

УДК 519.62

## Математическое моделирование и прогнозирование COVID-19 в Москве и Новосибирской области\*

О.И. Криворотко<sup>1,2,3</sup>, С.И. Кабанихин<sup>1,2,3</sup>, Н.Ю. Зятков<sup>1</sup>, А.Ю. Приходько<sup>1,2,3</sup>,  
Н.М. Прохошин<sup>2,3</sup>, М.А. Шишленин<sup>1,2,3</sup>

<sup>1</sup>Институт вычислительной математики и математической геофизики Сибирского отделения Российской академии наук, просп. Акад. Лаврентьева, 6, Новосибирск, 630090

<sup>2</sup>Математический центр в Академгородке, ул. Коптюга, 4, Новосибирск, 630090; ул. Пирогова, 1, каб. 4105, Новосибирск, 630090

<sup>3</sup>Новосибирский национальный исследовательский государственный университет (НГУ), ул. Пирогова, 2, Новосибирск, 630090

E-mails: krivorotko.olya@mail.ru (Криворотко О.И.), ksi52@mail.ru (Кабанихин С.И.), nikolay.zyatkov@gmail.com (Зятков Н.Ю.), a.prikhodko@ngs.ru (Приходько А.Ю.), nikita.prokoshin@gmail.com (Прохошин Н.М.), mshishlenin@ngs.ru (Шишленин М.А.)

Английская версия этой статьи печатается в журнале “Numerical Analysis and Applications” № 4, Vol. 13, 2020.

**Криворотко О.И., Кабанихин С.И., Зятков Н.Ю., Приходько А.Ю., Прохошин Н.М., Шишленин М.А.** Математическое моделирование и прогнозирование COVID-19 в Москве и Новосибирской области // Сиб. журн. вычисл. математики / РАН. Сиб. отд-ние. — Новосибирск, 2020. — Т. , № . — С. 395–412.

В работе исследованы задачи уточнения неизвестных параметров математических моделей SEIR-HCD и SEIR-D распространения коронавирусной инфекции COVID-19 по дополнительной информации о количестве выявленных случаев заболеваний, смертности, коэффициенте самоизоляции и проведенных тестах для города Москвы и Новосибирской области с 23.03.2020. В SEIR-HCD модели популяция разделена на семь, а в SEIR-D — на пять групп со схожими признаками и с вероятностями перехода между группами, зависящими от конкретного региона. Проведен анализ идентифицируемости математической модели SEIR-HCD, который выявил наименее чувствительные к дополнительной информации неизвестные параметры. Задачи уточнения параметров сведены к задачам минимизации целевых функционалов, которые решены с помощью стохастических методов (имитация отжига, дифференциальная эволюция, генетический алгоритм). Разработаны прогностические сценарии развития заболевания в Москве и Новосибирской области и проведен анализ применимости разработанных моделей.

DOI: 10.15372/SJNM20200405

**Ключевые слова:** математическое моделирование, эпидемия, COVID-19, модель SEIR-HCD, модель SEIR-D, сценарии развития, обратная задача, идентифицируемость, оптимизация, дифференциальная эволюция, имитация отжига, генетический алгоритм, Москва, Новосибирская область.

**Krivorotko O.I., Kabanikhin S.I., Zyatkov N.Yu., Prihodko A.Yu., Prokoshin N.M., Shishlenin M.A.** Mathematical modeling and forecasting of COVID-19 in Moscow and Novosibirsk region // Siberian J. Num. Math. / Sib. Branch of Russ. Acad. of Sci. — Novosibirsk, 2020. — Vol. , № . — P. 395–412.

We investigate the inverse problems of finding unknown parameters of the SEIR-HCD and SEIR-D mathematical models of the spread of COVID-19 coronavirus infection based on additional information about the

\*Работа выполнена при поддержке Российского научного фонда (проект № 18-71-10044), а именно, постановка, анализ идентифицируемости и численное решение задачи прогнозирования для математической модели SEIR-HCD (пункты 1, 3 и 4), а также при поддержке Математического Центра в Академгородке в рамках соглашения № 075-15-2019-1675 с Министерством науки и высшего образования Российской Федерации, а именно, постановка и численное решение обратной задачи для математической модели SEIR-D (пункты 2 и 3).

number of detected cases, mortality, self-isolation coefficient and tests performed for the city of Moscow and the Novosibirsk region since 23.03.2020. In the SEIR-HCD model, the population is divided into seven, and in SEIR-D - into five groups with similar characteristics and with transition probabilities depending on a specific region. An analysis of the identifiability of the SEIR-HCD mathematical model was made, which revealed the least sensitive unknown parameters as related to additional information. The task of determining parameters is reduced to the minimization of objective functionals, which are solved by stochastic methods (simulated annealing, differential evolution, genetic algorithm). Prognostic scenarios for the disease development in Moscow and in the Novosibirsk region were developed and the applicability of the developed models was analyzed.

**Keywords:** *mathematical modeling, pandemic, COVID-19, SIER-HCD, SIER-D, scenarios, inverse problem, identifiability, optimization, differential evolution, annealing simulation, genetic algorithm, Moscow, Novosibirsk region.*

## Введение

В декабре 2019 г. произошла вспышка пневмонии в Ухане, в результате которой был впервые обнаружен штамм COVID-19 при анализе нуклеиновой кислоты у пациента с пневмонией. К концу июня 2020 г. пандемия охватила 188 стран, в которых было выявлено более 10 миллионов случаев заражения, 505500 человек умерли. Российская Федерация по числу выявленных случаев находится на 3 месте в рейтинге стран после США и Бразилии с 484630 случаями заражения на 9 июня 2020 г. (см. таблицу 1). Несмотря на пройденный пик выявления заболевших в мире, уменьшение количества выявленных случаев не наблюдается в течение достаточно длительного периода [1]. Разработка сценариев развития заболевания в Москве (наибольшее число выявленных случаев) и Новосибирской области (средний показатель выявления) является важным шагом для регулирования мер по сдерживанию эпидемии в регионах.

**Таблица 1.** Статистика по распространенности коронавируса COVID-19 для пяти стран-лидеров по числу выявленных случаев на 9 июня 2020 г.

Страна	Количество зараженных		Количество смертей	Количество вылеченных
	Всего	Новые случаи		
США	1 961 185	17 250	111 007	524 855
Бразилия	707 412	15 654	37 134	396 737
Россия	485 253	8 595	6 141	241 917
Великобритания	288 834	1 213	40 680	1 257

В работах [2, 3] впервые описаны сценарии развития эпидемии коронавируса в Москве.

Изучение распространения эпидемии коронавируса COVID-19 в Китае с учетом инкубационного периода, пассажиропотоков и решений комиссии здравоохранения в рамках математических моделей, основанных на SEIR, проводилось в работах [4–9]. Обзор математических моделей и методов их решения приведен в работе [10], а обзор литературы по определению параметра заразности для различных стран — в [11].

В статье [12] предложена QSEIR модель, которая учитывает карантинные меры в Китае для противодействия эпидемии. С помощью статистического метода оценены параметры модели. Сделаны прогнозы развития эпидемии, основанные на различных стратегиях карантинных мер в Китае. В [13] также сделан прогноз по развитию пандемии в Китае на основе аппроксимации данных по базисным функциям.

В работе [14] на основе SIR-D модели проведена оценка параметра заразности, а также суточной смертности и темпов выздоровления. Найдены параметры по данным Китая

и сделан прогноз на три недели до 29 февраля. В [15] показано, что эпидемия в Китае должна достичь пика к концу февраля, а к концу апреля — постепенно снижаться. С помощью методов машинного обучения показано, что отмена карантина в Хубэе приведет ко второму пику эпидемии в этой провинции в середине марта и продлит эпидемию до конца апреля.

В [16] предложена математическая модель SEIR-HD с введением мер контроля.

В статье [17] обобщенная SEIR модель применена к распространению SARS-CoV-2 в Италии. С помощью метода роя частиц найдены параметры модели для повышения надежности прогнозов на 30 дней. Полученные результаты были сравнены с данными и прогнозами по Испании и Южной Корее. В предложенной модели учитывалась мобильность людей по отчетам Google COVID-19 Community Mobility Reports. В [18, 19] на основе модифицированных SEIR моделей получены сценарии развития эпидемии в Италии.

В работе [20] с помощью SIR-D модели получены сценарии развития пандемии в Бразилии с учетом закрытия школ и изоляции пожилых людей.

В [21] предложена SEIR-D модель с параметрами, зависящими от времени, и приведены результаты расчётов для округов Индии, а также Ломбардии (Италия) и Москвы (Россия).

Цель данной работы — исследовать зависимость достоверности прогнозов SEIR-моделей от количества и качества информации о распространении пандемии COVID-19.

Модель SIR (Susceptible, Infected, Recovered) является базовой для описания распространения инфекционных заболеваний и была предложена в 1920-х годах шотландскими эпидемиологами Андерсоном Кермаком и Уильямом Маккендриком. Согласно SIR, население делится на три группы:

- $S$  — восприимчивые (незараженные) индивидуумы с 3 лет;
- $I$  — инфицированные с симптомами;
- $R$  — вызлеченные.

С течением времени возможны переходы  $S \rightarrow I$  (заражение) и  $I \rightarrow R$  (выздоровление или смерть). SIR-модель не учитывает многие факторы, например: различную плотность населения в районах, различные способы передачи инфекции. Поэтому SIR неоднократно дорабатывалась [22–24]. Сегодня существует целое семейство моделей (и даже открытых кодов), разработанных на базе SIR-модели [25–28]:

- SIRS. Модель для описания динамики заболеваний с временным иммунитетом.
- SEIR (контактировавшие с инфекцией (Exposed)). Модель описывает распространение заболевания с учетом инкубационного периода.
- SIS. Модель распространения заболевания, к которому не вырабатывается иммунитет.
- MSEIR (наделенные иммунитетом от рождения (Maternally derived immunity)). Модель учитывает иммунитет детей, приобретенный внутриутробно.

В данной работе изучаются две модели: SEIR-HCD и SEIR-D, в которых добавляются

- $E$  — зараженные или находящиеся в инкубационном периоде индивидуумы;
- $H$  — госпитализированные, т. е. с тяжелым протеканием болезни;
- $C$  — находящиеся в критическом состоянии, подключенные к аппарату вентиляции легких;
- $D$  — умершие.

Выбор вариации моделей на базе SEIR не случайный, так как моделируемое заболевание COVID-19 обладает достаточно длительным инкубационным периодом (5–14 дней), в течение которого носители не проявляют симптомов заболевания, но являются инфицированными [29].

В работе сначала уточняются коэффициенты перехода из одной группы в другую посредством решения обратной задачи с использованием данных о количестве протестированных заразившихся  $I(t)$  и  $E(t)$ , а также умерших  $D(t)$ . После этого мы решаем прямую задачу с уже найденными коэффициентами перехода и рассчитываем сценарий развития эпидемии.

Мы также используем следующую информацию о заболевании: длительность инкубационного периода, продолжительность латентного периода (время с момента инфицирования до того момента, когда инфицированный становится распространителем инфекции), параметр контагиозности.

В результате обобщения модели появляется возможность рассчитать сценарий развития заболевания в конкретных городах России. Например, рассчитать время пика заболеваемости в Москве и Новосибирской области, выхода на плато, и каким ожидается спад эпидемии в каждом конкретном случае. В дальнейшем, при развитии моделей, планируется учитывать социальное дистанцирование, климатические условия, возрастные характеристики, особенности иммунитета человека, пассажиропотоки между городами РФ и другие.

## 1. Математическая модель SEIR-HCD

Математическая модель SEIR-HCD распространения коронавируса COVID-19 в конкретном регионе Российской Федерации впервые предложена в работе [30] и основана на системе из 7 нелинейных обыкновенных дифференциальных уравнений на отрезке  $t \in [t_0, T]$ :

$$\left\{ \begin{array}{l} \frac{dS}{dt} = -\frac{5-a(t-\tau)}{5} \left( \frac{\alpha_I S(t)I(t)}{N} + \frac{\alpha_E S(t)E(t)}{N} \right) + \gamma R(t), \\ \frac{dE}{dt} = \frac{5-a(t-\tau)}{5} \left( \frac{\alpha_I S(t)I(t)}{N} + \frac{\alpha_E S(t)E(t)}{N} \right) - (\kappa + \rho)E(t), \\ \frac{dI}{dt} = \kappa E(t) - (\beta + \nu)I(t), \\ \frac{dR}{dt} = \beta I(t) + \rho E(t) - \gamma R(t) + \varepsilon_{HR}H(t), \\ \frac{dH}{dt} = \nu I(t) + \varepsilon_{CH}C(t) - (\varepsilon_{HC} + \varepsilon_{HR})H(t), \\ \frac{dC}{dt} = \varepsilon_{HC}H(t) - (\varepsilon_{CH} + \mu)C(t), \\ \frac{dD}{dt} = \mu C(t) \end{array} \right. \quad (1)$$

с начальными условиями

$$\begin{aligned} S(t_0) &= S_0, \quad E(t_0) = E_0, \quad I(t_0) = I_0, \quad R(t_0) = R_0, \\ H(t_0) &= H_0, \quad C(t_0) = C_0, \quad D(t_0) = D_0. \end{aligned} \quad (2)$$

Вся популяция  $N = S + E + I + R + H + C + D$  состоит из 7 групп (схема модели представлена на рисунке 1 слева).

Описание параметров и интервалы поиска приведены в таблице 2.

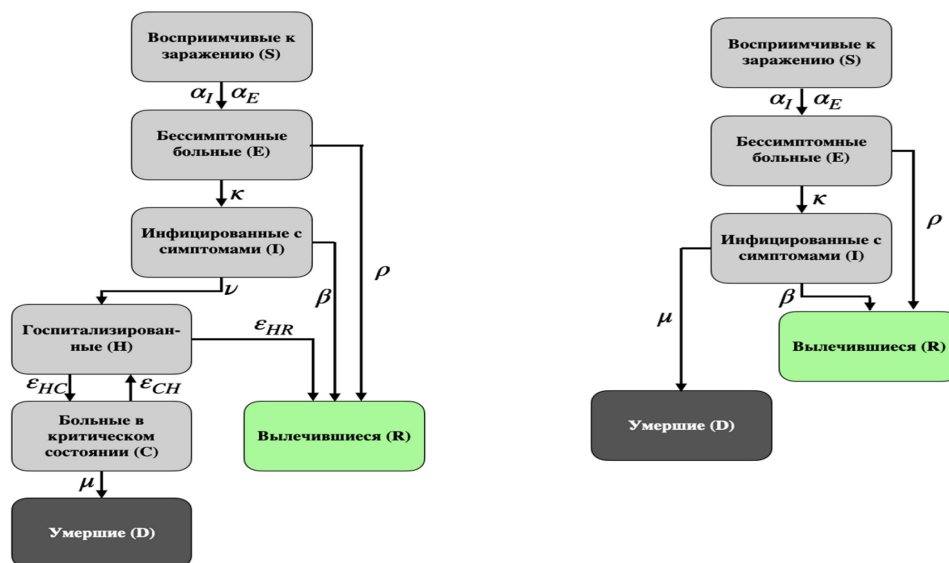


Рис. 1. Схемы математических моделей SEIR-HCD (1) слева и SEIR-D (5) справа

Таблица 2. Описание параметров моделей (1), (2) и (5), (6) и их пределы изменения

Символ	Описание	Пределы	
		SEIR-HCD	SEIR-D
$a(t)$	Индекс самоизоляции по данным Яндекс	(0, 5)	
$\tau$	Латентный период (характеризует запаздывание выделения вирионов или заразности)	2 дня	
$\alpha_I$	Параметр заражения между инфицированным и восприимчивым населением, который связан с контагиозностью вируса и социальными факторами	(0, 1)	
$\alpha_E$	Параметр заражения между бессимптомной и восприимчивой группами населения ( $\alpha_E \gg \alpha_I$ )	(0, 1)	
$\kappa$	Частота появления симптомов в открытых случаях, что приводит к переходу от бессимптомной к инфицированной популяции	(0, 1)	
$\rho$	Скорость восстановления выявленных случаев (случаи, которые выявлены, но выздоравливают без каких-либо симптомов)	0	(0, 1)
$\beta$	Скорость выздоровления зараженных случаев	(0, 1)	
$\gamma$	Скорость повторного заражения. Этот параметр является обратной величиной уровня иммунитета вируса (0 – устойчивый иммунитет, 0.001 – вероятность повторного заражения)	0	
$\nu$	Доля госпитализированных случаев с тяжелым протеканием заболевания	(0, 1)	–
$\varepsilon_{HR}$	Вероятность выздоровления индивидуумов, находящихся в тяжелом состоянии	0.225	–
$\varepsilon_{HC}$	Доля госпитализированных случаев, находящихся в критическом состоянии и требующих подключения аппарата ИВЛ	0.025	–
$\varepsilon_{CH}$	Вероятность отключения аппарата ИВЛ у пациента	(0, 1)	–
$c^{isol}$	Коэффициент влияния индекса самоизоляции на заражаемость	–	(0, 1)
$\mu$	Смертность в результате COVID-19	$(10^{-4}, 10^{-1})$	(0, 0.1)
$E_0$	Начальное количество бессимптомных инфицированных	(1, 800)	(0, 600)
$R_0$	Начальное количество вылеченных индивидуумов	–	(0, 600)

### 1.1. Постановка обратной задачи

Предположим, что известны дополнительные измерения о 3 функциях в фиксированные моменты времени:

$$E(t_k) = b_k f_k, \quad I(t_k) = (1 - b_k) f_k, \quad \Delta D(t_k) = g_k, \quad t_k \in (t_0, T), \quad k = 1, \dots, K, \quad (3)$$

где  $b(t) \in [0, 1]$  — доля бессимптомных больных в выявленных случаях (данные с сайта стопкоронавирус.рф),  $f_k$  — количество выявленных больных в день  $k$ ,  $g_k$  — количество умерших в результате заболевания в день  $k$ ,  $\Delta D(t_k) = D(t_k) - D(t_{k-1})$ ,  $K$  — количество дней в статистике.

Неизвестные параметры:  $q = (\alpha_E, \alpha_I, \kappa, \beta, \nu, \varepsilon_{CH}, \mu, E_0) \in \mathbb{R}^8$ .

Обратная задача (1)–(3) состоит в определении вектора параметров  $q$  по дополнительным измерениям (3).

Обратная задача сводится к задаче минимизации целевого функционала:

$$J(q) = \sum_{k=1}^K \left[ (c_k^{\text{test}} E(t_k; q) - b_k f_k)^2 + (c_k^{\text{test}} I(t_k; q) - (1 - b_k) f_k)^2 + (\Delta D(t_k; q) - g_k)^2 \right]. \quad (4)$$

Здесь  $c_k^{\text{test}} \in [0, 1]$  — отношение количества сделанных тестов к размеру здоровой популяции региона в день  $k$ . Функционал (4) построен по следующим соображениям:  $100b_k\%$  выявленных случаев не проявляют симптомов (относятся к группе  $E(t)$ ), в результате чего  $100b_k\%$  и  $100(1-b_k)\%$  выявленных случаев пропорциональны числу протестированных (с коэффициентом тестирования  $c_k^{\text{test}}$ ) бессимптомных и симптомных соответственно.

## 2. Математическая модель SEIR-D

В рамках модели SEIR-D распространение коронавируса COVID-19 описывается системой из 5 нелинейных обыкновенных дифференциальных уравнений на отрезке  $t \in [t_0, T]$  [31] (схема модели приведена на рис. 1 справа):

$$\begin{cases} \frac{dS}{dt} = -c(t - \tau) \left( \frac{\alpha_I S(t) I(t)}{N} + \frac{\alpha_E S(t) E(t)}{N} \right) + \gamma R(t), \\ \frac{dE}{dt} = c(t - \tau) \left( \frac{\alpha_I S(t) I(t)}{N} + \frac{\alpha_E S(t) E(t)}{N} \right) - (\kappa + \rho) E(t), \\ \frac{dI}{dt} = \kappa E(t) - \beta I(t) - \mu I(t), \\ \frac{dR}{dt} = \beta I(t) + \rho E(t) - \gamma R(t), \\ \frac{dD}{dt} = \mu I(t). \end{cases} \quad (5)$$

Здесь  $N = S + E + I + R + D$  — вся популяция.

Функция, использующая ограничения на передвижения граждан:

$$c(t) = 1 + c^{\text{isol}} \left( 1 - \frac{2}{5} a(t) \right), \quad c(t) \in (0, 2).$$

Начальные данные:

$$S(t_0) = S_0, \quad E(t_0) = E_0, \quad I(t_0) = I_0, \quad R(t_0) = R_0, \quad D(t_0) = D_0. \quad (6)$$

## 2.1. Постановка обратной задачи

Дополнительная информация:

- количество выявленных инфицированных за сутки  $f_k$ ,  $k = 1, \dots, K$ ,
- количество умерших за сутки  $g_k$ ,  $k = 1, \dots, K$ .

В модели (5) это означает

$$0.58f_k = \kappa E_{k-1}, \quad g_k = D(t_k) - D(t_{k-1}), \quad t_k \in (t_0, T), \quad k = 1, \dots, K. \quad (7)$$

Информации о реальном количестве бессимптомных инфицированных очень мало, но можно предполагать, что за день выявляют почти всех инфицированных с симптомами. В среднем 58 % среди выявленных за день ( $f_k$ ) — инфицированные с симптомами ( $I$ ). Выявленные в день  $k$  инфицированные с симптомами ( $0.58f_k$ ) — это некоторая доля инфицированных без симптомов в день  $k - 1$  ( $\kappa E(t_{k-1})$ ).

Неизвестные параметры модели:  $q = (\alpha_E, \alpha_I, \kappa, \rho, \beta, \mu, c^{\text{isol}}, E_0, R_0) \in \mathbb{R}^9$ .

Обратная задача (5)–(7) состоит в определении вектора параметров  $q$  по дополнительной информации (7).

Обратная задача (5)–(7) сводится к задаче минимизации функционала:

$$J(q) = \sum_{k=1}^K \left[ w_1 (\kappa E(t_{k-1}; q) - 0.58f_k)^2 + w_2 (D(t_k; q) - D(t_{k-1}; q) - g_k)^2 \right]. \quad (8)$$

Веса в функционале (8) выбираются следующим образом:

$$w_1 = \sum_{k=1}^K \frac{f_k}{K}, \quad w_2 = \sum_{k=1}^K \frac{g_k}{K}.$$

## 3. Численные эксперименты и результаты прогнозирования

В данном пункте приведено сравнение численных результатов в случае использования обратных задач (1)–(3) для SEIR-HCD модели и (5)–(7) для SEIR-D модели для Москвы и Новосибирской области. Применяются следующие критерии сравнения:

1. Точность предсказания пика эпидемии в Москве (дата и количество выявленных инфицированных).
2. Точность прогноза распространения заболевания в Москве по данным о ежедневных выявленных случаях  $f_k$  и смертности  $g_k$ ,  $k = 1, \dots, K$ .
3. Точность обработки исторических данных ежедневных выявленных случаев  $f_k$  и смертности  $g_k$  в Москве.
4. Требуемые вычислительные ресурсы (время работы алгоритмов, вычислительные мощности).
5. Сопоставление необходимой априорной информации для ограничений искомых параметров (см. табл. 2, 3 и 4 колонки).

### 3.1. Обработка данных

Индекс самоизоляции  $a(t)$  по Москве и Новосибирской области был учтен из карт Яндекс. Коэффициент тестирования вычислялся с использованием статистической информации из открытых источников по формуле

$$c_k^{\text{test}} = \frac{T_k}{Z_k},$$

где  $T_k$  — количество сделанных в регионе тестов в день  $k$ , полученное умножением количества тестов, сделанных в стране в день  $k$ , на долю населения региона в населении страны,  $Z_k$  — количество незараженных индивидуумов в регионе ко дню  $k$ , посчитанное как разность числа жителей региона и общего числа заболевших и умерших в регионе за весь период наблюдения до дня  $k$  включительно (данные о вылеченных не использовались). Для прогнозирования  $T_k$  и  $Z_k$  экстраполировались на дальнейший период времени.

### 3.2. Алгоритмы решения обратных задач

#### 3.2.1. Для модели SEIR-HCD

Для численного решения обратной задачи (1)–(3) использовалась следующая последовательность шагов:

1. Обработка данных  $f_k, g_k, c_k^{\text{test}}, a(t_k), b_k, k = 1, \dots, K$  для временных периодов а)–г) (см. п. 3.3).
2. Определение границ параметров для искомого вектора  $q$  (см. табл. 2, колонку 3).
3. Применение метода дифференциальной эволюции посредством модуля *scipy.optimize.differential\_evolution* библиотеки *SciPy* для определения минимума функционала (4) и получения оптимального вектора параметров  $q^*$  математической модели (1), (2):
  - 3.1. Решение прямой задачи (1), (2) и вычисление значения функционала (4) на каждой итерации метода. Для решения прямой задачи (1), (2) использовался модуль *scipy.integrate.odeint* библиотеки *SciPy*.
  - 3.2. Определение оптимального вектора параметров  $q^*$ .
4. Решение прямой задачи (1), (2) для временных периодов а)–г) с использованием найденных оптимальных параметров из  $q^*$  и построение прогнозов для соответствующих временных периодов а)–г). Количество выявленных за сутки инфицированных  $\hat{f}_k$  вычисляется по формуле  $\hat{f}_k = c_k^{\text{test}}(I(t_k) + E(t_k))$ .

#### 3.2.2. Для модели SEIR-D

Вычисления проводились с помощью языка программирования Python. Для численного решения обратной задачи (5)–(7) использовалась следующая последовательность шагов:

1. Обработка данных  $f_k, g_k, a(t_k), k = 1, \dots, K$ , для временных периодов а)–г).
2. Определение границ параметров для искомого вектора  $q$  (см. табл. 2, колонку 4).
3. Определение минимума функционала (8) и получение оптимальных векторов параметров  $q^*$  математической модели (5), (6) с помощью четырех различных методов:
  - 1) метод дифференциальной эволюции (модуль *scipy.optimize.differential\_evolution*),
  - 2) метод имитации отжига (модуль *scipy.optimize.dual\_annealing*),
  - 3) генетический алгоритм (<https://pypi.org/project/geneticalgorithm/> библиотека),
  - 4) метод роя частиц (<https://pythonhosted.org/pyswarm/> библиотека).



4. Из полученных методами 1)–4) пункта 3 векторов  $q^*$  выбирается лучший в смысле функционала (8):
  - 1) Решение прямой задачи (5), (6) и вычисление значения функционала (8) на каждой итерации метода. Для решения прямой задачи (5), (6) использовался модуль *scipy.integrate.ode\_ivp* библиотеки *SciPy*, метод *RK45*.
  - 2) Определение оптимального вектора параметров  $q^*$  математической модели (5), (6).
5. Решение прямой задачи (5), (6) для временных периодов а)–г) с использованием найденных оптимальных параметров из  $q^*$  и построение прогнозов для соответствующих временных периодов а)–г). Количество выявленных за сутки инфицированных  $\hat{f}_k$  вычисляется по формуле  $\hat{f}_k = \frac{\kappa E(t_{k-1})}{0.58}$ .

### 3.3. Численные расчеты для Москвы

В качестве начальных данных о распространении коронавируса COVID-19 в городе Москве ( $N_0 = 11\,514\,330$  человек) в рамках математической модели SEIR-HCD (1) была использована статистическая информация за 23 марта 2020 г.:

$$S_0 = 11\,514\,241 - q_8, \quad E_0 = q_8, \quad I_0 = 71, \quad R_0 = 9, \quad H_0 = 8, \quad C_0 = 1, \quad D_0 = 0.$$

В модели SEIR-D (5) использовались следующие начальные данные:

$$S_0 = 11\,514\,330 - I_0 - q_8 - q_9, \quad E_0 = q_8, \quad I_0 = 71, \quad R_0 = q_9, \quad D_0 = 0.$$

Данные обратных задач были получены из открытых источников для Москвы для временных периодов:

- |  |  |
|--|--|
| а) 23.03.2020–21.05.2020 ( $K = 60$ дней), | б) 23.03.2020–11.05.2020 ( $K = 50$ дней), |
| в) 23.03.2020–01.05.2020 ( $K = 40$ дней), | г) 23.03.2020–21.04.2020 ( $K = 30$ дней). |

Вычисления проводились с помощью языка программирования Python. Прогноз строился для следующих временных периодов в соответствии с данными а)–г), указанными выше:

- |                                     |                                     |
|-------------------------------------|-------------------------------------|
| а) 22.05.2020–21.06.2020 (30 дней), | б) 12.05.2020–11.06.2020 (30 дней), |
| в) 02.05.2020–01.06.2020 (30 дней), | г) 22.04.2020–21.05.2020 (30 дней). |

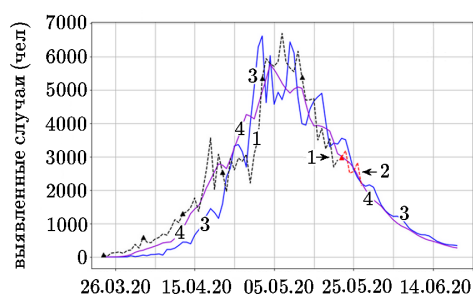
#### 3.3.1. Сравнение результатов

Было проведено сравнение результатов, полученных по двум математическим моделям типа SEIR-HCD и SEIR-D, параметры которых были уточнены для Москвы методами оптимизации (метод дифференциальной эволюции, метод имитации отжига, генетический алгоритм, метод роя частиц).

На рис. 2 представлено количество выявленных инфицированных случаев для двух моделей, верифицированных на реальных данных до 27 мая 2020 г. для временных периодов а)–г). Отметим, что даже при использовании меньшего количества статистических данных прогноз распространения эпидемии более точен в случае математической модели SEIR-HCD (кривая 3), т. е. с хорошей точностью построен прогноз для количества выявленных случаев в Москве (см. отношение результата моделирования SEIR-HCD, обозначенного кривой 3, к неиспользованным в решении обратной задачи статистическим

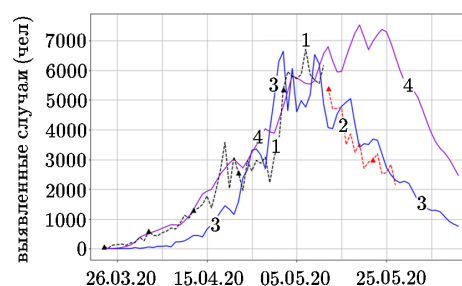
данным, обозначенным кривой 2). Математическая модель SEIR-D имеет большую погрешность в прогнозировании количества выявленных случаев, однако точнее описывает исторические данные (см. отношение результата моделирования SEIR-D, обозначенного кривой 4, к использованным в решении обратной задачи статистическим данным, обозначенным кривой 1).

Восстановленные параметры для моделей SEIR-HCD и SEIR-D приведены в табл. 3–6. Анализируя динамику поведения параметров для математической модели SEIR-HCD при разном наборе данных, отметим, что параметры  $\varepsilon_{CH}$  и  $\beta$  изменяются в больших, в отличие от остальных параметров, диапазонах, что было отмечено в рамках проведенного анализа чувствительности (см. п. 4). Таким образом, при решении обратной задачи необходимо применять регуляризацию и использовать априорную информацию [32], например, в виде статистических ограничений, что было использовано в расчетах.



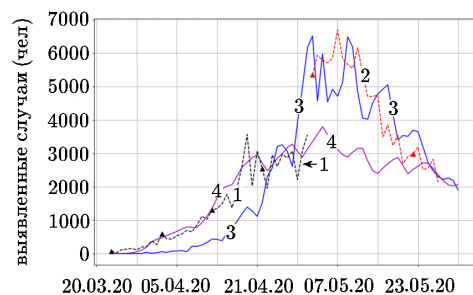
а) Измерения по 21.05.2020.

Параметры приведены в табл. 3



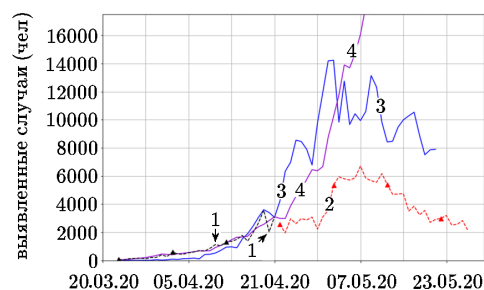
б) Измерения по 11.05.2020.

Параметры приведены в табл. 4



в) Измерения по 01.05.2020.

Параметры приведены в табл. 5



г) Измерения по 21.04.2020.

Параметры приведены в табл. 6

**Рис. 2.** Количество выявленных случаев в Москве при измерениях с 23.03.2020 по указанные под графиками даты с периодами моделирования 90, 80, 70 и 60 дней соответственно. Кривая 1 — реальные данные  $f_k$ , использованные при решении обратных задач, кривая 2 — реальные данные до 27.05.2020, не использованные для решения обратных задач. Кривая 3 — решение для математической модели SEIR-HCD, кривая 4 — решение для математической модели SEIR-D. Значения восстановленных параметров для математических моделей приведены в табл. 3–6

**Таблица 3.** Восстановленные параметры для периода измерений 23.03.2020–21.05.2020

Модель	$\alpha_E$	$\alpha_I$	$\kappa$	$\rho$	$\beta$	$\nu$	$\varepsilon_{CH}$	$\mu$	$c^{\text{isol}}$	$E_0$	$R_0$
SEIR-HCD	0.995	0.499	0.280	—	0.146	0.06	0.003	0.0076	—	795	—
SEIR-D	0.510	0.618	0.002	0.187	0.230	—	—	0.0052	0.9	783	187

**Таблица 4.** Восстановленные параметры для периода измерений 23.03.2020–11.05.2020

Модель	$\alpha_E$	$\alpha_I$	$\kappa$	$\rho$	$\beta$	$\nu$	$\varepsilon_{CH}$	$\mu$	$c^{\text{isol}}$	$E_0$	$R_0$
SEIR-HCD	0.999	0.5	0.290	–	0.16	0.037	0.001	0.011	–	798	–
SEIR-D	0.874	0.669	0.037	0.66	0.97	–	–	0.0163	0.346	598	48

**Таблица 5.** Восстановленные параметры для периода измерений 23.03.2020–01.05.2020

Модель	$\alpha_E$	$\alpha_I$	$\kappa$	$\rho$	$\beta$	$\nu$	$\varepsilon_{CH}$	$\mu$	$c^{\text{isol}}$	$E_0$	$R_0$
SEIR-HCD	0.994	0.491	0.28	–	0.165	0.04	0.0065	0.011	–	794	–
SEIR-D	0.849	0.716	0.011	0.58	0.36	–	–	0.0089	0.56	586	146

**Таблица 6.** Восстановленные параметры для периода измерений 23.03.2020–21.04.2020

Модель	$\alpha_E$	$\alpha_I$	$\kappa$	$\rho$	$\beta$	$\nu$	$\varepsilon_{CH}$	$\mu$	$c^{\text{isol}}$	$E_0$	$R_0$
SEIR-HCD	0.998	0.493	0.268	–	0.063	0.007	0.171	0.0473	–	798	–
SEIR-D	0.078	0.787	0.643	0.703	0.173	–	–	0.0039	0.63	17	332

В табл. 7–10 приведены варианты сравнения результатов моделирования  $\hat{y}$  с реальными данными  $y$  для данных типа а)–г) согласно следующим мерам с учетом длины временного периода  $M$ :

- среднеквадратичная ошибка:  $\text{MSE}(y, \hat{y}) = \sum_{k=1}^M \frac{(y_k - \hat{y}_k)^2}{M}$ ;
- среднее отклонение:  $\text{MAE}(y, \hat{y}) = \sum_{k=1}^M \frac{|y_k - \hat{y}_k|}{M}$ .

Введем обозначения:  $\text{Err}_{\text{fore}}^{\text{MAE}} = \text{MAE}(f, \hat{f})$  и  $\text{Err}_{\text{fore}}^{\text{MSE}} = \text{MSE}(f, \hat{f})$  — ошибки при соответствующем тестовом периоде  $N$ , данные из которого не использовались при решении обратных задач (на рис. 2 данные обозначены кривой 2),  $\text{Err}_{\text{desc}}^{\text{MAE}} = \text{MAE}(f, \hat{f})$  и  $\text{Err}_{\text{desc}}^{\text{MSE}} = \text{MSE}(f, \hat{f})$  — ошибки при соответствующем периоде обучения  $K$ , данные  $f_k$  и  $g_k$  из которого использовались для решения обратных задач (на рис. 2 данные обозначены кривой 1).

Для проверки критерия 1 введем  $\delta_t$  — отклонение даты пика эпидемии в модели от даты реального пика — 7 мая 2020 г. (в днях),  $\delta_f$  — отклонение величины пика эпидемии в модели от реальной величины пика — в 6703 выявленных случаев (в количестве людей).

**Таблица 7.** Сравнение моделей для данных 23.03.2020–21.05.2020,  $K = 60$ ,  $N = 6$ 

Модель	Критерий 1		Критерий 2		Критерий 3	
	$\delta_t$	$\delta_f$	$\text{Err}_{\text{fore}}^{\text{MAE}}$	$\text{Err}_{\text{fore}}^{\text{MSE}}$	$\text{Err}_{\text{desc}}^{\text{MAE}}$	$\text{Err}_{\text{desc}}^{\text{MSE}}$
SEIR-HCD	2	254	338	157690	769	931618
SEIR-D	3	930	198	54977	454	365087

**Таблица 8.** Сравнение моделей для данных 23.03.2020–11.05.2020,  $K = 50$ ,  $N = 16$ 

Модель	Критерий 1		Критерий 2		Критерий 3	
	$\delta_t$	$\delta_f$	$\text{Err}_{\text{fore}}^{\text{MAE}}$	$\text{Err}_{\text{fore}}^{\text{MSE}}$	$\text{Err}_{\text{desc}}^{\text{MAE}}$	$\text{Err}_{\text{desc}}^{\text{MSE}}$
SEIR-HCD	2	174	614	556795	780	974866
SEIR-D	12	817	3346	12745883	338	247688

**Таблица 9.** Сравнение моделей для данных 23.03.2020–01.05.2020,  $K = 40$ ,  $N = 26$ 

Модель	Критерий 1		Критерий 2		Критерий 3	
	$\delta_t$	$\delta_f$	$Err_{fore}^{MAE}$	$Err_{fore}^{MSE}$	$Err_{desc}^{MAE}$	$Err_{desc}^{MSE}$
SEIR-HCD	2	207	779	842300	714	889008
SEIR-D	3	2898	1546	3531484	233	105670

**Таблица 10.** Сравнение моделей для данных 23.03.2020–21.04.2020,  $K = 30$ ,  $N = 36$ 

Модель	Критерий 1		Критерий 2		Критерий 3	
	$\delta_t$	$\delta_f$	$Err_{fore}^{MAE}$	$Err_{fore}^{MSE}$	$Err_{desc}^{MAE}$	$Err_{desc}^{MSE}$
SEIR-HCD	5	7537	5242	30788912	342	183419
SEIR-D	31	348120	48996	6350947086	148	69867

### 3.4. Численные расчеты для Новосибирской области

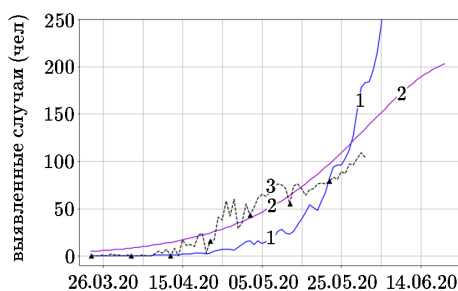
В качестве начальных данных о распространении коронавируса COVID-19 в Новосибирской области ( $N_0 = 2\,798\,170$  человек) для математической модели SEIR-HCD была взята статистическая информация за 23 марта 2020 г.:

$$S_0 = 2\,798\,170 - q_8, \quad E_0 = q_8, \quad I_0 = 0, \quad R_0 = 0, \quad H_0 = 0, \quad C_0 = 0, \quad D_0 = 0,$$

а для математической модели SEIR-D в следующем виде:

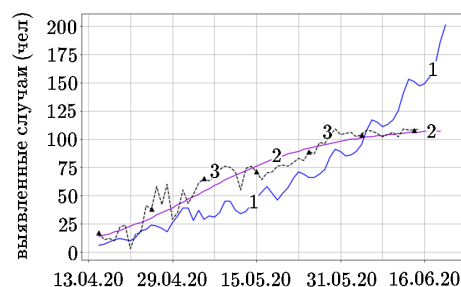
$$S_0 = 2\,798\,170 - q_8 - q_9, \quad E_0 = q_8, \quad I_0 = 0, \quad R_0 = q_9, \quad D_0 = 0.$$

На рис. 3 приведены результаты моделирования и прогнозирования количества выявленных случаев в Новосибирской области в случае измерений с 23 марта по 31 мая 2020 г. (рис. 3а) и по 15 июня 2020 г. (рис. 3б), полученные с использованием описанных алгоритмов для двух математических моделей, указанных выше. Отметим, что описание реальных данных по выявленным случаям, полученное с помощью математической модели SEIR-D (кривая 2), точнее, чем соответствующее описание с помощью модели SEIR-HCD (кривая 1).



а) Измерения по 31.05.2020.

Параметры приведены в табл. 11



б) Измерения по 15.06.2020.

Параметры приведены в табл. 12

**Рис. 3.** Количество выявленных случаев COVID-19 в Новосибирской области с 23.03.2020 по 21.06.2020 (90 дней) в случае измерений с 23.03.2020 по а) 31.05.2020 и б) 15.06.2020. Кривая 1 — решение для математической модели SEIR-HCD, кривая 2 — решение для математической модели SEIR-D, кривая 3 — реальные данные  $f_k$ , использованные при решении обратных задач. Значения восстановленных параметров для математических моделей приведены в табл. 11, 12.

Анализ моделирования смертности в Новосибирской области затруднителен для обеих моделей, так как имеющихся данных недостаточно для качественного прогнозирования смертности в регионе. Отметим, что влияние коэффициента индекса самоизоляции на заражаемость  $c^{\text{isol}}$  в Новосибирской области для математической модели SEIR-D нулевое, что, фактически, означает отсутствие влияния изоляции на заражаемость в регионе (см. табл. 11, 12).

**Таблица 11.** Восстановленные параметры для периода измерений 23.03.2020–31.05.2020, Новосибирская область

Модель	$\alpha_E$	$\alpha_I$	$\kappa$	$\rho$	$\beta$	$\nu$	$\varepsilon_{CH}$	$\mu$	$c^{\text{isol}}$	$E_0$	$R_0$
SEIR-HCD	0.001	0.224	0.108	–	0.013	0.006	0.055	0.072	–	1001	–
SEIR-D	0.999	0.999	0.042	0.952	0.999	–	–	0.0188	0	99	24

**Таблица 12.** Восстановленные параметры для периода измерений 23.03.2020–15.06.2020, Новосибирская область

Модель	$\alpha_E$	$\alpha_I$	$\kappa$	$\rho$	$\beta$	$\nu$	$\varepsilon_{CH}$	$\mu$	$c^{\text{isol}}$	$E_0$	$R_0$
SEIR-HCD	0.219	0.0001	0.118	–	0.046	0.001	0.001	0.1	–	6412	–
SEIR-D	0.271	0.999	$1.9 \cdot 10^{-5}$	$10^{-9}$	0.007	–	–	0.0009	0	93	96

#### 4. Анализ чувствительности SEIR-HCD

Используя программный пакет DAISY [33] по анализу структурной идентифицируемости модели (1)–(2), показано, что математическая модель распространения коронавируса в популяции является идентифицируемой. Однако необходимо установить чувствительность параметров  $q$  к функциям модели с целью контроля устойчивости полученного решения обратной задачи и качества прогнозирования [34–36].

В качестве первого шага мы попытаемся оценить те параметры, к которым решение модели наиболее чувствительно. Они, в свою очередь, определяются с помощью расчета полукотнотельной чувствительности.

Чтобы описать эту методологию, предположим, что мы хотим определить относительную чувствительность наблюдаемых модельных величин  $x = (E, I, D)^T$  (и, следовательно, модельного решения  $\bar{x} = (S, E, I, R, H, C, D)^T$ ) к конкретным параметрам  $q_k, k = 1, \dots, 8$ . Полукотнотельная чувствительность модельного решения к параметру  $q_k$  определяется выражением  $\frac{\partial x_i(t; q)}{\partial q_k} q_k$  и вычисляется путем формального дифференцирования модели ОДУ:

$$\begin{cases} \frac{d\bar{x}}{dt} = f(t, \bar{x}; q), \\ \bar{x}(0) = \bar{x}_0 \end{cases} \quad (9)$$

по  $q_k$  с изменением порядка дифференцирования по времени и параметрам.

Таким образом, мы получаем  $(7 \times 8)$ -мерную систему дифференциальных уравнений для функции чувствительности  $\bar{x}_q(t; q) = \partial \bar{x}(t; q) / \partial q$ :

$$\frac{d}{dt} \left( \frac{\partial \bar{x}(t)}{\partial q} \right) = \frac{\partial f}{\partial x} \frac{\partial \bar{x}(t)}{\partial q} + \frac{\partial f}{\partial q}$$

с начальными условиями

$$\frac{\partial \bar{x}(0)}{\partial q} = \frac{\partial \bar{x}_0}{\partial q} = 0.$$

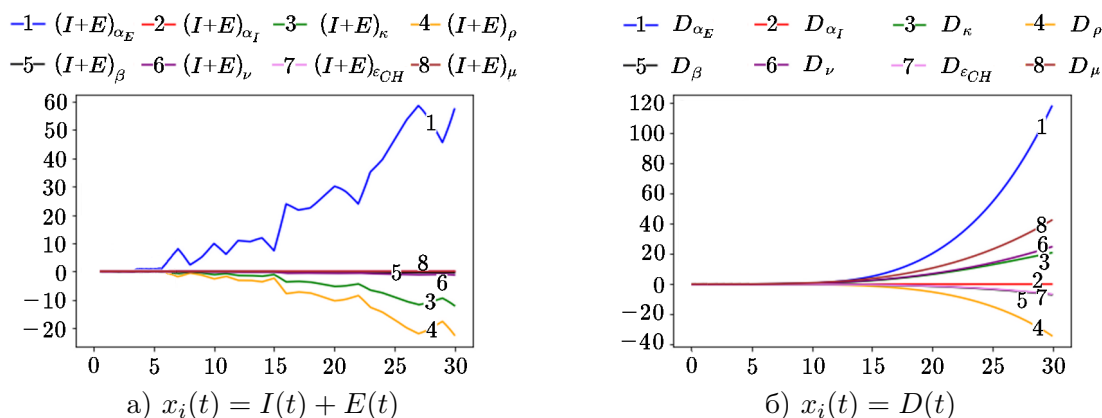
Последнее означает, что нулевая матрица при любом начальном условии не является функцией всех оцененных параметров. Здесь  $\partial f/\partial \bar{x}$  — якобиан системы ОДУ, а  $\partial f/\partial q$  — производная правой части по рассматриваемым параметрам.

Этот процесс дает информацию о чувствительности как функцию времени на интересующем интервале. Мы хотим иметь некоторую общую меру чувствительности решения к параметрам, поэтому для каждой комбинации состояние/параметр мы берем норму (в пространстве  $L_2$ ) по времени, а затем ранжируем полученные скаляры, чтобы определить наиболее чувствительные параметры (см. табл. 13). Чем меньше значение  $\|c_i(t) \frac{\partial x_i(t)}{\partial q_k}\|_2$ , тем меньше влияние параметра  $q_k$  на переменную  $x_i$ . Например, качество определения параметров  $\varepsilon_{CH}$  и  $\mu$  при решении обратной задачи не зависит от имеющихся измерений количества инфицированных  $I(t) + E(t)$  в отличие от более чувствительных к этим данным коэффициентов  $\alpha_E$ ,  $\rho$ ,  $\kappa$ .

**Таблица 13.** Полуотносительные чувствительности различных состояний модели к параметрам, отсортированные по убыванию

Переменная $x_i$	Параметр $q_k$	$\ c_i(t) \frac{\partial x_i(t)}{\partial q_k} q_k\ _2$	Переменная $x_i$	Параметр $q_k$	$\ c_i(t) \frac{\partial x_i(t)}{\partial q_k} q_k\ _2$
$D$	$\alpha_E$	225.81	$D$	$\beta$	13.60
$I + E$	$\alpha_E$	156.69	$D$	$\varepsilon_{CH}$	12.64
$D$	$\mu$	89.06	$I + E$	$\nu$	3.38
$D$	$\rho$	63.27	$I + E$	$\beta$	1.21
$I + E$	$\rho$	57.73	$D$	$\alpha_I$	0.22
$D$	$\nu$	53.78	$I + E$	$\alpha_I$	0.19
$D$	$\kappa$	47.44	$I + E$	$\mu$	0.00
$I + E$	$\kappa$	30.42	$I + E$	$\varepsilon_{CH}$	0.00

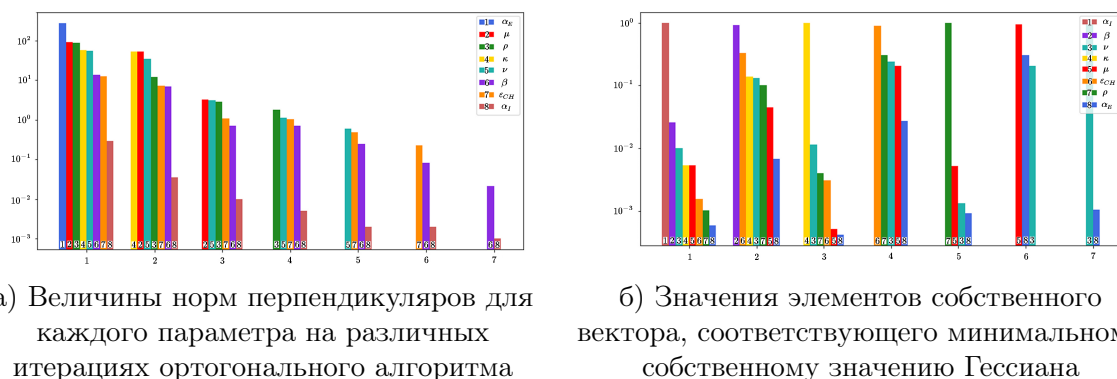
На рис. 4 представлены графики изменения от времени функции чувствительности  $c^{\text{test}}(t) \frac{\partial x_i(t)}{\partial q_k} q_k$  в зависимости от варьируемого параметра. Чем более изменчив параметр в динамике, тем чувствительность к данным измерениям выше, а, значит, определяться он будет более устойчиво. В обратной задаче (1)–(3) самыми идентифицируемыми параметрами оказались  $\alpha_E$ ,  $\rho$  и  $\kappa$ , а наименее чувствительными к данным — параметры  $\mu$ ,  $\varepsilon_{CH}$  и  $\alpha_I$ .



**Рис. 4.** Функция чувствительности  $c^{\text{test}}(t) \frac{\partial x_i(t)}{\partial q_k} q_k$  для временного периода с 23.03.2020 по 21.04.2020 (по горизонтали 30 дней)

На рис. 5 приведены результаты анализа чувствительности параметров математической модели (1) при различных итерациях ортогонального алгоритма и метода собствен-

ных значений соответственно [35]. Показано, что наиболее идентифицируемыми также оказались параметры заражения между бессимптомной и восприимчивой группами населения  $\alpha_E$ , смертности  $\mu$  и скорости восстановления выявленных случаев  $\rho$ .



**Рис. 5.** Анализ чувствительности для математической модели (1) ортогональным методом (а) и методом собственных значений (б) [35] (логарифмическая шкала)

**Таблица 14.** Последовательности параметров, полученные с помощью методов анализа чувствительности для математической модели (1), расположенные от наиболее до наименее чувствительного параметра

Ортогональный метод	Метод собственных значений
$\alpha_I, \beta, \varepsilon_{CH}, \nu, \rho, \mu, \kappa, \alpha_E$	$\alpha_I, \beta, \kappa, \varepsilon_{CH}, \rho, \mu, \nu, \alpha_E$

В табл. 14 приведены последовательности параметров чувствительности, полученные двумя методами. После анализа идентифицируемости можно заключить, что наименее чувствительными (более идентифицируемыми) параметрами модели к вариациям в данных (погрешностям) являются  $\alpha_E, \kappa$  и  $\mu$ , другими словами, эти параметры более устойчиво определяются при решении обратной задачи (1)–(3). Наиболее чувствительными (менее идентифицируемыми) к ошибкам в измерениях являются параметры  $\alpha_I, \varepsilon_{CH}$  и  $\beta$ , т.е. необходимо разработать алгоритм регуляризации, позволяющий контролировать качество определения чувствительных параметров.

## 5. Заключение

Были проанализированы две математические модели SEIR-HCD и SEIR-D, основанные на SEIR-структуре, для каждой из них сформулированы обратные задачи уточнения параметров для Москвы и Новосибирской области, проведен анализ идентифицируемости и чувствительности параметров к ошибкам в статистических данных, а также разработаны алгоритмы решения обратных задач. Построены сценарии развития распространения эпидемии COVID-19 в Москве и Новосибирской области для различных временных промежутков измерений и прогнозирования.

Прогнозирование пика развития эпидемии в Москве с ошибкой в 2 дня и на 174 выявленных случая меньше фактического было получено с помощью математической модели SEIR-HCD при использовании данных по выявленным случаям и смертности в Москве с 23 марта 2020 г. по 1 мая 2020 г. Также прогнозирование сценария развития эпидемии (количество новых выявленных случаев за день) с наименьшей ошибкой было

получено с помощью математической модели SEIR-HCD во всех случаях использования данных (30, 40, 50 дней статистики) кроме наибольшего периода (60 дней статистики), при котором меньшую ошибку продемонстрировала математическая модель SEIR-D. По-видимому, использование более грубой математической модели (с меньшим количеством однородных групп) оправдано в случае достаточно большого количества статистических данных и небольшого периода прогнозирования. Действительно, математическая модель SEIR-D наилучшим образом приближает исторические данные по выявленным случаям для Москвы.

Моделирование и прогнозирование развития коронавирусной инфекции в Новосибирской области продемонстрировано для двух наборов измерений: с 23.03.2020 г. по 31.05.2020 г. и по 15.06.2020 г. ввиду небольшой статистики вплоть до 15 апреля 2020 г. Показано, что пик эпидемии еще не достигнут. Для указанных промежутков измерений математическая модель SEIR-D наиболее точно моделирует ситуацию распространения коронавирусной инфекции в Новосибирской области.

Отметим, что в случае большого промежутка моделирования необходимо учитывать динамику введенных мер по борьбе с коронавирусной инфекцией (обязательное использование масок в закрытых помещениях, открытие границ для пассажироперевозок, открытие кинотеатров и др.), которая влияет на реальные данные, количественные и качественные показатели восстановленных параметров, сценарии развития и свойства моделей. В таком случае, например, неизвестные параметры  $q$  моделей следует рассматривать как функции, зависящие от времени (кусочно-постоянные, кусочно-линейные и т.п.). Характер изменения восстановленных параметров по времени поможет точнее объяснить развитие заболевания в регионах и то, как введение мер (или их ослабление) влияет на количество инфицированных случаев. В дальнейшем планируется рассматривать кусочно-постоянные функции от времени  $\alpha_E(t)$ ,  $\alpha_I(t)$ ,  $c^{\text{isol}}(t)$  и другие.

*Благодарности.* Авторы выражают признательность участникам рабочей группы РАН по моделированию COVID-19 за ценные советы и обсуждение работы, а также А.С. Козелкову и О.А. Кузенкову за ценные обсуждения и замечания по содержанию работы.

## Литература

1. Coronavirus COVID-19 Global Cases by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University. March 21, 2020. — <https://gisanddata.maps.arcgis.com/apps/opsdashboard/index.html#/bda7594740fd40299423467b48e9ecf6>
2. **Тамм М.В.** Коронавирусная инфекция в Москве: прогнозы и сценарии // ФАРМАКО-ЭКОНОМИКА. Современная Фармакоэкономика и Фармакоэпидемиология. — 2020. — Т. 13, № 1. — С. 43–51. — DOI: 10.17749/2070-4909.2020.13.1.43-51.
3. **Koltsova E.M., Kurkina E.S., Vasetsky A.M.** Mathematical modeling of the spread of COVID-19 in Moscow and Russian regions. — 2020. — arXiv:2004.10118 [q-bio.PE].
4. **Zlojutro A., Rey D., Gardner L.** Optimizing border control policies for global out-break mitigation // Scientific Reports. — 2019. — Vol. 9. — P. 2216. — <https://rdcu.be/bniOs>
5. **Chen Y., Cheng J., Jiang Y., and Liu K.** A time delay dynamical model for outbreak of 2019-nCoV and the parameter identification // J. of Inverse and Ill-posed Problems. — 2020. — Vol. 28, iss. 2. — P. 243–250.
6. **Tang B., Wang X., Li Q., Bragazzi N.L., Tang S., Xiao Y., Wu J.** Estimation of the transmission risk of 2019-nCoV and its implication for public health interventions // SSRN: <https://ssrn.com/abstract=3525558>



7. **Shi Pengpeng, Cao Shengli, Feng Peihua** SEIR Transmission dynamics model of 2019 nCoV coronavirus with considering the weak infectious ability and changes in latency duration // medRxiv. — 2020. — DOI: 10.1101/2020.02.16.20023655.
8. **Kiesha Prem, Yang Liu, Timothy W Russell, Adam J Kucharski, Rosalind M Eggo, Nicholas Davies, Mark Jit, Petra Klepac** The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: a modelling study // The Lancet Public Health. — 2020. — Vol. 5, № 5. — P. E261–E270.
9. **Fan Ru-guo, Wang Yi-bo, Luo Ming, Zhang Ying-qing, Zhu Chao-ping** SEIR-Based COVID-19 Transmission Model and Inflection Point Prediction Analysis // J. of University of Electronic Science and Technology of China. — 2020. — Vol. 49, № 3. — P. 369–374.
10. **Кабанихин С.И., Криворотько О.И.** Математическое моделирование эпидемии Уханьского коронавируса COVID-2019 и обратные задачи // Журн. вычисл. математики и мат. физики. — 2020. — Т. 11 (в печати).
11. **Ying Liu, Albert A Gayle, Annelies Wilder-Smith, Joacim Rocklöv** The reproductive number of COVID-19 is higher compared to SARS coronavirus // J. of Travel Medicine. — 2020. — Vol. 27, iss. 2. — taaa021, DOI: org/10.1093/jtm/taaa021.
12. **Liu Xiuli, Hewings Geoffrey J.D., Wang Shouyang, Qin Minghui, Xiang Xin, Zheng Shan, and Li Xuefeng** Modelling the situation of COVID-19 and effects of different containment strategies in China with dynamic differential equations and parameters estimation // medRxiv. — 2020. — DOI:10.1101/2020.03.09.20033498.
13. **Lijun Pei** Prediction of numbers of the accumulative confirmed patients (NACP) and the plateau phase of 2019-nCoV in China // Cognitive Neurodynamics. — 2020. — Vol. 14. — P. 411–424.
14. **Anastassopoulou C., Russo L., Tsakris A., Siettos C.I.** Data-based analysis, modelling and forecasting of the novel Coronavirus (2019-nCoV) outbreak // Plos One. — 2020. — DOI: 10.1101/2020.02.11.20022186.
15. **Yang Z., Zeng Z., Wang K., Wong S.S., Liang W., Zanin M., Liu P., Cao X., Gao Z., Mai Z., Liang J., Liu X., Li S., Li Y., Ye F., Guan W., Yang Y., Li F., Luo S., Xie Y., Liu B., Wang Z., Zhang S., Wang Y., Zhong N., He J.** Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions // J. Thorac Dis. — 2020. — Vol. 12, № 3. — P. 165–174. — DOI: 10.21037/jtd.2020.02.64.
16. **Ivorra B., Ferrández M.R., Vela-Pérez M., Ramos A.M.** Mathematical modeling of the spread of the coronavirus disease 2019 (COVID-19) taking into account the undetected infections. The case of China [published online ahead of print, 2020 Apr 30]. Commun Nonlinear Sci Numer Simul. — 2020. — Vol. 88:105303. — DOI: 10.1016/j.cnsns.2020.105303.
17. **Godio A., Pace F., Vergnano A.** SEIR modeling of the Italian epidemic of SARS-CoV-2 using computational swarm intelligence // International Journal of Environmental Research and Public Health. — 2020. — Vol. 17. — P. 3535. — DOI: 10.3390/ijerph17103535.
18. **Carcione J.M., Santos J.E., Bagaini C., and Ba J.** A Simulation of a COVID-19 Epidemic Based on a Deterministic SEIR Model. Front. Public Health. — 2020. — Vol. 8:230. — DOI: 10.3389/fpubh.2020.00230
19. **Ding Y., Gao L.** An evaluation of COVID-19 in Italy: A data-driven modeling analysis, Infectious Disease Modelling — 2020. — DOI: 10.1016/j.idm.2020.06.007.
20. **Askery Canabarro, Elayne Tenorio, Renato Martins, Lais Martins, Samurai Brito, Rafael Chaves** Data-driven study of the COVID-19 pandemic via age-structured modelling and prediction of the health system Failure in Brazil amid diverse intervention strategies // medRxiv. — 2020. — DOI: 10.1101/2020.04.03.20052498.
21. **Taarak Rapolu, Brahmani Nutakki, Sobha Rani T., Durga Bhavani S.** A time-dependent SEIRD model for forecasting the COVID-19 transmission dynamics // medRxiv — 2020. — DOI: 10.1101/2020.05.29.20113571.

22. **Klepac P., Pomeroy L.W., Bjørnstad O.N., Kuiken T., Osterhaus A.D.M.E., Rijks J.M.** Stage-structured transmission of phocine distemper virus in the Dutch 2002 outbreak // *Proc Biol Sci.* — 2009. — Vol. 276. — P. 2469–2476.
23. **Klepac P., Caswell H.** The stage-structured epidemic: linking disease and demography with a multi-state matrix approach model // *Theor Ecol.* — 2011. — Vol. 4. — P. 301–319.
24. **Siettos C.I., Russo L.** Mathematical modeling of infectious disease dynamics // *Virulence.* — 2013. — Vol. 4, № 4. — P. 295–306.
25. **Samuel M. Jenness, Steven M. Goodreau** Martina Morris EpiModel: An R package for mathematical modeling of infectious disease over networks // *J. of Statistical Software.* — 2018. — Vol. 84, № 8. — DOI: 10.18637/jss.v084.i08.
26. **Noll N.B., Aksamentov I., Druelle V., Badenhorst A., Ronzani B., Jefferies G., Albert J., Neher R.** COVID-19 Scenarios: an interactive tool to explore the spread and associated morbidity and mortality of SARS-CoV-2 // *medRxiv.* — 2020. — DOI: 10.1101/2020.05.05.20091363.
27. <https://covid19.biouml.org/COVID-19 Statistics and forecast at the regional level>
28. <https://covid19-projections.com/COVID-19 Projections Using Machine Learning>
29. Coronavirus disease 2019 (COVID-19) // Situation report. — May 31, 2020. — <https://covid19.who.int/>
30. **Unlu E., Leger H., Motornyi O., Rukubayihunga A., Ishacian T., Chouiten M.** Epidemic analysis of COVID-19 outbreak and counter-measures in France // *medRxiv.* — 2020. — DOI: 10.1101/2020.04.27.20079962.
31. **Sameni R.** Mathematical modeling of epidemic diseases; A case study of the COVID-19 coronavirus. — 2020. — arXiv:2003.11371.
32. **Kabanikhin S.I., Shishlenin M.A.** Quasi-solution in inverse coefficient problems // *J. of Inverse and Ill-posed Problems.* — 2008. — Vol. 16, № 7. — P. 707–715.
33. **Bellu G., Saccomani M.P., Audoly S., D’Angi’o L.** DAISY: A new software tool to test global identifiability of biological and physiological systems // *Computer Methods and Programs in Biomedicine.* — 2007. — Vol. 88, № 1. — P. 52–61.
34. **Raue A., Becker V., Klingmüller U., Timmer J.** Identifiability and observability analysis for experimental design in nonlinear dynamical models // *Chaos.* — 2010. — Vol. 20:045105. — DOI: 10.1063/1.3528102.
35. **Криворотько О.И., Андорная Д.В., Кабанихин С.И.** Анализ чувствительности и практическая идентифицируемость математических моделей биологии // *Сибирский журнал индустриальной математики.* — 2020. — Т. 23, № 1. — С. 107–125.
36. **Raue A., Karlsson J., Saccomani M.P., Jirstrand M., Timmer J.** Comparison of approaches for parameter identifiability analysis of biological systems // *Bioinformatics.* — 2014. — Vol. 30, № 10. — P. 1440–1448. — DOI: 10.1093/bioinformatics/btu006.

*Поступила в редакцию 20 июля 2020 г.*

*После исправления 30 июля 2020 г.*

*Принята к печати 30 июля 2020 г.*