

传染病动力学 SARD 新型模型 (五) 定量预测北半球 COVID-19 疫情

作者 章其初博士 2020 年 3 月 29 日

摘要

本文采用传染病动力学 SARD 新型模型，在 COVID-19 疫情早期，进行定量分析预测中国外北半球疫情。该 SARD 模型中，时间 t 的四个函数， $S(t)$ 称为易感者， $A(t)$ 累计感染者， $R(t)$ 累计痊愈者， $D(t)$ 累计死亡者。在模型的假设、定义和猜想下得到这四个函数的解析表达式。现有感染者 I 用公式 $I(t)=A(t)-R(t)-D(t)$ 计算得到。该 SARD 模型第一次试用于分析中国地区 COVID-19 疫情，模型计算疫情的各种病例数与通报数符合良好，这表明模型能准确定量分析和预测 COVID-19 疫情。这样 SARD 模型解决了以前传染病动力学 SIR、SEIR 等模型中疫情关键函数没有解析解的难题。

传染病动力学 SARD 新型模型定量分析预测中国外北半球 COVID-19 疫情要点如下。

(1) 累计感染者 A 病例总数 N 达 1 百 40 万，误差 7 万，即 1 百 33 万到 1 百 47 万，相对误差 5% (95%置信区间)；

(2) 新增感染者 B 病例拐点区在 3 月 30 到 4 月 1 日，最大病例数 67,000 上下；

(3) 现有感染者 I 病例拐点区将在 4 月 11 到 13 日期间，最大病例数 970,000 上下；

(4) 疫情在 7 月 1 日将走向尾声，现有感染者病例数将还剩 100,000 上下；

(5) 疫情后期死亡率 $\delta_M=5.2\%$ ，死亡病例数 73,000，误差 5,000 (95%置信区间)。

截止到 3 月 29 日，南半球病例数在现有数据占的比例很小。若南半球 5 月份开始温度降低，有可能大规模爆发疫情，本文没有包括这种情况在内。

1、前言

传染病动力学建模计算早在 20 世纪初就开始了。1927 年 Kermack 与 McKendrick 在研究流行于伦敦的黑死病时提出的 SIR 舱室模型，是传染病模型中最经典、最基本的模型，为传染病动力学的研究做出了奠基性的贡献。模型中把传染病流行范围内的人群分成三类：S 类为易感者 (Susceptible)，I 类为现有感染者 (Infective)，R 类为移出者 (Removal)。在若干假设下，模型导出 S、I、R 三个函数的一阶微分方程组。方程组三个函数 S、I、R 没有解析解。疫情预测只能根据已有数据得到的拟合曲线，并外延用来预测。随后，很多学者提出了其他很多模型，甚至更复杂模型，尝试建模计算预测疫情。其中广泛采用的 SEIR 模型，以及改进模型，其原理上与 SIR 类似。

新型冠状病毒 COVID-19 从 2019 年底在武汉被发现，中国和世界各国很多研究机构 and 高等院校流行病学、统计学及其他专业学者进行了大量建模计算，进行疫情分析和预测[1, 2]。

本文作者在 2020 年 2 月 COVID-19 疫情在湖北大规模爆发后，作为一名物理学者也加入了建模计算工作，建立了 SARD 新型模型，疫情的关键函数有解析表达式。采用 SARD 模型

第一次试用于分析中国地区 COVID-19 疫情,模型计算疫情关键病例数与与通报数符合良好,这表明模型能准确定量分析和预测 COVID-19 疫情[3~6]。这样 SARD 模型解决了传染病动力学 SIR、SEIR 等模型中疫情关键函数没有解析解的难题。

中国外北半球其他国家,截止到 2020 年 3 月 29 日,累计感染通报的病例数达到 63 万多。超过 1 万的国家有 11 个,占病例总数 80%。其中美国累计感染者达 14 万多,意大利超过 10 万,西班牙超过 8 万,德国超过 6 万,法国超过 5 万,伊朗超过 4 万,英国超过 2 万。本文在疫情早期,采用 SARD 新型模型定量分析和预测中国外北半球 COVID-19 疫情。

2、传染病动力学 SARD 新型模型

传染病动力学 SARD 新型模型中,四个时间 t 的病例函数分别为,易感者 $S(t)$, 累计感染者 $A(t)$, 累计痊愈者 $R(t)$ 和累计死亡者 $D(t)$ 。模型假设, $S(t)$ 减少与 $S(t)$ 和 $A(t)$ 乘积成正比, $A(t)$ 增加与 $S(t)$ 和 $A(t)$ 乘积成正比,假设 $S(t)$ 加 $A(t)$ 总数不变,即 $S(t) + A(t) = N$ 。模型定义, $R(t)$ 等于痊愈率函数 $\gamma(t)$ 乘 $A(t)$, $D(t)$ 等于死亡率函数 $\delta(t)$ 乘 $A(t)$ 。模型大胆猜想,痊愈率 $\gamma(t)$ 是逻辑函数,或 $R(t)$ 是逻辑函数;死亡率 $\delta(t)$ 是逻辑函数,或 $D(t)$ 是逻辑函数。图 1 给出 SARD 模型示意图。

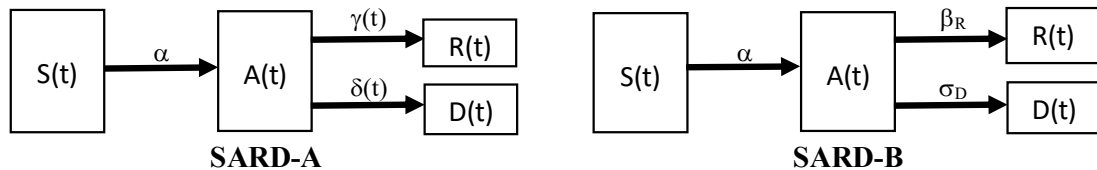


图 1. SARD 新型模型示意图

求解函数 $A(t)$ 和 $S(t)$ 的两个一阶微分方程组,得到累计感染者 $A(t)$ 的解析解为 Sigmoid 函数,习惯上称为如下逻辑函数。

$$A(t) = \frac{NA_0}{N + (N - A_0)e^{-\alpha(t-t_0)}}$$

式中 $A_0=A(t_0)$, 当 $t=t_0$ 时。公式中两个参数, α 是感染系数 (/天), N 是累计感染者总数。易感者 $S(t)=N-A(t)$ 。

痊愈率 $\gamma(t)$, 死亡率 $\delta(t)$ 的逻辑函数与上式类似, 这里略去。这样 SARD 模型中的四个函数 $S(t)$ 、 $A(t)$ 、 $R(t)$ 和 $D(t)$ 都有解析表达式。现有感染者 $I(t)$ 等于 $A(t)$ 减去 $R(t)$ 和 $D(t)$ 。每日新增感染者 B 等于当日 A 减去昨天的。

这里对 SARD 模型再作进一步说明。(1) 易感者变成感染者, 包括两个过程, 易感者与感染者接触 (对应接触概率), 接触后被感染成病人 (对应感染概率), 本文用感染系数 α 表征这两个过程的效果。(2) 感染者在其寿命期内具有传染性, 痊愈者无传染性。(3) 累计感染者 A 总数 N 是采用本 SARD 模型可计算得到。(4) 模型中提到的舱室应理解为, 易感者最终都变成感染者的群体, 分布在界定的一个地域范围内。

在 SARD 模型中, 由于疫情的关键函数全部有解析表达式, 因而传染病若干其他参数也

可以直接计算得到，如平均传染周期，基本再生数 R_0 等。

3、模型计算定量分析和预测中国外北半球 COVID-19 疫情

累计感染通报的病例数，中国外北半球在 2020 年 3 月 29 日达到 63 万 9 千多。南半球病例数在当前现有数据占的比例很小，所以本模型计算采用网站【头条】汇总的对应海外 COVID-19 疫情的数据作为中国外北半球数据。

3.1 累计感染者 A 和总数 N

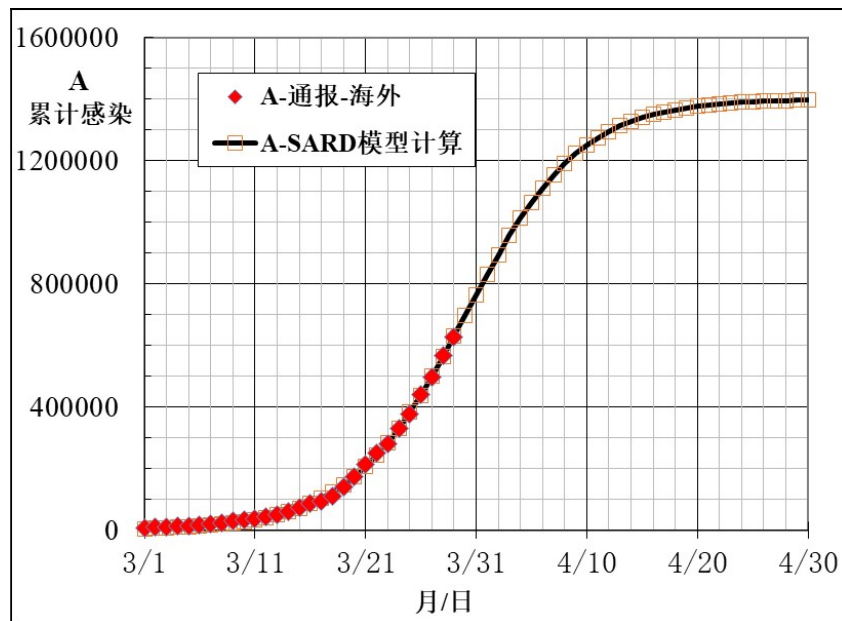


图 2. 累计感染者 A 的模型计算（实线）和通报病例数（离散点）

从图 2 中看出，累计感染者 A 的模型计算（实线）与通报（离散点）符合良好。

累计感染者 A 的模型数值计算时，选用最近 3 月 19~29 日 11 天疫情通报病例数，模型计算得到两个参数值，感染系数 $\alpha=0.194/\text{天}$ ；累计感染者 A 总数 $N=1,400,000$ ，误差 7 万人，相对误差 5%（95%置信区间）。

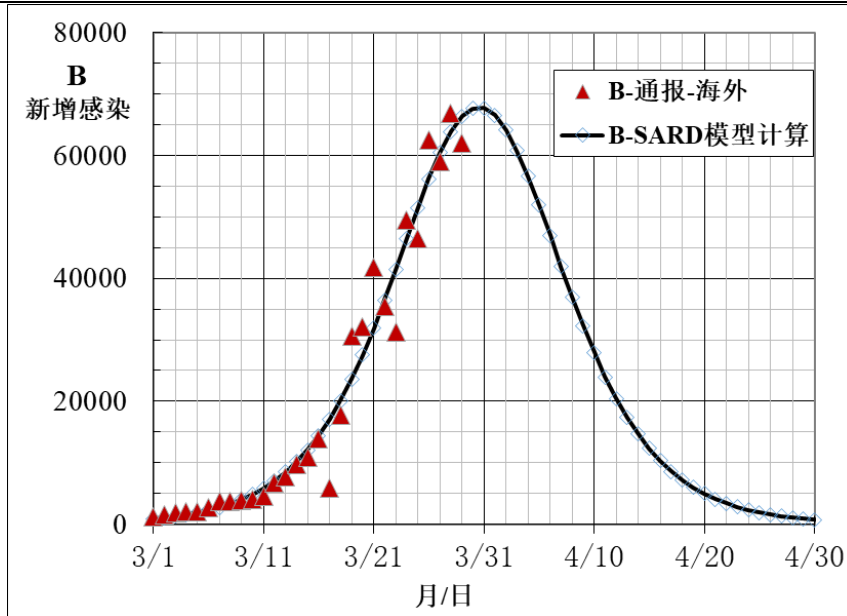


图 3. 新增感染者 B 的模型计算（实线）和通报病例数（离散点）

图 3 中显示新增感染者 B 通报数，图中离散点，分布在紧靠模型计算曲线（实线）两侧。每日新增感染者 B 病例拐点区在 3 月 30 到 4 月 1 日期间，3 天平均病例数 67,000。个别日期通报数涨落比较大，这是病例统计学上经常出现的想象。

3.2 累计痊愈者 R

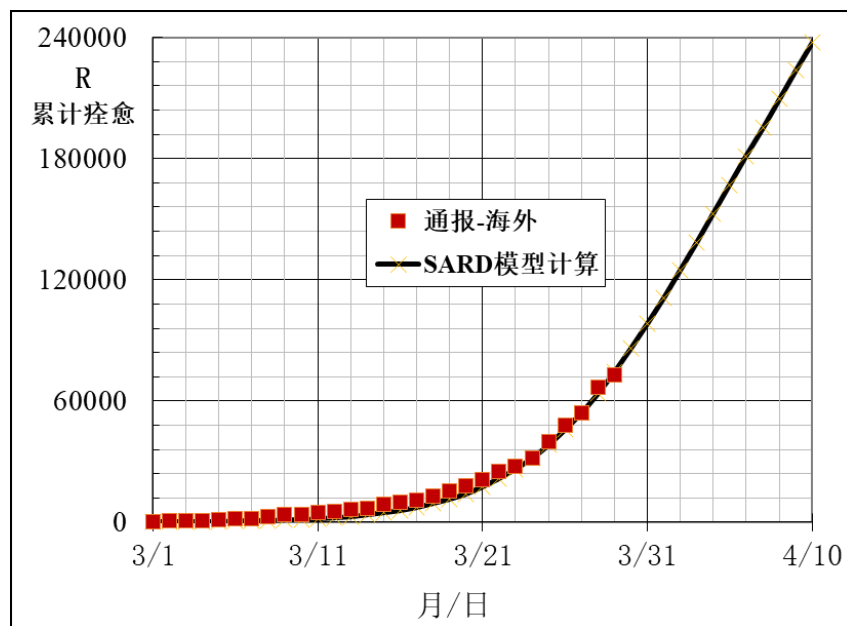


图 4 累计痊愈者 R 的模型计算（实线）和通报病例数（离散点）

从图 4 中看出，累计痊愈者 R 的模型计算（实线）与通报（离散点）符合良好，表明本 SARD 模型中，大胆猜想痊愈率 $\gamma(t)$ 是逻辑函数合理准确。模型计算得到两个参数，痊愈系数 $\beta=0.044/\text{天}$ ；疫情后期痊愈率 $\gamma_M=94.8\%$ ，误差 0.3%（95%置信区间）。

图中仅仅给出在 3 月 1 日到 4 月 10 日期间数据，使得显示比例大一些。

3.3 累计死亡者 D

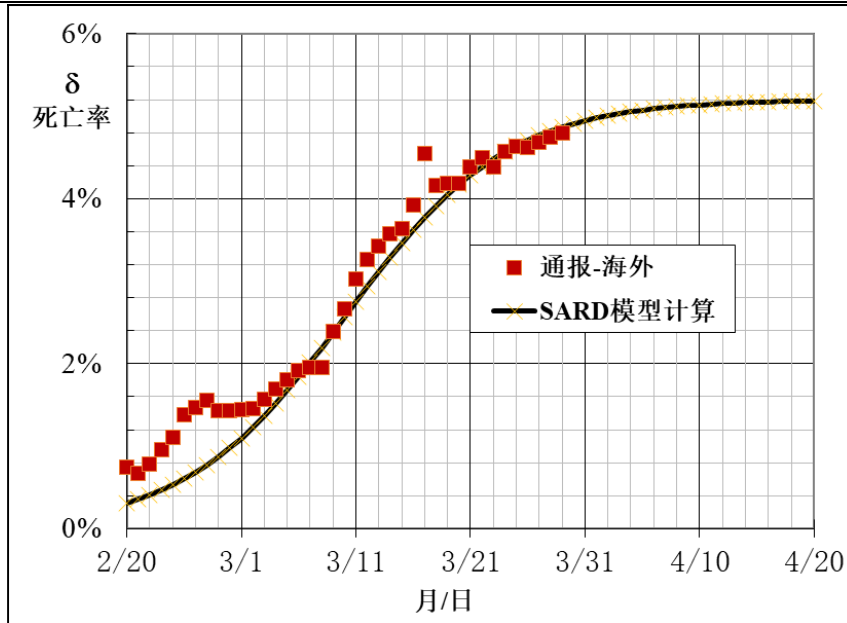


图 5 累计死亡率 δ 的模型计算（实线）和通报病例数（离散点）

从图 5 中看出，死亡率 $\delta(t)$ 的模型计算（实线）与通报（离散点）从 3 月 3 日起符合良好，表明本 SARD 模型中，猜想死亡率 $\delta(t)$ 是逻辑函数合理准确。模型计算得到两个参数，死亡系数 $\sigma=0.143/\text{天}$ ；疫情后期死亡率 $\delta_M=5.2\%$ ，误差 0.3%（95%置信区间）。

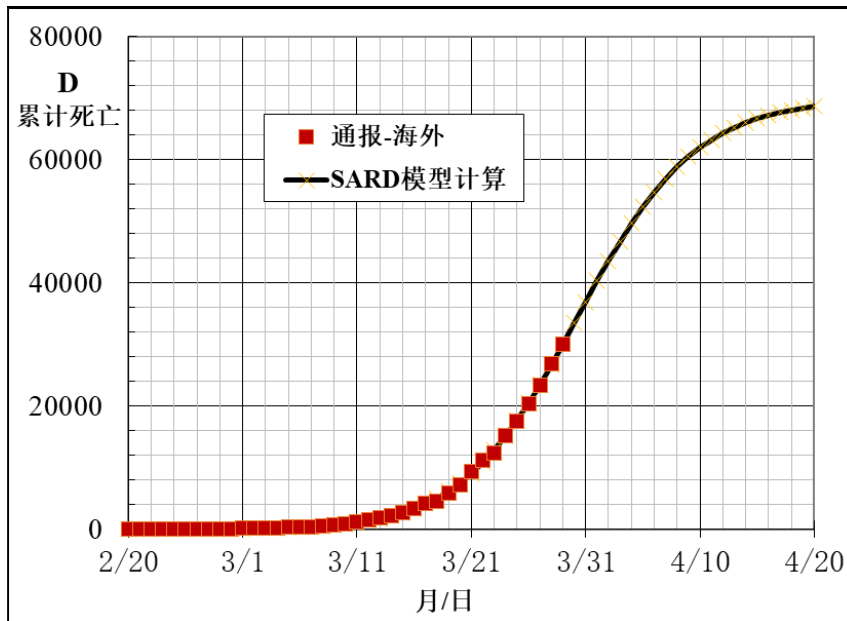


图 6 累计死亡者 D 的模型计算（实线）和通报病例数（离散点）

从图 6 中看出，累计死亡者 D 的模型计算（实线）与通报（离散点）符合良好，模型计算得到疫情后期死亡病例数 73,000，误差 4,400，相对误差 6%（95%置信区间）。

3.4 现有感染者 I

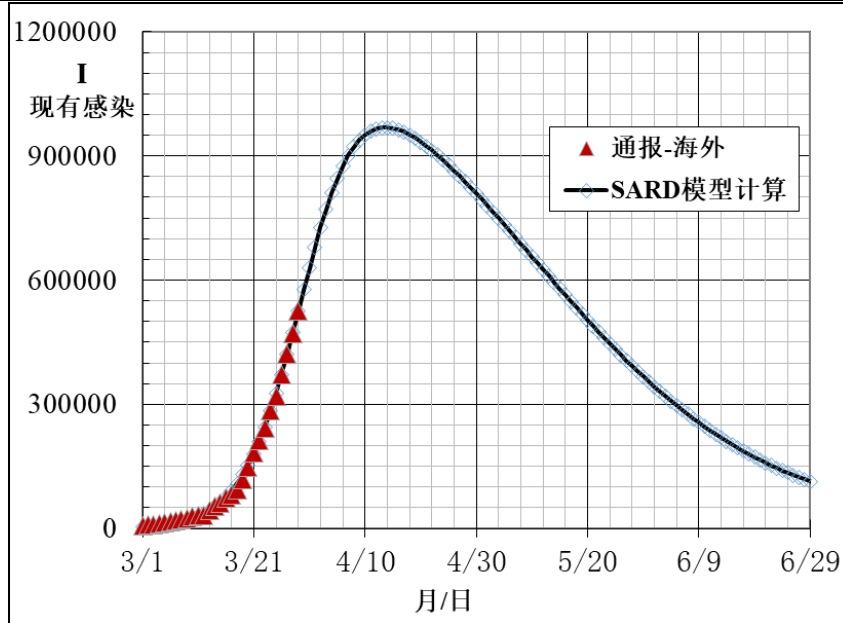


图 7 现有感染者 I 的模型计算（实线）和通报病例数（离散点）

从图 7 中看出，现有感染者 I 的模型计算（实线）与通报（离散点）符合良好。现有感染者 I 拐点区将在 4 月 11 到 13 日期间，三天平均数病例数 965,000。这表明现有感染者 I 达到最大值 96.5 万左右，才开始逐步慢慢下降。

疫情在 7 月 1 日将走向尾声，那时现有感染者病例数将还剩 10 万上下。从通报痊愈率看，比中国对应疫情期间的小。若疫情中后期，逐步提高治愈水平，图中的 I 曲线会下降快一些，疫情尾声将会提早听到，这意味着图中现有感染者病例在疫情中后期还会有一些修正。

3.5 疫情图

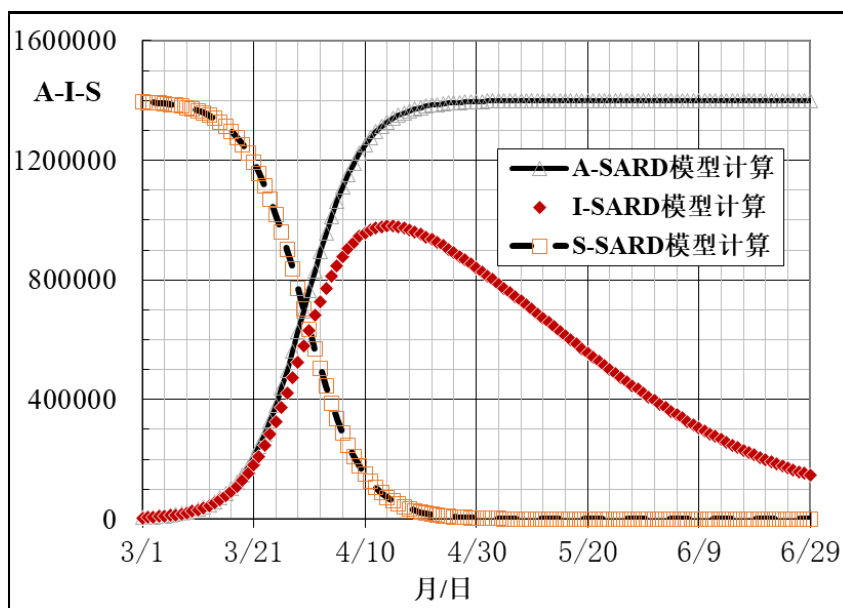


图 8. 累计感染者 A （实线）、现有确诊 I （离散点）和易感者 S （虚线）病例疫情图

图 8 是包含疫情三个关键函数 A 、 I 、 S 病例图，本文称为疫情图。图中易感者 S ，在 SARD 模型中，代表还没有感染，但会被感染转变成感染者。

由于疫情期间，采取各种措施，如隔离，提高痊愈率等，都会对疫情过程发生变化。所以采用本 SARD 模型，采用疫情最近若干天通报病例数，计算得到当前的疫情图，可更准确定量预测疫情。根据 SARD 模型，定量分析预测中国 COVID-19 疫情看来，在现有感染者 I 拐点区前一些日子的通报数，模型计算得到的疫情图，疫情中后期变化就很小了。

从图 8 看出，与以前作者本人发表在网上文章[4]给出的中国疫情比较，北半球其他国家作为整体，COVID-19 疫情来得慢一些，去得更慢一些，可以说相当慢了。

传染病 COVID-19 疫情在中国大爆发过程，现在看来，与报道的中国古老疫情经验历法符合良好。历次瘟疫总结的历法是瘟疫始于大雪 2019 年 12 月 8 日，发于冬至 2019 年 12 月 23 日，生于小寒 2020 年 1 月 5 日，长于大寒 2020 年 1 月 15 日，盛于立春 2020 年 2 月 4 日，弱于雨水 2020 年 2 月 19 日，衰于惊蛰 2020 年 3 月 5 日。其中日期对应这次 COVID-19 疫情。千百年来，发生在中国的历次瘟疫都是按这个规律。另外 2 月份有报道称，温度低、湿度大 COVID-19 病毒容易存活。看到这些报道后本人猜想，疫情大爆发会在北半球逐步从南到北蔓延。在 1 月和 2 月份时，欧洲和北美各国大多数人可能还在观望中国疫情大爆发，没有想到 COVID-19 疫情大爆发的脚步正在一步步向他们走去。

后来从报道数据分析看，确实是这样一个过程。北半球 COVID-19 疫情大爆发从北纬 30 度武汉开始，逐步向北推进。第二波韩国、伊朗，第三波意大利、西班牙、美国，第四波德国、法国，第五波英国、瑞士、比利时、荷兰等北欧诸国以及加拿大。另外 SARD 模型中感染系数与人口密度和流动性相关，所以疫情大爆发主要集中在人口密度高、流动性大的城市。

从以上分析得出，与中国相比，中国外北半球作为整体，COVID-19 疫情上升要慢一些，去得要更加慢一些，主要原因是用于疫情从南到北一波一波大爆发。另外疫情去得要更加慢一些的另一原因是，因对付 COVID-19 病毒没有特效药，中国治愈病人的水平要高一些，特别是这次中西医结合疗效特别明显，其他各国医务工作者不容易学到手。

最近有报道，香港大学研究 COVID-19 病毒，在摄氏 4 度，感染能力 14 天后仍无明显下降；在 22 度，第二日感染能力开始下降，14 天后无感染能力；在 37 度，一日后感染能力大幅下降；在 56 度，30 分钟后已没有感染能力。也就是说，平均传染周期，基本再生数 R_0 也与气温相关。这也验证了中国古老历法对传染病疫情的以上经验总结。

从 COVID-19 疫情大爆发至今已有 3 个多月了，南半球病例数在当前现有数据占的比例很小。若南半球 5 月份以后开始温度降低，有可能大规模爆发疫情，全球累计感染病例数将大大超过上述 1 百 40 万，本文没有包括这种情况在内。

4、结论

中国外北半球其他国家，截止到 2020 年 3 月 29 日，累计感染通报的病例数达到 63 万多。超过 1 万的国家合计 10 个，占病例总数 80%。其中第一，美国 14 万多，其他依次为意大利，西班牙，德国，法国，伊朗，英国，比利时，荷兰，土耳其等国。

本文采用传染病动力学 SARD 新型模型，在中国外北半球疫情早期，进行定量分析预测。模型计算得到的疫情要点如下。（1）累计感染者 A 病例总数 N 为 1 百 40 万，误差 7 万，即 1 百 33 万到 1 百 47 万，相对误差 5%（95%置信区间）；（2）新增感染者 B 病例拐点区在 3

月 30 到 4 月 1 日, 新增感染者最大数 6 万 7 千上下; (3) 现有感染者 I 病例拐点区将在 4 月 11 到 13 日期间, 现有感染者最大数 97 万上下; (4) 疫情在 7 月 1 日将走向尾声, 那时现有感染者病例数将还剩 10 万上下; (5) 疫情后期死亡率 $\delta_M=5.2\%$, 死亡病例数 7 万 3 千, 误差 5 千 (95% 置信区间)。

与中国相比, 中国外北半球作为整体, COVID-19 疫情上升要慢一些, 并且疫情去要更加慢一些, 可以说是相当慢。主要原因是, 北半球 COVID-19 疫情大爆发是一波一波传递的。大爆发从北纬 30 度武汉开始, 逐步向北推进。第二波韩国、伊朗, 第三波意大利、西班牙、美国, 第四波德国、法国, 第五波英国、瑞士、比利时、荷兰等北欧诸国以及加拿大。另外, 疫情去要更加慢一些与中国治愈病人的水平要高一些有关, 特别是中西医结合疗效明显, 痊愈系数更高一些。

南半球病例数在当前现有数据占的比例很小。若南半球 5 月份开始温度降低, 有可能大规模爆发疫情的话, 累计感染病例数将大大超过上述 1 百 40 万, 疫情走向尾声也将大大推迟。当前计算中没有包括这种情况在内。

参考

1. Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions
Zifeng Yang, Zhiqi Zeng, Ke Wang, Sook-San Wong, Nanshan Zhong et al.
J. Thorac Disease. Submitted Feb 27, 2020. Accepted for publication Feb 28, 2020
View this article at: <http://dx.doi.org/10.21037/jtd.2020.02.64>
2. A conceptual model for the coronavirus disease 2019 (COVID-19) outbreak in Wuhan, China with individual reaction and governmental action
Qianying Lin 1, Shi Zhao, Daozhou Gao, Weiming Wang, Lin Yang, Daihai He
International Society for Infectious Diseases VOLUME 93, P211-216, APRIL 01, 2020
View this article at: <https://doi.org/10.1016/j.ijid.2020.02.058>
3. 传染病动力学 SARD 新型模型试验用于 COVID-19 疫情分析和预测
章其初, Received 2020-02-17, Published on 2020-02-21
View this article at: <https://ausinan.org/cms/news/2020-02-21/sard-modeling-for-covid19>
4. 传染病动力学 SARD 新型模型分析预测 COVID-19 疫情 (二)
Qi Chu Zhang, Received 2020-02-24, Published on 2020-02-26
View this article at: ausinan.org/cms/news/2020-02-26/sard-modeling-for-covid19
5. 传染病动力学 SARD 新型模型 (三) 预测韩国 COVID-19 疫情
章其初, Received 2020-02-29, Published on 2020-03-03
View this article at: <https://ausinan.org/cms/news/2020-03-03/sard-modeling-for-covid19>
6. 传染病动力学 SARD 新型模型 (四) 预测澳大利亚 COVID-19 疫情
Qi Chu Zhang, Received 2020-03-27, Published on 2020-03-29
View this article at: ausinan.org/cms/news/2020-03-29/sard-modeling-for-covid19