

О математических моделях эпидемий

В.С. Рихвицкий

Объединенный институт ядерных исследований

23 июня 2020 г.

Аннотация

Рассматривается простейшая после моделей Мальтуса и Ферхюльста 3-state модель, обозначаемая в литературе аббревиатурой SIR (S – восприимчивые, I – с клиническими проявлениями инфекционной болезни, R – выздоровевшие), с коэффициентами, зависящими от времени. Модель сопоставляется текущим статистическим данным, находящимся в свободном доступе. Неизвестные параметры модели удается найти путем минимизации невязки между предсказанием и доступной статистикой.

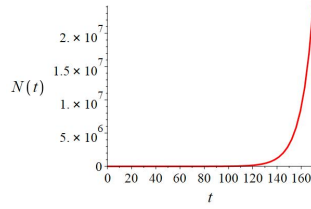
1 Модели Мальтуса и Ферхюльста

Будем рассматривать происходящее как поток случайных событий. Это самый простой подход. Принимая во внимание внимание только вероятность $\epsilon(t)$ заражения (за день) пишется уравнение Мальтуса

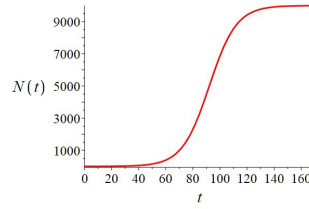
$$\frac{dN(t)}{dt} = \epsilon(t)N(t) \quad (1)$$

с решением

$$N(t) = N_0 e^{\int_0^t \epsilon(t) dt} \quad (2)$$



Динамика заражения
в модели Мальтуса



Динамика заражения
в модели Ферхюльста

Уравнение Ферхюльста уже учитывает ограниченность общей численности

$$\frac{dN(t)}{dt} = \epsilon(t)N(t) \left(1 - \frac{N(t)}{K}\right), \quad (3)$$

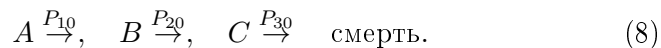
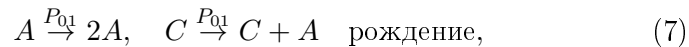
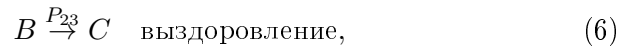
в котором вероятность добавления заразившегося уменьшается по мере роста числа зараженных (и переболевших), с решением

$$N(t) = \frac{K N_0}{N_0 + (K - N_0) e^{-\int_0^t \epsilon(t) dt}} \quad (4)$$

2 3-state (SIR) модель

В простой модели [1, 2] можно определить три состояния: A : неинфицированные («новички»), B : больные, C : переболевшие с численностями соответственно $A(t)$, $B(t)$, $C(t)$. И, соответственно, вероятности перехода (за день): P_{12} вероятность заболеть и P_{23} вероятность вылечиться. Это переходы между состояниями. Можно добавить вероятности P_{01} рождения из «ниоткуда» в неинфицированные и смерти P_{10} из неинфицированных в «никуда», P_{20} из инфицированных в «никуда» и P_{30} из переболевших в «никуда».

Эти же переходы и уравнения могут быть записаны на языке формальной химической кинетики



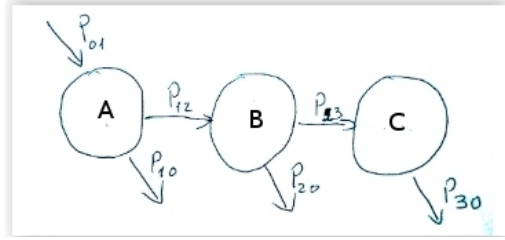


Рис. 1: Диаграмма состояний 3-state модели

Им соответствуют дифференциальные уравнения

$$\frac{dA(t)}{dt} = P_{01}(A(t) + C(t)) - P_{12}A(t)B(t) - P_{10}A(t), \quad (9)$$

$$\frac{dB(t)}{dt} = P_{12}A(t)B(t) - P_{23}B(t) - P_{20}B(t), \quad (10)$$

$$\frac{dC(t)}{dt} = P_{23}B(t) - P_{30}C(t). \quad (11)$$

3 Замечания о переменных коэффициентах

Можно сформулировать ряд утверждений. Согласно уравнениям (9,10), если $A(t_0) > 0$ и $B(t_0) > 0$, то и в дальнейшем при $t > t_0$ $A(t) > 0$, $B(t) > 0$. Пусть для простоты $P_{01} = P_{10} = P_{30} = 0$. Тогда для любых заданных гладких функций $A(t)$, $B(t)$ и $C(t)$ система (9-11) является разрешимой системой линейных уравнений относительно коэффициентов $P_{12}(t)$, $P_{20}(t)$, $P_{23}(t)$.

Поэтому, как ни странно, задача подгонки под экспериментальные данные (статистику) не в том, чтобы решить уравнения (9-11), а в том, чтобы найти $P_{12}(t)$ и $P_{23}(t)$. Но в реальности функции $A(t)$, $B(t)$, и $C(t)$ не заданы гладкими, а известны только в некоторых точках и то – с известной неопределенностью. Можно, конечно, гладко аппроксимировать их, но экстраполяция за пределы временного диапазона статистики становится совсем неопределенной.

Можно ужесточить задачу, потребовав, чтобы функции $P_{12}(t)$ и $P_{23}(t)$ были из заданного класса и, пользуясь заложенной в статистике неопределенностью, добиться наиболее правдоподобного выбора согласно критерию (12). Например, можно потребовать,

чтобы функции были кусочно-полиномиальные, с заданными точками времени – границами кусков, при этом этих точек времени должно быть немного.

4 Сравнение со статистикой

Параметры модели определяются в результате сравнения со статистическими данными [9]. Были приняты следующие допущения: в силу краткости периода влияние некоторых факторов считается незначительным, параметры приравнены нулю: $P_{01} = 0$, $P_{10} = 0$, $P_{30} = 0$. Остальные параметры находятся вариацией до достижения минимума критерия близости. Определяются такие константы: величина популяции $A(0)$, вероятности P_{23} , P_{20} . Вероятность заражения $P_{12} = \alpha(t)/K$ существенно зависит от времени, поскольку проводятся известные организационные мероприятия, но она ищется в виде кусочно-гладкого интерполирующего полинома (3-го порядка), заданного значениями в определенных узлах, эти значения и варьируются. Вероятность выздоровления $P_{23} = \beta(t)$ зависит от методики лечения и ее изменения.

Критерий близости

$$\chi^2 = \sum_i \frac{(A(i) - A_i)^2}{A(i)} + \sum_i \frac{(B(i) - B_i)^2}{B(i)} + \sum_i \frac{(C(i) - C_i)^2}{C(i)}. \quad (12)$$

Здесь i – номер дня, $A(i)$, $B(i)$, $C(i)$ – расчетные величины, A_i , B_i , C_i – данные статистики. Начальные значения для системы уравнений: $A(t_0) = K$, $B(t_0) = 1$, $C(t_0) = 0$. Величины K и t_0 также входят в число варьируемых параметров.

Результаты моделирования представлены на графиках (Рис. 2 - Рис. 6). Линиями изображены графики функций, на них наложены точки значений из статистики (на 23 июня 2020 г). Отклонения реальной статистики от расчетной, отнесенные к расчетному среднеквадратичному отклонению (равному корню квадратному из расчетного значения) представлены на графиках (Рис. 7 - Рис. 11).

Следует заметить, что найденный размер популяции не совпадает с численностью населения в регионе. По-видимому, это эффективный размер, с учетом средней дальности влияния отдельных людей в данной среде. Опыт наблюдения в течении многих дней обсчета показывает, что этот параметр со временем растет.

5 Обсуждение результатов

Графики показывают, что в рассмотренных регионах максимум скорости роста числа больных уже давно пройден – а это нагрузка на скорую помощь и приемный покой.

Почти везде количество больных уже находится на максимуме или около него. Это означает, что нагрузка на стационары тоже достигает максимума и дальше будет снижаться за счет выписки выздоровевших.

Список литературы

- [1] А.А.Лопатин, В.А.Сафронов, А.С.Раздорский, Е.В.Куклев. СОВРЕМЕННОЕ СОСТОЯНИЕ ПРОБЛЕМЫ МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ И ПРОГНОЗИРОВАНИЯ ЭПИДЕМИЧЕСКОГО ПРОЦЕССА, Проблемы особо опасных инфекций, вып. 105, 2010.
- [2] И.Д. Колесин, Е.М. Житкова. МАТЕМАТИЧЕСКИЕ МОДЕЛИ ЭПИДЕМИЙ. САНКТ-ПЕТЕРБУРГ, 2004.
- [3] Robert Schaback. On COVID-19 Modelling. arXiv:2005.07004 [physics.soc-ph]
- [4] Samuel Cho. Mean-Field Game Analysis of SIR Model with Social Distancing. arXiv:2005.06758 [q-bio.PE]
- [5] Kankan Sarkar. Modeling and forecasting of the COVID-19 pandemic in India. arXiv:2005.07071 [q-bio.PE]
- [6] J.E. Amaro, J. Dudouet, J.N. Orce. Global analysis of the COVID-19 pandemic using simple epidemiological models. arXiv:2005.06742 [q-bio.PE]
- [7] Muhammad Waqas, Muhammad Farooq, Rashid Ahmad, Ashfaq Ahmad. Analysis and Prediction of COVID-19 Pandemic in Pakistan using Time-dependent SIR Model. rXiv:2005.02353 [q-bio.PE]
- [8] Deeptak Biswas, Sulagna Roy. Analyzing COVID-19 pandemic with a new growth model for population ecology. arXiv:2004.12950 [q-bio.PE]
- [9] Источник данных: <https://coronavirus-monitor.ru/coronavirus-v-rossii/>, <https://gogov.ru/articles/covid-19>

The simplest epidemic process model

We consider the simplest 3-state model after the Malthus and Verhulst models, denoted in the literature by the abbreviation SIR (S - susceptible, I - with clinical tests of an infectious disease, R - recovered), with time-dependent coefficients. The model is mapped to current statistics that are in the public domain. Unknown model parameters can be found by minimizing the discrepancy between the prediction and available statistics.

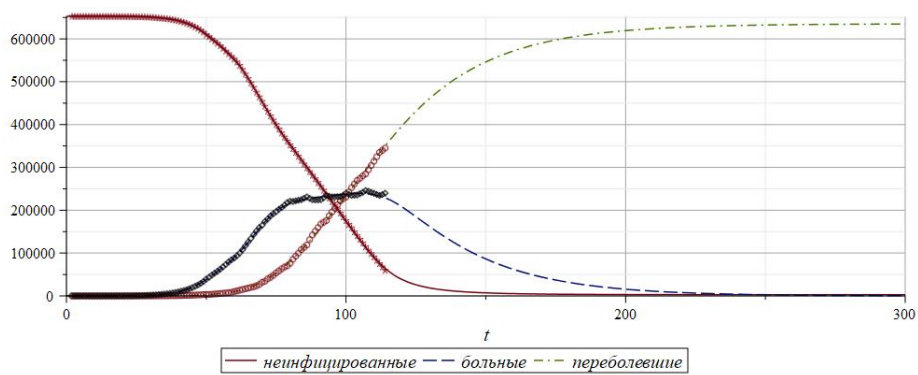


Рис. 2: Россия

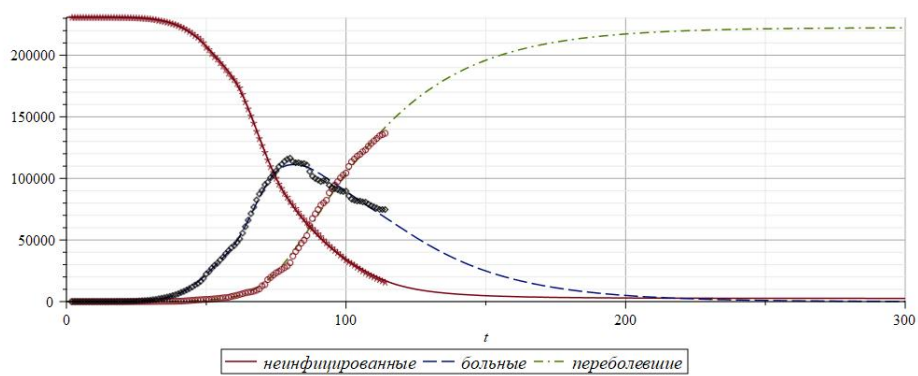


Рис. 3: Москва

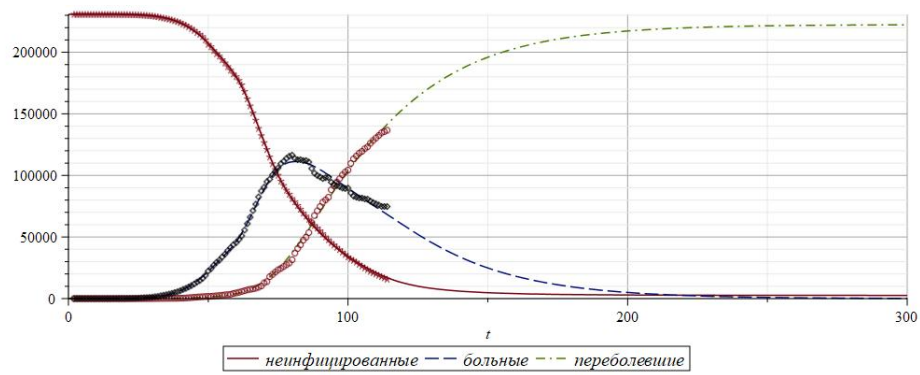


Рис. 4: Московская область

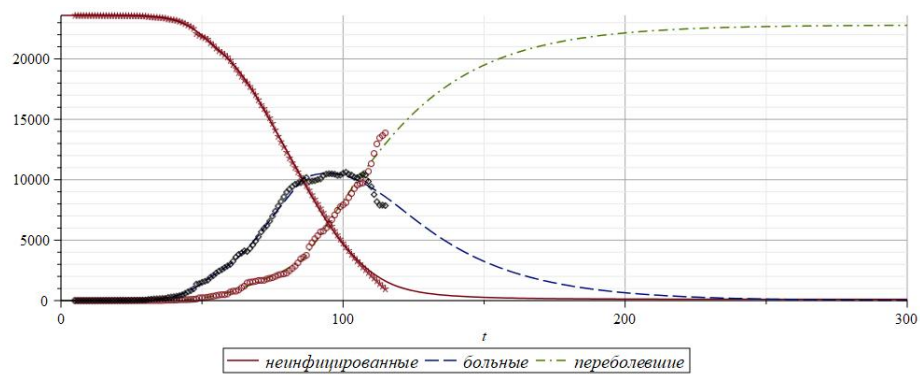


Рис. 5: Санкт-Петербург

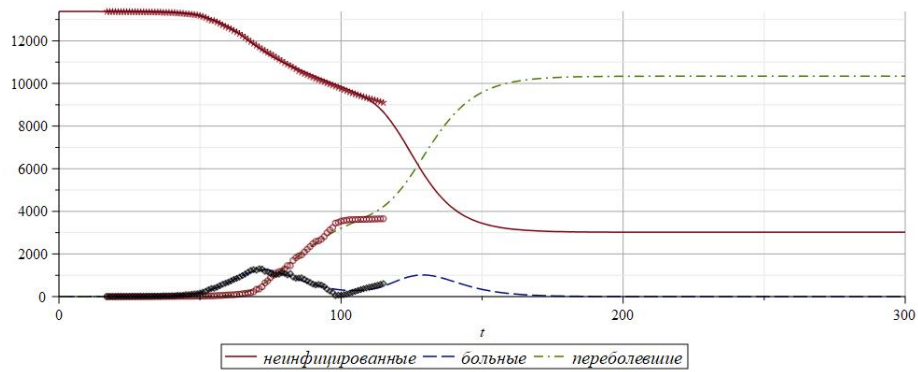


Рис. 6: Республика Татарстан

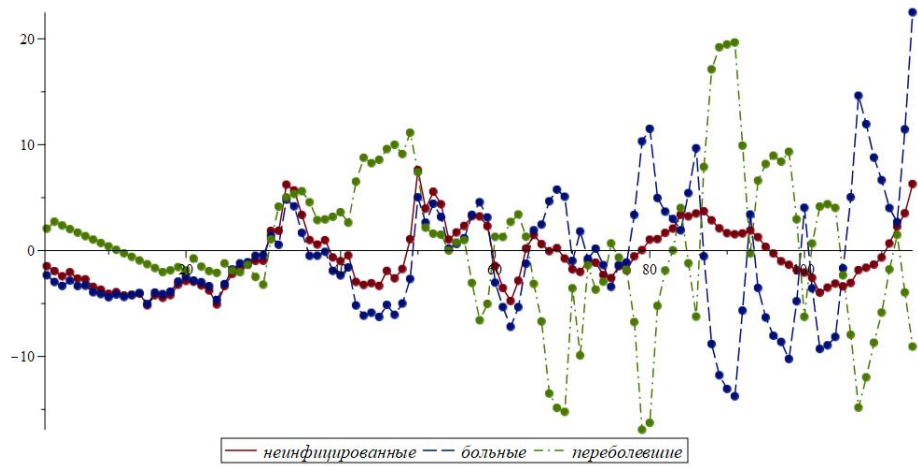


Рис. 7: Девиация. Россия

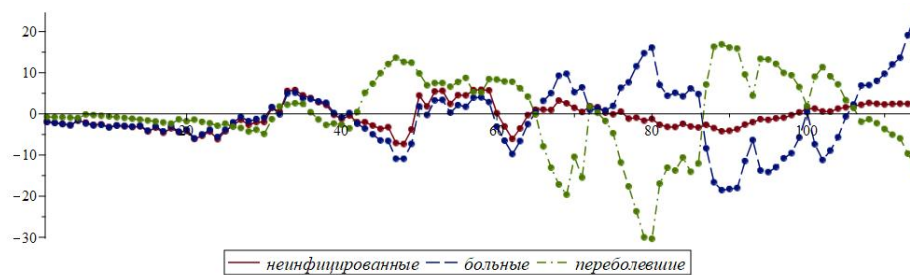


Рис. 8: Девиация. Москва

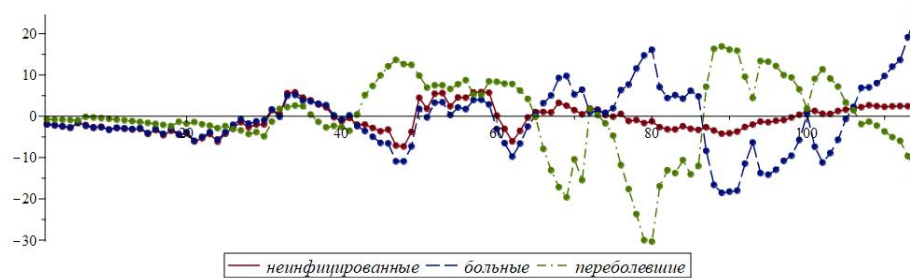


Рис. 9: Девиация. Московская область

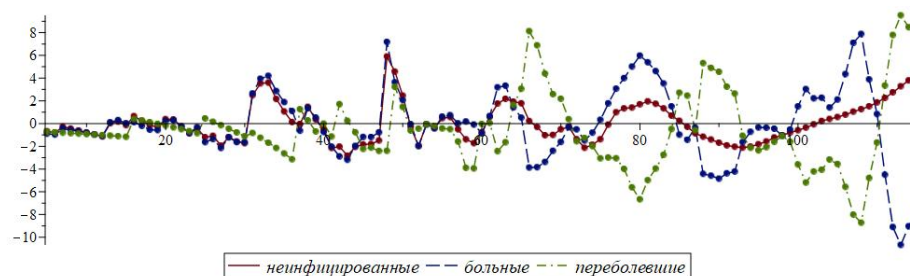


Рис. 10: Девиация. Санкт-Петербург

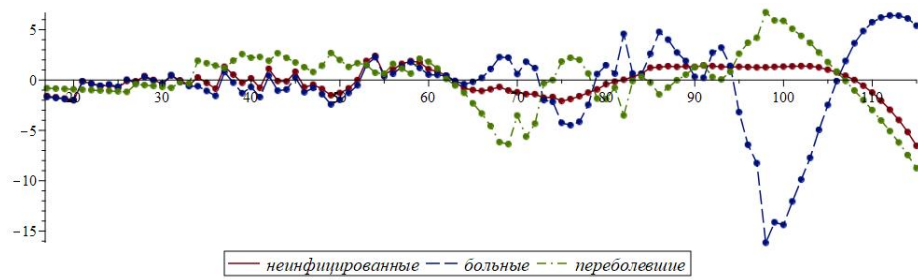


Рис. 11: Девиация. Республика Татарстан