

**ОЦЕНКА УРОВНЯ СМЕРТНОСТИ, ПОРОЖДЁННОЙ SARS-COV-2 В
РЕГИОНАХ РОССИИ, НА ОСНОВЕ ЭКОНОМЕТРИЧЕСКОЙ
МОДЕЛИ**

Степанов В. С.

Федеральное госбюджетное учреждение науки Центральный экономико-
математический институт Российской академии наук

**ASSESSING SARS-COV-2-RELATED MORTALITY RATE IN RUSSIAN
REGIONS, BASED ON THE ECONOMETRIC MODEL**

Stepanov V. S.

The Central Economics and Mathematics Institute of the Russian Academy of
Sciences

Резюме. Объектами исследования были официальные данные по заболеваемости и смертности от COVID-19 в российских регионах, а также региональные медико-демографические и экологические данные за последние годы. Математическими методами работы являются корреляционный и регрессионный анализ, методы проверки статистических гипотез. Сначала специфицируется модель регрессии, имеющая переменную структуру; свободный член в ней отличается от региона к региону в зависимости от сочетания значений ряда фиктивных переменных (*dummy*). На роль зависимой переменной Y^t выбиралась смертность, привязанная в сводной таблице оперативного штаба к суткам с отсчётом t , причём основной причиной смерти считалась COVID-19. Комплекс объясняющих переменных включал две факторные переменные, которые изменяются ежедневно, причём с лагами относительно t . Также в этот комплекс входил ещё ряд переменных, которые не изменялись во времени: показатель обеспеченности региона врачами ряда специальностей и четыре *dummy*-переменные. Одна из них кодировала южную принадлежность региона: ЮФО или СКФО. Три прочие характеризовали повышенное загрязнение атмосферного воздуха в населенных пунктах за последние годы, а также – радиационное загрязнение территории и здоровье жителей регионов по 10 классам болезней (для системы кровообращения, эндокринной и проч.). Значения этих *dummy* переменных были получены по открытым данным Росстата и др. Параметры модели оценивались методом наименьших квадратов по обучающей таблице, которая включала 40 регионов, где отсчёт t переменной Y^t входил за дату 1.xi.2021. В итоге построена статистическая модель с ошибкой аппроксимации 3%; для $\frac{3}{4}$ регионов из этой таблицы она оказалась 1.94 (± 1.5)%. Приводятся графики с ежедневными прогнозами смертности от COVID-19 в первой половине ноября 2021г для семи регионов в сравнении фактом. Модель может быть полезной при разработке медико-демографической политики в регионах, а также при построении уточнённых

моделей секторного типа, состоящих из систем дифференциальных уравнений (SEIRF, SIRD и др.).

Ключевые слова: модель регрессии, загрязнение воздуха, SARS-CoV-2, моделирование смертности, заболеваемость населения, общественное здоровье, прогноз смертности, регионы России.

Abstract. The objects of the study were the daily data on the population morbidity and mortality due to coronavirus disease 2019 (COVID-19) in Russian regions, as well as regional medical, demographic and environmental data recorded in recent years. COVID-19 is a contagious disease caused by the novel coronavirus (SARS-CoV-2). The mathematical methods consist of correlation and regression analysis, methods of testing statistical hypotheses. First, a multiple Variable Structure Regression should be specified. The intercept in the model differs from region to region, depending on the combination of values for dummy variables. The role of the dependent variable Y^t was chosen as the cumulative mortality published by the operational headquarters for the regions that has been linked to day t , so that COVID-19 was considered the main cause of death. The complex of explanatory variables included two factorial variables that changed daily, and had a lag relative to t value. Also, this complex included a number of variables that did not change with the growth of t : the explanatory variable with the region's availability with doctors of certain specialties; and four dummy variables. One of them coded the region's belonging to the two southern Russian Federal Districts. Three other variables characterized the increased air pollution in settlements recorded in recent years, as well as the level of radiation pollution of the region's territory and the population health estimated for 10 classes of diseases (for the circulatory system, endocrine system, etc.). The values of such dummy variables were obtained from open data from the Federal State Statistics Service (Rosstat)

etc. The model parameters were estimated by the least squares method using the training table, which included 40 Russia's regions, the t parameter for variable Y^t was assessed starting from November, 1, 2021. As a result, a statistical model was built with an approximation error equal to 3%. For $\frac{3}{4}$ regions of the regions examined this error was 1.94 (± 1.5)% for the value Y^t that has been fixed on the 1st Nov. The plots show daily prediction for mortality rate due to COVID-19 in the first half of November for seven Russian regions compared with actual data. The model can be useful in development of medical and demographic policy in geographic regions, as well as generating adjusted compartment models that based on systems of differential equations (SEIRF, SIRD, etc.).

Keywords: regression model, air pollution, SARS-CoV-2, mortality modeling, morbidity, public health, mortality forecast, Russia's regions.

1 Прогнозирование смертности от COVID-19 на территориях стран и их
2 регионов является актуальной проблемой. Это важно при разработке
3 оперативных профилактических мер в регионах, а также – решений по
4 снижению смертности; последнее – стратегическая цель правительства РФ.

5 Вопросы математического моделирования в эпидемиологии сжато
6 освещаются в [1]; небольшой обзор подходов к моделированию
7 заболеваемости и смертности от COVID-19 приводится в [2]. Дополняя его,
8 отметим, что эконометрическое моделирование этой смертности выполнено в
9 [3]; к временным рядам также применялся подход Бокса и Дженкинса [4].
10 Сопоставление точности моделей, созданных известными научными
11 центрами, сделано в [5]. Из весьма успешных работ по моделированию
12 смертности в регионах можно выделить [6-8].

13 Ряд исследователей в качестве инструментария выбирают
14 регрессионный анализ [2, 3, 9-11]. Так, в [9] рассматривалась “избыточная
15 смертность” по данным ЗАГС и строились линейные модели; в результате на
16 тех или иных отрезках пандемии наиболее информативными оказались
17 различные факторы. Высокая информативность одного из показателей,
18 связанных со смертностью от COVID-19, была выявлена в [10]. Этот
19 показатель рассчитывается как доля летальных случаев среди случаев с
20 известным исходом (умерших и выздоровевших), которые ранее входили в
21 список зарегистрированных лиц. Он был назван Э.М. Гольштейном
22 “коэффициентом летальности”, поэтому ниже он именуется нами как
23 “индикатор Гольштейна” (ИГ) (использование “летальности” является здесь
24 нестандартным, согласно О.М. Драпкиной, et al. (2020)). Также важными
25 оказались распространённость активных случаев и обеспеченность населения
26 врачами-специалистами [2]. Влияние загрязнённости атмосферного воздуха
27 на заболеваемость изучалось в [11].

28 *Целью* исследования является построение (с оценкой по точности)
29 модели линейной регрессии, которая связывает смертность от COVID-19 в

30 регионах с комплексом объясняющих факторных переменных. На основе
31 этой модели можно выполнять краткосрочный прогноз смертности и более
32 эффективно планировать медицинские ресурсы.

33 Материалы и методы

34 Регионами, попавшими в таблицу с обучающими данными, были: 14
35 областей ЦФО, 2 региона из СЗФО и 4 – из ПФО; в состав южных округов
36 входило 11, а 9 оставшихся находятся на Урале, Дальнем Востоке и в
37 Сибири. В выбранных регионах проживает почти $\frac{1}{2}$ населения страны; в
38 итоге таблица включала $n = 40$ строк. В таблицу не вошли города
39 федерального уровня, ряд промышленных регионов (Липецкая, Вологодская
40 и многие другие), а также – территории из Арктики, ряд регионов СКФО,
41 Сибири, Дальнего Востока, Урала.

42 Некоторые объясняющие факторы были качественными
43 (ординальными) и номинальными по своей природе, поэтому их градации
44 кодируются фиктивными переменными (dummy). В качестве факторных
45 переменных были выбраны (индекс t здесь для простоты не пишется):

46 X_1 – распространённость активных случаев COVID-19 на сутки $t - 10$,
47 относительно суток с отчётом t , к которым в [12] был привязан показатель
48 смертности (т.е. общее количество подтверждённых случаев COVID-19 за
49 вычетом выздоровевших и умерших к суткам $t - 10$ в соотношении со
50 среднегодовой численностью населения в 2020г) (*Active Cases*);

51 X_2 – показатель обеспеченности населения врачами-специалистами (для трёх
52 групп: инфекционисты (I), санитарно-противоэпидемическая ($ПЭ$) и
53 анестезиологи-реаниматологи (AP), но без учёта токсикологов; здесь бралась
54 их линейная комбинация: $X_2 = 0.423 \cdot I + 0.267 \cdot ПЭ + 0.118 \cdot AP$, где I –
55 количество врачей-инфекционистов на 10 тыс. жителей за 2020г и т.д.) [13]
56 (*Doctors*);

57 X_3 – осреднённое значение для ИГ, измеренное в логарифмической шкале, с
 58 использованием 7-дневного окна (при обучении это 23.x-29.x), $Log(Mortality$
 59 $[12]$;
 60 d_1 – dummy переменная тех регионов, где хуже популяционное здоровье
 61 (*Type 1*);
 62 d_2 – тоже для варианта, когда уровень здоровья был сравнительно высок
 63 (*Type 2*);
 64 d_3 – тоже для южных регионов, т.е. регионов из ЮФО, СКФО (*South Region*),
 65 d_4 – тоже, когда наблюдался повышенный уровень загрязнения атмосферного
 66 воздуха в населенных пунктах региона за ряд последних лет (*Air Pollution*).
 67 Эти четыре переменные и X_2 влияют лишь на изменение свободного члена в
 68 (1).

69 Для материала обучения были взяты значения фактора X_1 на дату
 70 21.x.2021, а значения Y' включались здесь на 10 суток позднее, т.е. на
 71 1.xi.2021. При этом на этапе “обучение” использовалось значение Y' ,
 72 осреднённое за 5 суток: с 30.x по 3.xi.2021. При обработке данных по ИГ
 73 использовалась операция осреднения по окну шириной 7 суток, после чего
 74 брался натуральный логарифм, в итоге получались значения X_3 . Факторы
 75 X_2, d_1, \dots, d_4 не изменялись по t ; у нас не было актуальной информации.

76 Затем оценивались параметры модели регрессии с переменной
 77 структурой:

$$78 \quad Y' = \beta_0 + \beta_1 X_1^{t-10} + \beta_2 X_2 + \beta_3 X_3^{t-7} + \beta_4 d_1 + \beta_5 d_2 + \beta_6 d_3 + \beta_7 d_4 + \varepsilon, \quad (1)$$

79 где Y' – смертность от COVID-19 (отношение количества умерших от этого
 80 заболевания за всё время пандемии к суткам t к среднегодовой численности
 81 населения в 2020г, умноженное на 100 тыс. [12, 13]). Здесь
 82 $X_1, X_2, X_3, d_1, d_2, d_3, d_4$ – объясняющие факторы, β_0 отражает влияние на Y'
 83 неучтённых переменных, ε – случайная величина с $E(\varepsilon) = 0$ и дисперсией σ^2 .
 84 Неизвестные параметры $\beta_0, \beta_1, \dots, \beta_7$ и σ^2 оцениваются по таблице с
 85 обучающими данными, причём оценки b_0, \dots, b_7 для β -параметров и σ^2

86 находятся методом наименьших квадратов (МНК) средствами *MS-Excel*.
87 Каждая строка этой таблицы содержит набор $Y^t, X_1, \dots, X_3, d_1, \dots, d_4$ для того
88 или иного региона; факторы X_1, X_3 входили здесь с разными лагами
89 относительно даты привязки показателя смертности в [12]. Отчасти близкий
90 подход применялся ранее при анализе данных по онкологии [14].

91 Нами выделялись регионы типа 1, в которых в 2015-2020гг
92 наблюдались повышенная заболеваемость болезнями системы
93 кровообращения, эндокринной системы и обмена веществ, онкологии,
94 органов дыхания и пищеварения, мочеполовой системы, “врождёнными
95 аномалиями...”. Часть таких регионов сильно пострадала после катастрофы
96 на ЧАЭС; в других регионах или вблизи них действуют АЭС (с не самыми
97 передовыми технологиями). На территории некоторых, или сравнительно
98 недалеко от них, ещё во времена СССР был ряд технологических аварий или
99 взрывов [15]. Учёт ситуации с радиационным загрязнением важен в плане
100 онкологии [14]. Для 17-ти регионов этого типа выбирались значения
101 переменных $d_1 = 1, d_2 = 0$. К этому типу можно отнести почти $\frac{1}{2}$ областей из
102 ЦФО: Брянская, Владимирская, Ивановская, Курская, Орловская, Рязанская,
103 Смоленская; также – 7 регионов из СЗФО, ПФО, ЮФО: Архангельская,
104 Астраханская, Волгоградская, Оренбургская, Ростовская, Саратовская
105 области и республика Удмуртия, а также некоторые регионы из УФО, СФО,
106 ДФО: Тюменская область, Алтайский край и Еврейская автономия.

107 Для регионов типа 2, наоборот, характерны довольно низкие значения
108 заболеваемости для вышеупомянутых классов, а также – относительно малые
109 радиационные загрязнения и/или загрязнения из-за техногенных аварий. Для
110 них принималось $d_1 = 0, d_2 = 1$; к этому типу были отнесены пять регионов:
111 три южных (республики Адыгея, Кабардино-Балкария и Ставрополье), а
112 также – Хабаровский край и Московская область; здесь проживает больше
113 молодого населения, а также в них сравнительно мала доля лиц старших

114 возрастов. Все прочие регионы из материала обучения имели значения
 115 $d_1=d_2=0$.

116 Ещё одна переменная кодировала южные регионы на карте страны:
 117 $d_3=1$. Наконец, высокий уровень загрязнения воздуха регулярно выявлялся в
 118 Архангельской, Белгородской, Воронежской, Курской, Новосибирской
 119 Ростовской, Саратовской и Смоленской областях, а также в Алтайском,
 120 Забайкальском, Приморском и Хабаровском краях, республиках Кабардино-
 121 Балкария, Крым и Северная Осетия–Алания [13]. Для всех этих 15-ти
 122 регионов, принималось, что $d_4=1$.

123 Математические методы включали корреляционный и регрессионный
 124 анализ, методы проверки статистических гипотез. Использовались различные
 125 тесты (через t -статистику Стьюдента для гипотез $H_0: \beta_j=0$ и DW -статистику
 126 Дарбина-Уотсона для выявления автокорреляции), а также традиционный
 127 дисперсионный анализ Р.Фишера. Для корректного использования тестов
 128 также применялись критерии согласия распределения остатков с
 129 нормальным законом. Так как дисперсия σ^2 неизвестна, то были взяты тесты
 130 χ^2 Пирсона и Лилиефорса – после анализа эмпирической функции
 131 распределения остатков. Условие гомоскедастичности проверялось на
 132 диаграммах рассеяния остатков регрессии. Весь статистический анализ и
 133 обработка данных выполнялись в *MS-Excel*.

134 Результаты и обсуждение

135
 136 В верхней части табл. 2 приводятся b -коэффициенты, или *MНК*- оценки
 137 β .

138 (ТАБЛИЦА 1)

139 В 3-й строке в скобках приводятся модули t -статистик, из которых
 140 после проверки соответствующих гипотез следует, что β -коэффициенты
 141 отличны от 0. В последней строке приводится F -статистика Р.Фишера и
 142 число степеней свободы v . Все достигнутые уровни значимости (p -values) в
 143 тестах Стьюдента и Фишера были ниже 0.001. Также в 5-й строке приводятся

144 стандартная ошибка s , значение DW -статистики, несмещённая оценка
145 коэффициента детерминации R^2 , а также две статистики для критериев
146 согласия: χ^2 и статистика, содержащая расстояние Колмогорова D_n . Наконец,
147 в 4-й строке приводятся средние коэффициенты *эластичности*. Каждый из
148 них показывает, на сколько процентов изменится смертность Y' в среднем по
149 таблице обучения, если X_j увеличить на 1% и при этом зафиксировать
150 значения прочих факторов. Сильнее всего на переменную Y' влияют
151 обеспеченность врачами и распространённость активных случаев, а
152 несколько слабее – d_1 и X_3 (логарифм скользящего среднего для ИГ). Так,
153 если значение фактора X_2 с обеспеченностью врачами увеличить на 1%
154 (соблюдая нормативы по балансу между специализациями), то смертность
155 снизится на 0.5%.

156 Модель получилась довольно точной для многих регионов материала
157 обучения, когда отсчёт t привязан в [12] к суткам 1.xi.2021. А именно для 30
158 регионов ошибка аппроксимации $MARE = 1.94\%$ для модели (1) с b -
159 коэффициентами и со стандартной ошибкой 1.5%. Иногда она была сравнима
160 по точности с моделью из [7]. Для всего материала обучения $MARE$ (mean
161 absolute percentage error) была 3%; при этом модель объяснила 96% разброса
162 переменной Y' вокруг её среднего значения. На рис. 1 показано, как модель
163 (1) с b -коэффициентами работает на “экзамене” при изменении X_1 и X_3 по оси
164 $0-t$. Модель можно улучшить, если уточнить данные по врачам-
165 специалистам; другой вариант её улучшения – регулярно проводить новое
166 обучение: через 5-7 дней.

167 (РИСУНОК 1)

168 Каждый точечный график связан с прогнозом Y' – распространённости
169 смертности от COVID-19 за сутки t из промежутка $[1.xi, 17.xi]$ – для одного
170 из семи регионов. Здесь дате 1.xi.2021 для показателя Y' будет
171 соответствовать отсчёт t ; дате 2.xi – $t+1, \dots$. Линиями без маркеров на рис. 1
172 изображается изменение смертности по факту.

173 Например, чтобы получить прогнозное значение Y' на 5 ноября для
174 Подмосковья, надо взять его значения dummy-переменных d_1, \dots, d_4 , а также –
175 переменной с врачами X_2 . Все эти пять переменных не изменяются по t ; они
176 лишь скорректируют для региона значение b_0 (см. табл.1). От суток к суткам
177 изменяется количество активных случаев; оно легко находится из [12]; для
178 прогноза на 5.xi.2021 надо брать накопленные данные на 25.x, когда значение
179 фактора $X_1 = 8.267$ на 1 тыс. жителей (с учётом того, что здесь проживало
180 около 7700 тыс. граждан). Аналогично изменяется и фактор X_3 ; он
181 определяется по набору из нескольких значений ИГ. Для его расчёта здесь
182 берутся данные за 22.x, ..., 28.x по количеству выздоровевших r пациентов в
183 регионе и умерших d за *каждые сутки* [12], затем находятся значения
184 $ИГ = 100 \cdot d / (d + r)$ за 7 суток: с 22.x по 28.x. (Предполагается, что за эти дни
185 имелся хотя бы 1 летальный случай). Потом берётся их среднее значение, а
186 от него – натуральный логарифм, равный 0.3863; это и есть значение X_3 на
187 25.x, т.е. середину 7-дневного окна. В результате находится сумма:
188 $9.141 \cdot X_1 + 12.76 \cdot X_3 - 96.99 \cdot X_2 + (\text{сумма четырёх произведений с}$
189 $\text{dummy...}) + 106.462$, которая считается через b -коэффициенты из табл.1.
190 Сумма членов, что по предположению не зависит от t , выделена более
191 блекло, при этом выражение в скобках равно -21.82 , т.к. $d_2 = 1$, а все d_1, \dots, d_3
192 нулевые (см. текст ниже). Итак, для региона будет Y'
193 $= 9.14X_1 + 12.76X_3 + 37.9186$, т.к. здесь $X_2 = 0.48792$. В итоге на дату 5.xi.2021
194 получаем прогноз показателя Y' , равный 118. Для сравнения смертность от
195 онкологических заболеваний в регионе к концу 2020 года была 172 [13].

196 Заметим, что распространённость активных случаев COVID-19 в
197 последней декаде октября была примерно одинаковой в Ленинградской,
198 Московской, Тверской и Костромской областях [12]. Среди них, в первой
199 декаде ноября в Тверской области была самая высокая смертность от
200 COVID-19; у неё также одно из наихудших в РФ мест по коэффициенту

201 общей смертности в 2019-20гг. Значения Y' были здесь близки в
202 Костромской и Ленинградской области, а в Подмосковье умирали реже.

203 Сравнивая средние значения для этой четвёрки регионов по набору
204 более чем 10 медико-демографических показателей, взятых за 2015-2020гг
205 [13], можно предположить, почему это так. Вблизи нашей столицы
206 проживает явно моложе население и мало людей старше 70 лет (по
207 отношению к тем, кто младше 25), также здесь стабильно высокие душевые
208 расходы бюджета на здравоохранение. Из-за перечисленного здесь невелик
209 фактор X_3 ; его снизила и более широкая вакцинация взрослых. Также
210 X_3 бывает меньше в регионах, где активнее проводится тестирование на
211 антитела к *SARS-CoV-2*, отчего возрастает доля нетяжелых случаев и поэтому
212 больше выздоравливающих [10]. Ещё в этом регионе в 2019-20гг была
213 минимальная общая смертность среди областей ЦФО, низкие уровни
214 заболеваемости упомянутых классов заболеваний (исключая болезни
215 дыхательной системы; а также – сравнительно невысокие заболеваемости
216 новообразованиями и для органов пищеварения). Также здесь относительно
217 мала онкологическая смертность, нет острых проблем с радиоактивным
218 загрязнением территории и загрязнением атмосферного воздуха; поэтому для
219 Подмосковья принималось, что значения dummy-переменных: $d_1 = d_3 = d_4 = 0$
220 , $d_2 = 1$.

221 Для читателя с более продвинутой математической подготовкой может
222 быть интересно следующее. Нами проверялось выполнение предпосылок
223 теоремы Гаусса-Маркова, в которой обоснованы хорошие свойства МНК
224 оценок. Для этого сначала рассматривались коэффициенты корреляций Y' с
225 каждым из факторов, а также для пар последних. Явление
226 мультиколлинеарности не обнаружено. Гипотеза о нормальности по
227 критерию χ^2 на 5%-м уровне не отвергалась. График эмпирической функции
228 распределения для остатков сравнивался с графиком распределения Гаусса с
229 нулевым средним и стандартом s , где s – несмещенная оценка σ . Наибольшее

230 расхождение графиков $D_n = 0.100$, так что из-за малости $n^{1/2} \cdot D_n$, аналогичный
231 вывод сделан тестом Лиллиефорса [14]. Итак, распределение остатков не
232 противоречит закону Гаусса. Анализ диаграмм рассеяния остатков
233 относительно каждого X_j , а также относительно оценки Y' , не выявил явного
234 нарушения условия гомоскедастичности. Тест об отсутствии автокорреляции
235 остатков не отвергался. В итоге все предпосылки теоремы были выполнены.

236 В результате была построена статистическая модель для краткосрочного
237 прогноза смертности в регионах от COVID-19 на текущие сутки и для неё
238 найдены коэффициенты эластичности. Модель может оказаться полезной при
239 разработке мер борьбы против пандемии в ряде регионов, а также при
240 создании улучшенных секторных моделей (типа SEIRF, SIRD и др.).

РИСУНКИ

Рис. 1. Прогноз смертности Y на $1, \dots, 17.xi$ на 100 тыс. жителей

Fig. 1. predicting the mortality Y on the 1st... 17th Nov on 10^5 Pop

(Tyumen Region is one less autonomous areas over here)

НАРИСУНОЧНЫЕ ПОДПИСИ

Воронежская - Voronezh Region

Костромская - Kostroma Region

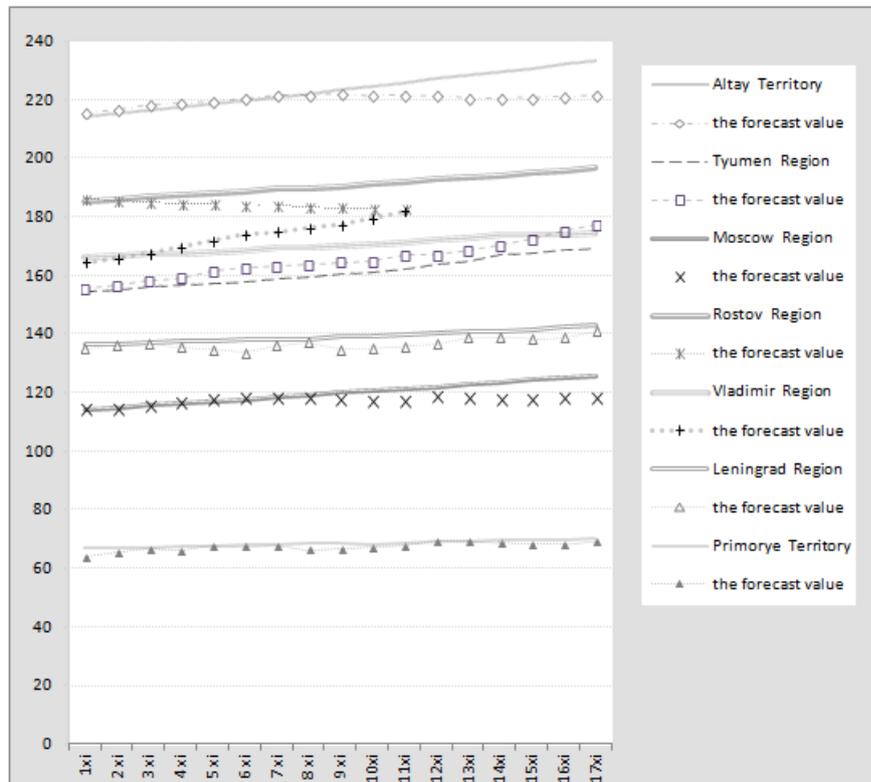
Белгородская - Belgorod Region

Московская - Moscow Region

Ростовская - Rostov Region

Орловская - Orel Region

её оценки – relevant estimates



ТАБЛИЦЫ

Таблица 1. Результаты оценивания, коэффициенты эластичности, статистики
 Table 1. Least square estimates for all parameters, coefficients of elasticity, statistics

Переменная Variable		X_1	X_2	X_3	d_1	d_2	d_3	d_4
Её краткое имя relevant short name		<i>Active Cases</i>	<i>Doctors</i>	<i>Log(mortality)</i>	<i>Type 1</i>	<i>Type 2</i>	<i>South Regions</i>	<i>Air Pollution</i>
1 МНК-оценки	b_0	b_1	b_2	b_3	b_4	b_5	b_6	b_7
2 <i>LS estimates</i>	106.46	9.141	- 96.988	12.76	47.494	- 21.82	26.965	19.116
3 Статистика t <i>t-statistics</i>	(20.7)	(24)	(16.4)	(10.7)	(21.8)	(7.9)	(11.2)	(9.5)
4 Эластичность <i>Elasticity coefficients</i>		0.35	- 0.48	0.15	0.14	- 0.04	0.05	0.05
5 F -stat. = 252	$v = 3$ 2	$s = 5.6$ 06	$R^2 = 96$ %	$n^{1/2} \cdot D_n = 0.6$ 3	$\chi^2 = 0.$ 9		$DW = 2.$ 10	

ТИТУЛЬНЫЙ ЛИСТ_МЕТАДААННЫЕ

Оценка уровня смертности, порождённой *SARS-CoV-2* в регионах России, на основе эконометрической модели

Estimation of the mortality rate caused by *SARS-CoV-2* in Russia's regions, based on the econometric model

БЛОК 1

Степанов Владимир Сергеевич, кандидат физико-математических наук,
старший научный сотрудник

Vladimir S. Stepanov, Candidate of Physical and Mathematical Sciences,
Senior researcher

Федеральное государственное учреждение науки
Центральный экономико-математический институт
Российской академии наук

The Central Economics and Mathematics Institute of the Russian Academy of
Sciences

Россия, 117418, Москва, Нахимовский проспект, 47, к.810

47, Nakhimovsky Prospect, Moscow, 117418, Russia

+7(916) 872-52-07

+7(499) 129-16-44 (зав.канцелярией)

stepanov@cemi.rssi.ru, vladstep0355@gmail.com

ORCID: 0000-0002-4478-376X

БЛОК 2

Степанов Владимир Сергеевич, кандидат физико-математических наук
старший научный сотрудник

Vladimir S. Stepanov, Candidate of Physical and Mathematical Sciences,
Senior researcher

Степанов В.С.

ЦЭМИ РАН, Москва, лаборатория 2.01

БЛОК 3

Оценка уровня смертности, порождённой COVID-19 в регионах РФ, на
основе эконометрической модели

Estimation of the mortality rate caused by COVID-19 in Russia's regions, based on
the econometric model

Ключевые: *модель регрессии, загрязнение воздуха, COVID-19,
моделирование смертности, заболеваемость населения, общественное
здоровье, эпидемиологический прогноз, регионы*

Keywords: *regression model, air pollution, COVID-19, mortality modeling,
morbidity, public health, epidemiological forecast, regions*

Сокращенное название статьи: прогнозирование смертности от covid-19;
prediction of the covid-19 mortality

Менее 9 страниц, 1 рисунок (с графиками), 1 таблица

КРАТКОЕ СООБЩЕНИЕ

14 декабря 2021 - дата отправления работы

СПИСОК ЛИТЕРАТУРЫ

№	Авторы, название публикации и источника, где она опубликована, выходные данные	ФИО, название публикации и источника на английском	Полный интернет-адрес (URL) цитируемой статьи и/или DOI
15	Александров Ю.А. Основы радиационной экологии. Йошкар-Ола: Марийский госуниверситет, 2007. 268 с.	Aleksandrov Yu.A. Fundamentals of radiation ecology. Yoshkar-Ola: Mariyskii gosuniversitet, 2007. 268 p. (In Russ.)	https://www.elibrary.ru/item.asp?id=25433650
7	Гаврилов Д.В., Абрамов Р.В., Кирилкина А.В., Ившин А.А., Новицкий Р.Э. Модель прогнозирования пандемии COVID-19 на основе машинного обучения в отдельных регионах Российской Федерации // Фармакоэкономика. Современная фармакоэкономика и фармакоэпидемиология. 2021. Т. 14, № 3. С. 342–356. Гольдштейн Э.М. Факторы,	Gavrilov D.V., Abramov R.V., Kirilkina A.V., Ivshin A.A., Novitskiy R.E. COVID-19 pandemic prediction model based on machine learning in selected regions of the Russian Federation. <i>Farmakoekonomika. Sovremennaya farmakoekonomika i farmakoepidemiologiya = Farmakoekonomika. Modern Pharmacoepidemiology and Pharmacoepidemiology</i> , 2021, vol. 14, no. 3, pp. 342–356 (In Russ.) Goldstein E.M. Factors affecting mortality for the novel coronavirus infection in different regions of the Russian Federation. <i>Zhurnal mikrobiologii, epidemiologii i immunobiologii = Journal of microbiology, epidemiology and immunobiology</i> , 2020,	https://www.pharmacoeconomics.ru/jour/article/view/541 [10.17749/2070-4909/farmakoekonomika.2021.108]

10.	<p>влияющие на смертность новой коронавирусной инфекцией в разных субъектах Российской Федерации // Журнал микробиологии, эпидемиологии и иммунобиологии. 2020. Т. 97, № 6. С. 604–607.</p>	<p><i>vol.97, no 6, pp. 604–607.</i></p> <p>Druzhinin P.V., Molchanova E.V. Mortality rates in Russian regions in the context of the COVID-19 pandemic. <i>Regionologiya = Regionology. Russian Journal of Regional Studies</i>, 2021, vol. 29, no 3, pp. 666-685. (In Russ.)</p>	<p>https://microbiol.elpub.ru/jour/article/view/952</p> <p>[10.36233/0372-9311-2020-97-6-11]</p>
9	<p>Дружинин П.В., Молчанова Е.В. Смертность населения российских регионов в условиях пандемии COVID-19 // Регионология. 2021. Т. 29, № 3. С. 666–685.</p>	<p>Zaitseva N.V., May I.V., Reis J., Spenser P.S., Kiryanov D.A., Kamaltdinov M.R. On estimating the additional incidence of COVID-19 among populations exposed to polluted ambient air: methodical approaches and some practical results. <i>Analiz riska zdorov'yu = Health Risk Analysis</i>, 2021, no. 3, pp. 14–28. (In Russ.)</p>	<p>http://regionsar.ru/sites/default/files/2021_3/reg_2021_3.pdf</p> <p>[10.15507/24131407.116.029.202103.666-685]</p>
11	<p>Зайцева Н.В., Май И.В., Рейс Ж., Спенсер П.С., Кирьянов Д.А., Камалтдинов М.Р. К оценке дополнительной заболеваемости населения COVID-19 в условиях загрязнения атмосферного воздуха: методические</p>	<p>Makarov V.L., Bakhtizin A.R., Sushko E.D., Ageeva A.F. COVID-19 epidemic modeling – advantages of an agent-based approach. <i>Ekonomicheskiye i sotsial'nyye peremeny: fakty, tendentsii, prognoz = Economic and Social Changes: Facts, Trends, Forecast</i>, 2020, vol. 13, no. 4, pp. 58–73. (In Russ.)</p> <p>Official statistical data on the novel coronavirus in Russia's regions</p> <p>Statistical yearbooks for Russia's regions</p>	<p>https://journal.fcisk.ru/2021/3/2</p> <p>[10.21668/health.risk/2021.3.02]</p> <p>[10.21668/health.risk/2021.3.02.eng]</p>

	<p>подходы и некоторые практические результаты // Анализ риска здоровью. 2021. № 3. С. 14–28.</p>	<p>published by the Federal State Statistical Service</p>	
6	<p>Макаров В.Л., Бахтизин А.Р., Сушко Е.Д., Агеева А.Ф. Моделирование эпидемии COVID-19 – преимущества агент-ориентированного подхода // Экономические и социальные перемены: факты, тенденции, прогноз. 2020. Т. 13, № 4. С. 58–73.</p>	<p>Stepanov V.S. Dependence between mortality in regions and prevalence of active SARS- CoV-2 carriers and resources available to public healthcare organizations. <i>Analiz riska zdorov'yu = Human Risk Analysis</i>, 2020, no 4, pp. 12-23. (In Russ.)</p> <p>Stepanov V.S. The relationship of cancer prevalence with age of the population living under adverse environmental factors. <i>Vrach i informatsionnyye tekhnologii = Medical doctors and information technology</i>, 2021, no 3, pp. 38-49. (In Russ.)</p>	<p>http://esc.vscs.ac.ru/article/28636</p> <p>[10.15838/esc.2020.4.70.3]</p>
12	<p>Официальная информация о коронавирусе в России: Оперативные данные</p>	<p>Yanchevskaya E.Ya., Mesnyankina O.A. Mathematical modelling and prediction in infectious disease epidemiology. <i>RUDN Journal of Medicine</i>, 2019, vol. 23, no 3, pp. 328–334.</p>	<p>https://стопкоронавирус.рф/information/</p> <p>https://rosstat.gov.ru</p>
13	<p>Ежегодные статистические сборники Росстата по регионам РФ</p> <p>Степанов В.С. Зависимость уровня смертности в регионах от распространенности активных носителей SARS-</p>	<p>--</p>	<p>https://journal.fcisk.ru/2020/4/</p>

2	<p>CoV-2 и ресурсов организаций здравоохранения // Анализ риска здоровью. 2020. № 4. С. 12-23.</p>		<p>[10.21668/health.risk/2020.4.02] [10.21668/health.risk/2020.4.02.eng]</p>
14	<p>Степанов В.С. Связь онкологической заболеваемости с возрастом населения, проживающего при неблагоприятных факторах окружающей среды // Врач и информационные технологии. 2021. № 3. С. 38–49.</p>	<p>--</p>	<p>https://vit-j.ru/upload/uf/729/wta2tiq4bmcauv60nvq1089o47x5ilbu/vit_2021_03.pdf</p> <p>[10.25881/18110193_2021_3_38]</p>
1	<p>Янчевская Е.Ю., Меснянкина О.А. Математическое моделирование и прогнозирование в эпидемиологии инфекционных заболеваний // Вестник Российского университета дружбы народов. Серия: Медицина. 2019. Т. 23, № 3. С. 328–334.</p> <p>Friedman J., Liu P., Troeger C.E., Carter A., Reiner R.C. Jr., Barber R.M., Collins J., Lim</p>	<p>--</p>	<p>http://journals.rudn.ru/medicine/article/view/22521</p> <p>[10.22363/2313-0245-2019-23-3-328-334]</p>

5	<p>S.S., Pigott D.M., Vos Th., Hay S.I., Murray C.J.L., Gakidou Em. Predictive performance of international COVID-19 mortality forecasting models. <i>Nature Communications</i>, 2021, vol. 12, article no 2609, 38 p.</p>		<p>https://pubmed.ncbi.nlm.nih.gov/33972512/</p> <p>[10.1038/s41467-021-22457-w]</p>
8	<p>Kurkina E.S., Koltsova E.M. Mathematical modeling of the propagation of COVID-19 pandemic waves in the world. <i>Computational Mathematics and Modeling</i>, 2021, vol. 32, no. 2, pp.46-79.</p>		<p>https://link.springer.com/article/10.1007%2Fs10598-021-09523-0</p> <p>[10.1007/s10598-021-09523-0]</p>
3	<p>Lifshits M.L., Neklyudova N.P. COVID-19 mortality rate in Russian regions: forecasts and reality. <i>R-economy</i>, 2020, vol. 6, no 3, pp. 171–182.</p>		<p>https://journals.urfu.ru/index.php/r-economy/article/view/4725</p> <p>[10.15826/recon.2020.6.3.015]</p>
	<p>Wang Y., Xu Ch., Yao S., Zhao Y., Li Y., Zhao X. Estimating the prevalence and mortality of coronavirus disease 2019 (COVID-19) in the USA, the UK, Russia, and India. <i>Infection and Drug Resistance</i>, 2020, vol. 13, pp. 3335–3350.</p>		<p>https://www.elibrary.ru/title_about.asp?id=35306</p> <p>https://pdfs.semanticscholar.org/fa1a/2ca51ccbf725ed</p>

4			0309ae3d4addcd1740ed2f.pdf?_ga=2.57456106.135647449.1639497288-53455721.1639497288 [10.2147/IDR.S265292]
---	--	--	---