

多波段 Richards 模型及其数据拟合分析

张悦, 彭源源, 杨志春

重庆师范大学 数学科学学院, 重庆 401331

摘要:目的 在传染病传播的过程中, 往往会经过反复感染, 出现传播多次高峰。基于这个特点, 提出了一种多波段 Richards 模型, 对传染病的感染人数进行数值模拟。方法 利用多波段 Richards 模型结合最小二乘法对西太平洋区域的累计感染病例进行拟合, 得到感染率、拐点、最终感染人数等参数估计值并判断其是否收敛, 从而确定传染病传播的波段数以及每个波段的拐点日期。结果 结果表明: 西太平洋区域在 2020-01-04—2020-09-13 暴发的传染病分为 3 个波段, 其中 03-15 和 07-04 为分割 3 个波段的两个拐点, 由第一波段求得的基本再生数约为 8.53, 说明这次传染病的传播能力很强。结论 借助多波段 Richards 模型可以有效拟合多波段传染病的感染情况, 能够确定传染病在多阶段暴发中的关键转折点, 以及预测未来一段时间的感染规模, 为制定传染病防控措施提供参考。

关键词:多波段 Richards 模型; 基本再生数; 拐点; 最终感染数

中图分类号: O29 文献标识码: A doi: 10.16055/j.issn.1672-058X.2024.0006.012

Multi-band Richards Model and Its Data Fitting Analysis

ZHANG Yue, PENG Yuanyuan, YANG Zhichun

School of Mathematical Sciences, Chongqing Normal University, Chongqing 401331, China

Abstract: Objective In the process of infectious disease transmission, it tends to go through repeated infections with multiple transmission peaks. Based on this characteristic, a multi-band Richards model was proposed to numerically simulate the number of infected people of infectious diseases. **Methods** The multi-band Richards model was used in conjunction with the least squares method to fit the cumulative infection cases in the Western Pacific region. Estimated values of parameters including infection rate, inflection point, and final number of infections were obtained, and convergence was evaluated to determine the number of bands and the inflection dates for each band in the spread of the disease. **Results** The results showed that the infectious disease outbreak in the Western Pacific region from January 4th, 2020 to September 13th, 2020 was divided into 3 bands, with the inflection points on March 15th and July 4th separating the three bands. The basic reproduction number obtained from the first band was about 8.53, indicating a strong transmission capacity of the disease. **Conclusion** With the help of the multi-band Richards model, the infection of multi-band infectious diseases can be effectively fitted, and the key turning points of infectious diseases in multistage outbreaks can be identified, as well as the scale of infections in the coming period can be predicted, which can provide a reference for the development of preventive and control measures for infectious diseases.

Keywords: multi-band Richards model; basic reproductive number; turning point; number of final infections

收稿日期: 2023-03-05 修回日期: 2023-05-18 文章编号: 1672-058X(2024)06-0094-05

基金项目: 国家自然科学基金(11971081, 10471061); 重庆市教委科学技术研究项目(KJZD-M202000502); 重庆市研究生科研创新项目(CYS23414); 重庆市教委雏鹰计划研究项目(CY220503)。

作者简介: 张悦(1998—), 女, 重庆永川人, 硕士研究生, 从事微分方程与动力系统研究。

通讯作者: 杨志春(1971—), 男, 重庆江津人, 教授, 博士生导师, 从事微分方程与动力系统研究。Email: yangzhch@126.com。

引用格式: 张悦, 彭源源, 杨志春. 多波段 Richards 模型及其数据拟合分析[J]. 重庆工商大学学报(自然科学版), 2024, 41(6): 94-98.

ZHANG Yue, PENG Yuanyuan, YANG Zhichun. Multi-band Richards model and its data fitting analysis [J]. Journal of Chongqing Technology and Business University (Natural Science Edition), 2024, 41(6): 94-98.

1 引言

传染性疾​​病深刻地影响着人类的生命健康和社会的经济发展,特别是近些年来出现的新型传染病。由于在传染病流行初期对其缺乏机理上的认识,有效疫苗的开发在短期(甚至一、二年内)存在困难,传染病暴发传播过程无法重复试验等因素,在传染病流行初期,利用数学模型来研究传染病的传播机制和控制策略至关重要^[1-3]。

1927年,Kermack等^[4]建造了著名的SIR仓室模型,以研究1658年至1666年伦敦黑死病的流行以及孟买1906年鼠疫的流行规律,接下来,又在1932年创建了SIS仓室模型,并在分析所建立模型的基础上,提出了区分疾病流行与否的“阈值理论”,为传染病动力学的研究奠定了基础。近年来,仓室模型成为许多专家学者研究的重点。Meng等^[5]提出具有非线性发病率和双重流行病假设的SIS传染病模型,结果表明:较大的随机扰动可导致传染病灭绝;钟南山院士团队开展了基于SEIR仓室优化模型和AI对公共卫生干预下的我国传染病流行趋势预测研究^[6];Xing等^[7]提出了一个非自主的SVIR复发流行模型,结果表明:为了控制传染病的流行,应该在降低疾病复发率的同时提高疫苗接种率。

上述仓室模型对传染病的传播机理有一定合理解释,但模型的多个变量可能导致模型参数鲁棒性能不强,反演计算中的参数估计容易陷入局部优化,计算过度依赖变量多个初值导致稳定性差等问题。一些简单的现象学模型,如指数增长模型、Logistic模型、Gompertz模型等,用于传染病初期数据分析,则可能显得简单有效。1959年,Richards提出的广义Logistic模型^[8],因其模型参数较少,相比Logistic曲线而言,克服了实际传播过程中S型曲线关于拐点的对称性缺陷,被较成功地应用于一些流行病如SARS、埃博拉等病毒传播的数据分析^[9-11]。

近些年出现的传染病并不只是单波大流行,而是呈现多阶段性,因此选择一种可以预测多波段传染病的模型至关重要。Blonigan等^[12]提出利用贝叶斯方法来推断和预测多波传染病,并在传染病案例中进行演示,提供了传染病演变的短期预测;Zuhairah等^[13]利用多波段Richards模型对9个国家的传染病数据进行分析,结果表明,Richards模型可利用变点检测来分析多波数据,从而更加准确的预测最后一波的初始数据;Hsieh等^[14]提出了一种多波段的Richards模型,通过实验证明了此模型是应对具有多个转折点的未来疾病暴发的宝贵工具。现在大部分的研究还停留在单波段传染病上,对于多波段传染病的研究方法还不完善。文献^[12]利用贝叶斯方法去预测分析了多波段传染病,

但没有得到传染病的感染率、拐点时刻、基本再生数等关键数据,不能为人们提供关键信息;文献^[13]中利用多波段Richards模型对9个国家多波段的感染数据进行了对比分析,但没有对单个区域传染病的感染情况进行具体分析,描述不够详尽;分析文献^[14]中利用多波段Richards模型仅探究了爆发的两个波段,没有对两个以上波段的感染进行研究,与现实情况有所差距。

在上述讨论的推动下,将建立改进后的多波段Richards模型,利用最小二乘法对西太平洋区域发生的三波段传染病进行拟合,通过参数反演法求得每个波段的感染率、拐点时间、最终感染数、基本再生数等关键数据,再根据这些数据进行具体分析,这对于全球传染病防控具有重要参考价值。

2 模型与方法

2.1 Richards 模型

1959年,Richards^[8]提出了以下生物生长模型,也被用作传染病增长模型:

$$y'(t) = ry \left[1 - \left(\frac{y}{K} \right)^a \right] \quad (1)$$

其中, $y(t)$ 是 t 时间的累积感染病例数, r 是传染率(感染人均增长率), K 是感染承载量或最终感染数;当 $a=1$ 时,模型为经典的Logistic模型, a 为偏离标准Logistic曲线的指数。

由式(1),可计算 $y(t)$ 二阶导数,得: $y''(t) = r^2 y \left[1 - \left(\frac{y}{K} \right)^a \right] \left[1 - \left(\frac{y}{K} \right)^a - a \left(\frac{y}{K} \right)^a \right]$ 。假设 t_i 为 $y(t)$ 的拐点时刻,即 $y''(t_i) = 0$ 。由上式不难得到式(1)过拐点的初始条件:

$$y(t_i) = K / [1 + a]^{1/a} \quad (2)$$

令 $z = y^{-a}$,则 $z'(t) = -arz + ar/K^a$,求解可得:

$$z(t) = Ce^{-art} + 1/K^a \quad (3)$$

$$y(t) = K / [1 + K^a Ce^{-art}]^{1/a}$$

把式(2)代入式(3),解得Richards模型式(1)的解析解为

$$y(t) = K / [1 + ae^{-ar(t-t_i)}]^{1/a} \quad (4)$$

注:现将采用式(4)作为Richards模型拟合曲线,曲线清楚地给出了拐点时刻 t_i 、感染率 r 、最终感染值 K 、偏离Logistic曲线指数 a 等关键参数。与文献^[15]中Richards模型的一般初始条件下的近似拟合曲线式(3)比较,式(4)给出的拟合曲线更有助于数据拟合中关键参数的直接获取。另外,在文献^[10-11]中,尽管采用了含拐点时刻的拟合曲线 $y(t) = K / [1 + e^{-r(t-t_i - \ln a/r)}]^{1/a}$ 来分析2003年发生的SARS病例数据,但不难验证表达式并非Richards模型式(1)的解析解。当然,当 a 接近于1时,曲线近似为Logistic曲线,

作为累计病例的拟合曲线也可能适用。

基本再生数 R_0 是刻画流行病暴发的重要参数,它表示在发病初期,当所有人均为易感者时,一个病人在其平均患病期内所传染的人数。通常, $R_0=1$ 可作为决定疾病是否消亡的一个阈值,即当 $R_0<1$ 时疾病将消除,当 $R_0>1$ 时疾病将流行。基于 Richards 模型的基本再生数计算可采用^[10] $R_0=e^{rT}$,其中 T 为平均潜伏期。假设西太平洋区域传染性 T 的持续时间为 5.2 d。

式(4)的变化曲线依赖于参数 $\theta=(a,r,t_i,K)$,可以利用最小二乘原理来拟合这些参数。设 $(t_j,y_j)(j=1,2,\dots,n)$ 有实验数据,使得函数 $y(t,\theta)$ 在点 $(t_j,y_j)(j=1,2,\dots,n)$ 处的函数值与观测数据偏差的平方和达到最小,即满足如下条件的函数 $(t,\hat{\theta})$ 。使 $\min \sum_{j=1}^n (y(t_j,\theta) - y_j)^2 = \sum_{j=1}^n (y(t_j,\hat{\theta}) - y_j)^2$,其中 θ 是待定参数,而 $\hat{\theta}$ 就是最小二乘法所确定的最佳参数。

2.2 多波段 Richards 模型的应用步骤

由多波感染引起的每个 S 形节段为 1 个波段,并且每个波段是通过拐点来区分的,这些拐点表示每个 S 型片段末端先减速后加速得点,即相应的发病率曲线的局部最小值。对于暴发的 n 波段流行病, $n-1$ 个局部最小值将 n 个波段分离。

对于多波段 Richards 模型的参数估计需要 5 个步骤:

步骤 1 首先使用 Richards 模型结合最小二乘法拟合实际每日累计感染人数,对于单波暴发的疾病随着轨迹接近感染承载量 K ,参数 θ 将收敛。

步骤 2 如果参数估计保持收敛,直到不再有新增确诊病例,则传染病只有一个波段。反之,如果估计值开始偏离当前的固定值,这表示第二波段开始的拐点已经出现。

步骤 3 定位拐点 t_{\min} ,将流行病的两个 S 型波段分离作为发病率曲线的局部最小值。

步骤 4 从第二波段开始后的第二天,即 $t_{\min}+1$ 天

开始,用 Richards 模型拟合累计确诊病例。当曲线接近第二波段的感染承载量 K 时,参数 θ 将再次收敛。

步骤 5 如果发生更多波段,重复步骤 2—步骤 4,直到暴发结束。

因此,通过分别考虑流行病曲线的连续 S 型片段,可以估计最终感染数,并定位拐点,从而估计每个波段的累计病例数。

3 数据拟合及分析

数据来源:所用数据均来自世界卫生组织公开数据。收集了西太平洋区域传染病暴发的累计确诊人数、新增确诊人数、累计死亡人数和新增死亡人数等信息,其中西太平洋区域包括中国、日本、韩国、柬埔寨、老挝、越南、缅甸、菲律宾、马来西亚、新加坡、文莱、澳大利亚、新西兰等 27 个国家。

西太平洋区域在 2020-01-04—2020-09-13 时期暴发的传染病感染情况,如图 1 所示。

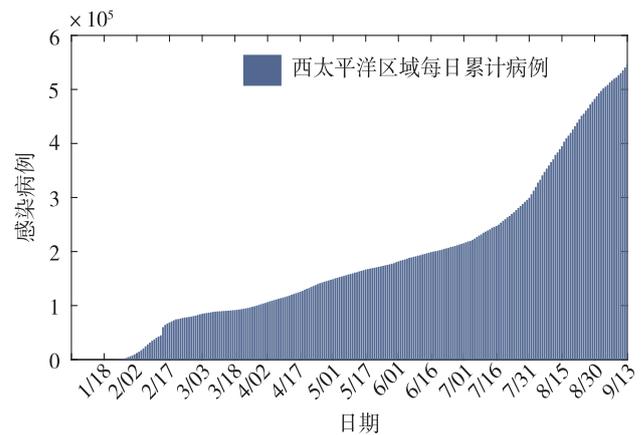


图 1 西太平洋区域每日累计确诊病例图 (2020-01-04—2020-09-13)

Fig. 1 Daily cumulative confirmed cases in the Western Pacific region from January 4, 2020 to September 13, 2020

下面将利用多波段 Richards 模型拟合西太平洋区域 2020-01-04—2020-09-13 的每日累计确诊数据,给出每个波段的关键数据表(表 1、表 2、表 3),并得到疾病流行趋势拟合图,见图 2。

表 1 西太平洋区域传染病暴发第一波段的参数估计(2020-01-04—2020-03-15)

Table 1 Parameter estimation for the first band of epidemic outbreaks in the Western Pacific Region from January 4 to March 15, 2020

日期	a	r	t_i	K	$K(95\%CI)$	N
1/04-2/03	0.141 4	1.080 7	32.078	51 320.370	(22 515-80 126)	17 358
1/04-2/13	0.003 2	25.804 5	39.853	133 612.054	(-43 930-311 150)	60 091
1/04-2/23	1.132 9	0.204 2	37.791	82 882.575	(78 019-87 746)	77 976
1/04-3/04	0.851 4	0.234 9	37.433	84 734.234	(82 888-86 581)	86 278
1/04-3/14	0.385 3	0.389 9	36.632	89 082.162	(87 626-90 539)	90 693
1/04-3/15	0.357 7	0.412 3	36.567	89 365.037	(87 946-90 784)	91 054
1/04-3/16	0.329 0	0.439 5	36.497	89 659.588	(88 272-91 047)	91 580

表 2 西太平洋区域传染病暴发第二波段的参数估计 (2020-03-16—2020-07-04)

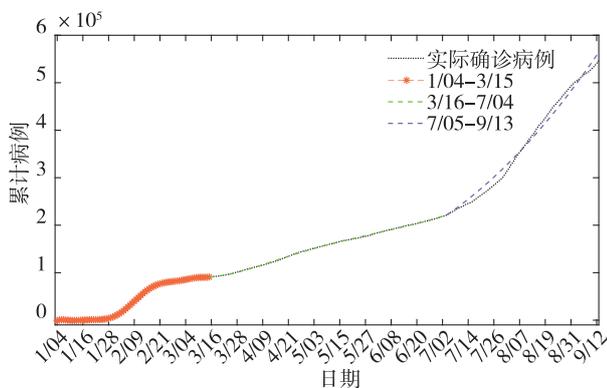
Table 2 Parameter estimation for the second band of epidemic outbreaks in the Western Pacific Region from March 16, 2020 to July 4, 2020

日期	<i>a</i>	<i>r</i>	<i>t_i</i>	<i>K</i>	<i>K</i> (95%CI)	<i>N</i>
3/16-4/06	112.607 8	0.010 4	49.007	154 996.443	(-22 358 000 000 000-22 358 000 000 000)	112 576
3/16-4/28	92.590 4	0.011 4	50.610	164 250.808	(-6 039 800 000 000-6 040 100 000)	145 387
3/16-5/20	11.432 5	0.011 4	41.546	175 206.805	(170 880-179 530)	169 955
3/16-6/11	2.658 6	0.013 9	37.371	218 878.397	(205 780-231 980)	194 470
3/16-7/03	0.111 4	0.108 4	26.434	310 175.427	(261 840-358 510)	219 543
3/16-7/04	0.023 6	0.476 1	25.505	319 715.279	(264 310-373 280)	221 794
3/16-7/05	0.002 3	4.674 8	26.134	325 242.108	(265 590-385 170)	223 915

表 3 西太平洋区域传染病暴发第三波段的参数估计 (2020-07-05—2020-09-13)

Table 3 Parameter estimation for the third band of epidemic outbreaks in the Western Pacific Region from July 5, 2020 to September 13, 2020

日期	<i>a</i>	<i>r</i>	<i>t_i</i>	<i>K</i>	<i>K</i> (95%CI)	<i>N</i>
7/05-7/18	262.789	0.010 0	37.144	329 524.782	(-13 891 000 000 000-13 891 000 000 000)	256 788
7/05-8/01	85.816	0.012 0	46.015	400 117.627	(-52 906 000 000 000-52 906 000 000 000)	312 771
7/05-8/15	115.236	0.015 1	46.214	438 891.908	(-115 290 000 000 000-115 290 000 000)	403 399
7/05-8/29	32.107	0.015 4	48.457	489 826.181	(433 970-545 680)	482 583
7/05-9/13	10.545	0.015 5	48.224	549 720.017	(532 560-566 880)	546 552



注:采用多波段 Richards 模型得到拐点分别为 02-10、03-15、04-11、07-04、08-22

图 2 西太平洋地区累计确诊病例的拟合曲线 (2020-01-04—2020-09-13)

Fig. 2 Fitted curve for cumulative cases in the Western Pacific from January 4, 2020 to September 13, 2020

截至 03-15, 第一波段的实际确诊病例数 $N = 91\ 054$ 例, 与我们估计的感染承载量 $K = 89\ 365.037$ 仅相差 1.85% (表 1), 此时 $t_i = 36.567$ 意味着第一波感染由严重转向轻缓的拐点发生在 02-10 左右。利用截至 03-04, 即拐点发生后 23 d 的流行病数据, 可以获得较为满意的累计病例数。在 03-15 之后不久, 参数估计的差异表明第二波感染已经开始, 即 03-15 为分割两个 S 型波段的拐点。

第一波段感染率 $r = 0.412\ 3$, 通过公式 $R_0 = e^{rT}$ 得

到西太平洋区域爆发的传染病的基本再生数约为 8.53。目前各方面证据显示这类传染病的传染力非常强, 远远高于 2003 年的 SARS, 获得的基本再生数也印证了这个说法。

第二波段的参数估计 $t_i = 25.505$ 意味着第二波段感染拐点是 04-11。07-04 之后不久, 参数估计值的差异表明第三波感染已经开始, 即 07-04 为第三波疫情开始的拐点。此外, 利用截至 06-11 的感染承载量 K , 得到其 95% 置信区间为 (205 780—231 980), 因此可以提前近一个月预测 07-04 的累计病例数 221 794 例。

第三波段的参数估计得到 2020-09-13 的感染承载量 (最终感染数) $K = 549\ 720.016\ 5$, 这与西太平洋区域暴发的实际病例数 $N = 546\ 552$ 非常接近。此时, 估计拐点 $t_i = 48.224$ 意味着第三波段感染由严重转向轻缓的拐点是 08-22。并且, 可以通过使用拐点后 7 d, 即 08-29 的数据来获得实际累积病例数的准确预测 $K = 489\ 826.181$, 95% 置信区间为 (433 970—545 680)。其中, 包含了 09-12 实际累计病例数 540 562 例, 也就是说能提前半个月估计到 09-12 的实际累计病例数。因此, 通过这种方法, 可以较早且较准确的获得传染病感染规模, 这说明多波段 Richards 模型具有预测功能, 这将为疾病防控提供参考。

由图 2 可知, 多波段 Richards 模型很好地拟合了数

据,能够回顾性地确定西太平洋地区传染病暴发期间疾病传播的 5 个关键转折点。

4 结 论

利用多波段 Richards 模型结合最小二乘法较好地拟合了西太平洋区域传染病的传播情况,并得到基本再生数、拐点时刻、感染率、感染承载量(最终感染数)及置信区间等关键参数估计值,根据这些参数分析得到西太平洋区域在 2020-01-04—2020-09-13 时期暴发的传染病分为 3 个波段,03-15 和 07-04 为分割 3 个波段的拐点,02-10、04-11、08-22 分别为 3 个波段中每日新增感染人数由上升转为下降的转折点。说明从 2020-01-04 开始到 2020-02-10 号第一波感染才慢慢减速趋于平稳。由第一波段求得的基本再生数可得,这次传染病的传播能力很强,远高于 2003 年的 SARS。利用多波段 Richards 模型能够回顾性地确定传染病在多波段暴发中的关键转折点,能够通过估计承载能力实时预测疾病流行当前波段的严重程度,以及预测未来一段时间传染病的感染情况,为制定传染病防控措施提供参考。

多波段 Richards 模型在传染病初期运用具有一定实用性,但该模型在传染机理分析、传播过程中的潜伏期和随机效应、传染途径的网络化特征、传染病防控因素等方面还没有刻画,因而模型拟合效果仍待改进,今后将进一步研究在上述因素作用下的传染病分析。最后,指出本文主要基于理论分析和数值计算,由于疾病传播的复杂性,所得结论仅作为理论参考。

参考文献(References):

- [1] ANDERSON R M, MAY R M. Infectious diseases of humans dynamics and control[M]. Oxford: Oxford University Press, 1991.
- [2] PRESS W H, LEVIN R C. Modeling, post COVID-19[J]. Science, 2020, 370(6520): 1015.
- [3] 丁黄艳, 铁禧玥. 基于 SIR 传染病模型的印度新冠疫情波及影响分析[J]. 重庆工商大学学报(自然科学版), 2022, 39(5): 70-77.
DING Huang-yan, TIE Xi-yue. Influence analysis of COVID-19 in India based on SIR infectious disease model[J]. Journal of Chongqing Technology and Business University (Natural Science Edition), 2022, 39(5): 70-77.
- [4] KERMAK W O, MCKENDRICK A G. A contribution to the mathematical theory of epidemics[J]. Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character, 1927, 115 (772): 700-721.
- [5] MENG X, ZHAO S, FENG T, et al. Dynamics of a novel nonlinear stochastic SIS epidemic model with double epidemic hypothesis[J]. Journal of Mathematical Analysis and Applications, 2016, 433(1): 227-242.
- [6] YANG Z, ZENG Z, WANG K, et al. Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions[J]. Journal of Thoracic Disease, 2020, 12(3): 165-174.
- [7] XING Y, LI H X. Almost periodic solutions for a SVIR epidemic model with relapse[J]. Mathematical Biosciences and Engineering, 2021, 18(6): 7191-7217.
- [8] RICHARDS F J. A flexible growth function for empirical use[J]. Journal of Experimental Botany, 1959, 10(2): 290-301.
- [9] ZHOU G, YAN G. Severe acute respiratory syndrome epidemic in Asia[J]. Emerging Infectious Diseases, 2003, 9 (12): 1608-1610.
- [10] HSIEH Y H, MA S. Intervention measures, turning point, and reproduction number for dengue, Singapore, 2005 [J]. The American Journal of Tropical Medicine and Hygiene, 2009, 80(1): 66-71.
- [11] HSIEH Y H, CHEN C W. Turning points, reproduction number, and impact of climatological events for multi-wave dengue outbreaks [J]. Tropical Medicine & International Health, 2009, 14(6): 628-638.
- [12] BLONIGAN P, RAY J, SAFTA C. Forecasting multi-wave epidemics through Bayesian inference[J]. Archives of Computational Methods in Engineering, 2021, 28(6): 4169-4183.
- [13] ZUHAIROH F, ROSADI D. Real-time prediction for multi-wave COVID-19 outbreaks[J]. Communications for Statistical Applications and Methods, 2022, 29(5): 499-512.
- [14] HSIEH Y H, CHENG Y S. Real-time forecast of multiphase outbreak[J]. Emerging Infectious Diseases, 2006, 12 (1): 122-127.
- [15] 崔恒建, 胡涛. 新型冠状病毒肺炎疫情预测预报的非线性回归方法[J]. 中国科学: 数学, 2021, 51(8): 1267-1278.
CUI Heng-jian, HU Tao. Nonlinear regression in COVID-19 forecasting[J]. Scientia Sinica (Mathematica), 2021, 51(8): 1267-1278.

责任编辑:田 静