

Для цитирования: Рослякова Н.А. Коронавирус: уточнение прогнозов и новые оценки [Электронный ресурс]. URL: [http://roslyakova24.ru/statja\\_1.pdf](http://roslyakova24.ru/statja_1.pdf)

### Коронавирус: уточнение прогнозов и новые оценки

Несомненно, тема коронавируса и характера его распространения в разных странах является весьма интересной. Более того моделирование процессов заболеваемости и смертности от коронавируса в разных странах позволяют получить некоторые ориентиры для анализа систем управления кризисными ситуациями.

Ранее нами была предпринята попытка оценить динамику заболеваемости и смертности на данных (по 14 апреля, включительно) по миру в целом, Китаю и России на базе модели логистической кривой (кривая Ферхюльста) [1]. Более того на основе динамики Китая, где процесс распространения коронавируса протекал с заметным опережением, относительно России удалось обосновать достаточное количество наблюдений для получения относительно точных оценок.

Данная работа имеет целью уточнить полученные для России, Китая и мира прогнозы и представить аналогичные прогнозы для ряда других стран. Более того, во-первых, в этой работе мы оперируем более полной статистикой: изначально данные были найдены только с 20 января, сейчас для Китая и мира собрана статистика с 1 января, источником послужил сайт [2]. Во-вторых, исследование дополнилось 10 свежими наблюдениями (с 15 по 24 апреля), что позволит уточнить оценки по пикам распространения болезни и смертности. И, в-третьих, данная работа будет дополнена рядом параметров, которые получаются уже на основе оцененных коэффициентов модели Ферхюльста.

Формула для оценки параметров логистической кривой:

$$Y(t) = \frac{k_2}{1+10^{a+bt}} \quad (1)$$

основные предположения заключаются в том, что изначальный уровень события нулевой, это определяет отсутствие параметра  $k_1$ .  $Y(t)$  значение функции (переменная смертность (death) или заболеваемости (morbid));  $t$  – экзогенная переменная времени (периоды от 1 до  $i$ );  $k_2$  – верхняя асимптотами логистической кривой; параметры  $a$  и  $b$  определяют крутизну наклона и положение точки перегиба логистической кривой. Подробнее о выводе данной спецификации в [1].

#### Данные по заболеваемости

С помощью нелинейного оценивания программного комплекса Statistica оцениваются параметры уравнения (1)  $a$ ,  $b$ ,  $k_2$ . Полученные оценки представлены в таблице 1, также они дополнены количеством наблюдений ( $n$ ), поскольку эпидемия в разных странах стартовала не одновременно и коэффициентом детерминации ( $R$ ).

Таблица 1 – Модельные оценки уровня заболеваемости в разных странах

Страны	Параметры модели			R	n
	$k_2$	$a$	$b$		
Мир	3292555	4.7640	-0.0462	0.9991	116
Китай	82006	3.8431	-0.0950	0.9991	115
США	940761	5.4869	-0.0667	0.9991	96
Франция	123375	4.6971	-0.0653	0.9995	92
Канада	47779	4.8061	-0.0606	0.9984	91
Германия	149208	4.6503	-0.0690	0.9992	89
Великобритания	159358	4.7720	-0.0642	0.9996	86
Италия	189058	3.3426	-0.0545	0.9986	86

Испания	207605	4.3454	-0.0681	0.9986	85
Швеция	20892	3.5454	-0.0481	0.9995	85
Россия	134090	6.4056	-0.0748	0.9999	85
Бельгия	46388	4.0810	-0.0612	0.9990	82
Иран	93267	2.1550	-0.0489	0.9993	66
Бразилия	74464	3.1949	-0.0576	0.9993	60
Швейцария	28111	2.5431	-0.0721	0.9992	60
Норвегия	7211	2.1336	-0.0653	0.9991	59
Мексика	25465	3.0632	-0.0519	0.9991	57
Эквадор	12987	2.6664	-0.0603	0.9959	56
Индия	34657	3.2831	-0.0658	0.9995	54
Украина	11646	3.3835	-0.0689	0.9993	53
Перу	26877	3.5898	-0.0815	0.9987	50
Турция	116648	2.5254	-0.0734	0.9993	45
Казахстан	3976	2.0906	-0.0530	0.9982	42

Из таблицы 1 видно, что все модели имеют высокую объясняющую способность ( $R$  более 0.99). Также следует учитывать, что страны в таблице упорядочены по продолжительности распространения коронавируса в них. То есть, для стран, расположенных в начале списка уже состоялся выход на плато, поэтому полученные оценки максимально точно описывают имеющуюся картину. Для стран, расположенных в конце таблицы 1, напротив, эпидемия находится на начальном этапе, поэтому и точность прогноза по логистической кривой может оказаться ниже. При этом отметим, что отладка методики на данных Китая позволила определить, что для получения относительно точных оценок необходимо 31-35 наблюдений и более. Для всех рассматриваемых стран уже имеется большее количество наблюдений.

Если интерпретировать коэффициенты моделей, то интерес представляет параметр  $b$ , который определяет максимальная скорость распространения эпидемии (она пропорциональна  $|b|$ ). Чем больше данный параметр по модулю, тем стремительнее распространяется эпидемия. В таблице 1 по данному показателю выделяются: Китай, Перу, Росси, Турция, Швейцария. Однако здесь может быть два возможных вектора. В одной стороны, может иметь место действительно быстрое распространение коронавируса. А, с другой стороны, лидерство в данном параметре может быть обусловлено совершенством процедуры выявления больных и высокой степенью охвата населения тестовыми процедурами. Вероятно, такая ситуация может иметь место в России и Швейцарии, которые вышли в лидеры (среди относительно крупных по населению стран) по числу проведённых тестов на 1 млн. чел. Также могут представлять интерес оценки потенциального охвата населения некоторой страны эпидемией коронавируса, для чего полученную оценку  $k_2$  соотнесём с численностью населения страны и отдаленность пика эпидемии от его начала, которая вычисляется через отношение параметров  $a$  и  $b$ , эти оценки приравнены в таблице 2.

Таблица 2 – Параметры, рассчитанные на основе полученных моделей заболеваемости

Страны	$k_2$ /на млн. населения	$-a/b$	$n$
Мир	423.2	103	116

Китай	58.4	40	115
США	2860.3	82	96
Франция	1838.7	72	92
Канада	1260.7	79	91
Германия	1793.4	67	89
Великобритания	2360.9	74	86
Италия	3135.3	61	86
Испания	4445.5	64	85
Швеция	2028.3	74	85
Россия	913.4	86	85
Бельгия	4033.7	67	82
Иран	1121.0	44	66
Бразилия	354.3	55	60
Швейцария	3268.7	35	60
Норвегия	1335.4	33	59
Мексика	201.1	59	57
Эквадор	768.5	44	56
Индия	25.6	50	54
Украина	279.3	49	53
Перу	819.4	44	50
Турция	1402.0	34	45
Казахстан	212.6	39	42

Из таблицы 2 можно видеть страны, которые наиболее подвержены распространению коронавируса. Здесь выделяются: США, Испания, Бельгия, Швейцария, Италия. При этом, согласно моделям, для них и большинства из рассмотренных стран пик распространения миновал. Это можно видеть если сопоставить параметр  $(-a/b)$ , который отражает период, в который ожидается максимальный прирост заболеваемости, с количеством имеющихся наблюдений ( $n$ ). Однако для некоторых стран пик ещё не пройден и должен последовать буквально в ближайшие дни. Среди таких стран Россия и Мексика. Уточнённый прогноз для России будет представлен ниже.

Изначально для динамики заболеваемости в России была получена модель (см. источник [1]):

$$Morbid_{Ru} = \frac{81543}{1+10^{5.145-0.0768t}} \quad (2),$$

сейчас на более полных данных возможно выписать уточнённую модель:

$$Morbid_{Ru} = \frac{134090}{1+10^{6.406-0.0748t}} \quad (2')$$

существенное увеличение параметра  $k_2$  может быть вызвано, во-первых, сменой источника статистики. Так для модели (2) источник статистики [3] фиксирует заболеваемость в России с 14-15 февраля, тогда как модель (2'), отстроена по более полным данным источника [2], который фиксировал заболеваемость в России ещё 1-2 февраля. Во-вторых, влияние оказали дополненные наблюдения о периоде с 15 по 24 апреля. Ниже на рисунке 1 представлены прогнозы по моделям (2) и (2'), а также фактические данные.

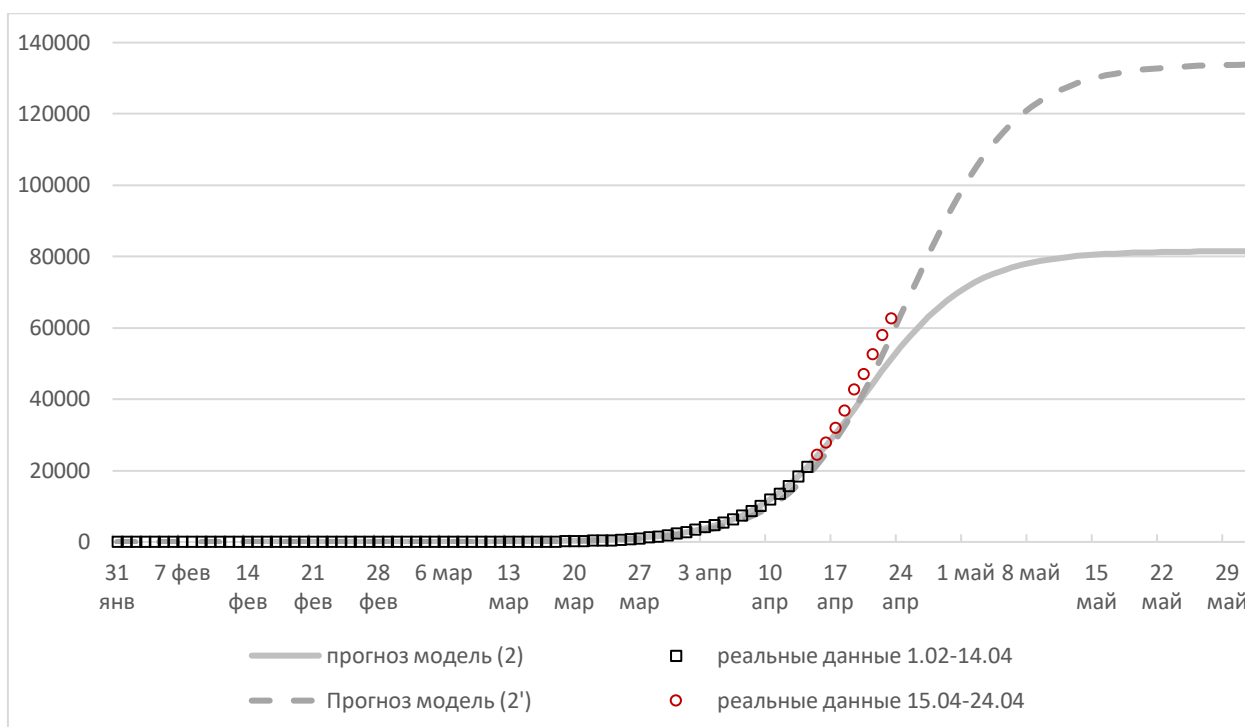


Рисунок 1 – Первичный (2) и уточнённый (2') прогноз заболеваемости в России

Из рисунка 1 хорошо видно, что основной вид кривой, связанный с параметром  $b$  (скоростью распространения явления) в двух моделях соответствует. Также несущественно сместилась и точка перегиба. Тем не менее, следует отметить, что новые наблюдения скорректировали прогноз в сторону повышения, что было характерно и при уточнении других прогнозов (например, по Китаю см. [1]).

#### Данные по смертности

Аналогичные оценки были получены и для параметра смертность от коронавируса (death), полученные оценки представлены в таблице 3.

Таблица 3 – Модельные оценки уровня смертности в разных странах

Страны	Параметры модели			R	n
	$k_2$	a	b		
Мир	243226	5.0610	-0.0528	0.9997	106
Китай	3292	2.5340	-0.0662	0.9995	98
Франция	22844	4.5072	-0.0795	0.9993	71
Иран	5600	2.1665	-0.0510	0.9982	66
Италия	25613	2.3928	-0.0581	0.9984	63
США	61881	3.5901	-0.0746	0.9994	56
Испания	21853	2.4125	-0.0739	0.9983	52
Великобритания	20418	3.1805	-0.0802	0.9994	51
Швейцария	1307	2.3853	-0.0676	0.9978	51
Канада	3308	3.1886	-0.0732	0.9996	47
Германия	6161	2,5394	-0,0692	0.9988	47
Швеция	2587	2.5083	-0.0663	0.9968	45
Бельгия	7181	2.8147	-0.0826	0.9997	45
Норвегия	198	2.1853	-0.0683	0.9991	44
Индия	891	2.9319	-0.0793	0.9992	44

Эквадор	618	2.1614	-0.0695	0.9979	43
Бразилия	4461	2.4224	-0.0719	0.9988	39
Украина	253	2.2992	-0.0734	0.9994	38
Турция	3091	2.0943	-0.0701	0.9988	38
Перу	1534	2.3893	-0.0585	0.9981	37
Мексика	2273	2.5812	-0.0697	0.9982	36
Россия	1260	2.4222	-0.0776	0.9996	30
Казахстан	22	1.4735	-0.0969	0.9942	27

Из таблицы 3 можно видеть страны с наибольшим ожидаемым количеством погибших от коронавируса. В их числе: США, Франция, Италия, Испания, Великобритания. Относительно данных моделей также справедливо мнение, что с повышением количества наблюдений будет корректироваться прогноз, и более существенные изменения будут характерны для стран, расположенных ближе к концу таблицы, так как переломного момента в динамике там ещё не наступило. Если оценивать рост смертности, то выделяются: Казахстан, Бельгия и Великобритания. Однако здесь стоит отметить, что Казахстан находится в начальной фазе распространения эпидемии, что обуславливает самый высокий коэффициент. По мере накопления наблюдений он будет снижаться. С другой стороны, показатели для Бельгии и Великобритании могут считаться более показательными, так как там накоплено существенно больше наблюдений. То есть, можно говорить о действительно большем росте смертности относительно похожих стран, например, Швеция и Швейцария, где смертность фиксируется соответствующее количество дней, 45 и 51 день, соответственно. Оценки смертности на 1 млн. населения страны и отдаленность пика эпидемии представлены в таблице 4.

Таблица 4 – Параметры, рассчитанные на основе полученных моделей смертности

Страны	$k_2$ /на млн. населения	-a/b	n
Мир	31.3	96	106
Китай	2.3	38	98
Франция	340.4	57	71
Иран	67.3	42	66
Италия	424.8	41	63
США	188.1	48	56
Испания	467.9	33	52
Великобритания	302.5	40	51
Швейцария	152.0	35	51
Канада	87.3	44	47
Германия	74.1	37	47
Швеция	251.2	38	45
Бельгия	624.4	34	45
Норвегия	36.7	32	44
Индия	0.7	37	44
Эквадор	36.6	31	43
Бразилия	21.2	34	39

Украина	6.1	31	38
Турция	37.2	30	38
Перу	46.8	41	37
Мексика	18.0	37	36
Россия	8.6	31	30
Казахстан	1.2	15	27

Выделяются страны, где не преодолён пик роста по смертности – Россия, Мексика и Перу. Для этих стран пик смертности прогнозируется в ближайшие дни. Исключительно высокий уровень смертности наблюдается в Бельгии, Испании, Италии, Франции и Великобритании.

Также следует отметить одну деталь по поводу данных о смертности в Китае. Так, смертность росла до 16 апреля, а 17 апреля статистика была дополнена 1290 случаями, то есть произошёл резкий скачок. Этот факт был объяснен как уточнение статистики, то есть были переклассифицированы 1290 случаев смерти в прошлые периоды и записаны в статистические данные как смерть от коронавируса за 17 апреля [4]. Поскольку нет возможности распределить эти случаи смерти по конкретным датам (хотя это представляется возможным и логичным), то мы приняли решение ограничить статистику по Китаю 16 апреля и не учитывать данный выброс. Ранее в источнике [1] по усечённым данным была получена модель для Китая:

$$Death\_Ch = \frac{3177}{1+10^{1.847-0.0722t}} \quad (3),$$

сейчас на основе более полных данных оценена новая модель:

$$Death\_Ch = \frac{3292}{1+10^{2.534-0.0662t}} \quad (3')$$

Сопоставление двух прогнозов и реальных данных по Китаю представлены на рисунке 2.

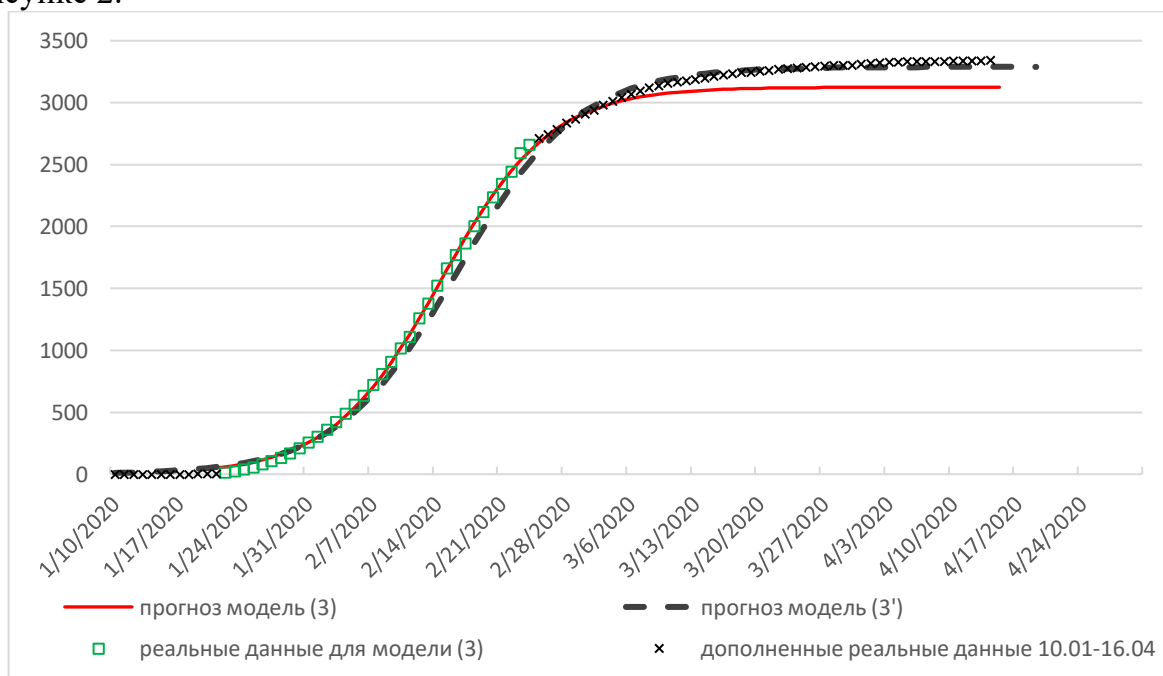


Рисунок 2 – Первичный (3) и уточнённый (3') прогноз смертности в Китае

Из рисунка 2 видно, что на основе первых 34 наблюдений (реальные данные за 20.01-24.02) удалось построить довольно точную модель (3). Дополнение данными позволило уточнить оценки, которые также к в случае с

заболеваемостью для России больше касались корректировки параметра  $k_2$ . Аналогичные уточнения для России представлены ниже.

В работе (1) был получен прогноз уровня смертности (4) (см. рисунок 3):

$$Death_{Ru} = \frac{622}{1+10^{2.773-0.0809t}} \quad (4),$$

который был уточнён на основе данных за 15-24 апреля и получена новая модель:

$$Death_{Ru} = \frac{1260}{1+10^{2.422-0.0776t}} \quad (4')$$

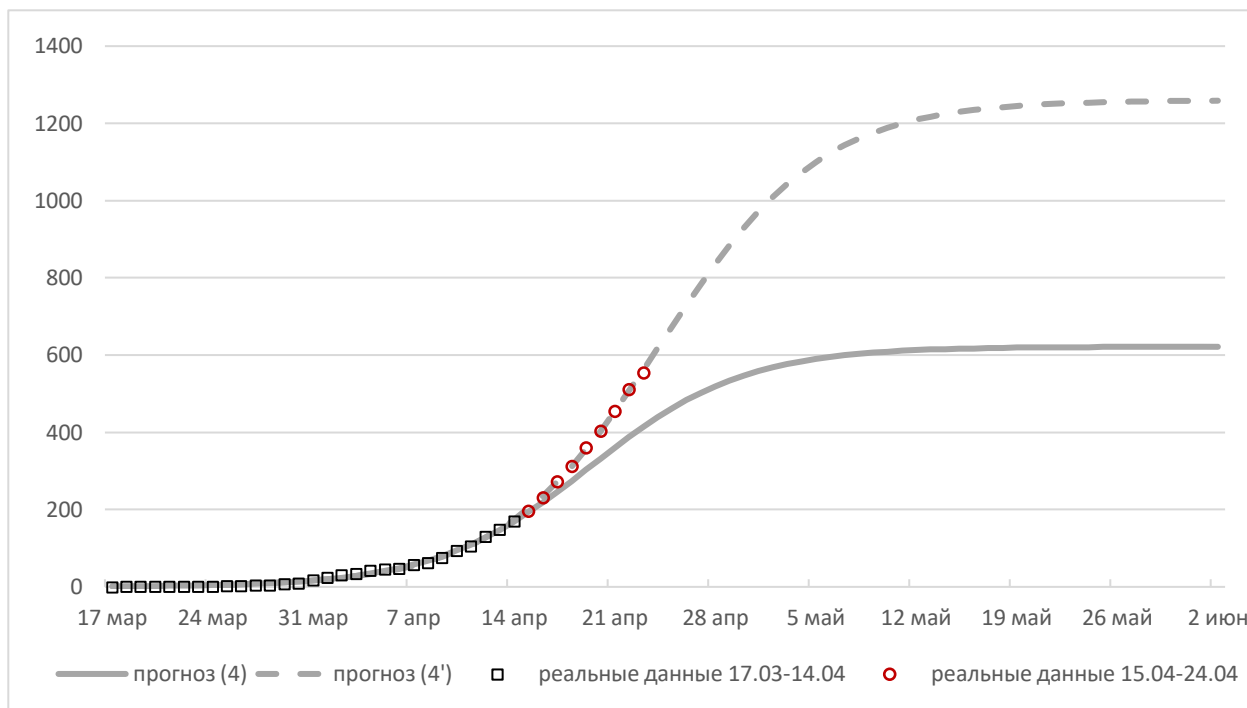


Рисунок 3 – Первичный (4) и уточнённый (4') прогноз смертности в России

Здесь мы также наблюдаем повышение предельного уровня смертности, однако нужно заметить, что по состоянию на 24 апреля имелось только 30 наблюдений, что в случае оценок для Китая приводило к погрешности 10-13%. То есть реальные показатели смертности могут оказаться ещё несколько выше.

### Приложения

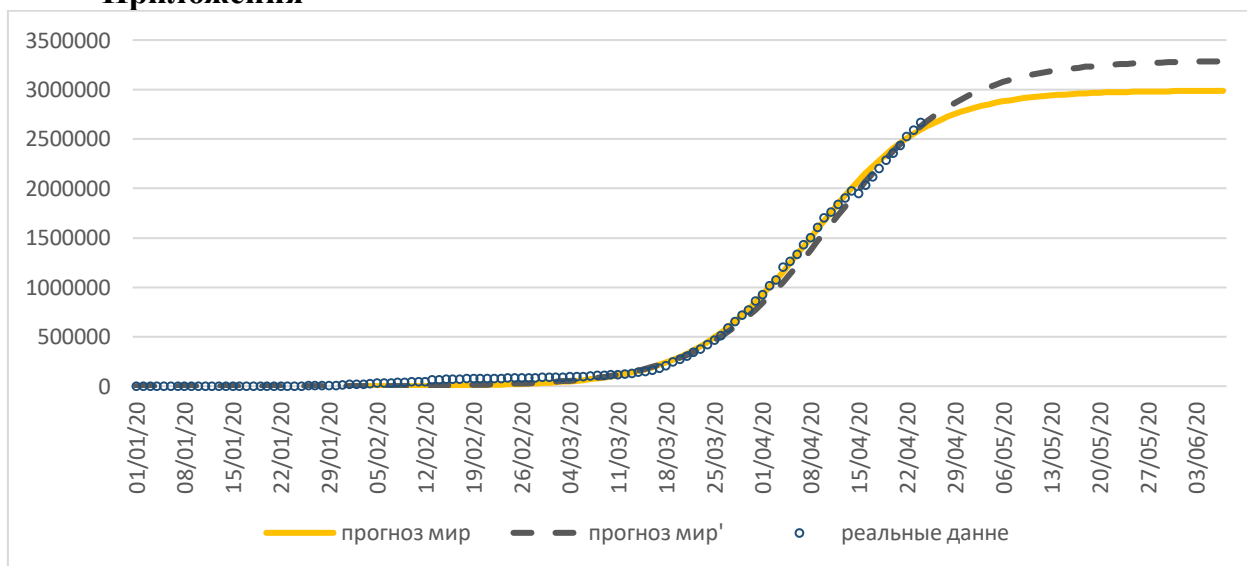


Рисунок 4 – Первичный (прогноз мир) и уточнённый (прогноз мир') прогноз заболеваемости в мире в целом

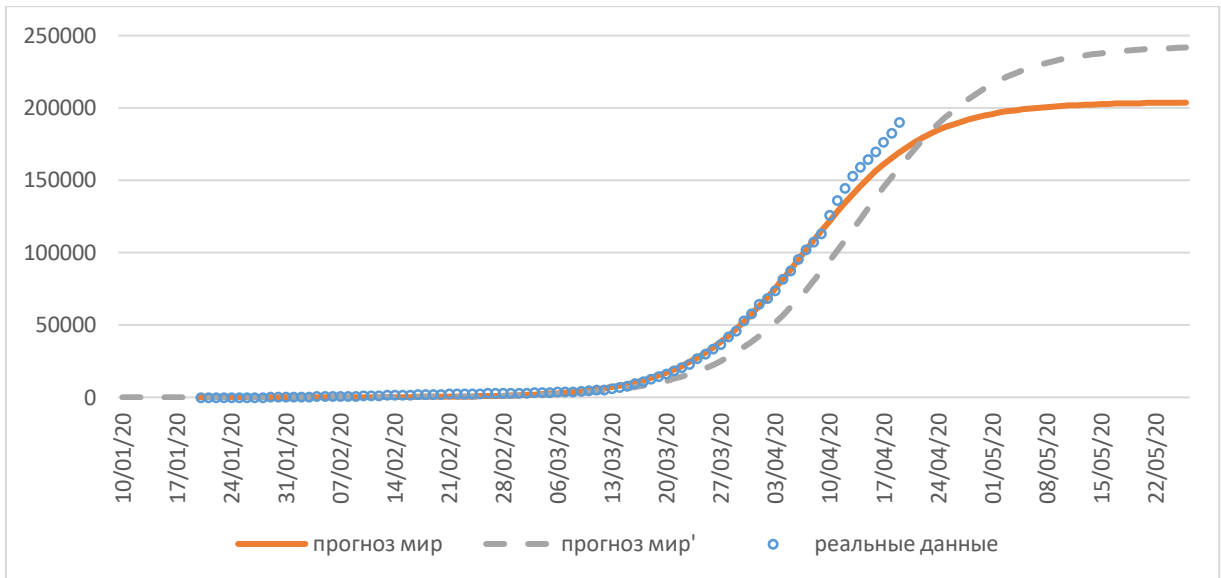


Рисунок 5 – Первичный (прогноз мир) и уточнённый (прогноз мир') прогноз смертности в мире в целом

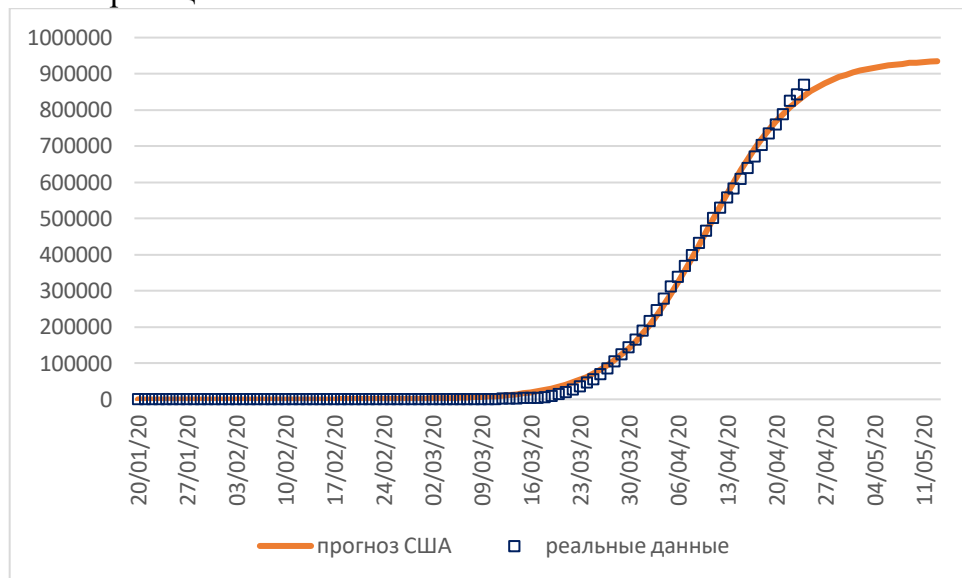


Рисунок 6 – Прогноз заболеваемости в США

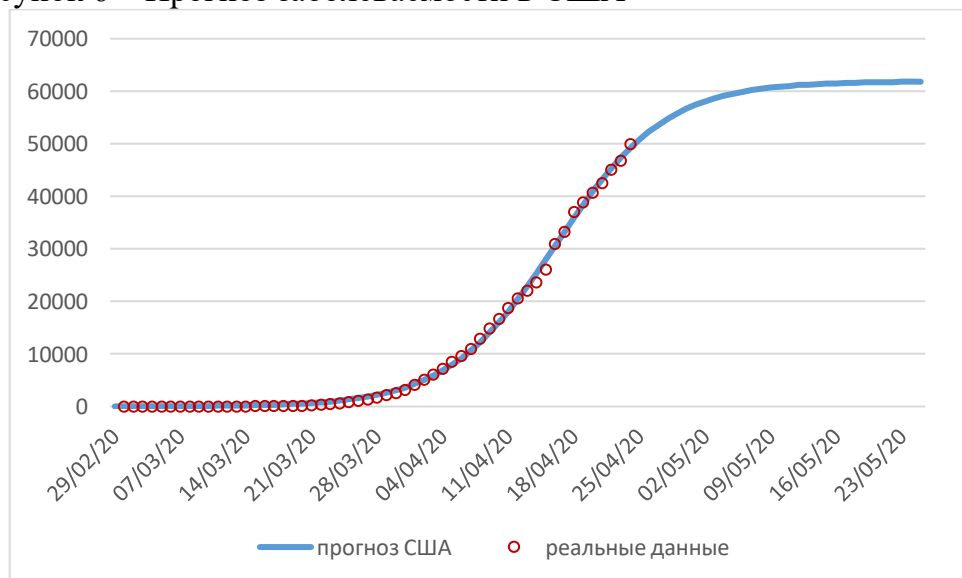


Рисунок 7 – Прогноз смертности в США



Литература:

1. Нижегородцев Р.М., Рослякова Н.А., Горидько Н.П. Прогноз распространения коронавируса в России: свет в конце туннеля [Электронный ресурс, 27.04.2020]. URL: [http://inf-econ.ru/\\_ld/1/105\\_\\_-1-.pdf](http://inf-econ.ru/_ld/1/105__-1-.pdf)
2. Statistics and Research. Coronavirus Disease (COVID-19) [Электронный ресурс, 24.04.2020]. URL: <https://ourworldindata.org/coronavirus>
3. Онлайн-мониторинг коронавируса в мире. Данные ВОЗ по ситуации с коронавирусом в мире на сегодня [Электронный ресурс, 24.04.2020]. URL: <https://coronavirus-monitoring.ru/>
4. Апарина Э. [Электронный ресурс, 27.04.2020]. URL: <https://www.spb.kp.ru/daily/27119/4200091/>