



Е.С.Куркина, Д.И.Зинченко,
Е.М.Кольцова

**Математическое моделирование и
прогнозирование распространения
COVID-19: многоволновая модель
развития эпидемии в Великобритании**

Рекомендуемая форма библиографической ссылки

Куркина Е.С., Зинченко Д.И., Кольцова Е.М. Математическое моделирование и прогнозирование распространения COVID-19: многоволновая модель развития эпидемии в Великобритании // Проектирование будущего. Проблемы цифровой реальности: труды 5-й Международной конференции (3-4 февраля 2022 г., Москва). — М.: ИПМ им. М.В.Келдыша, 2022. — С. 182-192. — <https://keldysh.ru/future/2022/16.pdf>
<https://doi.org/10.20948/future-2022-16>

Размещено также [видео выступления](#)

Математическое моделирование и прогнозирование распространения COVID-19: многоволновая модель развития эпидемии в Великобритании

Е.С. Куркина^{1,2}, Д.И. Зинченко² Э.М. Кольцова²

¹ *Московский государственный университет им. М.В. Ломоносова, ВМК*

² *Российский химико-технологический университет им. Д.И. Менделеева*

Аннотация. В работе с помощью дискретного логистического уравнения Фейгенбаума проводится математическое моделирование распространения волн эпидемии коронавируса в Великобритании. Модель применяется для расчетов общего числа случаев заболевания, общего количества смертей, выздоровлений и активных случаев. Используется методика выявления локальных волн эпидемии и определения параметров модели. Для разных случаев и исходов заболевания параметры подбираются согласованно. Показано, что модель позволяет хорошо описывать статистические данные и делать реалистичные прогнозы.

Ключевые слова: математическое моделирование, эпидемия COVID-19, открытые системы, дискретное логистическое уравнение, суперпозиция эпидемиологических волн, прогнозирование

Mathematical modeling and forecasting of the spread of the COVID-19: Multiwave model of epidemic development in the UK

E.S. Kurkina¹, D. I. Zinchenko², E.M. Koltsova²

¹ *Lomonosov Moscow State University, Department of Computational Mathematics & Cybernetics*

² *Mendeleev University of Chemical Technology of Russia*

Abstract. In this paper, using the Feigenbaum logistic map, mathematical modeling of the spread of waves of the coronavirus epidemic in UK is carried out. It is shown that the chosen model is optimal for describing the complex dynamics of the spread of infections in open, heterogeneous systems with inaccurately specified statistical data. A method for detecting local epidemic waves and determining the model parameters has been used. It is shown that the model

makes it possible to describe statistical data well and make realistic forecasts. The model is used to model the total number of cases, the total number of deaths, recoveries, and the calculation of active cases.

Keywords: mathematical modeling, COVID-19 epidemic, open systems, logistic map, superposition of epidemiological waves, forecasting

Введение

Эпидемия коронавируса COVID-19 была впервые выявлена в конце декабря 2019 г. в Китае. Сейчас, спустя два с половиной года общее число инфицированных составляет более 452,5 млн, причем за первые полтора года заболело 118,4 млн чел., а за второй год – уже 334,1 млн. [1] Вирус продолжает мутировать: за последний год было выявлено два штамма, вызывающих опасения – «дельта» (ранее называемый индийским) и «омикрон». [2]

Для того, чтобы успешно и заранее рассчитывать нагрузку на здравоохранение и вводить необходимое количество ограничительных мер, не перегружающее население психологически и экономически (перевод сотрудников на дистанционный режим работы и отправление в отпуск, когда это невозможно, закрытие и ограничение посещения общественных мест и др.), требуется анализировать динамику развития заболеваемости и делать реалистичные прогнозы. Это можно сделать с адекватной математической моделью.

Распространение эпидемии в разных странах происходит по-разному: из-за влияния различной социальной динамики, разных штаммов (разного сочетания нескольких штаммов) и т.д., но для любого распространения инфекции характерно несколько стадий: экспоненциального роста, логистического роста, пиковой заболеваемости, спада эпидемии; возможна эндемическая стадия, когда ежедневная заболеваемость держится на небольшом, но постоянном уровне, что может привести к новой вспышке.

Математическое моделирование распространения эпидемии COVID-19 в разных странах и регионах мира началось с момента первой вспышки болезни и продолжает проводиться в настоящее время. В самом начале эпидемии проводились исследования, использующие модель логистического роста Ферхюльста [3-6], компартментальные модели типа SIR, SEIR, SIRS [7-14] и др. более сложные модели [15-17]. Они моделировали одну волну распространения заболевания. За последний год появились исследования, использующие принцип многоволнового развития эпидемии [18-20], в том числе учитывающие фактор различной социальной реакции населения [20] и влияние перекрестного иммунитета между штаммами [19]. Однако модели, представленные в этих работах, являются так же модификациями SIR.

Реальные системы (страны, города, регионы) являются открытыми и неоднородными. Статистические данные ежедневных приростов заболе-

вающих часто показывают не один четкий пик волны заболевания, а несколько пиков различной высоты, идущих подряд, «горные гряды», или длинные плоские «плато». Сложные пики говорят о том, что в реальности одна волна эпидемии представляет собой суперпозицию нескольких локальных волн, перехлестывающих друг друга, что представляет большую сложность для моделирования.

Статистические данные также имеют большую погрешность, и каждый дополнительный параметр в модели только увеличивает ошибку. Используя разработанную в [22-26] методику, основанную на дискретном логистическом уравнении, мы, дополнив ее, исследовали распространение эпидемии COVID-19 в европейской стране, применительно к которой данная модель еще не применялась – Великобритании.

1.1 Математическая модель распространения инфекции

Для описания и прогнозирования сложной динамики распространения эпидемий типа коронавируса COVID-19 мы использовали дискретное логистическое уравнение, предложенное Мэйем в 1976 г. [21] и ставшее популярным после работ Фейгенбаума, Шарковского и других ученых. Оно имеет вид:

$$y_{n+1} = \alpha y_n (1 - y_n/N), y_1 = y_0 - \text{задано.} \quad (1)$$

где y_n – общее число рассматриваемых случаев в n -й день; y_1 – начальное число случаев; α – параметр, характеризующий скорость роста популяции; N – емкость популяции. При этом к концу эпидемии общее число заболевших выходит на стационарное значение:

$$\bar{y} = \left(\frac{\alpha-1}{\alpha}\right) \cdot N; y_n \rightarrow \bar{y}, \text{ при } n \rightarrow \infty. \quad (2)$$

Для проведения расчетов модель приводится к безразмерному виду:

$$x_{n+1} = \alpha x_n (1 - x_n), 0 < x_n \leq 1, x_n = y_n/N, x_0 = y_0/N. \quad (3)$$

Для каждой выбранной системы имеются несколько рядов статистических данных – для инфицированных вирусом, выздоровевших, умерших, болеющих в текущий момент (активных). Смерть или выздоровление являются исходами заражения инфекцией. Зараженный находится в «инкубационном периоде». Поэтому рост кумулятивного числа смертей и выздоровлений описывается тем же дискретным уравнением, только со своими параметрами. Для значений емкостей также должно выполняться соотношение:

$$N = N_d + N_r. \quad (4)$$

Для каждого ряда статистических данных отдельно проводятся расчеты с использованием уравнений (5). После рассчитывается число активных случаев q_n по формуле (6) и сравнивается со статистическими данными.

$$u_{n+1} = ru_n(1 - u_n/N_r) - \text{случаи выздоровления в } n + 1\text{-й день}; \quad (5)$$

$$v_{n+1} = dv_n(1 - v_n/N_d) - \text{смертельные случаи в } n + 1\text{-й день.}$$

$$q_n = y_n - v_n - u_n, n = 1, 2, \dots \quad (6)$$

Для расчета распространения инфекции в виде нескольких волн для каждой волны используется несколько уравнений вида (3) со своей емкостью $N^{(i)}$, набором показателей $\alpha^{(i)}$, началом отсчета времени $t_0^{(i)}$ и начальным значением заболевших $x_0^{(i)}$, i – номер волны.

Новая волна начинается, когда расчетные данные начинают отставать от статистических.

$$x_{n+1}^{(i)} = \alpha x_n^{(i)} (1 - x_n^{(i)}), x_0^{(i)} = \frac{y_0^{(i)}}{N^{(i)}}, i = 1..M. \quad (7)$$

1.2. Определение параметров дискретного логистического уравнения по статистическим данным

Рассмотрим сначала калибровку модели для описания одной волны.

1. Как правило, начальное значение y_0 задается равным 1 ($y_0 = 1$).
2. Дискретный интервал времени $\Delta t = t_{n+1} - t_n$ был взят равным 12 часам, т.е. мы пересчитываем данные по модели каждые 12 часов, а с фактическими данными сверяем через 24 часа. Временной интервал влияет на значение коэффициента роста α в модели. Чем он меньше, тем меньше показатель α . Действительно, рассмотрим стадию экспоненциального роста. Тогда общее число случаев в каждый n -й интервал времени описывается формулой:

$$y_{n+1} = \alpha y_n = \alpha^n y_0. \quad (8)$$

Если число случаев заражения вирусом (статистические данные) растет с показателем α , и в модели мы используем временной интервал в 1 день, то $\alpha = \alpha_s$. Если же временной интервал равен 12 часов, то $\alpha = \sqrt{\alpha_s}$, если 8 часов (треть суток) – то $\alpha = \sqrt[3]{\alpha_s}$ и т.д. В уравнении (1) параметр α не должен приближаться к предельному значению 2. Оптимальность выбранного интервала времени в 12 часов подтвердили расчеты, проведенные для разных стран в течение всего прошедшего года.

3. Параметр α . В самом начале развития волны эпидемии на стадии экспоненциального роста, как правило, можно выделить два интервала

[1, ..., M1] и [M1+1, ..., M2], в которых коэффициенты скорости роста различаются. На первом интервале показатель – α_1 , на втором – α_2 . Ширина первого интервала примерно соответствует максимальному инкубационному периоду, ширина второго интервала определяется длительностью стадии экспоненциального роста, когда статистические данные хорошо описываются уравнением (8). (Иногда первый интервал начинается со второй или m -й точки, если статистические данные в этих точках сильно выбиваются из общего графика, то есть содержат большую погрешность.)

Параметр α_2 в уравнении (3) определяется в начале волны на стадии экспоненциального роста по статистическим данным.

$$\frac{X_{s,n+1}}{X_{s,n}} \simeq \alpha_{s,2}, \quad (9)$$

где $X_{s,n}$ – статистические данные, число случаев в n -ый день; α_s – среднее значение отношения (9) за несколько дней. Поскольку в модели число случаев мы пересчитываем через 12 часов, то коэффициент α равен:

$$\frac{x_{n+1}}{x_n} = \alpha = \sqrt{\alpha_s}, \quad x_{n+1} = \alpha x_n, \quad (10)$$

Со временем показатель α может уменьшаться, и мы это отслеживаем из анализа статистических данных и связываем с информацией о введении ограничительных мер. Последнее значение α используется для прогнозирования.

Для коэффициентов справедливо неравенство: $\alpha_1 > \alpha_2$. Интервалы определяются по графикам кумулятивных случаев в логарифмическом масштабе. На этих интервалах статистические данные хорошо ложатся на прямые. По статистическим данным значения коэффициентов $\alpha_{s,1}$, $\alpha_{s,2}$ и начальная точка x_0 находятся методом наименьших квадратов. Поскольку в модели число случаев x_n пересчитывается через 12 часов, то коэффициенты скорости роста в модели равны:

$$\alpha_1 = \sqrt{\alpha_{s,1}}, \alpha_2 = \sqrt{\alpha_{s,2}}. \quad (11)$$

Этот подход хорош в самом начале эпидемии в первую волну, или если предыдущая волна затухла и эпидемия долго находится в эндемической стадии.

3.1 Если же пик волны сложный – «плато» или «горная гряда», то значит, локальные волны идут внахлест, а не последовательно. Поэтому мы уже не можем опираться на рассчитанный методом меньших квадратов первый день x_0 , поскольку мы не можем точно сказать, что и когда привело к новой вспышке – ослабление социальных мер, приезд людей из других городов/стран, развитие нового штамма на исследуемой территории.

Поэтому в таком случае нельзя начинать новую локальную волну только в тот момент, когда расчеты начнут отставать от фактических данных. Разработанный по работам [22-26] программный код для таких случаев был усовершенствован.

Первый день локальной волны тогда выступает свободным подгоночным параметром, подбираемым вручную. Это допущение исходит из того, что мы не можем знать, в какой момент началась новая вспышка. Часто для вспышек со сложным пиком первый день x_0 для $n + 1$ локальной волны брался в середине n -ой волны, что отражает идею того, что $n + 1$ волна уже начинает развиваться, когда n -ая волна в разгаре. Пользуясь этим допущением в сумме с принципом согласования емкостей, на практике нам оказалось гораздо легче описывать волны – стало проще подбирать параметры так, чтобы качественно описать сложные пики. При этом одной волне стало соответствовать не более 1-2-х, реже – 3-х значений параметра α . В случаях волн, идущих внахлест, емкости подбираются для соседних волн вместе, поочередно подправляя значения, исходя из логических соображений: для «плато» емкости локальных волн будут равными друг другу; для «горных гряд» оценивается относительное положение пиков, например, 2:1, 3:1 и т.д.

4. *Параметр N* можно определить в пике эпидемии, если продифференцировать уравнение (2):

$$N \cong X_s^* \left(\frac{2\alpha}{\alpha-1} \right), \text{ или } N \cong \Delta_s^* \frac{4\alpha}{(\alpha-1)^2}, \quad (12)$$

где X_s^* и Δ_s^* – средние значения числа случаев и ежедневных приростов в окрестности пика, соответственно.

До пика, начиная со стадии экспоненциального роста, мы задаем разные предполагаемые значения емкости N , и для разных значений N мы прогнозируем разные сценарии развития эпидемии. Для каждого сценария мы рассчитываем, когда будет пик, каков будет прирост в пике, сколько будет случаев каждый день, когда волна закончится, и сколько заболевших она даст.

Результаты моделирования

Вот уже почти два года мы проводим исследования и делаем прогнозы по динамике распространения эпидемии в разных странах и регионах, начиная с марта 2020 г. [22-25]. Мы моделировали развитие эпидемии в Бразилии, Индии, США, Испании, Италии, Швеции, России, Москве и др. [26]. В настоящей работе впервые представлены результаты моделирования не только общего числа случаев заболевания, но и общего количества смертей, выздоровлений и активных случаев в Великобритании. Статистические данные по развитию эпидемии заражения коронавирусом COVID-19 мы брали с сайта Worldometer [1].

Великобритания. На 20.03.22 Великобритания – пятая по числу случаев заражения COVID-19. [1] За два года распространения заболевания в Великобритании наблюдалось пять больших вспышек. В модели они описаны в виде суперпозиции шести-девяти локальных волн: шесть для случаев смертей, девять – для случаев выздоровления и общих случаев. Различное количество волн связано с тем, число летальных случаев невелико и сильно сдвинуто во времени (до месяца). Расчетные и фактические данные для случаев заражения коронавирусом, для смертельных случаев и случаев выздоровления представлены на рис. *a–f*.

Эпидемия начала развиваться в Великобритании в начале марта 2020 г. Первый год описывается в модели тремя-четырьмя локальными волнами. Три волнами были описаны смертельные случаи. Отметим, что наибольшее количество смертей отнесено к первой, весенней волне, что связывается с невозможностью диагностирования, неумением лечить болезнь и игнорированием пандемии правительством Великобритании.

Вторая волна пошла в рост в начале сентября 2020 г., а третья – в начале декабря 2020 г. Спад зимней волны хорошо описан моделью.

Второй год эпидемии более насыщен: можно выделить две глобальных вспышки, связанные с появлением двух новых штаммов вируса – «дельта» и затем «омикрон».

Четвертая глобальная вспышка идет в рост в июне 2021 г. Видно, что четыре локальные волны «захлестывают друг друга». Так, к примеру, пятая локальная волна общего числа заражений начинается на 373-й расчетный день, а полностью затухнет к 600-му дню. В то время как шестая локальная волна начинается уже на 425-й расчетный день и затухает к 630-му дню. В итоге вклад в пик модели 515-го дня (15.07.21) делают обе локальные волны. Аналогичным образом моделируются и последующие локальные волны.

Пятая вспышка начинается в декабре 2021 г., и она оказывается гораздо более яркой, чем предыдущая.

На графиках видно, что все пики хорошо качественно и количественно описываются, погрешность по каждой из расчетных кривых составляет не более 2%, причем не более 1% за последний год расчетов.

Динамика прироста числа смертных случаев и случаев выздоровления также хорошо описываются выбранной моделью (см. рис. *c–f*). Невооруженным глазом видно, что даты пиков прироста смертельных случаев сдвинуты по отношению к датам пиков приростов выздоровевших (они отмечены на графиках), что не согласуется с SIR-моделью: в то время как пик по кривой общих заражений уже пройден, на кривой выздоровлений данные находятся на пике, а данные на кривой смертей к пику только стремятся.

Зная, как изменяется общее число заболевших, умерших и выздоровевших в модели, можно рассчитать динамику изменения активных случаев по дням:

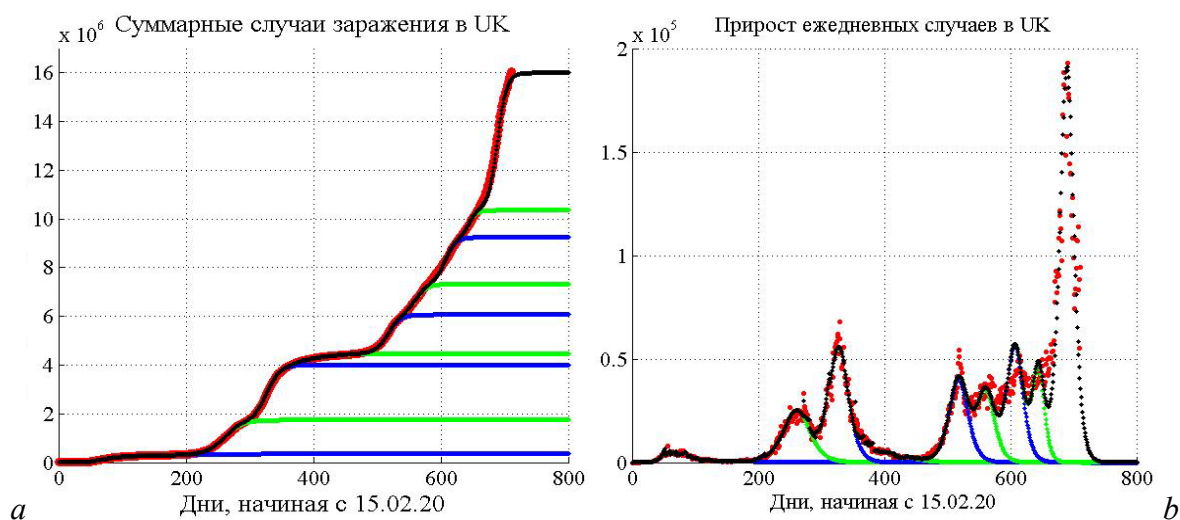
$$q_n = y_n - v_n - u_n, n = 1, 2, \dots,$$

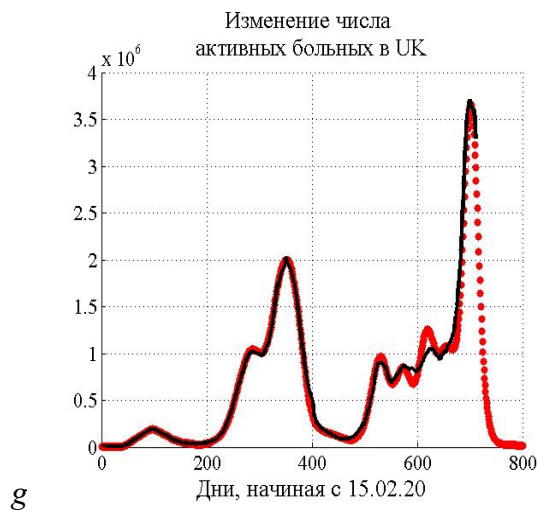
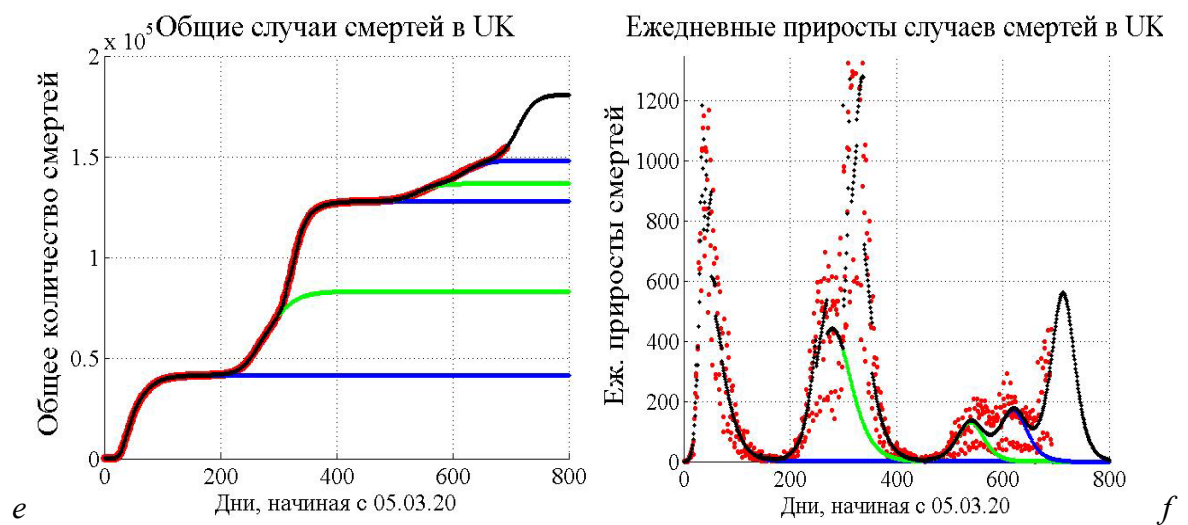
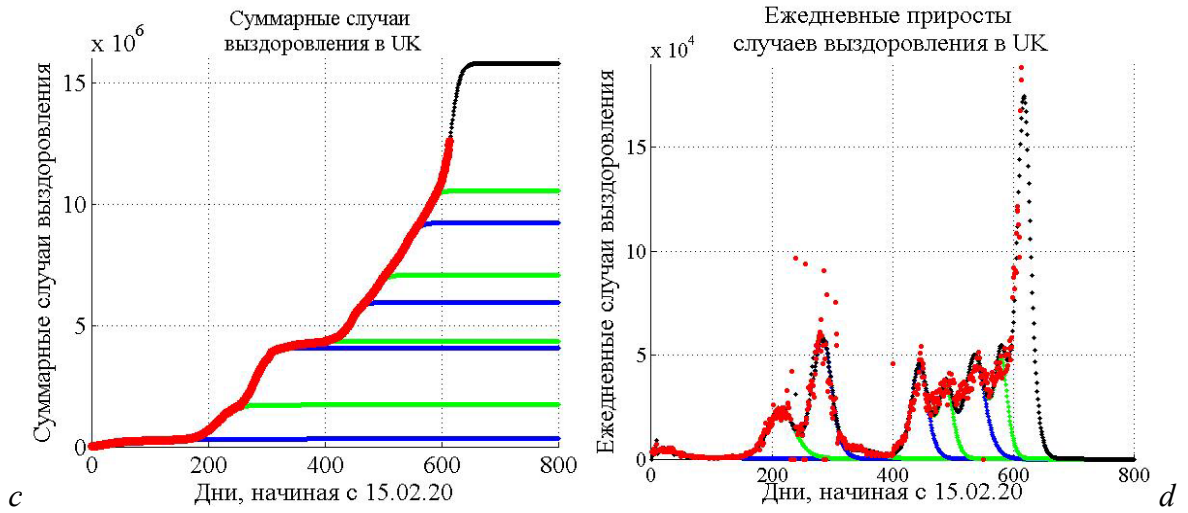
где q_n — число активных случаев в n -й день, y_n , v_n , u_n — число общих случаев, число смертей и выздоровлений соответственно.

Емкости, разделенные по интервалам согласования

Интервал	Заражено, млн чел. I	Выздоровело, млн чел. R	Умерло, млн чел. D	Разница, млн чел. I - R - D
До 500-го дня	211	205.55	5.45	4.45
С 500-го дня до конца расчета	164	160	1.7	2.6
Всего	375	365.55	7.15	7.05

Соответствие расчетного числа активных случаев со статистическими данными на отрезке наблюдения и стремление прогнозного значения активных случаев к нулю по окончании последней волны является критерием правильности проведенного моделирования. В таблице представлена сводная таблица емкостей, причем показано, что согласование емкостей по соотношению (4) проводилось на двух промежутках, исходя из допущения, что заболевание успело достаточно утихнуть. На рис. *g* мы видим прекрасное совпадение расчетных значений активных случаев с фактическими данными.





Динамика распространения эпидемии коронавируса в Великобритании между 15.02.2020 и 25.01.2022. Фактические данные и результаты моделирования, прогноз. Красным обозначены фактические данные. Синим обозначены нечетные локальные волны, зеленым – четные

Литература

1. <https://www.worldometers.info/coronavirus/>
2. <https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/>
3. *Martelloni G., Martelloni G.* Analysis of the evolution of the Sars-Cov-2 in Italy, the role of the asymptomatics and the success of Logistic model // *Chaos, Solitons and Fractals* 140 (2020) 110150
4. *Zheng X.P., Li J., Zhu B.* Prediction of epidemic trends in COVID-19 with logistic model and machine learning technics // [Chaos, Solitons and Fractals 139, 110058 \(2020\).](#)
5. *Wu K., Darcet D., Wang Q., Sornette D.* Generalized logistic growth modeling of the COVID-19 outbreak in 29 provinces in China and in the rest of the world. [DOI:10.1101/2020.03.11.20034363](https://doi.org/10.1101/2020.03.11.20034363)
6. *Guo J.* theoretical epidemic laws based on data of COVID-19 pandemic // medRxiv. – 2020.
7. *Kermack W.O., McKendrick A.G.* A contribution to the mathematical theory of epidemics // *Proceedings of the Royal Society of London Series A, Containing Papers of a Mathematical and Physical Character* 1927; 115(772):700–721.
8. *Marinov T.T., Marinova R.S.* Dynamics of COVID-19 using inverse problem for coefficient identification in SIR epidemic models // *Chaos, Solitons & Fractals*. X5, 100041 (2020).
9. *Li M.Y., Muldowney J.S.* Global stability for the SEIR model in epidemiology // *Mathematical biosciences* 125(2), 155-164 (1995).
10. *Odagaki T.* Analysis of the outbreak of COVID-19 in Japan by SIQR model // *Infectious Disease Modelling* 5, 691e698 (2020).
11. *Neves A.G.M., Guerrero G.* Predicting the evolution of the COVID-19 epidemic with the A-SIR model: Lombardy, Italy and São Paulo state, Brazil // *Physica D* 413, 132693 (2020).
12. *Comunian A., Gaburro R., Giudici M.* Inversion of a SIR-based model: A critical analysis about the application to COVID-19 epidemic // *Physica D* 413, 132674 (2020).
13. *Contreras S., Villavicencio H.A. et al.* A multi-group SEIRA model for the spread of COVID-19 among heterogeneous populations // *Chaos, Solitons and Fractals* 136, 109925 (2020).
14. *Agrawal M., Kanitkar M., Vidyasagar M.* SUTRA: An Approach to Modelling Pandemics with Asymptomatic Patients, and Applications to COVID-19. [arXiv:2101.09158](https://arxiv.org/abs/2101.09158)
15. *Ramos A.M., Ferrández M.R., Vela-Pérez M., Kubik A.B., Ivorra B.* A simple but complex enough θ -SIR type model to be used with COVID-19 real data. Application to the case of Italy // *Physica D* 412, 132839 (2020).
16. *Avila-Ponce U. de León, Pérez Á.G.C., Avila-Vales E.* An SEIARD epidemic model for COVID-19 in Mexico: Mathematical analysis and state-level forecast // *Chaos, Solitons and Fractals* 140, 110165 (2020).

17. *Higazy M.* Novel fractional order SIDARTHE mathematical model of COVID-19 pandemic // *Chaos, Solitons and Fractals* 138, 110007 (2020).
18. *Ghosh K., Ghosh A.K.* Study of COVID-19 epidemiological evolution in India with a multi-wave SIR model. [arXiv:2202.04917](https://arxiv.org/abs/2202.04917).
19. *Arruda E.F. et al.* Reinfection and low cross-immunity as drivers of epidemic resurgence under high seroprevalence: A model-based approach with application to Amazonas, Brazil. [arXiv:2106.15928](https://arxiv.org/abs/2106.15928).
20. *Kastalskiy I.A. et al.* Social stress drives the multi-wave dynamics of COVID-19 outbreaks // *Scientific reports* 11(1), 1-12 (2021).
21. *May R.* Simple mathematical models with very complicated dynamics // *Nature* 261, 459-467 (1976).
22. *Koltsova E.M., Kurkina E.S., Vasetsky A.M.* Mathematical Modeling of COVID-19 coronavirus spread in Moscow // *Computational nanotechnology* 7(1), 99-105 (2020).
23. *Koltsova E.M., Kurkina E.S., Vasetsky A.M.* Mathematical Modeling of COVID-19 coronavirus spread in several European, Asian countries, Israel and Russia // *Problems of economics and law practice*. 2020. Moscow. N2. p.154-165.
24. *Koltsova E.M., Kurkina E.S., Vasetsky A.M.* Mathematical modeling of the spread of COVID-19 in Moscow and Russian regions. [arXiv:2004.10118](https://arxiv.org/abs/2004.10118).
25. *Koltsova E.M., Kurkina E.S., Vasetsky A.M.* Mathematical modeling of the spread of waves of the COVID-19 coronavirus epidemic in different regions of the world. [arXiv:2007.02283](https://arxiv.org/abs/2007.02283).
26. *Куркина Е.С., Кольцова Е.М.* Математическое моделирование распространения волн эпидемии коронавируса COVID-19 в разных странах мира // *Прикладная математика и информатика* №66. – М.: Изд-во факультета ВМК МГУ, 2021. С.41-66.