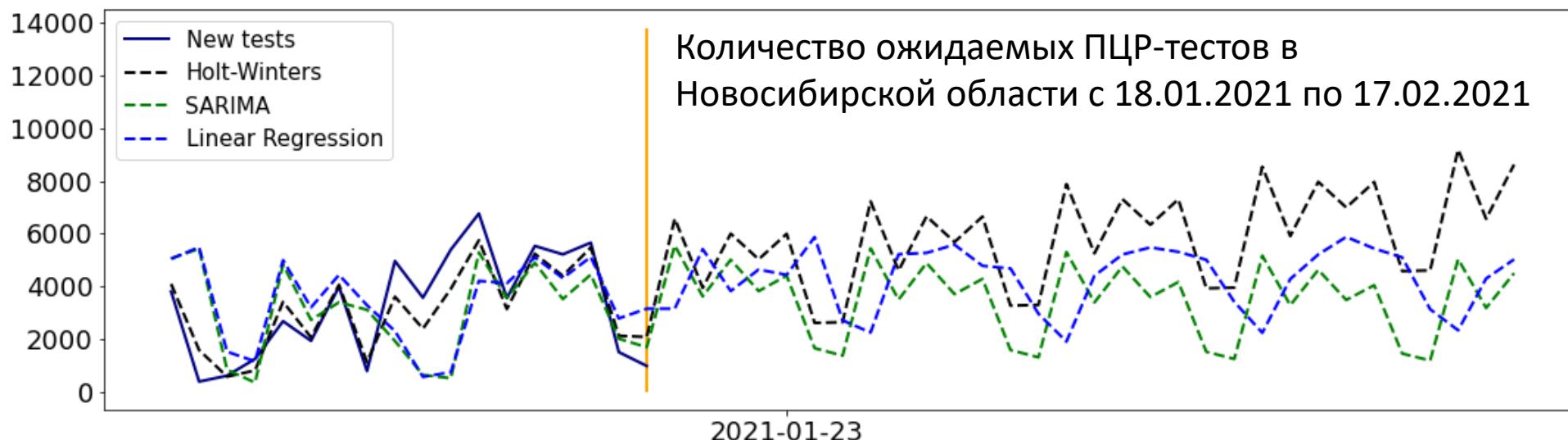
$$\tilde{T}(t) = b_0 + b_1 t + \epsilon(n).$$

$$\begin{aligned} l_x &= b(T_x - s_{x-L}) + (1-b)(l_{x-1} + g_{x-1}), \\ g_x &= c(l_x - l_{x-1}) + (1-c)g_{x-1}, \\ s_x &= d(T_x - l_x) + (1-d)s_{x-L}, \\ \hat{T}_{x+m} &= l_x + mg_x + s_{x-L+1+(m-1)\bmod L}. \end{aligned}$$

2.1. Модели, основанные на анализе временных рядов. Совмещение алгоритмов усвоения данных и машинного обучения (метод кросс-валидации и алгоритм экстраполяции временного ряда)



1. Выбираем тренировочный набор данных (часть временного ряда), на котором обучаем модель.
2. Делаем прогноз на n шагов вперед и вычисляем погрешность (квадратичное отклонение реальных данных на тестовой выборке).
3. Расширяем обучающую выборку до $t+n$ и снова строим прогноз на $t+2n$, на котором минимизируем квадратичную ошибку.
4. Продолжаем до тех пор, пока не исчерпаем всю длину временного ряда.



Экстраполяция сезонных временных рядов используется для построения сценариев распространения исследуемого инфекционного заболевания после уточнения параметров моделей.

2.2. SEIR-HCD моделирование вспышки COVID-19

$$\left\{ \begin{array}{l} \frac{dS}{dt} = -\frac{5-a(t-\tau)}{5} \left(\alpha_E(t) \frac{S(t)E(t)}{N} + \alpha_I(t) \frac{S(t)I(t)}{N} \right) + \frac{1}{t_{imm}} R(t), \\ \frac{dE}{dt} = \frac{5-a(t-\tau)}{5} \left(\alpha_E(t) \frac{S(t)E(t)}{N} + \alpha_I(t) \frac{S(t)I(t)}{N} \right) - \frac{1}{t_{inc}} E(t), \\ \frac{dI}{dt} = \frac{1}{t_{inc}} E(t) - \frac{1}{t_{inf}} I(t), \\ \frac{dR}{dt} = \frac{\beta(t)}{t_{inf}} I(t) + \frac{1-\varepsilon_{HC}(t)}{t_{hosp}} H(t) - \frac{1}{t_{imm}} R(t), \\ \frac{dH}{dt} = \frac{1-\beta(t)}{t_{inf}} I(t) + \frac{1-\mu(t)}{t_{crit}} C(t) - \frac{1}{t_{hosp}} H(t), \\ \frac{dC}{dt} = \frac{\varepsilon_{HC}(t)}{t_{hosp}} H(t) - \frac{1}{t_{crit}} C(t), \\ \frac{dD}{dt} = \frac{\mu(t)}{t_{crit}} C(t). \end{array} \right.$$

Прямая задача: найти вектор (S, E, I, R, H, C, D) по начальным данным $S(0) = S_0, E(0) = E_0, I(0) = I_0, R(0) = R_0, H(0) = H_0, C(0) = C_0, D(0) = D_0$.

Обратная задача: найти $\mathbf{q}(t) = (\alpha_E(t), \alpha_I(t), \varepsilon_{HC}(t), \mu(t), E_0, I_0)$ по данным $\mathbf{I}(\mathbf{t}_k) = (1 - b_k)\mathbf{f}_k, \mathbf{C}(\mathbf{t}_k) = \mathbf{C}_k, \mathbf{D}(\mathbf{t}_k) = \mathbf{g}_k$, $\beta(t)$ – концентрация антител IgG с учетом доп.ограничений.



Свойства SEIR-HCD модели

Базовый индекс репродукции вируса (начало эпидемии):

$$\mathcal{R}_0(t) = \frac{5 - a(t - \tau)}{5} \left(\alpha_E t_{inc} + \alpha_I t_{inf} \frac{1 + \varepsilon_{HC}(1 - \mu)}{1 - \varepsilon_{HC}(1 - \mu)} \right)$$

Преимущество SIR-моделей

- Простота реализации.
- Основа: закон сохранения масс с установленными взаимосвязями с различными состояниями системы.
- Описывает большие популяции (регионы, страны).
- Фундаментальность (для моделирования распространения другой инфекции в другом регионе достаточно уточнить параметры).

Недостатки SIR-моделей

- Описывает только 1 пик.
- Изменение параметров (мутации вируса, ограничительные меры, вакцинация) приводит к необходимости заново решать обратную задачу и рассчитывать сценарии с новыми параметрами.

Параметр	Описание	Границы
$a(t)$	Индекс самоизоляции (данные Яндекса)	(0, 5)
$\alpha_E(t)$	Параметр заражения между бессимптомной $E(t)$ и восприимчивой $S(t)$ группами населения ($\alpha_E \gg \alpha_I$)	(0, 1)
$\alpha_I(t)$	Параметр заражения между инфицированным $I(t)$ и восприимчивым $S(t)$ населением	(0, 1)
$\beta(t)$	Доля инфицированных, имеющая антитела IgG к SARS-CoV-2	(0, 1)
$\varepsilon_{HC}(t)$	Доля госпитализированных случаев $H(t)$, которым требуется подключение ИВЛ	(0, 1)
$\mu(t)$	Доля смертельных случаев	(0, 0.5)
τ	Латентный период	2 дня
t_{inc}	Длительность инкубационного периода	2-14 дней
t_{inf}	Длительность периода инфицирования	2,5-14 дней
t_{hosp}	Длительность периода госпитализации	4-5 дней
t_{crit}	Длительность использования ИВЛ	10-20 дней
t_{imm}	Средняя продолжительность гуморального иммунитета после выздоровления	180 дней

Модель «реакции-диффузии» (1937)

$$\frac{\partial u}{\partial t} = G(u) + d \frac{\partial^2 u}{\partial x^2}$$

$$u(x, 0) = u_0(x)$$

$u(x, t)$ - вектор плотности распределения групп в популяции,
 d - коэффициент диффузии,
 $G(u) = u(1 - u)$ - функция, характеризующая характер распространения заболевания в популяции,
удовлетворяющая закону сохранения масс и условиям:

$$G(0) = G(1) = 0, \quad G(u) > 0, \text{ если } 0 < u < 1,$$

$$G'(0) > 0, \quad G'(u) < G'(0), \text{ если } 0 < u \leq 1.$$

Динамика популяции по переменной t описывается скоростью $v^* = 2\sqrt{G'(0)d}$.



Колмогоров А.Н.
25.04.1903-20.10.1987



Петровский И.Г.
18.01.1901-15.01.1973

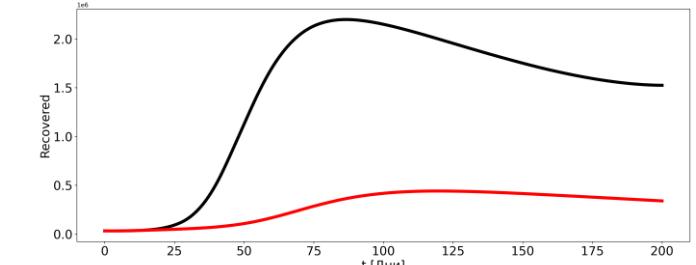
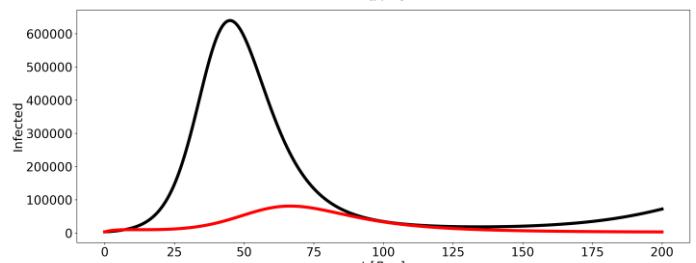
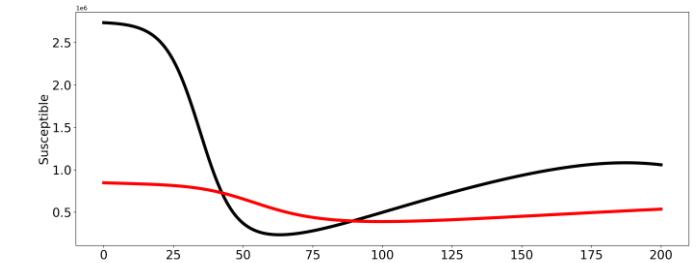
Введем в SEIR-HCD модель пространственные функции

$$S(t) = \int_{\Omega} s(t, x) dx, I(t) = \int_{\Omega} i(t, x) dx, R(t) = \int_{\Omega} r(t, x) dx, \dots$$

$$\begin{aligned} \frac{du}{dt} &= G(t, u, \mathbf{q}), & u(t_0) &= u_0. \longrightarrow \frac{du}{dt} = G(t, u, \mathbf{q}) + \nabla(N \mathbf{v} \nabla u), \\ u &= (S, E, I, R, H, C, D) & u(x, t_0) &= u_0(x), \quad u_x(0, t) = u(1, t) = 0. \\ & & \mathbf{v} &= (\mathbf{v}_s, \mathbf{v}_e, \mathbf{v}_i, \mathbf{v}_r, 0, 0, 0). \end{aligned}$$

Учет пространственной неоднородности позволяет более точно моделировать эпидемию от очага распространения (крупного города в стране, столицы в регионе и пр.) при известных начальных условиях. Однако использование модели для описания второй и последующих волн эпидемии требует добавления уравнений, введения множественных известных источников распространения заболевания $u(x, 0)$ (обратная задача определения источника) и вычислительных ресурсов.

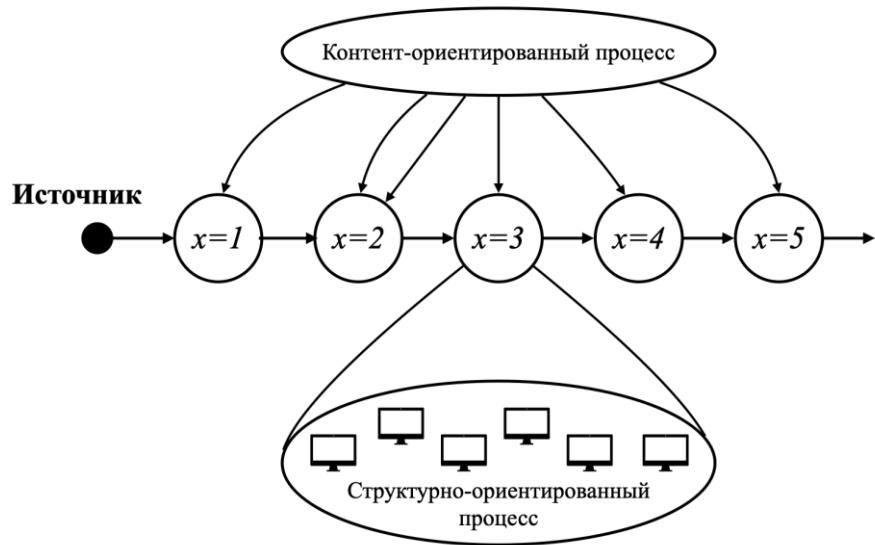
- Viguerie A., Veneziani A., Lorenzo G. et al. Diffusion-reaction compartmental models formulated in a continuum mechanics framework: application to COVID-19, mathematical analysis, and numerical study. Comput. Mech., 2020, V. 66, p. 1131-1152.
- Aristov V.V., Stroganov A.V., Yastrebov A.D. Simulation of spatial spread of the COVID-19 pandemic on the basis of the kinetic-advection model. Physics, 2021, V. 3(1), p. 85-102.



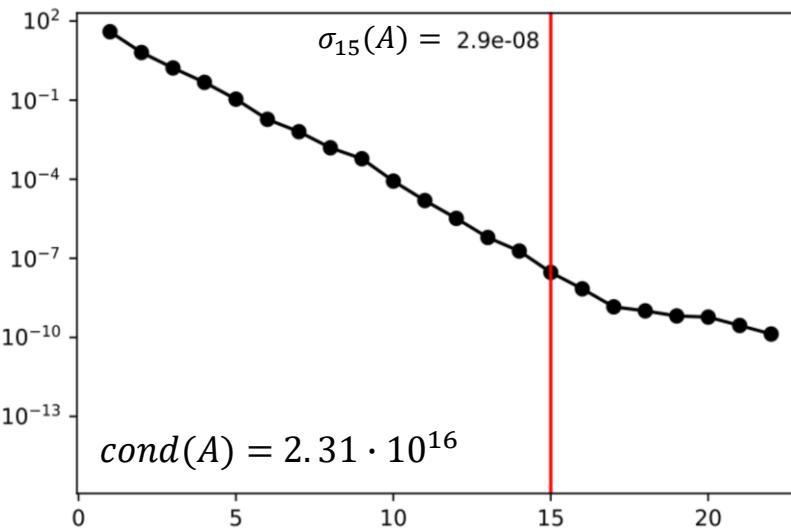
Обратная задача для диффузионно-логистической модели распространения информации в онлайн социальных сетях

$$\begin{cases} \frac{\partial u}{\partial t} = d \frac{\partial^2 u}{\partial x^2} + G(u, t) \\ u(x, 1) = \varphi(x), \quad l_1 \leq x \leq l_2 \\ \frac{\partial u}{\partial x}(l_1, t) = \frac{\partial u}{\partial x}(l_2, t) = 0, \quad t \geq 1 \end{cases} + \sum_{i=1}^{N_1} u(x_i, t_k) = f_k, \quad k = 1, \dots, N_2$$

Обратная задача: найти $\varphi(x)$ при заданных d и $G(u, t)$.



$u(x, t)$	плотность активных пользователей
d	популярность информации
$G(u, t) = \left(1 - \frac{u}{K}\right)r(t)u$	Описывает поведение популяции, численность которой не превышает значения K
r	скорость роста числа активных пользователей
K	пропускная способность
$\varphi(x)$	начальная функция плотности



Сингулярные числа матрицы линеаризованной обратной задачи $Aq=f$
 $A \in R^{22 \times 50}, \varphi \in R^{50}, f \in R^{22}$

Необходимо применять регуляризацию

- Wang F., Wang H., Xu K., Wu J., Jia X. Characterizing Information Diffusion in Online Social Networks with Linear Diffusive Model // IEEE Proceedings of ICDCS. 2013. P. 307-316.
- Звонарева Т.А., Криворотко О.И. Сравнительный анализ градиентных методов определения источника диффузионно-логистической модели. Журнал вычислительной математики и математической физики, 2022, т. 62, № 4, стр. 694-704.

SIR-модели, марковские процессы и стохастические дифференциальные уравнения

1. Время и состояния системы дискретны

$$S(t), I(t), R(t) \in \{0, 1, 2, \dots, N\}, \quad t \in \{0, \Delta t, 2\Delta t, \dots\}$$

Нелинейный марковский процесс, однородный по времени

$$\begin{aligned} p_{(s,i)}(t + \Delta t) = & p_{(s+1,i-1)}(t) \frac{\alpha}{N} (i-1)(s+1)\Delta t + \\ & + p_{(s,i+1)}(t) \beta(i+1)\Delta t + p_{(s,i)}(t) \left(1 - \frac{\alpha s}{N} - \beta i\right) \Delta t \end{aligned}$$

$$\Delta t \rightarrow 0$$

2. Состояния системы дискретны, а время – непрерывно

$$S(t), I(t), R(t) \in \{0, 1, 2, \dots, N\}, \quad t \in (0, T)$$

Прямое уравнение Колмогорова

$$\begin{aligned} \frac{dp_{(s,i)}}{dt} = & p_{(s+1,i-1)} \frac{\alpha}{N} (i-1)(s+1) + p_{(s,i+1)} \beta(i+1) + \\ & + p_{(s,i)} \left(1 - \frac{\alpha s}{N} - \beta i\right) \end{aligned}$$

Эндемическое равновесие не достигается, поскольку существует конечная вероятность того, что число инфицированных людей в системе упадет ниже единицы.

3. Время и состояния системы непрерывны

$$S(t), I(t), R(t) \in (0, N), \quad t \in (0, T)$$

Система СДУ

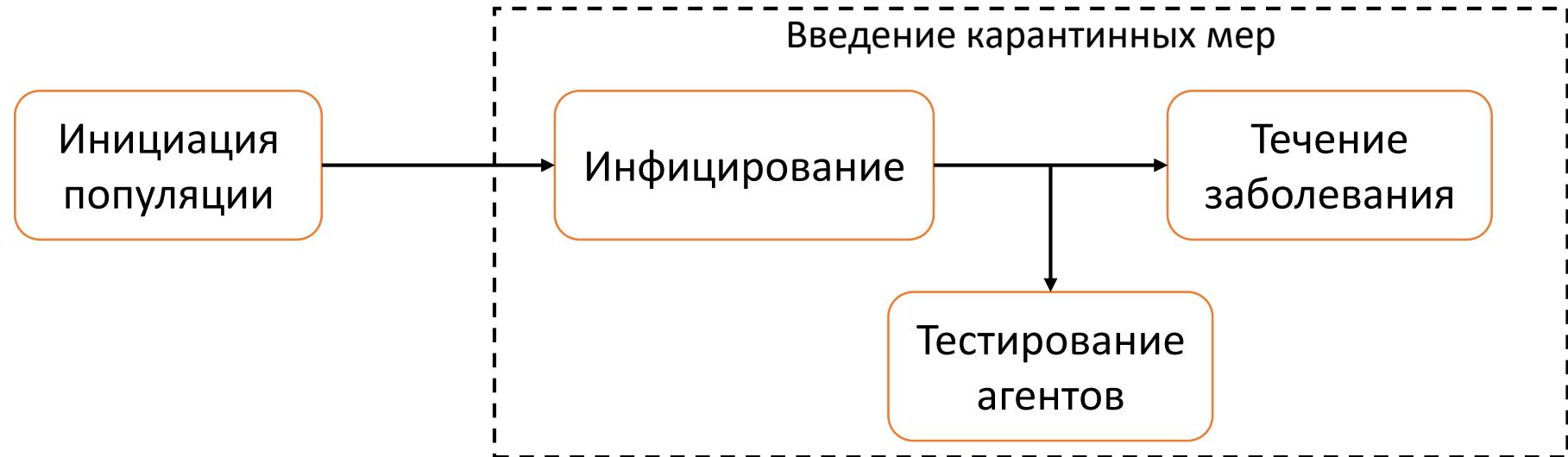
$$\begin{cases} dS = -\frac{\alpha}{N} S I dt + B_{11} dW_1 + B_{12} dW_2, \\ dI = \left(\frac{\alpha}{N} S I - \beta I\right) dt + B_{21} dW_1 + B_{22} dW_2, \end{cases}$$

$$S, I \sim N(\mu, \sigma^2)$$

$$B = \sqrt{V}, V = \begin{pmatrix} \frac{\alpha}{N} S I & -\frac{\alpha}{N} S I \\ -\frac{\alpha}{N} S I & \frac{\alpha}{N} S I + \beta I \end{pmatrix}$$

W_1 и W_2 - независимые винеровские процессы

2.3. Агентно-ориентированная модель распространения COVID-19



Person A	
-uid	23928
-age	55.1
-dead	0
-susceptible	1
-infected	0
-diagnosed	0
-...	
-date_infected	NaN
-date_diagnosed	NaN

Person B	
-uid	41135
-age	13.5
-dead	0
-susceptible	0
-infected	1
-diagnosed	0
-...	
-date_infected	44
-date_diagnosed	NaN

Person C	
-uid	76851
-age	83.2
-dead	1
-susceptible	0
-infected	1
-diagnosed	1
-...	
-date_infected	46
-date_diagnosed	53

Инициация популяции [1]

- Агенты разбиваются на возрастные группы по 10 лет [2];
- Популяция разделяется на 4 слоя контактов (домохозяйства, образовательные учреждения, рабочие и общественные места), которые заполняются агентами (рис. А). Среднее количество человек в семье в РФ по данным ООН [3] – 2,7.

В зависимости от возраста агенты контактируют друг с другом в контактных сетях, которые представлены **полносвязными** графами, количество вершин которых является пуассоновской случайной величиной с средним (рис. В):

- для домохозяйства – *размер семьи*,
- для общественных мест и образовательных учреждений – 20,
- для работы – 8.

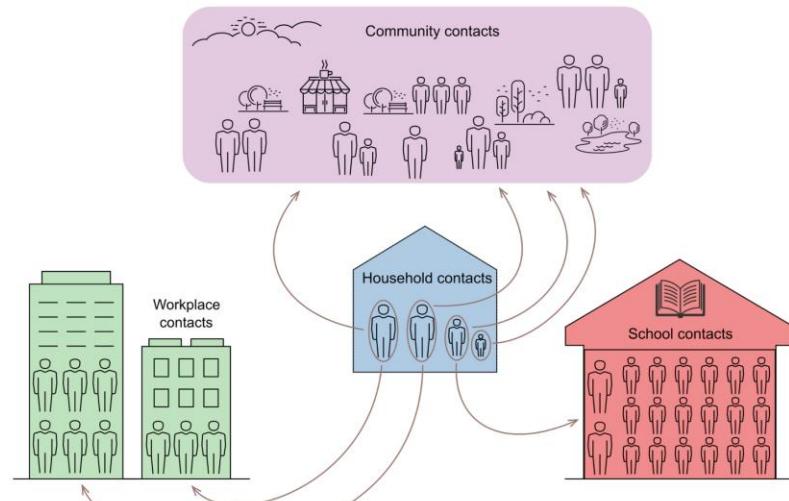
Агенты в возрасте 6-21 лет могут контактировать в образовательных учреждениях с агентами своего возраста, агенты в возрасте 22-65 лет – на работе.

[1] Kerr C. et al. Covasim: an agent-based model of COVID-19 dynamics and interventions. Medverix. 2020. DOI: 10.1101/2020.05.10.20097469.

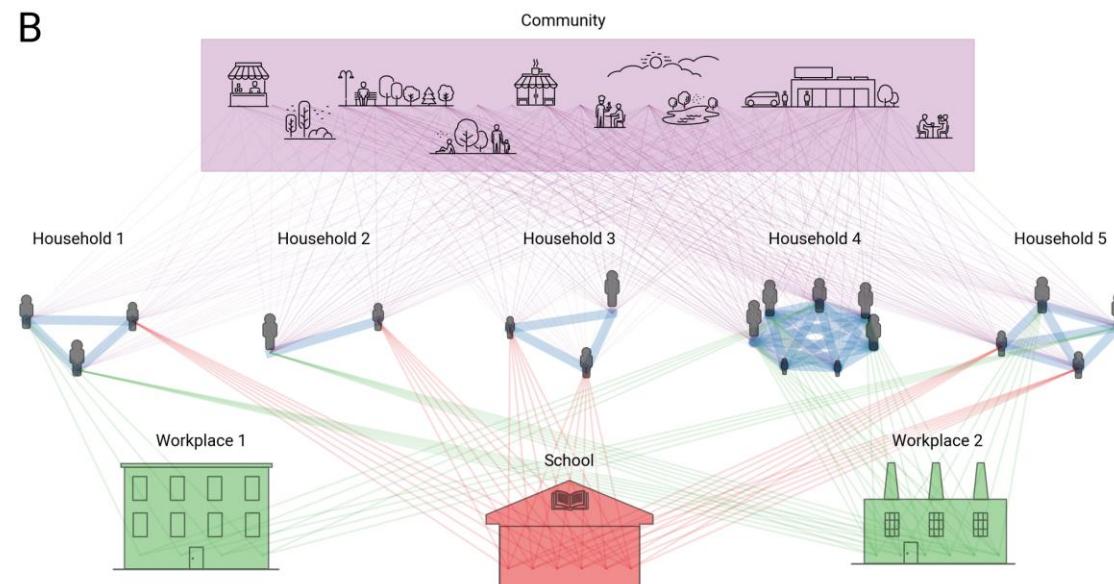
[2] Федеральная служба государственной статистики, Новосибирская обл.
<https://novosibstat.gks.ru/folder/31729>

[3] Household Size <https://population.un.org/Household/\#/countries/840>

А



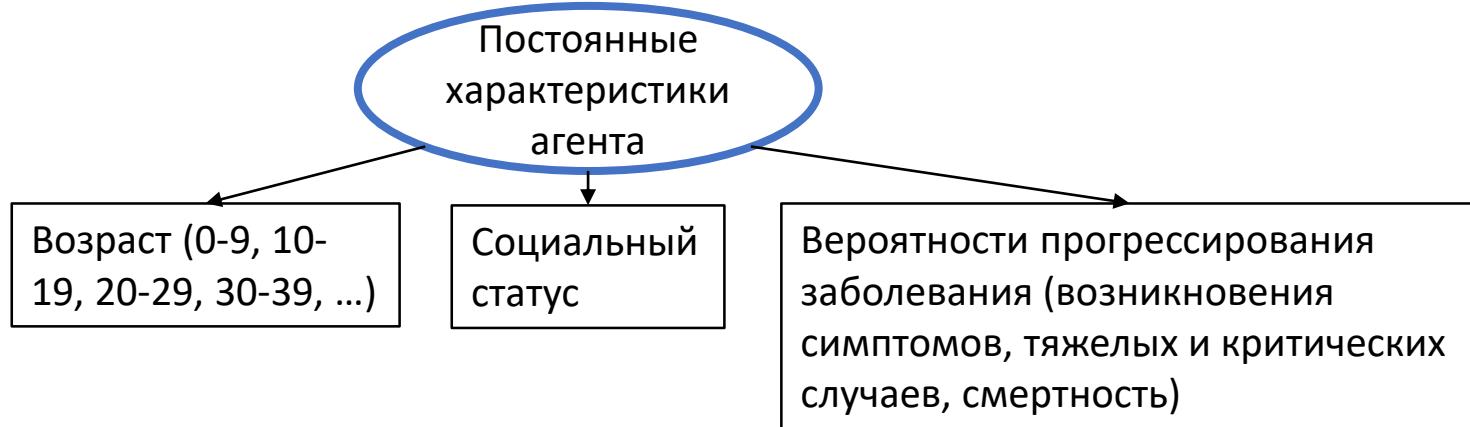
В



Инфицирование

В рамках модели предполагается, что вирус передается между агентами, которые соединены между собой ребром в графе. За передачу вируса при близком контакте с зараженным отвечает параметр β и является кусочно-постоянным. В зависимости от структуры контакта, параметр β умножается на соответствующую константу (см. таблицу). Симптомные и бессимптомные агенты передают вирус одинаково.

Таким образом, вероятность передачи вируса для каждой контактной сети случайная и различная.



Вероятность заражения зависит от пребывания агента в семье, на работе, на учебе, в публичных местах.

Структура	Скорость заражения
Домохозяйства	$\beta \cdot 3.0$
Школы и университеты (6-22 года)	$\beta \cdot 0.6$
Работа (22-65 лет)	$\beta \cdot 0.6$
Публичные места	$\beta \cdot 0.3$

Течение заболевания и тестирование

Каждый агент может находиться в 9 состояниях заболевания. Переход из одной стадии заболевания в другую контролируется параметрами p_{sym} , p_{sev} , p_{crit} , p_{death} , зависящими от возрастной группы агента. Агенты старшего возраста более уязвимы для ранних штаммов вируса (данные параметры оцениваются из статистических данных).

Тестирование проводится согласно ежедневным статистическим данным по количеству проведенных тестов в регионе. Шанс быть протестированным на COVID-19 $p = \frac{p_{sym}}{p_{asym}}$, зависит от эпидемиологического статуса агента. Здесь p_{sym} — вероятность тестирования для больных с симптомами, p_{asym} — вероятность тестирования для больных без симптомов.

Параметр p восстанавливается в ходе решения обратной задачи.

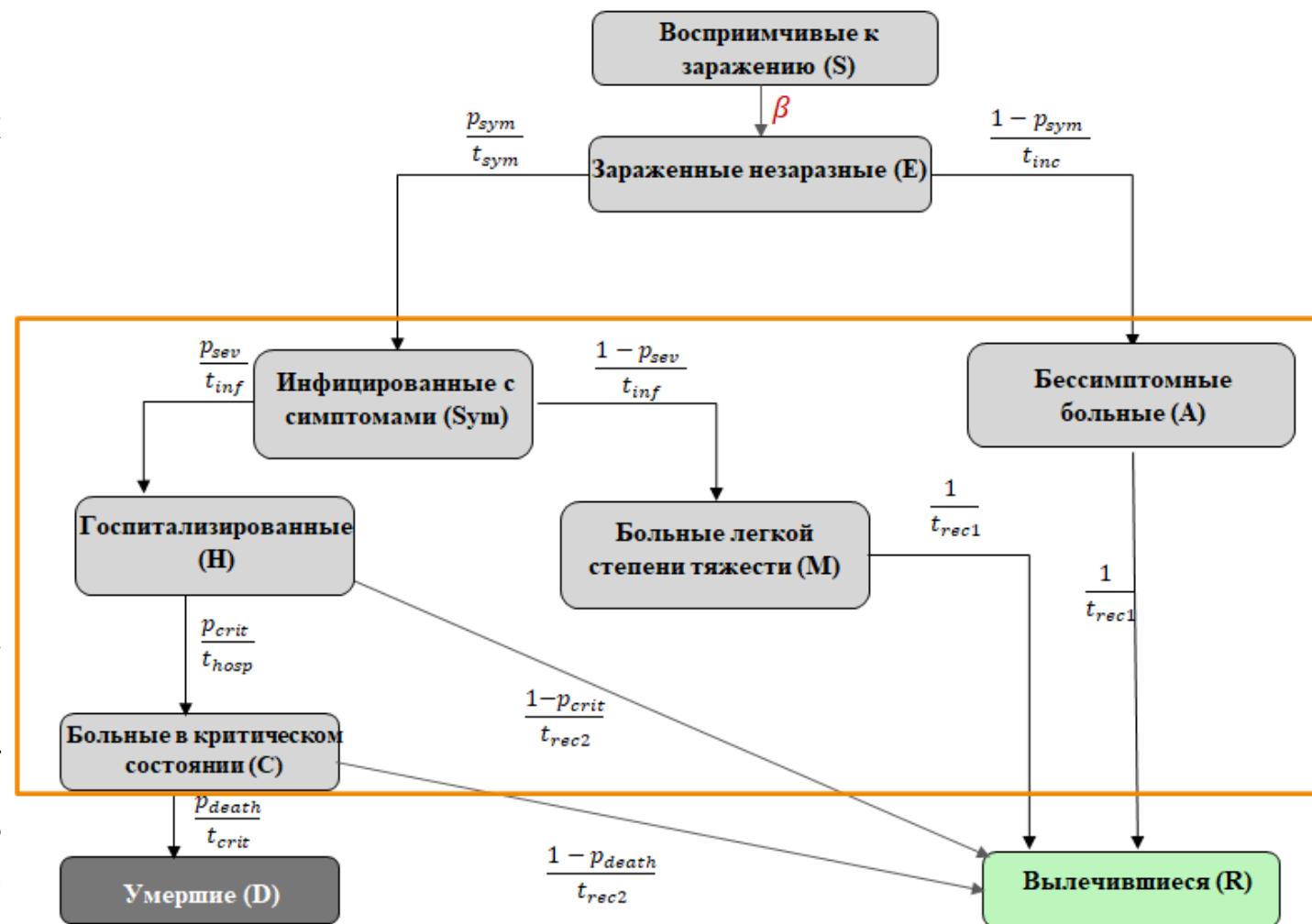
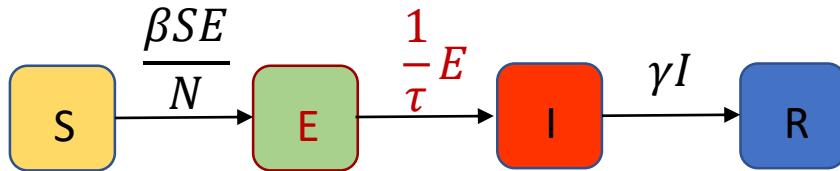


Диаграмма состояний агентов в модели Covasim. Оранжевой рамкой обозначены те состояния, находясь в которых агент имеет возможность получить положительный тест на COVID-19.

Длительность стадий заболевания

Случайная логнормальная величина с фиксированными параметрами среднего и дисперсии: $f(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-(\ln x - \mu)^2/2\sigma^2}$

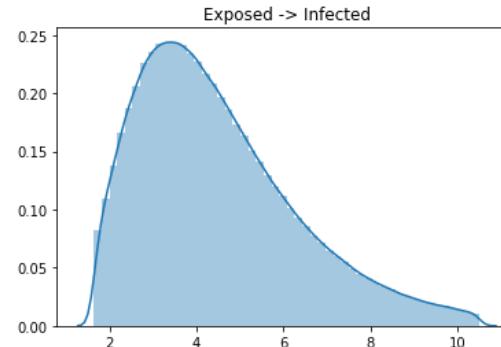
В простейшем случае для SEIR модели



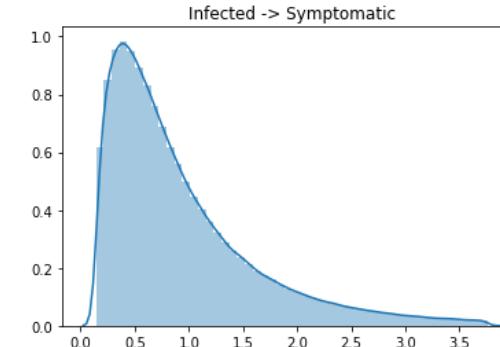
$$\frac{dE}{dt} = \frac{\beta SE}{N} - \frac{1}{\tau} E.$$

$$E(t) + \tau \frac{dE}{dt} = \frac{\beta \tau SE}{N} \approx E(t + \tau).$$

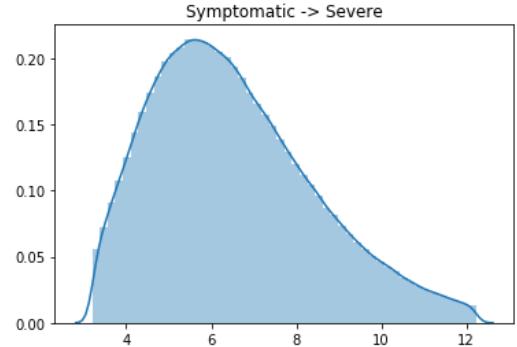
Вероятность $\frac{1}{\tau}$ перехода из группы E в группу I характеризует обратную величину к среднему времени τ пребывания индивидуума в группе E (в линейном приближении).



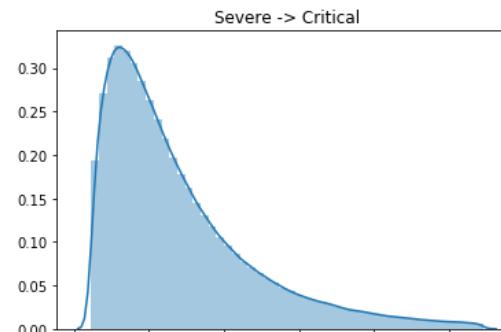
$$t_{inc} \sim \text{LogN}(4.6, 4.8)$$



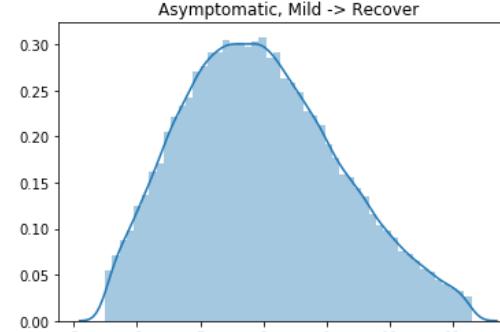
$$t_{sym} \sim \text{LogN}(1.0, 0.9)$$



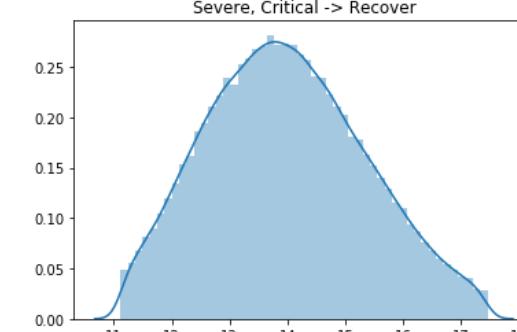
$$t_{inf} \sim \text{LogN}(6.6, 4.9)$$



$$t_{hosp} \sim \text{LogN}(3.0, 7.4)$$



$$t_{rec1} \sim \text{LogN}(8.0, 2.0)$$



$$t_{rec2} \sim \text{LogN}(14.0, 2.4)$$

Течение заболевания: пример заражения восприимчивого агента

Пусть контакт между двумя индивидуумами происходит в домохозяйстве.

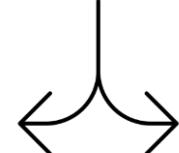
Восприимчивый к
заражению (S)



Инфицированный с
симптомами (Sym)

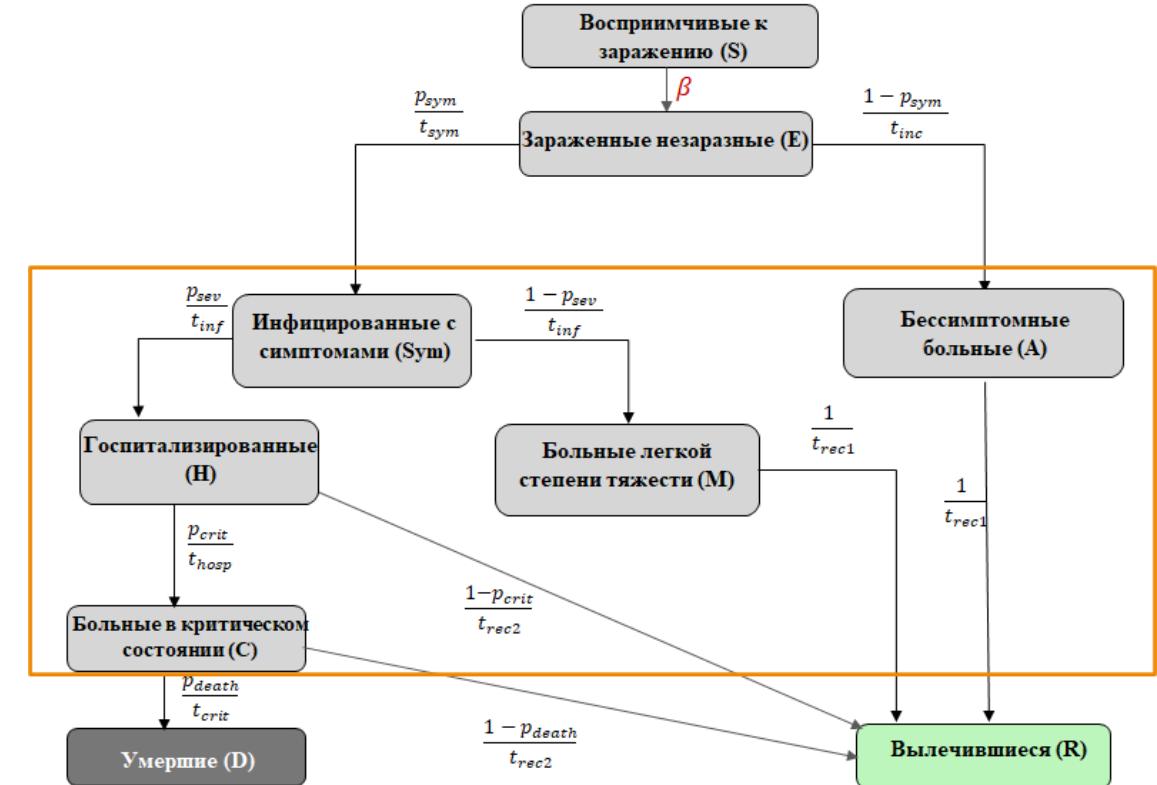


$$3 \cdot \beta$$



$$p(Sym) = 3 \cdot \beta \cdot \frac{p_{sym}}{t_{sym}}$$

$$p(A) = 3 \cdot \beta \cdot \frac{1 - p_{sym}}{t_{inc}}$$



Разыгрывается случайная величина γ . Если $\gamma > p_{sym}$, то инфицированный становится симптомным больным через t_{sym} дней после контакта, иначе переходит в бессимптомного носителя инфекции через t_{inc} дней.

Введение карантинных мер

В рамках модели возможно ввести ограничительные меры глобально для всех слоев или для каждого по отдельности. Это можно сделать двумя способами:

- изменить значение параметра β (характеризующее уменьшение передачи вируса например при соблюдении масочного режима, соблюдения социальной дистанции);
- удалить ребра в графах (характеризующее самоизоляцию, переход на дистанционное обучение или работу).

В нашей модели используются оба этих подхода.

Предположения и ограничения модели

- Изначально иммунитета нет ни у одного агента;
- Есть начальная доля инфицированных в инкубационном состоянии, количество которых характеризуется параметром $E(0)$;
- Умереть могут только агенты, которые находятся в реанимации (больные в критическом состоянии);
- После выздоровления агент не может заразиться вновь в течение 4-5 месяцев;
- Состояние агента обновляется в конце дня с учетом усреднения возможного пребывания в разных структурах контактов (например, 8 часов на работе, 4 часа в общественном месте и 12 часов дома за весь моделируемый день).

Обратная задача для агентной модели

Определить вектор параметров $q(t) = (E(0), \beta, \beta_d, \beta_c, p)$ по дополнительной информации о количестве **ежедневно проведенных ПЦР-тестов, выявленных и умерших случаев от COVID-19** в конкретном регионе, а также усредненных статистических показателей для региона (размер семьи, средняя продолжительность состояния стадии заболевания и т.п.).

β_d - день измерения параметра β , β_c - значение, на которое изменится параметр β ,
 p - параметр уровня тестирования.

Глава 2. Выводы

История заболевания (временные ряды эпидемиологических данных) составляет обучающие множества для методов машинного обучения.

SIR модели действительны в случае закрытой системы (для регионов с небольшим притоком населения, удаленных районов) и в случае, когда население практически бесконечно (например, континенты, страны).

В стохастических SIR моделях эндемическое равновесие не достигается, поскольку существует конечная вероятность того, что число инфицированных особей в системе упадет ниже единицы. Таким образом, в данной системе патоген может не размножаться.

В SIR моделях среднего поля число зараженных агентов может принимать нецелые значения, и число агентов в модели может быть меньше единицы, что позволяет патогену размножаться в модели.

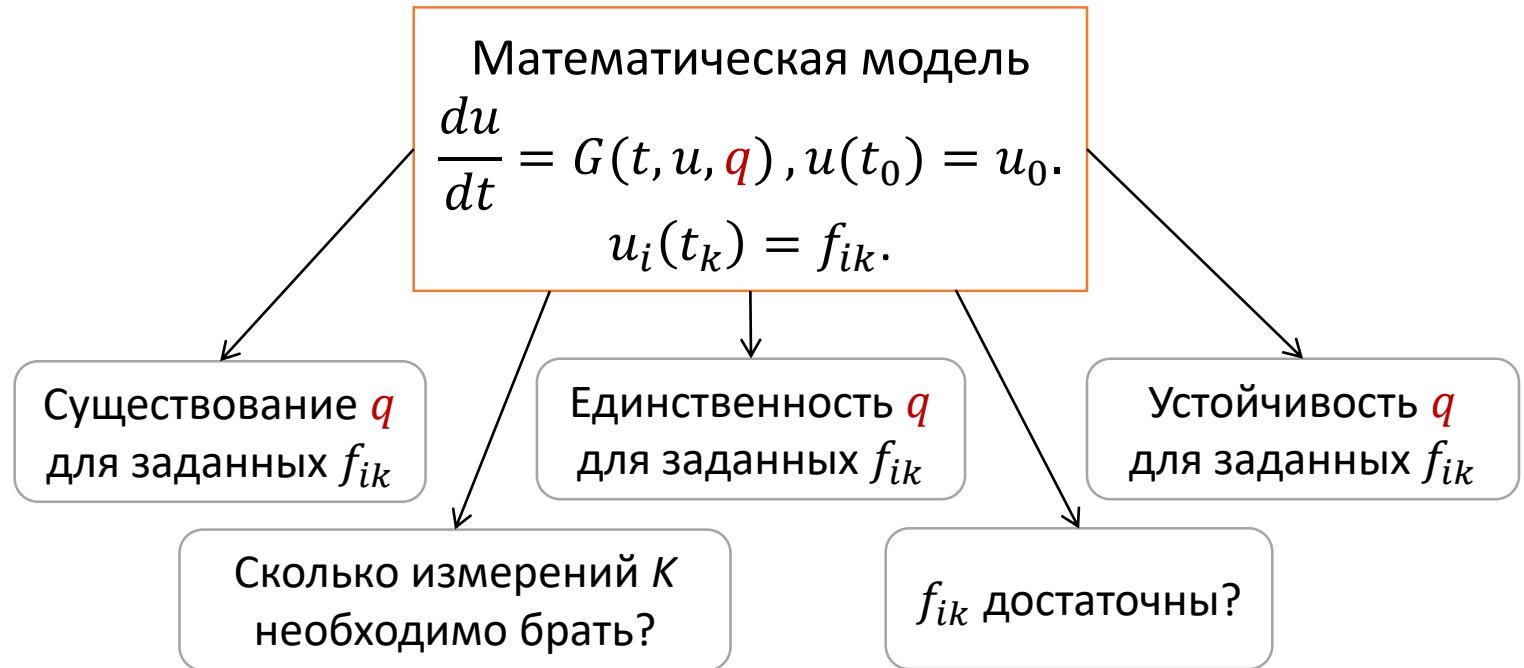
Сложность применения АОМ возрастает с увеличением моделируемой популяции и учетом внешних воздействий на систему. Усреднение в течение 1 дня моделирования в пределе приводит к SIR модели.

При усложнении SIR-модели, вводя возрастные различия в популяции и пространственные перемещения, мы получим первое приближение АОМ.

Оглавление

1. Результаты, выносимые на защиту
 - 1.1. Идентифицируемость агентных, дифференциальных и стохастических моделей эпидемиологии с учетом экономических и социальных процессов
 - 1.2. Комбинированные алгоритмы решения прямых и обратных задач
 - 1.3. Комплекс программ моделирования и построения сценариев распространения эпидемии COVID-19 в Новосибирской области
2. Обзор и анализ математических моделей эпидемиологии, участвующие в комбинированной модели
 - 2.1. Модели, основанные на анализе временных рядов
 - 2.2. Дифференциальные модели
 - 2.3. Агентно-ориентированные модели
3. Идентифицируемость математических моделей
 - 3.1. Дифференциальные модели (структурная и практическая идентифицируемость, анализ чувствительности)
 - 3.2. Чувствительность стохастических моделей
4. Алгоритмы решения обратных задач
 - 4.1. Локальная оптимизация – градиентные методы
 - 4.2. Глобальная оптимизация (эволюционные алгоритмы, тензорная оптимизация, байесовские подходы)
5. Комплекс программ
6. Численная реализация на примере моделирования распространения COVID-19 в Новосибирской области

3.1. Идентифицируемость дифференциальных моделей



Анализ идентифицируемости - это группа методов математической статистики, которые используются для определения того, насколько хорошо параметры модели оцениваются по количеству и качеству экспериментальных данных.

Идентифицируемость математических моделей для ОДУ

Структурная

- Зависит от структуры модели
- Измерения гладкие

Методы

- Передаточной функции
- Дифференциальной алгебры
- Разложение в ряд Тейлора

Какая комбинация параметров q является идентифицируемой по заданным измерениям f_{ik} ?
Единственность решения обратной задачи.

Практическая

- Зависит от уровня шума в данных
- Необходима аппроксимация параметров

Методы

- Монте-Карло
- Корреляционной матрицы

Какое минимальное число измерений K необходимо для устойчивого решения обратной задачи?

Анализ чувствительности

- Не использует реальные данные
- Необходима аппроксимация параметров
- Зависит от количества измерений

Методы

- Собственных значений
- Ортогональный



$$s_{ij}(t) = \frac{\partial G_i(t, u, q_{ex})}{\partial q_j}$$

1. Криворотко О.И., Кабанихин С.И., Сосновская М.И., Андорная Д.В. Анализ чувствительности и идентифицируемости математических моделей распространения эпидемии COVID-19. Вавиловский журнал генетики и селекции, 2021, т. 25(1), стр. 82-91.
2. Криворотко О.И., Андорная Д.В., Кабанихин С.И. Анализ чувствительности и практическая идентифицируемость математических моделей биологии. Сибирский журнал индустриальной математики, 2020, т. 23, № 1, стр. 107-125.

Анализ чувствительности SEIR-HCD модели (параметры постоянны)

Модель ОДУ:

$$\frac{du}{dt} = G(t, u, \mathbf{q}), u(t_0) = \mathbf{u}_0.$$

Здесь $u = (S, E, I, R, H, C, D)^T$,
 $\mathbf{q} = (\alpha_E, \alpha_I, \beta, \varepsilon_{HC}, \mu, E_0)^T$

Дифференцируем по q_j

$$\begin{aligned} \frac{d}{dt} \left(\frac{\partial u(t)}{\partial q_j} \right) &= \frac{\partial G}{\partial u} \frac{\partial u(t)}{\partial q_j} + \frac{\partial \mu}{\partial q_j}, \\ \frac{\partial u(0)}{\partial q_j} &= \frac{\partial u_0}{\partial q_j}. \end{aligned}$$

Матрица чувствительности:

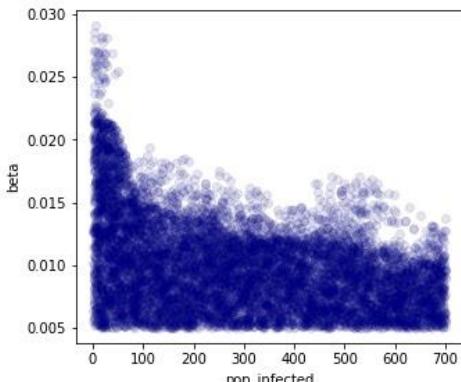
$$\frac{\partial y_l(t_k)}{\partial q_m}, l = 1, 2, 3; m = 1, \dots, 6; k = 1, \dots, K.$$

Здесь $y = (f_k, C_k, g_k)^T$ - данные о ежедневно выявленных, критических и умерших случаях.

y_i	q_j	$\left\ \frac{\partial y_i(t)}{\partial q_j} q_j^* \right\ $	y_i	q_j	$\left\ \frac{\partial y_i(t)}{\partial q_j} q_j^* \right\ $	y_i	q_j	$\left\ \frac{\partial y_i(t)}{\partial q_j} q_j^* \right\ $
f	α_E	$2.865 \cdot 10^{14}$	g	α_E	$7.110 \cdot 10^{11}$	g	μ	$2.982 \cdot 10^4$
f	α_I	$2.396 \cdot 10^{14}$	g	α_I	$5.947 \cdot 10^{11}$	g	β	$2.164 \cdot 10^4$
f	E_0	$1.854 \cdot 10^{14}$	g	E_0	$4.601 \cdot 10^{11}$	C	μ	$3.695 \cdot 10^2$
C	α_E	$2.386 \cdot 10^{12}$	C	ε_{HC}	$4.833 \cdot 10^4$	f	β	$2.03 \cdot 10^{-6}$
C	α_I	$1.996 \cdot 10^{12}$	C	β	$3.428 \cdot 10^4$	f	ε_{HC}	$1.6 \cdot 10^{-7}$
C	E_0	$1.544 \cdot 10^{12}$	g	ε_{HC}	$3.041 \cdot 10^4$	f	μ	$3 \cdot 10^{-8}$

С помощью
ортогонального метода
анализа матрицы
чувствительности
получена
последовательность
параметров от наиболее
идентифицируемого до
наименее:
 $\alpha_E, E_0, \alpha_I, \varepsilon_{HC}, \mu, \beta$

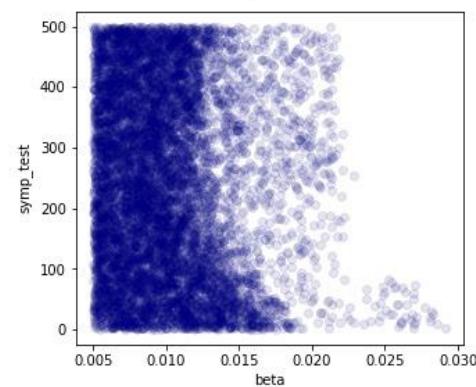
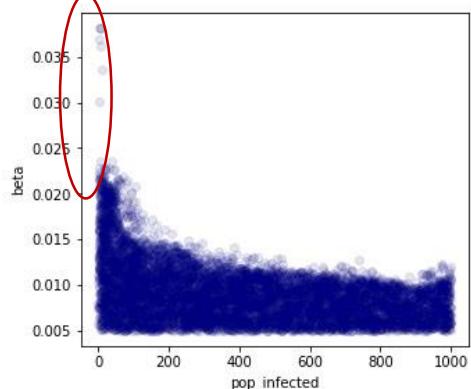
Анализ чувствительности для агентной модели распространения COVID-19



Данные аппроксимируются регрессионной моделью

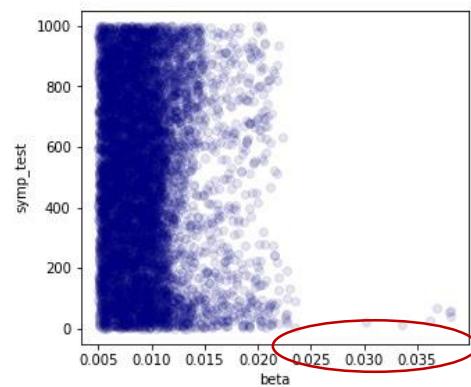
$$g_i(\mathbf{q}) = \sum_{j=1}^k h_j(\mathbf{q}) \mathbf{b}_j + u(\mathbf{q})$$

Где параметры регрессии определяются из задачи максимизации функции правдоподобия.

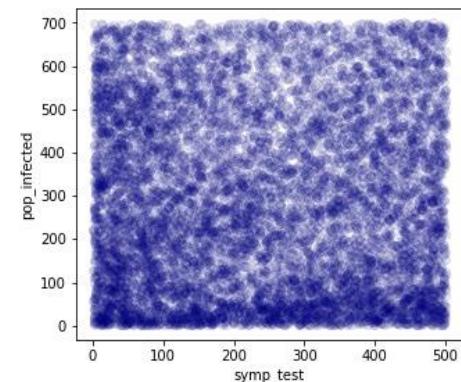


$\mathbf{q} = (\beta, E(0), p), f$ – выявленные и умершие случаи

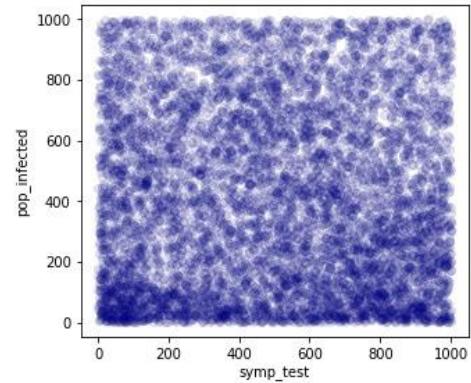
Параметр контагиозности β (скорость передачи инфекции) чувствителен к дополнительным измерениям f .



\bar{f} - выявленные, критические и умершие случаи



Добавление информации о критических случаях позволяет уточнить границы изменения параметра β . Количество исходных инфекционных случаев $E(0)$ и шанс быть протестированным p не зависели от используемых данных.



Глава 3. Выводы

1. Анализ чувствительности моделей SIR позволяет упорядочить параметры моделей в порядке их влияния на измеряемые данные.

2. Анализ чувствительности агентных моделей позволяет упорядочить параметры моделей в порядке их влияния на измеряемые данные, а также уточнить границы изменения этих параметров на основе анализа доступных данных.

Для SEIR-HCD модели распространения COVID-19:

Параметр, описывающий скорость выздоровления после инфицирования (гуморальный иммунитет) β наименее чувствителен к данным о ежедневно выявленных, критических и умерших случаях.

Для агентной модели распространения COVID-19:

Параметр, описывающий скорость передачи инфекции β наиболее чувствителен к данным о ежедневно выявленных, критических и умерших случаях. В результате удалось уменьшить границы его изменения в 2 раза по сравнению с рекомендацией ВОЗ.

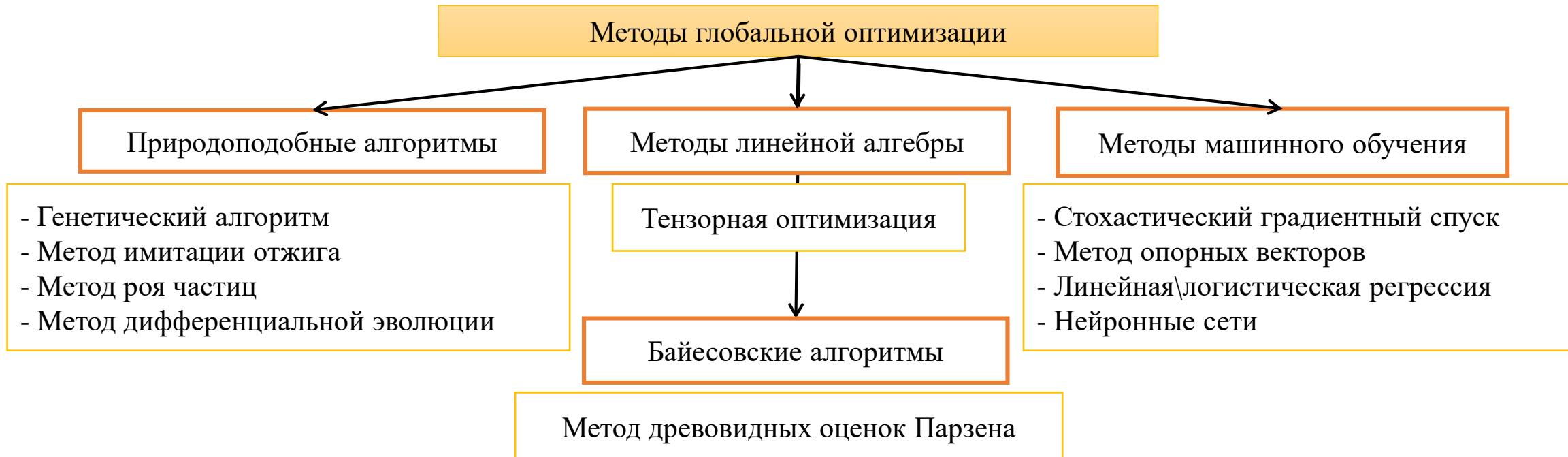
Оглавление

1. Результаты, выносимые на защиту
 - 1.1. Идентифицируемость агентных, дифференциальных и стохастических моделей эпидемиологии с учетом экономических и социальных процессов
 - 1.2. Комбинированные алгоритмы решения прямых и обратных задач
 - 1.3. Комплекс программ моделирования и построения сценариев распространения эпидемии COVID-19 в Новосибирской области
2. Обзор и анализ математических моделей эпидемиологии, участвующие в комбинированной модели
 - 2.1. Модели, основанные на анализе временных рядов
 - 2.2. Дифференциальные модели
 - 2.3. Агентно-ориентированные модели
3. Идентифицируемость математических моделей
 - 3.1. Дифференциальные модели (структурная и практическая идентифицируемость, анализ чувствительности)
 - 3.2. Чувствительность стохастических моделей
4. Алгоритмы решения обратных задач
 - 4.1. Локальная оптимизация – градиентные методы
 - 4.2. Глобальная оптимизация (эволюционные алгоритмы, тензорная оптимизация, байесовские подходы)
5. Комплекс программ
6. Численная реализация на примере моделирования распространения COVID-19 в Новосибирской области

4. Алгоритмы решения обратных задач

$$J(\mathbf{q}) = \sum_{k=1}^K \sum_s w_s \frac{|c_d^{i,s} - c_m^{i,s}(\mathbf{q})|^2}{(c_d^{s,t})^2} \rightarrow \min_{\mathbf{q}}$$

Методы глобальной оптимизации + Градиентные (локальные) методы = $q^* = \underset{\mathbf{q} \in Q}{\operatorname{argmin}} J(\mathbf{q})$



4.1. Градиентные методы: выражение градиента целевого функционала

ОДУ

$$\frac{du}{dt} = \mu(t, u, \mathbf{q}), u(t_0) = \mathbf{u}_0 \quad + \quad u_i(t_k) = \mathbf{f}_{ik}$$

$$(q_{n+1}, \mathbf{u}_{0_{n+1}}) = (q_n, \mathbf{u}_{0_n}) - \alpha_n J'(q_n, \mathbf{u}_{0_n}), \quad \alpha_n > 0$$

$$J(q, \mathbf{u}_0) = \sum_{k=1}^K \sum_i \omega_i (u_i(t_k; q, \mathbf{u}_0) - f_{ik})^2 \rightarrow \min_{q, \mathbf{u}_0}$$

$$J'(q, \mathbf{u}_0) = \left(- \int_{t_0}^T \Psi^T(t) \mu_q(u(t), q) dt, -\Psi(t_0) \right)$$

$$\begin{cases} \dot{\Psi} = -\mu_u^T(u(t), q)\Psi, \\ \Psi(T) = 0, \\ [\Psi_i]_{t=t_k} = 2\omega_i(u_i(t_k; \mathbf{q}) - \mathbf{f}_{ik}), k = 1, \dots, K, \\ \quad i \in I \subset \{1, \dots, N\}. \end{cases}$$

S.I. Kabanikhin, O.I. Krivorotko. Identification of biological models described by systems of nonlinear differential equations. J. Inverse Ill-Posed Probl. 2015. V. 23(5). P. 519-527.

УЧПР

$$\begin{cases} \frac{\partial u}{\partial t} = d \frac{\partial^2 u}{\partial x^2} + g(u, \mathbf{q}), \\ u(x, t_0) = \mathbf{u}_0(x) \end{cases} \quad + \quad \sum_i u(x_i, t_k) = \mathbf{f}_k$$

$$J(q, \mathbf{u}_0) = \sum_{i \in I} \sum_{k=0}^K \omega_i |u(x_i, t_k; q, \mathbf{u}_0) - f_{ik}|^2 \rightarrow \min_{q, \mathbf{u}_0}$$

$$J'(q, \mathbf{u}_0) = - \left(\int_{t_0}^T \int_{\Omega_l} \Psi^T \mu_q dt dx, -\Psi(x, t_0) \right)^T$$

$$\begin{cases} \frac{\partial \Psi}{\partial t} = -\mu_u^T(u(x, t), q) - d \frac{\partial^2 \Psi}{\partial x^2}, t > 0, x \in \Omega, \\ \Psi(T) = 0, \frac{\partial \Psi}{\partial x} \Big|_{\partial \Omega_l} = 0, \\ [\Psi_i]_{t=t_k} = 2\omega_i \left(\sum_i u(x_i, t_k; \mathbf{q}) - \mathbf{f}_k \right), k = 1, \dots, K. \end{cases}$$

O. Krivorotko, S. Kabanikhin, Sh. Zhang, V. Kashtanova. Global and local optimization in identification of parabolic systems. J. Inverse Ill-Posed Probl. 2020. V. 28(6). P. 899-913.

4.1. Многоуровневый градиентный метод

$$q^{m+1} = \theta_{m+1} z^m + (1 - \theta_{m+1}) y^m, \quad \theta_{m+1} \in \operatorname{Arg} \min_{\theta \in [0,1]} J(\theta z^m + (1 - \theta) y^m),$$

$$y^{m+1} = q^{m+1} - \zeta_{m+1} J'(q^{m+1}), \quad \zeta_{m+1} \in \operatorname{Arg} \min_{\zeta \geq 0} J(q^{m+1} - \zeta J'(q^{m+1})),$$

$$z^{m+1} = z^m - \eta_{m+1} J'(q^{m+1}), \quad \eta_{m+1} = \frac{1}{2L_{m+1}} + \sqrt{\frac{1}{4L_{m+1}} + \eta_m^2}, \quad \eta_0 = 0, \quad L_{m+1} = \frac{\|J'(q^{m+1})\|^2}{2(J(q^{m+1}) - J(y^{m+1}))}.$$

$$\|J'(y) - J'(x)\|_2 \leq L \|y - x\|_2, \quad J(y^{m+1}) \leq J(q^m) - \frac{1}{2L} \|J'(q^m)\|_2^2$$

$$J(q^m) \leq \frac{4L \|q^0 - q_{ex}\|^2}{(m+1)^2}$$

Теорема. Пусть для решения задачи $f(x) \rightarrow \min_x f$, где $f(x)$ удовлетворяет условию Липшица, используется градиентный спуск:

$$x^{m+1} = x^m - \frac{1}{L} f'(x^m).$$

Тогда при условии существования решения x_*

$$\min_{m=1,\dots,N} \|f'(x^m)\|_2 \leq \sqrt{\frac{2L(f(x^0) - f(x_*))}{N}}.$$

Если $f(x)$ - μ -сильно выпуклая функция в 2-норме ($\mu \geq 0$), то

$$\min_{m=1,\dots,N} f(x^m) - f(x_*) \leq \frac{L \|x^0 - x_*\|_2^2}{2} \min \left\{ \frac{1}{N}, e^{-\frac{\mu}{L} N} \right\}.$$

4.2. Природоподобные алгоритмы

Стохастическая сходимость:

- А.А. Жиглявский, А.Г. Жилинскas. Методы поиска глобального экстремума. М.: Наука, 1991:

Теорема. Пусть $J(q)$ – непрерывная функция в области A , для которой существует точка глобального минимума q^* , и пусть для любого $\varepsilon > 0$ выполнены следующие условия:

$$\sum_{k=1}^{\infty} q_k(\varepsilon) = \infty, \quad \text{где } q_k(\varepsilon) = \inf_{Q_i, 1 \leq i \geq k} P_k(B(\varepsilon)),$$

где P_k - распределение вектора q' .

Тогда последовательность случайных векторов $q_1, q_2 \dots$, полученная в результате применения природоподобного алгоритма (генетическим, имитацией отжига, роя частиц, дифференциальной эволюции), попадает в множество $A(\varepsilon)$ с вероятностью 1 почти наверное.

Метод древовидных оценок Парзена

1. Выбираем J^* как квартиль γ , т.е. $p(J < J^*) = \gamma$ и моделируем $p(J)$.
2. Все остальные значения разделяются на 2 группы: те, которые доставляют значение функционала меньше $J^*(\mathbf{q}^i)$ и остальные.
Т.о. строятся плотности вероятности значений \mathbf{q}^i – это $l(\mathbf{q})$ и плотности остальных значений - $g(\mathbf{q})$. Получаем вероятность

$$p(\mathbf{q}|J) = \begin{cases} l(\mathbf{q}), & J < J^*, \\ g(\mathbf{q}), & J \geq J^*. \end{cases}$$

3. Используя формулу полной вероятности, задача оптимизации сводится к задаче максимизации $l(\mathbf{q})/g(\mathbf{q})$.

Сходимость: $E \ln \frac{|q_k - q^*|}{|q_{k-1} - q^*|} \rightarrow -c < 0, k \rightarrow \infty$.

Преимущества:

- ✓ Определяет область глобального экстремума
- ✓ Не чувствителен к начальному приближению
- ✓ Быстро сходится к области «точного» решения
- ✓ Простая структура (имитирует простейшую задачу физики, биологии)

Ограничения:

- Стохастическая сходимость
- Настройка параметров методов трудоемкая

4.3. Тензорная оптимизация

$$J(q) \rightarrow g_\alpha(q)$$

$$g_\alpha(q) = h(J(q) - \alpha)$$

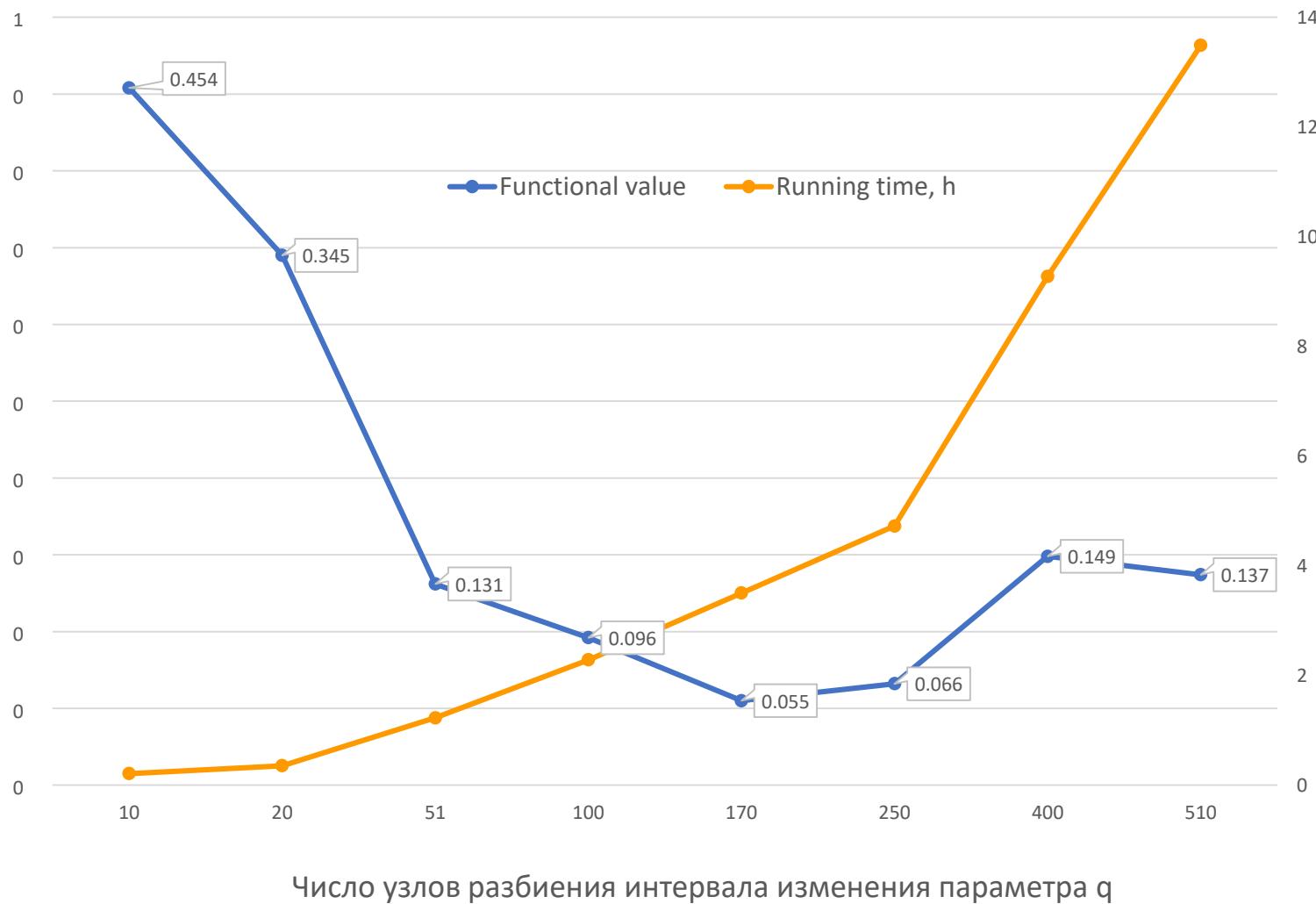
$h(a)$ – непрерывная монотонно убывающая функция, α – некоторое начальное приближение к минимуму функционала $J(q)$.

Исходная задача минимизации функционала $J(q)$, $q \in \mathbb{R}^d$ заменяется на эквивалентную задачу поиска максимума по модулю функции g .

Алгоритм:

1. Вводится сетка с n узлами по каждому из $d=6$ направлений, и тензор значений функции g на этой сетке обозначается через $A(i_1, \dots, i_d)$.
2. Затем составляются подматрицы $A(R_k, C_k)$ матриц разверток $A_k(i_1 \dots i_k, i_{k+1} \dots i_d) = A(i_1, \dots, i_d)$. Они приближаются с помощью матричного крестового метода.
3. Далее вычисляются пространственные координаты точек, соответствующих элементам матриц.
4. Найденные точки локально оптимизируются и добавляются в множество локальных оптимумов M .
5. Определяются ближайшие элементы матриц к элементам этого множества и строковые и столбцовые номера этих элементов записываются в множества R_k и C_k соответственно.
6. Обновляется сдвиг функционала $\alpha = \min_{q \in M} J(q)$.
7. Повторяем эти действия, пока не выполнится критерий остановки.

4.3. Вычислительная сложность тензорной оптимизации



Характеристики вычислительного узла в ССКЦ СО РАН:

2 процессора Intel Xeon Gold
6248R (3 GHz, 24 ядра)
384 GB оперативной памяти

При увеличении размерности тензора вычислительное время растет экспоненциально.

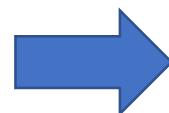
Глава 4. Выводы

1. Комбинация методов глобальной и локальной оптимизации позволяет уточнить минимум целевого функционала. А именно, методы глобальной оптимизации определяют область глобального минимума, а локальные методы уточняют решения задачи минимизации в указанной области.
2. Получены выражения для градиентов целевых функционалов для систем ОДУ и уравнений в частных производных параболического типа, связанные с решением соответствующих сопряженных задач.
3. При увеличении размерности вектора неизвестных параметров оптимально использовать метод тензорной оптимизации в качестве метода глобальной минимизации целевого функционала.

Методы, используемые при построении алгоритма решения обратных задач

Обработка статистических данных

- Методы регрессионного анализа для обработки данных.
- Методы машинного обучения (нейронные сети, линейная регрессия) для экстраполяции данных с целью построения прогноза.



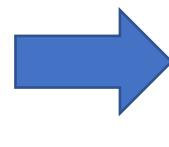
Получено свидетельство о государственной регистрации программы для ЭВМ №2021614740 от 29.03.2021

Анализ идентифицируемости

- Методы дифференциальной алгебры и разложения в ряды Тейлора анализа структурной идентифицируемости.
- Ортогональный метод и метод сингулярного разложения получения упорядоченной последовательности чувствительных параметров для систем Оду.
- Методы Монте-Карло и корреляционной матрицы анализа практической идентифицируемости.
- Байесовский подход к анализу чувствительности стохастических систем.



Использовано программное обеспечение DAISY (<https://daisy.dei.unipd.it/>)



Получено свидетельство о государственной регистрации программы для ЭВМ №2019661851 от 10.09.2019



Использована и адаптирована библиотека <https://github.com/alan-turing-institute/mogp-emulator>

Получены последовательности идентифицируемых параметров и оценки их областей для моделей внутриклеточной динамики ВИЧ, ко-инфекции туберкулеза и ВИЧ, распространения COVID-19.

Построение агентной модели

- Метод случайных графов с учетом особенностей COVID-19



Улучшена библиотека COVASIM
<https://github.com/InstituteforDiseaseModeling/covasim>

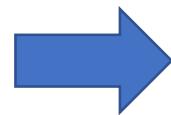
Получено свидетельство №2021614740 от 29.03.2021

Методы, используемые в представленных работах

Методы решения задач минимизации

Глобальные

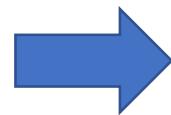
- Метод тензорной оптимизации (использована библиотека <https://github.com/oseledets/TT-Toolbox>).
- Генетический алгоритм, методы дифференциальной эволюции, имитации отжига, роя частиц (собственная реализация, а также использование библиотеки *SciPy.Optimize Python*).



Для модели эпидемии туберкулеза с возникновением лекарственно-устойчивых штаммов получено внедрение в учебный процесс ЕНУ им. Л.Н. Гумилева в 2020 году.

Локальные

- Градиентные методы – собственная реализация.
- Метод Нелдера-Мида (использована библиотека *SciPy.Optimize Python*).



Получено выражение для градиента целевого функционала в общем случае.
Получена оценка остаточного члена в разложении производной Фреше целевого функционала.

Синий цвет – результаты диссертационной работы.

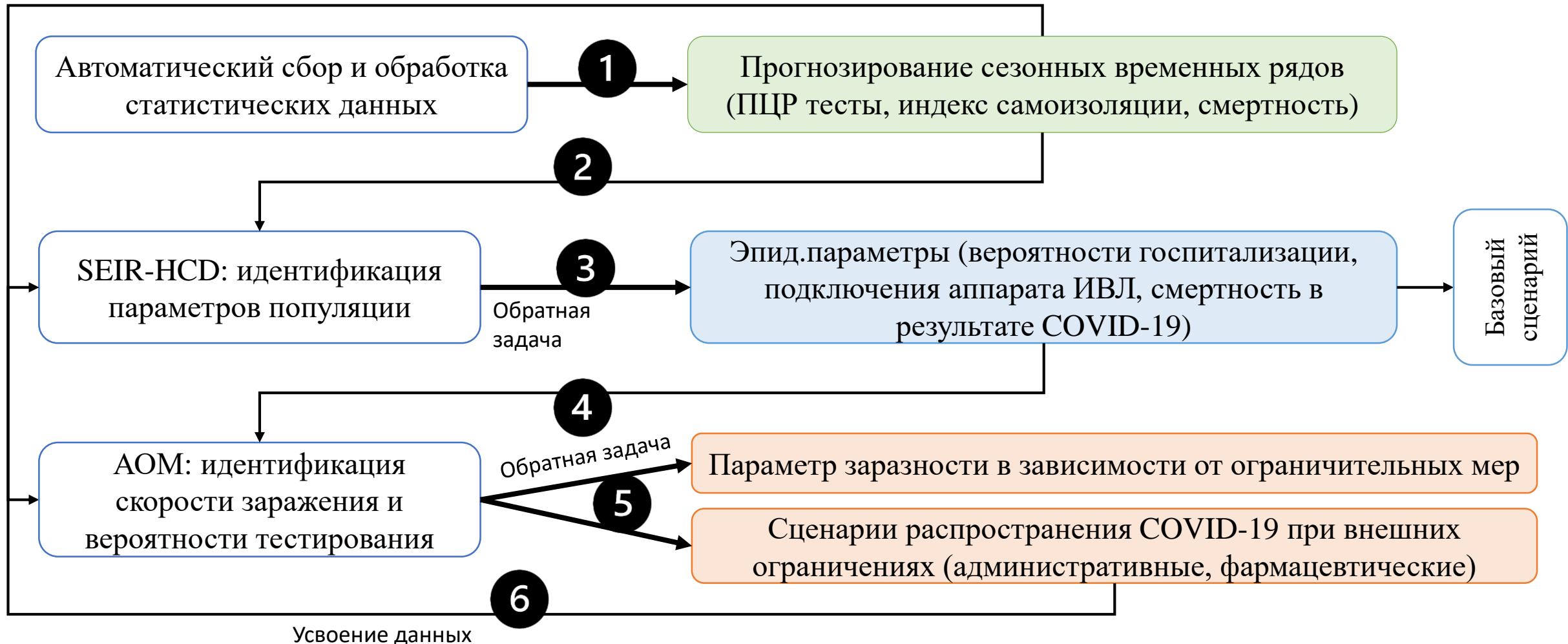
Оранжевый цвет – открытые программные пакеты.

Желтый цвет – улучшение открытых программных комплексов.

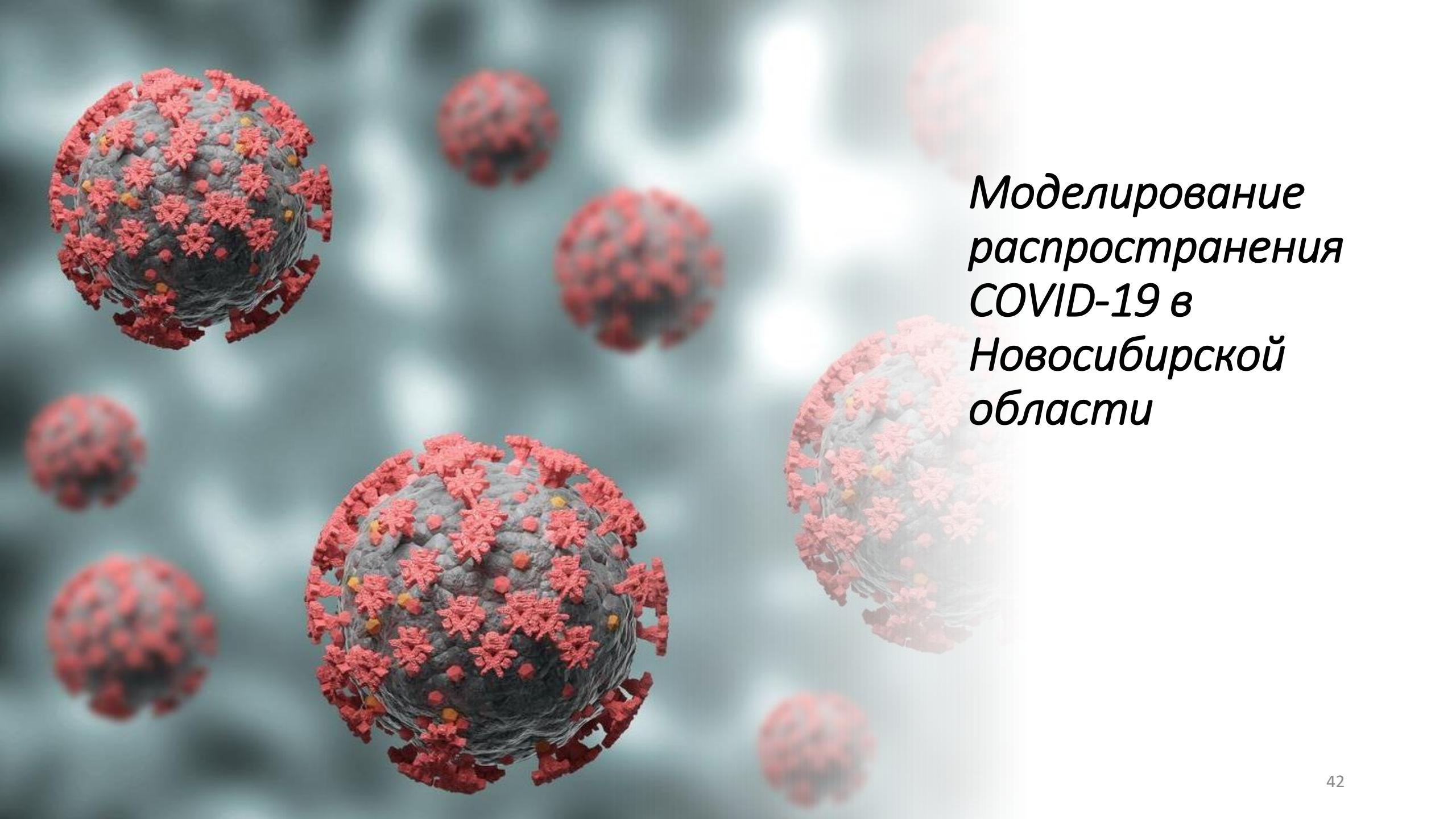
Оглавление

1. Результаты, выносимые на защиту
 - 1.1. Идентифицируемость агентных, дифференциальных и стохастических моделей эпидемиологии с учетом экономических и социальных процессов
 - 1.2. Комбинированные алгоритмы решения прямых и обратных задач
 - 1.3. Комплекс программ моделирования и построения сценариев распространения эпидемии COVID-19 в Новосибирской области
2. Обзор и анализ математических моделей эпидемиологии, участвующие в комбинированной модели
 - 2.1. Модели, основанные на анализе временных рядов
 - 2.2. Дифференциальные модели
 - 2.3. Агентно-ориентированные модели
3. Идентифицируемость математических моделей
 - 3.1. Дифференциальные модели (структурная и практическая идентифицируемость, анализ чувствительности)
 - 3.2. Чувствительность стохастических моделей
4. Алгоритмы решения обратных задач
 - 4.1. Локальная оптимизация – градиентные методы
 - 4.2. Глобальная оптимизация (эволюционные алгоритмы, тензорная оптимизация, байесовские подходы)
5. Комплекс программ
6. Численная реализация на примере моделирования распространения COVID-19 в Новосибирской области

Комбинированная модель распространения COVID-19: цикл 1 дня (<https://covid19-modeling.ru/>)



Ежедневный сценарий распространения COVID-19 обновляется на основе комбинации SIR и агентного подходов.

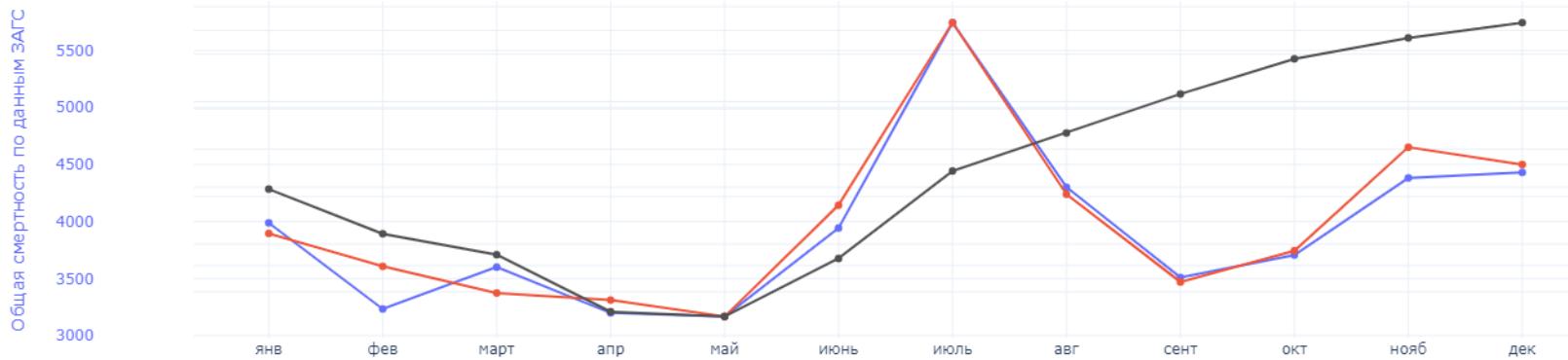


*Моделирование
распространения
COVID-19 в
Новосибирской
области*

Эпидемиологические показатели в Новосибирской области



Смертность в 2021



Черная линия – смертность от COVID-19, красная линия – захоронения больных, синяя линия – общая смертность

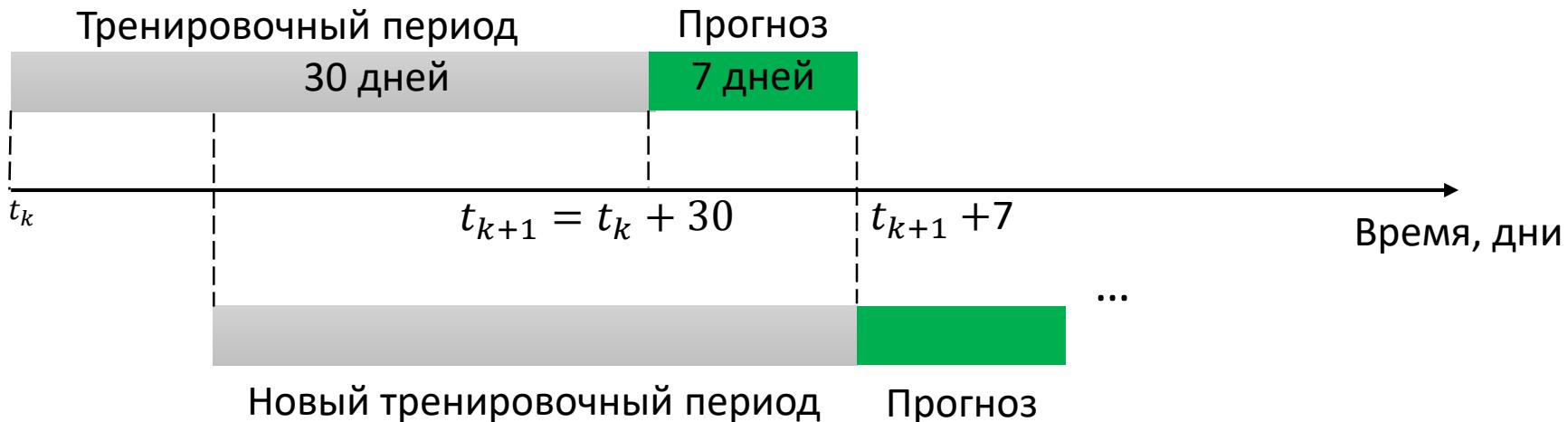
<http://covid19-modeling.ru/>

Сглаженные данные с 12.03.2020 по 12.02.2022
(стопкоронавирус.рф)

- Ежедневное количество проведенных ПЦР-тестов (красная линия),
- Ежедневно выявленные случаи (синяя линия),
- Ежедневное количество смертей от COVID-19 (чёрная линия).

Количество захоронений от COVID-19, полученные муниципальной информационной системой «Ритуал» (мэрия города Новосибирска) коррелирует с данными по общей смертности (ЗАГС).

Алгоритм усвоения данных



Тренировочный период (30 дней) – уточнение параметров q .

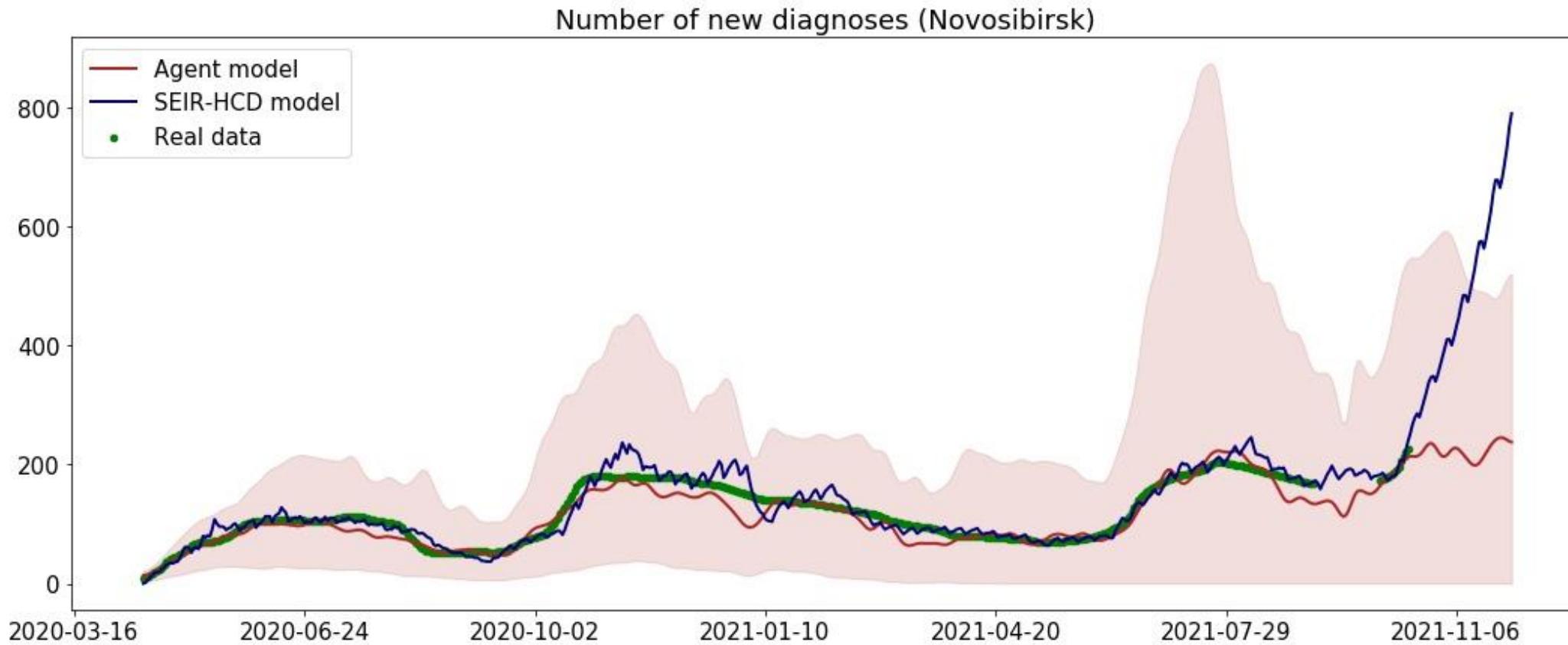
Прогноз (7 дней) рассчитывается при найденных q .

Новый период сдвигается на 7 дней и снова решается обратная задача (уточнение параметров q).

Алгоритм построения доверительных интервалов (для агентной модели)

- Решаем прямую задачу с идентифицированными параметрами 1000 раз (вычислительное время решения прямой задачи для агентной модели на кластере = 10 сек);
- Для каждого дня считаем квантили уровня 0.1, 0.5, 0.9 (строим функцию распределения случайной величины выявленных случаев и выбираем ее значения в точках 0.1, 0.5 и 0.9);
- Получаем 3 массива точек с шагом один день, по которым строим медианное значение (квантиль уровня 0.5, сплошная линия на графиках ниже), верхнюю и нижнюю границу (закрашенная область).

Моделирование распространения COVID-19 в Новосибирской области с 12.03.2020 по 30.11.2021: ежедневно выявленные случаи (на момент расчетов 16.10.2021)



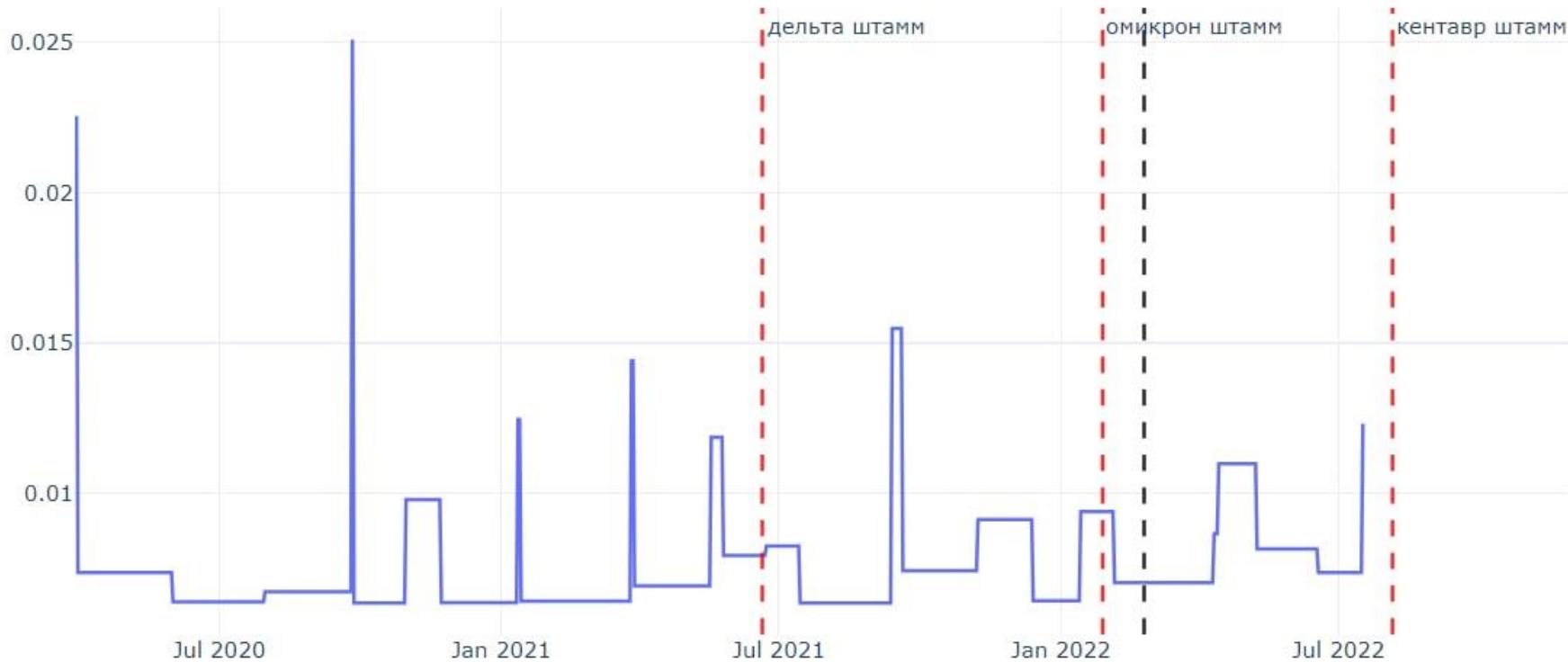
Красная линия – агентная модель, синяя линия - SEIR-HCD модель, жирная зеленая линия - реальные данные с 12.03.2020 по 16.10.2021. Прогноз сделан на 45 дней с 17.10 по 30.11. Закрашенная красная область - 80% доверительный интервал.

Сравнение прогноза на 24.10.2021:

- 250 человек – агентная модель
- 392 человека – SEIR-HCD,
- 281 человек – реальные данные

Влияние ограничительных мер и появления новых штаммов на изменение параметра заразности в НСО с 04.2020 по 08.2022.

Изменение параметра заразности в Новосибирской области



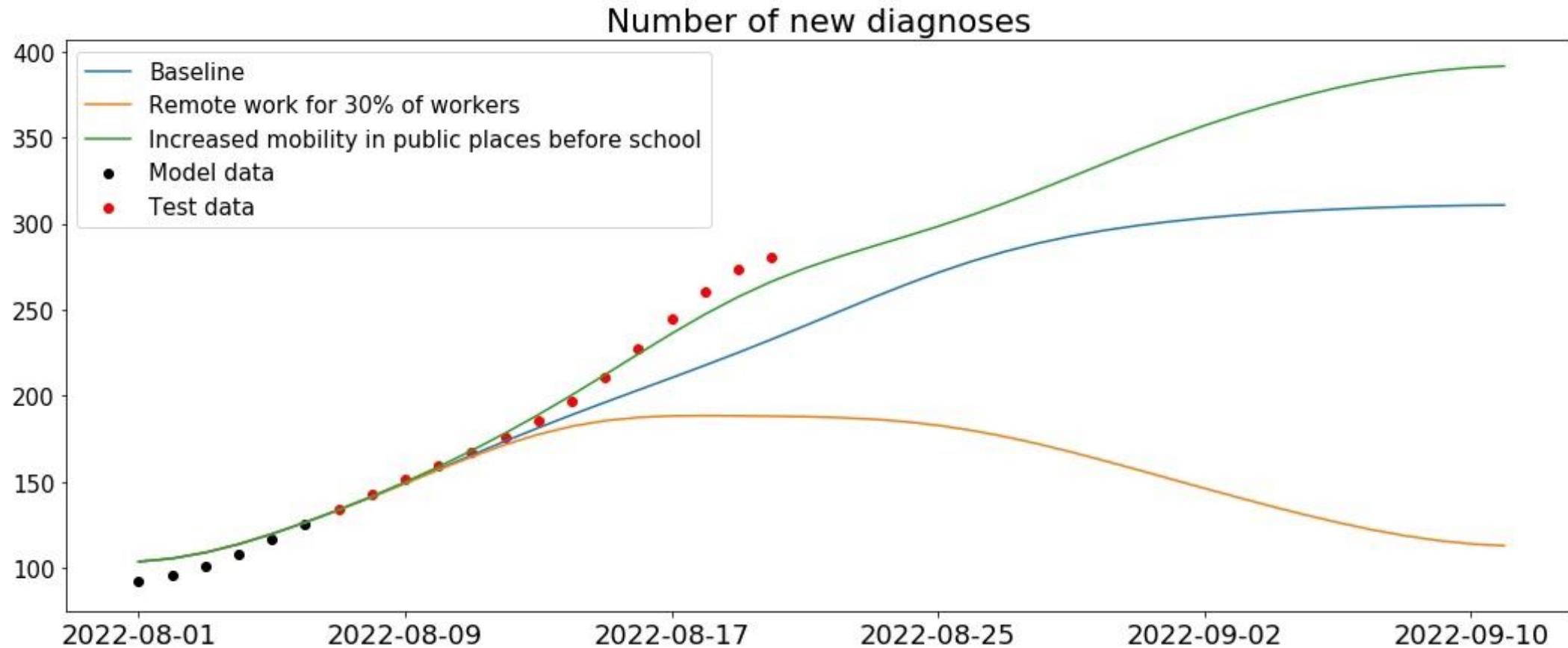
Синяя линия – параметр заразности β ,

Пунктирные красные линии – появление штамма в РФ,

Черная линия – отмена масочного режима в регионе 24.02.2022.

Отмена масочного режима в весенний период 2022 года повлекло за собой увеличение параметра заражения, после чего в летний период его значение стабилизировалось. Увеличение параметра заразности в конце июля свидетельствует о более раннем функционировании нового штамма «Кентавр» в регионе.

Сценарии распространения COVID-19 для новых выявленных случаев в Новосибирской области с 07.08 по 12.09.2022



Голубая линия – базовый сценарий с 07.08.2022 по 12.09.2022.

Оранжевая линия – сценарий, когда 30% рабочих мест переходят в удаленный режим с 21.08.2022.

Зеленая линия – сценарий увеличения мобильности граждан в общественных местах с 21.08.2022.

Черные точки - реальные данные по 06.08.2022, участвующие в решении обратной задачи (<https://covid19-modeling.ru/>).

Красные точки – тестовые данные с 07.08.2022 по 22.08.2022 (для верификации модели).

Теоретические результаты, выносимые на защиту

Впервые изучена идентифицируемость агентных, дифференциальных и стохастических моделей эпидемиологии с учетом экономических и социальных процессов

SEIR-HCD

$$\begin{aligned} q_1 &= (\alpha_E, \alpha_I, \beta_{IgG}, \varepsilon_{HC}, \mu, E_0)^T \\ \text{data} &= (f_k, C_k, g_k)^T, k = 1, \dots, K \end{aligned}$$

Последовательность чувствительных параметров от наибольшего к наименьшему
 $\alpha_E, E_0, \alpha_I, \varepsilon_{HC}, \mu, \beta_{IgG}$

Доля инфицированных, у которых заболевание протекает в легкой форме β_{IgG} , определяется из информации о проценте людей с антителами IgG

АВМ

$$\begin{aligned} q_2 &= (\beta, E(0), \tilde{p})^T \\ \text{data} &= (f_k, C_k, g_k)^T, k = 1, \dots, K \end{aligned}$$

На основе байесовского подхода и метода Монте Карло уточнены границы изменения параметров
 $\beta \in [0.005, 0.05] \rightarrow \hat{\beta} \in [0.005, 0.022]$
 $E_0 \in [1, 1000] \rightarrow \hat{E}_0 \in [1, 1000]$
 $\tilde{p} \in [1, 1000] \rightarrow \hat{\tilde{p}} \in [1, 1000]$

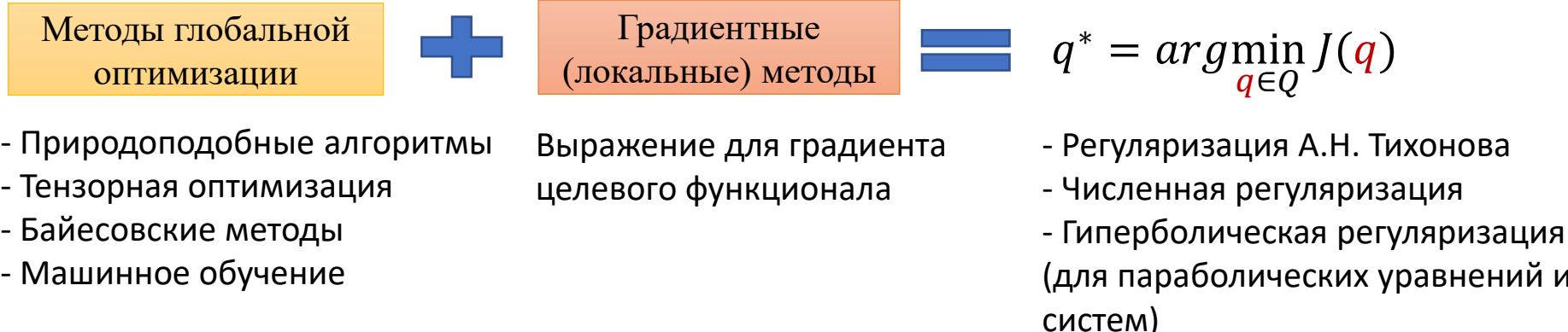
Удалось уменьшить границы изменения параметра заразности β более чем в два раза



На этой основе построены новые градиентные методы решения обратных задач с использованием априорной информации

Разработанные алгоритмы, выносимые на защиту

Впервые разработаны комбинированные алгоритмы численного решения прямых и обратных задач, в которых агентные, дифференциальные и стохастические модели взаимозависимы, и включают в себя градиентные, тензорную оптимизацию и природоподобные алгоритмы (дифференциальной эволюции, генетический алгоритм, методы имитации отжига, роя частиц и древовидных оценок Парзена).



Комплекс программ, выносимый на защиту

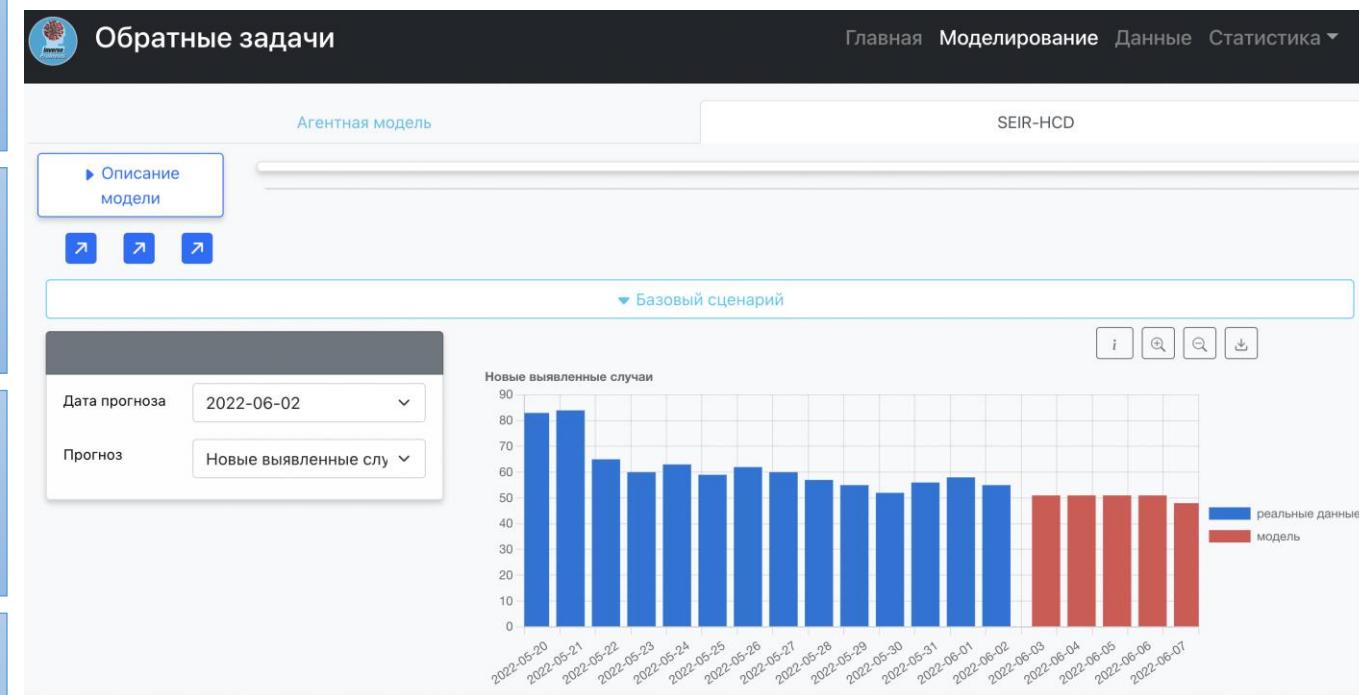
Впервые разработан комплекс программ моделирования и построения сценариев развития COVID-19 в Новосибирской области (<https://covid19-modeling.ru/>) с учетом эпидемической обстановки в прилегающих регионах, а также наиболее интенсивных транспортных потоков Новосибирск-Москва, Алтай, Казахстан, Красноярск, Омск

Свидетельство о государственной регистрации
программы для ЭВМ №2019661851 от 10.09.2019
*Анализ идентифицируемости математических моделей
имmunологии и эпидемиологии*

Свидетельство о государственной регистрации
программы для ЭВМ №2021614740 от 29.03.2021
*Агентное моделирование распространения COVID-19 в
Новосибирской области*

Свидетельство о государственной регистрации
программы для ЭВМ №2022664019 от 21.07.2022
*Моделирование распространения COVID-19 на основе
модели игры среднего поля*

Свидетельство о государственной регистрации
программы для ЭВМ №2022664244 от 27.07.2022
*Автоматизированная программа для прогнозирования
распространения COVID-19 в регионах Российской
Федерации на основе усвоения данных в модели SEIR- HCD*



Конференции в декабре 2022 года по проблемам моделирования в эпидемиологии

Sobolev Institute of Mathematics,
Novosibirsk State University,
Mathematical Center in Akademgorodok,
Institute of Computational Mathematics and
Mathematical Geophysics
hold

XIV International Youth Scientific school-conference

**Theory and numerical methods for solving
inverse and ill-posed problems**

Novosibirsk, Akademgorodok,
December 24 - 27, 2022

Format of the Conference is mixed

http://conf.nsc.ru/tcmiip2022/general_info

S.L. Sobolev Institute of Mathematics SB RAS
Novosibirsk State University
Mathematical Center in Akademgorodok
hold a conference

Modern challenges of inverse problems

Novosibirsk, Akademgorodok,
December 19 - 23, 2022

Format of the Conference is mixed

http://conf.nsc.ru/mdip2022/ru/general_info



Russian Science
Foundation



RUSSIAN
FOUNDATION
FOR BASIC
RESEARCH

Спасибо за внимание!



Результаты моделирования:
<https://covid19-modeling.ru/>



О.И. Криворотко
krivorotko.olya@mail.ru
+7-983-303-3083

*Препринт Института математики СО РАН № 300
«Математические модели распространения COVID-19»*

