

一种COVID-19类传染病爆发的数学建模预测方法

郑佳恩¹, 郑若滨²

¹深圳外国语学校龙华校区, 广东 深圳

²华为技术有限公司2012实验室, 广东 深圳

收稿日期: 2025年9月10日; 录用日期: 2025年9月30日; 发布日期: 2025年10月13日

摘要

本文提出了一种基于离散时间递推方法的改进SIR (Susceptible-Infective-Removed)模型, 用于预测COVID-19 (Coronavirus disease 2019)类传染病的传播趋势。模型将人群划分为易感者(S)、感染者(I)和移除者(R)三类, 通过引入动态感染率(β)和移除率(γ), 建立了简化的、可实时计算的离散时间递推方程, 以替代计算复杂的传统常微分方程。本文的建模以美国的实际疫情数据为案例, 验证了模型在短期预测和长期趋势分析中的有效性。结果表明, 模型能够较好地拟合疫情发展曲线, 并准确捕捉关键传播指标(如基本繁殖数 R_0 和有效繁殖数 R_e)的变化规律。特别地, 当 $R_e > 1$ 时疫情呈爆发趋势, 而 $R_e < 1$ 时疫情逐渐消退。此外, 模型还定量分析了不同防控措施(如佩戴口罩、减少社交活动)对疫情传播的抑制效果, 并探讨了群体免疫阈值的数学依据。本文的数学模型为公共卫生决策提供了科学支持, 帮助评估防控措施效果并预测疫情发展趋势, 同时展示了数学建模在应对突发传染病中的实用价值。

关键词

COVID-19传染病, SIR模型, 离散时间, 递推方程, 爆发预测

A Mathematical Modeling and Prediction Method for COVID-19-Like Infectious Disease Outbreaks

Jia'en Zheng¹, Ruobin Zheng²

¹Longhua Campus, Shenzhen Foreign Language School, Shenzhen Guangdong

²2012 Lab, Huawei Technologies Co., Ltd, Shenzhen Guangdong

Received: September 10, 2025; accepted: September 30, 2025; published: October 13, 2025

Abstract

This paper proposes an improved SIR (Susceptible-Infective-Removed) model based on a discrete-time recursive method for predicting the spread of COVID-19-like infectious diseases. The model divides the population into three categories: susceptible (S), infected (I), and removed (R). By introducing dynamic infection rates (β) and removal rates (γ), a simplified, real-time, discrete-time recursive equation is established, replacing the traditional, computationally complex ordinary differential equations. Actual epidemic data from the United States are used as a case study to validate the model's effectiveness in short-term forecasting and long-term trend analysis. Results demonstrate that the model fits the epidemic curve well and accurately captures the dynamics of key transmission indicators, such as the basic reproduction number (R_0) and the effective reproduction number (R_e). Specifically, when $R_e > 1$, the epidemic exhibits an explosive trend, while when $R_e < 1$, the epidemic gradually subsides. Furthermore, the model quantitatively analyzes the effectiveness of different prevention and control measures (such as wearing masks and reducing social activities) on epidemic spread and explores the mathematical basis for the herd immunity threshold. This mathematical model provides scientific support for public health decision-making, helps evaluate the effectiveness of prevention and control measures, and predicts epidemic trends. It also demonstrates the practical value of this mathematical model in responding to infectious disease emergencies.

Keywords

COVID-19 Epidemic, SIR Model, Discrete Time, Recurrence Equation, Outbreak Prediction

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

从人类历史的长河来看, 传染病和流行病一直是人类社会面临的重大挑战。在中世纪欧洲, 鼠疫(黑死病)的爆发导致约三分之一的人口丧生。1918年, 第一次世界大战结束后不久, 流感大流行席卷全球, 造成了近一亿人口的死亡。近年来, SARS (Severe acute respiratory syndrome)、流感和 COVID-19 (新冠肺炎)等流行病的爆发, 也对公共安全、国家经济、社会稳定构成了严重威胁。

在面对传染病的威胁时, 医疗专业人员和公共卫生部门必须迅速做出关键决策, 以阻止并预防流行病的传播。此时, 需要考虑以下一些重要问题:

- 流行病有多严重? 会大规模爆发吗?
- 将有多少人会感染?
- 最有效的控制或治疗策略是: 隔离或接种疫苗?

人类必须快速且高效地做出这些决策。否则, 疫情可能会迅速蔓延, 从而对全球健康构成严重威胁。

历史上, 就曾出现过因采取了错误的控制措施而使情况进一步恶化的例子, 例如 17 世纪的鼠疫和一战时的西班牙流感。从历史案例中, 确实可以获取许多有关流行病传播的信息, 但这些案例只能提供大致的指导方向。但是, 卫生部门不可能执行流行病学的实验, 因为人类必须面对严肃的道德方面的考虑, 还存在数据不完整或完全缺失等问题。而数学建模为解决这些复杂的公共卫生问题提供了一种有效的途径。Bernoulli、Ronald Ross (诺贝尔奖获得者), 以及 Anderson Gray McKendrick 和 William Ogilvy Kermack

基于一些相对较简单的假设开发了“房室(Compartment)模型” [1] [2]来预测流行病的行为。

本文重点研究 Ronald Ross 等人房室模型中的一个变体, “SIR”模型[3] [4], 但该模型基于连续时间下的常微分方程, 求解复杂, 运算量大。为了将其实际运用到现实世界中, 本文提出一种 COVID-19 类传染病爆发的数学建模预测方法, 运用离散时间递推方法简化了常微分方程的求解, 以模拟 COVID-19 在不同人群中的传播情况。本文通过数学建模尝试预测 COVID-19 类传染病的部分行为, 并进一步与真实数据进行了初步的拟合验证。

2. COVID-19 类传染病数学建模

2.1. 模型假设

为了建立数学模型, 我们假设每名涉疫人员分别属于以下三种不同的房室(Compartment)状态: 如易感染者状态、感染者状态和移除者状态。其中:

(1) 易感者状态: 指未得病者, 但缺乏免疫能力, 与感染者接触后容易受到感染;

(2) 感染者状态: 指染上传染病的人, 可以传播给 S 类成员, 将其变为 I 类成员;

(3) 移除者状态: 指被隔离人员, 或因病愈而具有免疫力的人, 或病故人员。传统“SIR”模型[3] [4]缺乏对传染病潜伏期的考虑, 因此衍生出各种派生复杂潜伏期状态的模型。为了简化模型复杂度, 本文提出移除者状态包括因传染病检测呈阳性而被隔离人员, 即模型考虑了被发现传染病进入潜伏期的涉疫人员, 这样既能较好还原涉疫情况, 又能简化模型的复杂度。如免疫期小于传染病传播期, 移除者可以重新变为易感者。

在数学建模过程中, 进一步考虑以下假设:

- 时间 t 的单位为天, 便于医务工作人员实际统计。则在 t 时刻, 相应状态类型的人群数量分别对应的函数为 $S(t)$ 、 $I(t)$ 、 $R(t)$ 。
- 在传染病传播期内, 涉疫地区的人口总数为 N 且保持不变, 包括易感染者(S)、感染者(I)和移除者(R), 即 $N = S(t) + I(t) + R(t)$ 。
- 假设免疫期大于传染病传播期, 移除者(R)包括痊愈者和死亡者, 此类人群均不会再次感染病毒。

2.2. 状态图与常微分方程表达

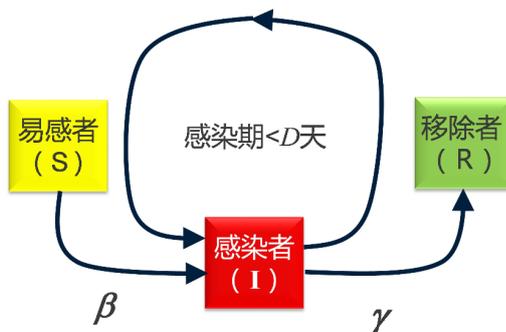


Figure 1. Model state diagram

图 1. 模型状态图

COVID-19 类传染病的模型状态图如图 1 所示。图中, 每个方框代表模型假设中所述的一种状态。涉疫人员可以在不同房室之间移动。其中, 易感人群以 β/N 的速率从易感者(S)状态向感染者(I)状态迁移; 从感染者(I)状态到感染者(I)状态的箭头表示一名涉疫人员在平均感染期 D 天内保持感染者(I)状态; 被感

染人群以 γ 的速率从感染者(I)状态向移除者(R)状态迁移。上述系统通常用以下常微分方程组(1)~(3)表示 [4] [5]:

$$\frac{dS}{dt} = -\frac{\beta \cdot I \cdot S}{N} \tag{1}$$

$$\frac{dI}{dt} = \frac{\beta \cdot I \cdot S}{N} - \gamma \cdot I \tag{2}$$

$$\frac{dR}{dt} = \gamma \cdot I \tag{3}$$

其中, 感染率 β 表示在单位时间内一个人获得足够传染病传播接触机会(即足以实现传染病传播的接触机会)的平均接触次数。 $D = 1/\gamma$ 为平均传染时间, 移除率 γ 表示在单位时间内一名感染者(I)被隔离、痊愈或死亡后变为移除者(R)的平均速率。

在传染病流行期间, 通常 β 或 γ 会随时间而变化。因此, 在流行病学中 Reproduction Number (繁殖数) [6] 应该表达为时间的函数——有效繁殖数 $R_e(t)$ [7], 在下文中亦可简称为 R_e 。 $R_e(t)$ 定义为在易感人群中一个典型感染者在某个时间 t 直接引发的新感染病例数。 $R_e(t)$ 估算的难点在于实时, 计算量巨大, 而这正是本文计算方法所要解决的问题之一。为此, 本文引入一个中间变量——实时繁殖数 R_t , R_t 定义为在某个时间 t 被一个典型感染者直接引发的感染病例数。当最初第一个感染者出现时, $R_t(t=0)$ 退化为 R_0 , R_0 定义为当一个典型感染者被引入到一个所有个体都是易感者状态的群体中时, 该感染者直接引发的感染病例数。所述繁殖数的具体意义将在 4.3 章节进一步讨论。根据以上定义, 三个繁殖数相应的计算公式如下:

$$\text{实时繁殖数 } R_t = \frac{\beta(t)}{\gamma(t)} \tag{4}$$

$$\text{有效繁殖数 } R_e = \frac{\beta(t)}{\gamma(t)} \cdot \frac{S(t)}{N(t)} \tag{5}$$

$$\text{基本繁殖数 } R_0 = \frac{\beta(t=0)}{\gamma(t=0)} \tag{6}$$

2.3. 离散时间递推方程表达

上述基于连续时间下的常微分方程(1)~(3), 求解复杂, 运算量大。为了将其实际运用到现实世界中, 本文运用离散时间递推方法简化常微分方程的求解, 以模拟 COVID-19 在不同人群中的传播情况。

为此, 将 dt 转换为时段 $\Delta t = t_{i+1} - t_i$, 从而将常微分方程(1)~(3)转换为离散时间递推方程(7)~(9)。为了与国家疾控中心原始监测数据兼容, 令 $\Delta t = t_{i+1} - t_i = 1$ 天为单位时段, 而不采用更精细的时段。其中, 下标 $i+1$ 和 i 表示两个相邻 Δt 的日子, 例如 $i+1$ 表示今天, i 表示昨天。

$$S_{i+1} = S_i - \left[\beta_i \cdot \left(\frac{I_i}{N} \right) \cdot S_i \right] \Delta t \tag{7}$$

$$I_{i+1} = I_i + \left[\beta_i \cdot \left(\frac{I_i}{N} \right) \cdot S_i \right] \Delta t - (\gamma_i \cdot I_i) \Delta t \tag{8}$$

$$R_{i+1} = R_i + (\gamma_i \cdot I_i) \Delta t \tag{9}$$

其中, S_i 表示在时段 i 易感者的数量; I_i 表示在时段 i 感染者的数量; R_i 表示在时段 i 移除者的数量。 β_i 表示在时段 i 的 β 数值; γ_i 表示在时段 i 的 γ 数值。

$$\text{新增感染人数}_{i+1} = \beta_i \cdot \left(\frac{I_i}{N}\right) \cdot S_i \cdot \Delta t \tag{10}$$

$$\text{新增移除人数}_{i+1} = \frac{I_i}{D} \cdot \Delta t = \gamma_i \cdot I_i \cdot \Delta t \tag{11}$$

图 2 给出了新增感染人数的预测公式(10)的医学解释。

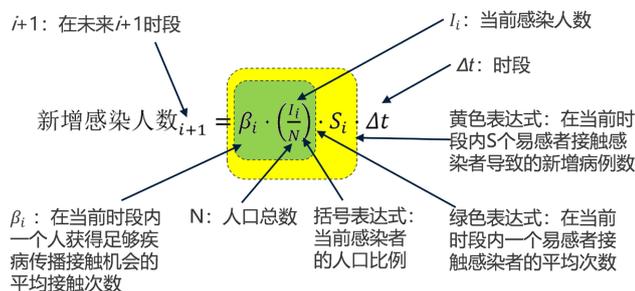
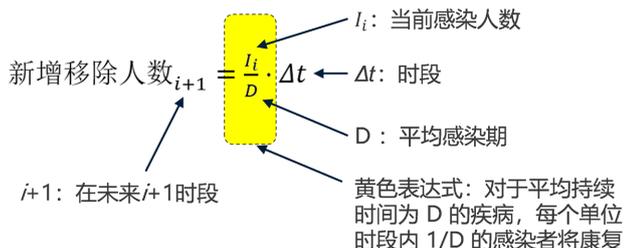


Figure 2. Medical explanation of Formula (10)

图 2. 公式(10)的医学解释

图 3 给出了新增移除人数的预测公式(11)的医学解释。其中, $D = 1/\gamma$ 为平均传染时间, 例如, 如果感冒平均持续 $D = 7$ 天, 则每天有 $\gamma = 1/7$ 的当前感染者会恢复健康。COVID-19 是有潜伏期的, 人在潜伏期内抗体检测呈阳性或出现新冠肺炎的临床特征涉疫人员就会被隔离, 因此平均传染时间 D 可以设定为 COVID-19 的平均潜伏期。



例如, 如果感冒持续5天, 我们假设每天有 1/5 的当前感染者会恢复健康。在这 1/5 中, 可能有些人患病时间超过5天, 而其他人患病时间短于5天。

Figure 3. Medical explanation of Formula (11)

图 3. 公式(11)的医学解释

图 4 给出了未来易感者人数预测公式(7)的医学解释, 为了预测未来 $i+1$ 时段的易感者人数, 可以将当前易感者人数减去当前时间段内新增感染者人数。

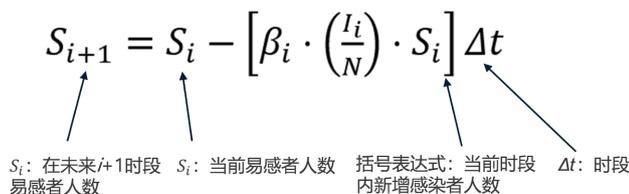


Figure 4. Medical explanation of Formula (7)

图 4. 公式(7)的医学解释

图5给出了未来感染者人数预测公式(8)的医学解释, 为了预测未来 $i+1$ 时段感染者人数, 可将当前时段内新增感染人数添加到当前感染者人数中, 然后减去当前时段内新增移除者人数。

$$I_{i+1} = I_i + \left[\beta_i \cdot \left(\frac{I_i}{N} \right) \cdot S_i \right] \Delta t - (\gamma_i \cdot I_i) \Delta t$$

I_{i+1} : 在未来 $i+1$ 时段感染者人数
 I_i : 当前感染者人数
 括号表达式: 当前时段内新增感染人数
 括号表达式: 当前时段内新增移除者人数
 Δt : 时段

Figure 5. Medical explanation of Formula (8)
图5. 公式(8)的医学解释

图6给出了未来移除者人数预测公式(9)的医学解释, 为了预测未来 $i+1$ 时段移除者人数, 可将当前时段内新增移除者人数添加到当前移除者人数。

$$R_{i+1} = R_i + (\gamma_i \cdot I_i) \Delta t$$

R_{i+1} : 在未来 $i+1$ 时段移除者人数
 R_i : 当前移除者人数
 括号表达式: 当前时段内新增移除者人数
 Δt : 时段

Figure 6. Medical explanation of Formula (9)
图6. 公式(9)的医学解释

2.4. 离散时间递推方程求解 R_t 和 $R_e(t)$

如何确定传染病爆发初期的 R_t 和 $R_e(t)$ 值一直是困扰卫生防疫政策制定者的一个难点。本文中的递推公式(10)可以被用来估算 β_i , 即通过将公式(10)中预测的新增感染人数 $i+1$ 替换为国家疾控中心公布的每日新增病例数来估算 β_i :

$$\beta_i = \frac{\text{公布的每日新增病例数}}{\Delta t} \cdot \frac{1}{\left(\frac{I_i}{N} \right) \cdot S_i} \tag{12}$$

由公式(4)计算得到实时 R_t 值:

$$R_{t_i} = \frac{\beta_i}{\gamma_i} = \frac{\text{公布的每日新增病例数}}{\Delta t} \cdot \frac{1}{\left(\frac{I_i}{N} \right) \cdot S_i \cdot \gamma_i} \tag{13}$$

由公式(5)计算出实时 R_e 值:

$$R_{e_i} = \frac{\beta_i \cdot S_i}{\gamma_i \cdot N} \tag{14}$$

得到 β_i 后, 即可运用公式(7)~(9)预测未来 $i+1$ 时段(即下一天)的 S 、 I 、 R , 以及运用公式(10)预测未来 $i+1$ 时段新增感染者人数(即每日新增病例数)。

3. 原始数据来源

根据中国传染病防疫中心观察表明[8], 新冠肺炎的临床特征在感染后 1~14 天内出现, 大多数患者在

3~7 天内出现症状, 平均潜伏期为 5 天, 这个数据在疫情的初期就能获得。所以, COVID-19 平均传染时间 D 可以设定为 5 天, $\gamma = 1/D = 0.2$ 。

本文选取美国俄亥俄州 COVID-19 疫情作为实验对象, 数据来源于 GitHub 收录的美国《纽约时报》自 2020 年以来 COVID-19 数据报道的可下载档案[9]。本文仅使用了其中的 date (日期)和 cases (每日累计病例数), 每日新增病例数可通过前后日累计病例数相减得到。

4. 基于数学模型进行 COVID-19 数据分析

4.1. 原始数据的平滑处理

为了去除 COVID-19 原始数据中可能出现的噪声毛刺, 可以使用前后 3 天每日病例变化的移动平均值来生成平滑处理后的每日病例数, 参见图 7。

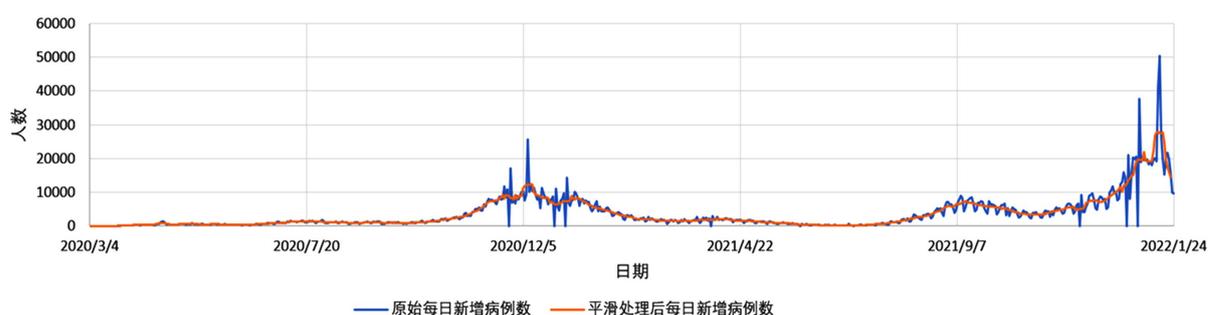


Figure 7. Smoothing of the raw daily number of new cases

图 7. 原始每日新增病例数的平滑处理

4.2. COVID-19 数据计算处理

1. 输入初始值: 美国俄亥俄州的人口总数 $N = 11,690,000$;
COVID-19 平均传染时间 $D = 5$ 天, 则 $\gamma = 1/D = 0.2$;
单位时段 $\Delta t = 1$ 天。

输入 COVID-19 原始数据: 2020 年 3 月 9 日至 2022 年 1 月 25 日的每日累计病例数。

2. 计算: 根据公式(12)~(14), 计算得到实时 β 、 R_t 和 $R_e(t)$ 每日数值, 如图 8 所示。



Figure 8. Real-time calculation of reproductive rate

图 8. 繁殖率的实时计算

3. 输出: 根据公式(7)~(9), 可得到易感者人数 S、感染者人数 I、移除者(隔离者或康复者)人数 R 的每日预测值, 如图 9 所示。由于感染者人数 I 远小于易感者人数 S, 为此特地将感染者人数 I 单独示于图 10, 以便于观察其变化趋势细节。

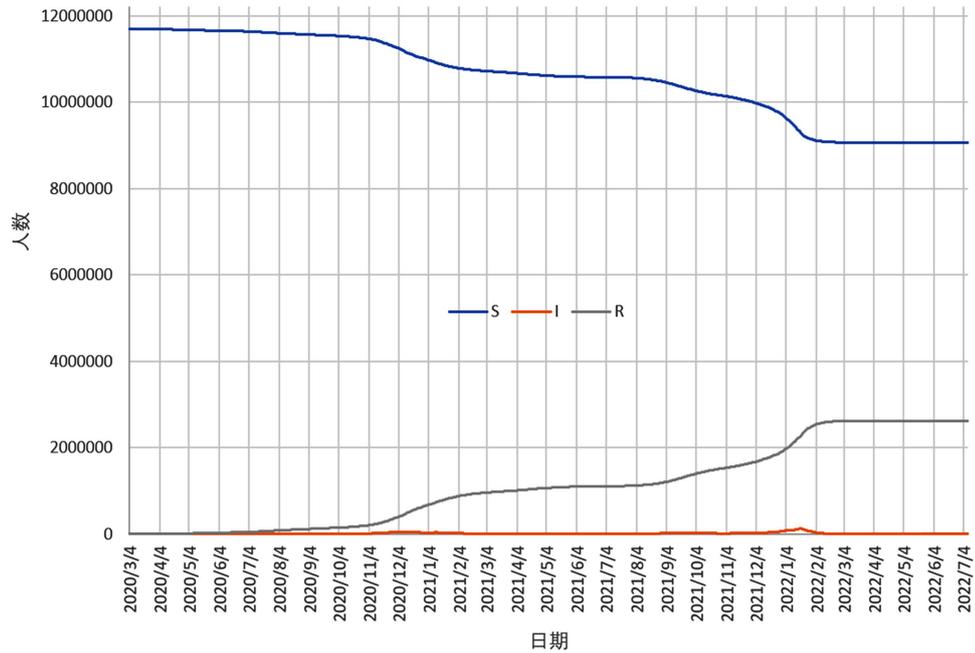


Figure 9. Simulation of prediction of the number of S, I, and R
图 9. S、I、R 人数模拟预测

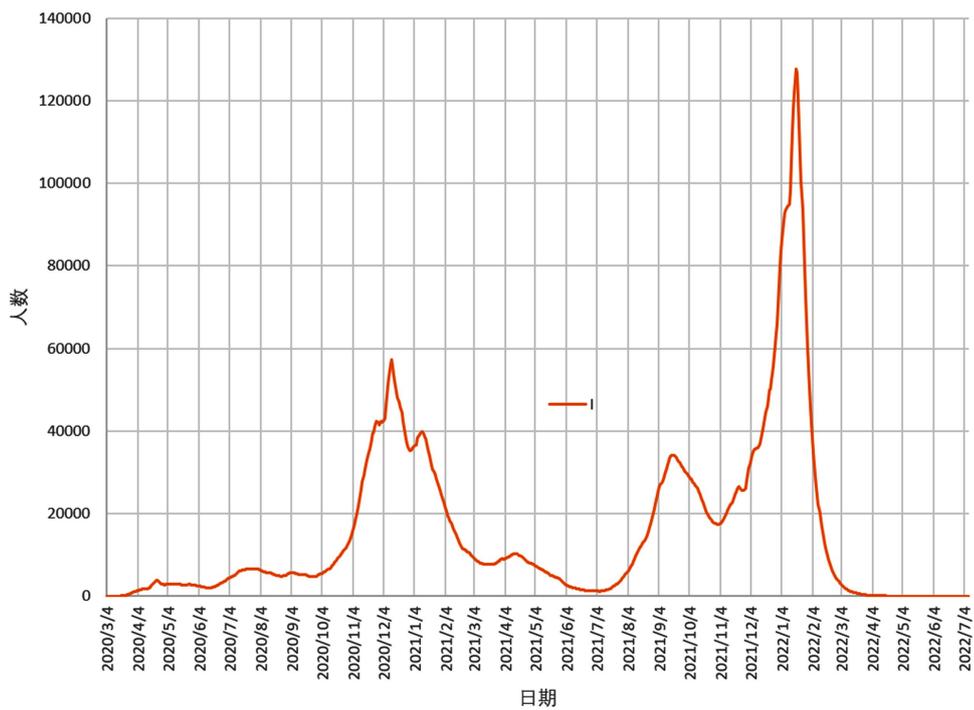


Figure 10. Simulation of prediction of the number of infected cases
图 10. I (感染者人数)模拟预测

4.3. R_t 与 R_c 的数据分析

上述 4.2 章节的数学模型输入的是 COVID-19 原始数据, 因此反应了美国疾控中心 CDC 所制定的防疫措施所产生的综合结果。为了理解 $R_c(t)$ 医学意义, 我们需要剔除美国疾控中心 CDC 所制定的防疫措施所带来的影响, 以推断单次疫情爆发的可能性。为此, 做如下假设性输入。

1. 输入初始值: 美国俄亥俄州的人口总数 $N = 11,690,000$;
COVID-19 平均传染时间 $D = 5$ 天, 则 $\gamma = 1/D = 0.2$;
 R_0 选取 2.8, 与 4.2 章节保持一致;
单位时段 $\Delta t = 1$ 天。

不再输入 COVID-19 原始数据, 以剔除美国疾控中心 CDC 所制定的防疫措施所带来的影响。

2. 计算: 根据公式(13)~(14), 计算得到 R_t 和 $R_c(t)$ 每日数值, 如图 11 所示。

3. 输出: 根据公式(7)~(9), 可得到图 11 所示的单次疫情爆发模型的易感者人数 S 、感染者人数 I 、移除者(隔离者或康复者)人数 R 的每日预测值。

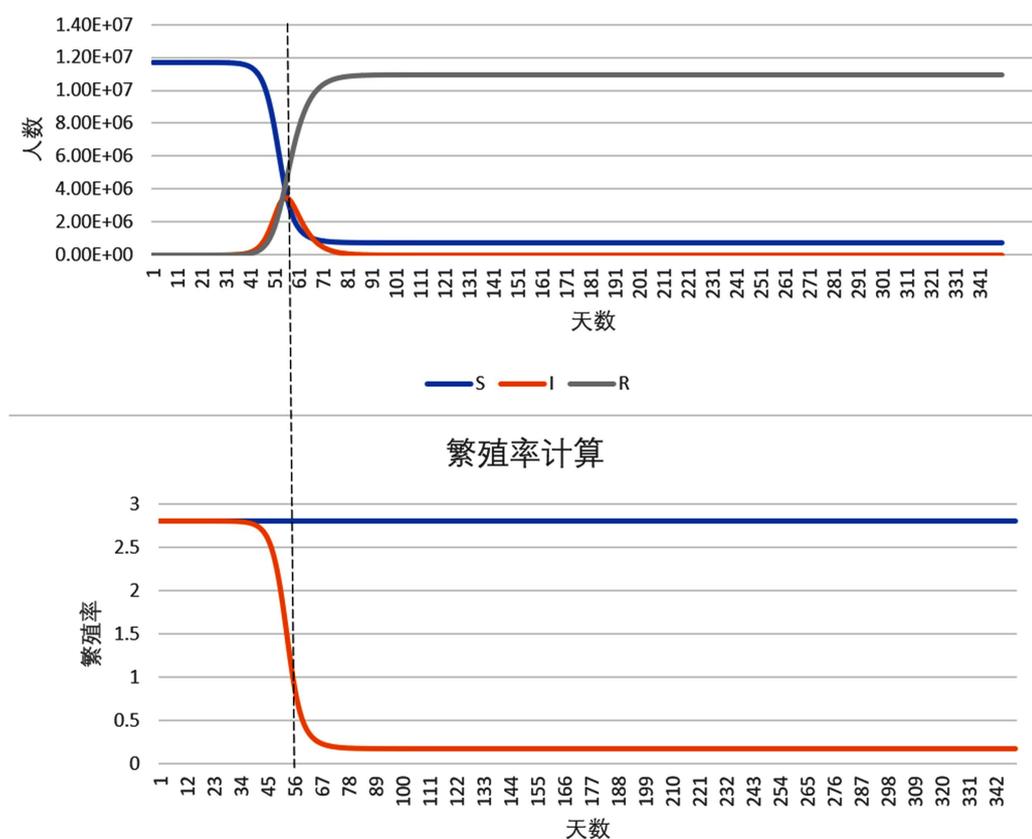


Figure 11. Single outbreak prediction (no epidemic prevention measures, $R_0 = 2.8$)

图 11. 单次疫情爆发预测(无防疫措施, $R_0 = 2.8$)

4. 分析与讨论:

从图 11 可以看出: 当 $R_c(t) > 1$ 时, 方程(2)中 dI/dt 为正值, 感染者人数 R 呈指数增长, 意味着疫情将会流行并爆发; 而当 $R_c(t) = 1$ 时, 方程(2)中 dI/dt 等于 0, 感染者人数 R 曲线达到“稳定状态”; 当 $R_c(t) < 1$ 时, 方程(2)中 dI/dt 为负值, 感染者人数 R 呈指数衰减, 意味着疫情将逐渐消失。另外, 从图 11

还可以看出：如果不采取任何疫情防控措施， R_t 会随时间保持恒定在 R_0 。

所以，如果我们能数学建模实时计算出 $R_c(t)$ 值，则可以及时判断疫情变化趋势，从而为疫情防控策略的制定提供科学依据。

例如，根据 4.2 章节计算得到的初始 $R_c(t=2020 \text{ 年 } 3 \text{ 月 } 9 \text{ 日}) = R_0 = 2.8 > 1$ ，可知美国俄亥俄州的疫情将会爆发。如果美国俄亥俄州不采取任何防控措施，根据图 11 可知在疫情出现 56 天后感染者人数 $R = 3,454,806$ 人达到最高点的，在疫情出现 144 天后疫情消失。虽然可以预测出疫情会自然消失，但感染者人数过于巨大，全州医院资源会被单一传染病所耗尽，死亡病例也会巨大，所以必须制定疫情防控措施，如戴口罩、减少人员流动甚至隔离等。

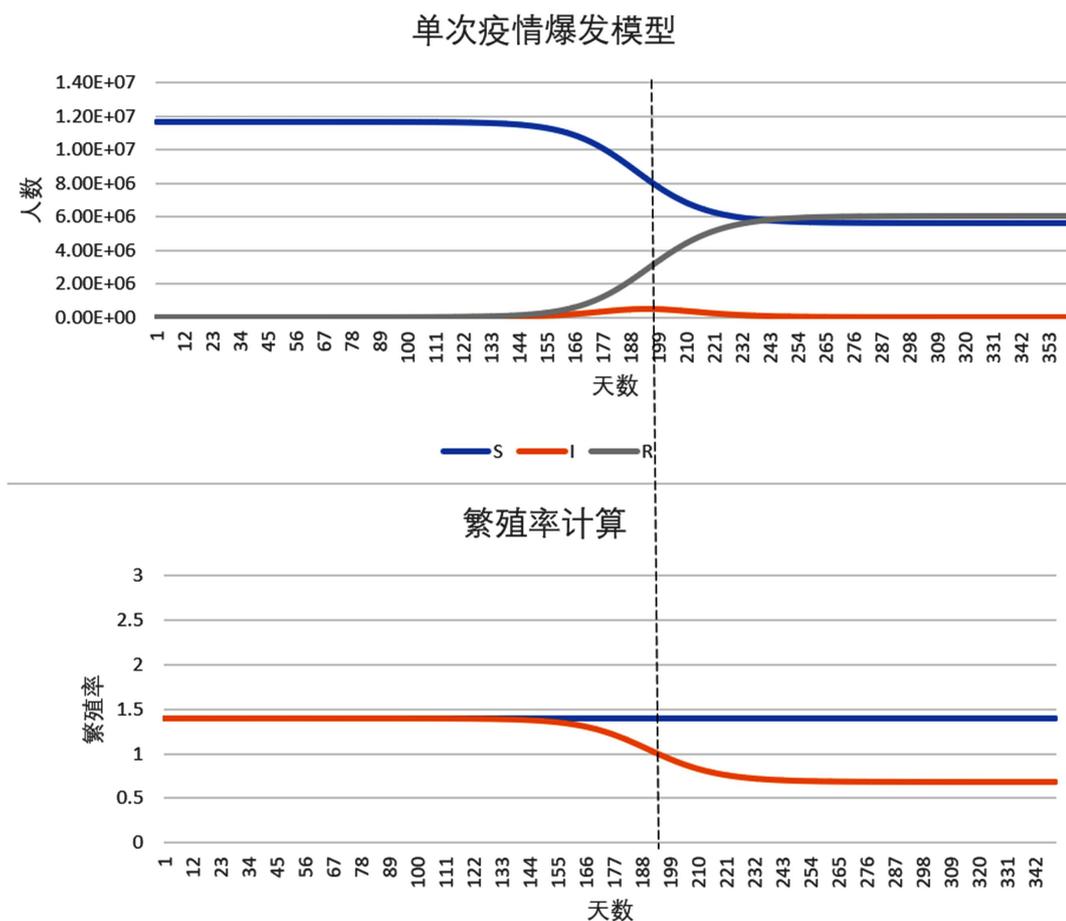


Figure 12. Single outbreak prediction (wearing masks, $R_0 = 1.4$)

图 12. 单次疫情爆发预测(戴口罩, $R_0 = 1.4$)

防疫管控措施以口罩为例，如果口罩可以阻止 50% 的病毒传播，那么倘若每个人都戴上口罩，则可以将 R_c 值降低 50%；减少 50% 的社交活动可以使 R_c 值进一步减半。

假设美国俄亥俄州在疫情初期就借鉴中国的防疫措施，坚持戴口罩令，将 R_0 降低 50% 至 1.4，则单次疫情爆发模型的易感者人数 S 、感染者人数 I 、移除者(隔离者或康复者)人数 R 的每日预测值，如图 12 所示，可知在疫情出现 155 天后感染者人数 $R = 542,558$ 人达到最高点的，在疫情出现 411 天后疫情消失。对比图 11 无防疫措施 $R_0 = 2.8$ 的单次疫情爆发模型，虽然疫情感染者人数最高点出现时间由 56 天推迟到了 155 天，疫情消失时间由 114 天推迟到了 411 天，但感染者人数由 3,454,806 人大幅降低至

542,558 人, 足足降低了一个数量级。

注意到如果本文的所有模型假设都适用于 2020~2023 年的 COVID-19 疫情, 那么常微分方程组(1)~(3)就意味着只会会有一个峰值。但实际情况是如图 10 所示, 虽然美国俄亥俄州采取了某些疫情防控措施, 但在 2020~2022 年期间连续出现了三次疫情大爆发, 第三次疫情大爆发的感染者人数甚至高于前两次一个数量级。似乎表明: 美国俄亥俄州的疫情防控策略是相对失败的, 又或者与病毒变异为传染性更高的变种而不得不放开采取“群体免疫”的策略有关。其中“群体免疫”将在第 6 章节讨论。

为此, 本文将模型计算出的 $R_e(t)$ 曲线(参见图 8)与美国俄亥俄州在疫情期间的关键事件(如政策发布/解除、新变种出现、疫苗接种推广等)进行叠加对比分析, 如图 13。以 2020 年 3 月 9 日为起点, 美国俄亥俄州在(一)阶段无任何防控措施; 2020 年 3 月 23 日, 居家令生效进入(二)阶段, R_e 值有所下降但出现反复; 2020 年 5 月初, 口罩令生效, 进入(三)阶段, R_e 值稳步下降; 2020 年 6 月 2 日, 开始放松防疫管控, 居家令、口罩令解除, 商店重新开业, 进入(四)阶段, 结果 R_e 值出现小幅上升趋势; 2020 年 7 月 22 日, 口罩强制令再次生效, 进入(五)阶段, 酒吧重新关闭, R_e 值出现下降趋势; 但在万圣节、感恩节、圣诞节和新年期间, 进入(六)阶段, 上升峰值与节日季的开始有关; 2020 年 11 月 11 日, 口罩强制令再次生效以压制节日季疫情的反弹; 美国政府于 2021 年 3 月 11 日, 宣布美国全国疫苗接种令, 进入(七)阶段; 截至 2021 年 4 月 19 日, 约有 50% 美国国民已经接种了最少一剂疫苗, 之后 R_e 值稳步下降; 2021 年 5 月 13 日, CDC 第二次放松防疫管控, 进入(八)阶段; 数据显示: 2021 年 5 月, 俄亥俄州实验室测序病例中不到 1% 被确诊为 Delta 变异株, 但至 2021 年 7 月 17 日, 86.4% 的俄亥俄州实验室测序病例为 Delta 变异株, Delta 变异株经历了从出现到大流行, R_e 值也随之由低谷升至阶段性峰值; CDC 于 2021 年 7 月 27 日又改口建议戴口罩, 进入(九)阶段, R_e 值开始下降; 到 2021 年 11 月入冬后, Omicron 变异株爆发, 进入(十)阶段, R_e 值再度高企。截至 2022 年 1 月初, 美国已有近 2.1 亿人口完全接种了新冠疫苗, 约占总人口的 63.7%, 之后开始了第三次放松防疫管控。

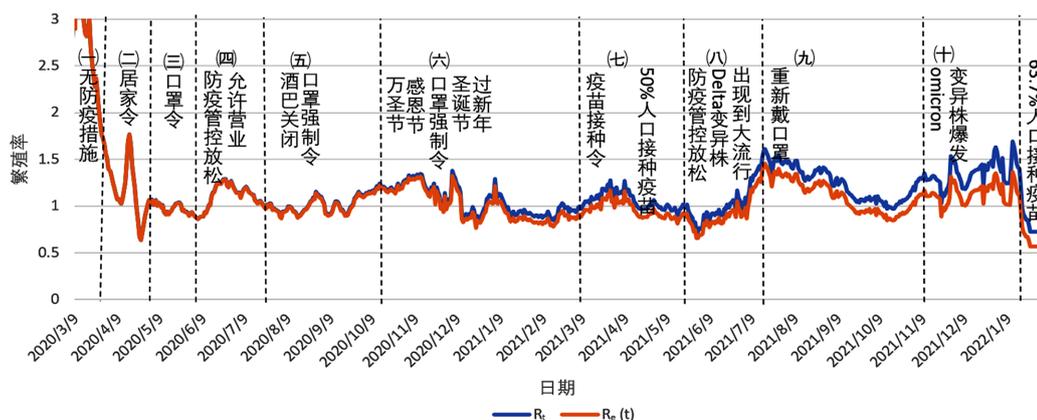


Figure 13. The relationship between reproduction rate and key events
图 13. 繁殖率与关键事件的关系

5. 模型预测与原始数据的比较验证

5.1. 验证一：每日预测验证

1. 比较对象:

- 美国疾控中心 CDC 公布的每日新增病例数, 时间范围: 2020 年 3 月 9 日至 2022 年 1 月 25 日;
- 基于本文数学模型的新增感染人数预测值 I_{i+1} , 时间范围: 2020 年 3 月 9 日至 2022 年 1 月 25 日至

2022年7月8日。

2. 比较区间: 2020年3月9日至2022年1月25日。

比较结果按公式(15)量化为残差, 如图14所示。图14在2022年1月25日之后仅有预测曲线, 也就是本文的模型2022年1月25日之后是不喂入原始数据进行每日预测的调整的, 即2022年1月25日之后模型进入长时段预测过程。

3. 结论: 两者在趋势上基本吻合, 在疫情爆发的高点有较大的残差。

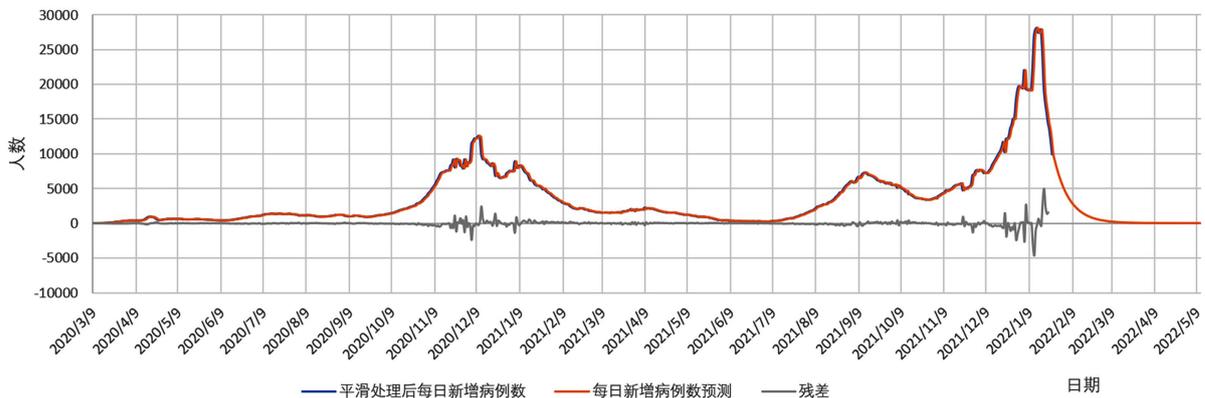


Figure 14. Daily new case verification

图14. 每日新增病例验证

5.2. 验证二: 长时段趋势预测验证

1. 比较对象:

- 美国疾控中心 CDC 公布的每日新增病例数, 时间范围: 2022年1月26日至2022年3月16日;
- 基于本文数学模型的新增感染人数预测值 I_{t+1} , 时间范围: 2022年1月26日至2022年7月8日。本文的模型在这段时间内是不喂入原始数据进行每日预测的调整的, 即模型进入长时段预测过程。

2. 比较区间: 2022年1月26日至2022年3月16日。

比较结果, 如图15所示。2022年1月25日之后仅有预测曲线, 也就是本文的模型2022年1月25日之后是不喂入原始数据进行每日预测调整的, 即2022年1月25日之后模型进入长时段预测过程。

3. 结论: 两者在趋势上基本吻合, 预测值大概比原始数据提前了一天, 所以残差为负。

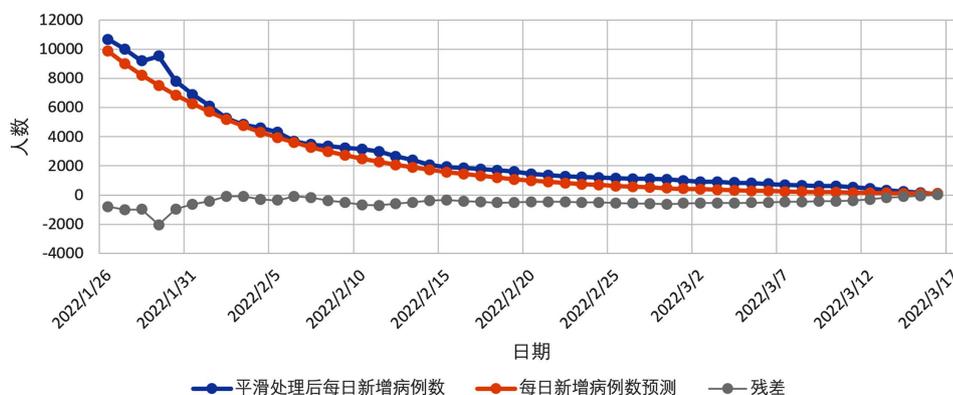


Figure 15. Daily new case verification

图15. 每日新增病例验证

6. 群体免疫的数学模型解释

防疫政策决策者会遇到另一个难题：“群体免疫”是真的吗？有科学依据吗？通过本文的数学建模大体能回答这个问题。

回顾 4.3 章节关于 R_e 的讨论，为了预防传染病大规模流行爆发，要求： $R_e \leq 1$ 。

假设 ρ 表示易感人群接种疫苗的比例，接种疫苗的人将从易感者状态转移到移除者状态，因此易感者人数 $S(t)$ 变为： $(1-\rho)S(t)$ 。从而，根据公式(5)可得：

$$\frac{\beta(t)}{\gamma(t)} \cdot \frac{(1-\rho)S(t)}{N(t)} = (1-\rho) \cdot R_e \leq 1 \quad (15)$$

当存在 $\rho \geq \rho_c$ 时上式成立， $\rho_c = 1 - 1/R_e$ ， ρ_c 即为群体免疫阈值[4]。

对于 $R_e \approx 3$ 的 COVID-19 传染病，为了防止一个易感人群的疫情爆发，“仅”须为 $\rho_c = 1 - 1/R_e = 66.7\%$ 的易感人口接种疫苗即可预防疫情流行。如果疫苗的有效性只有 80%，那么必须至少为 $66.7\%/80\% = 83.4\%$ 的易感人口接种疫苗才能预防疫情流行。

因此，COVID-19 群体免疫阈值的确存在，即使不是 100% 人口都获得免疫力，针对 COVID-19 传染病的疫苗接种也可以有效预防疫情流行。

7. 结语

本文提出一种 COVID-19 类传染病爆发的数学建模预测方法，运用离散时间递推方法简化了常微分方程的求解，以模拟 COVID-19 在不同人群中的传播情况；也探讨了提取非线性常微分方程解的部分有用属性(如 R_e)，即使系统没有明确的公式解。本文利用数学建模对 COVID-19 类传染病的爆发及其防控产生一些基本见解——这些见解似乎不可能仅从感染数据中辨别出来。本文尝试为公共卫生干预提供了一种理论数学模型，从而在一定程度上帮助医务工作者模拟传染病传播的动态过程，预测疫情的发展趋势，评估不同防控措施的效果。本文的计算机模拟，虽然部分弥补了实际数据的不足，为应对突发公共卫生事件提供响应方案的部分快速事前评估，但由于本文的数学建模是对原始复杂系统的一种简化处理，并且本文的验证范围仅对美国部分地区实施了验证，难免存在局限性与不可预测因素，例如因疑似阳性被隔离人员数据与真实情况存在一定误差，所以有待在后续的工作中持续进行优化。

参考文献

- [1] Kermack, W. and McKendrick, A. (1991) Contributions to the Mathematical Theory of Epidemics—I. *Bulletin of Mathematical Biology*, **53**, 33-55.
- [2] Kermack, W. and McKendrick, A. (1991) Contributions to the Mathematical Theory of Epidemics—II. The Problem of Endemicity. *Bulletin of Mathematical Biology*, **53**, 57-87. [https://doi.org/10.1016/s0092-8240\(05\)80041-2](https://doi.org/10.1016/s0092-8240(05)80041-2)
- [3] Anderson, R.M. (1991) Discussion: The Kermack-Mckendrick Epidemic Threshold Theorem. *Bulletin of Mathematical Biology*, **53**, 1-32. <https://doi.org/10.1007/bf02464422>
- [4] Weiss, H.H. (2013) The SIR Model and the Foundations of Public Health. *Materials Matemàtics*, **2013**, 17.
- [5] Hethcote, H. (2000) The Mathematics of Infectious Diseases. *SIAM Review*, **42**, 599-653.
- [6] Basic Reproduction Number. https://en.wikipedia.org/wiki/Basic_reproduction_number
- [7] Marinov, T.T., Marinova, R.S., Marinov, R.T. and Shelby, N. (2023) Novel Approach for Identification of Basic and Effective Reproduction Numbers Illustrated with COVID-19. *Viruses*, **15**, Article No. 1352. <https://doi.org/10.3390/v15061352>
- [8] 杨孝坤, 等. 新型冠状病毒感染不同阶段的传染性研究进展[J]. 中华流行病学杂志, 2021, 42(1): 33-38.
- [9] The New York Times. GitHub-Nytimes/COVID-19-Data: An Ongoing Repository of Data on Coronavirus Cases and Deaths in the U.S. <https://github.com/nytimes/covid-19-data>