
传染病动力学 SARD 新型模型分析预测 COVID-19 疫情 (二)

章其初 2020 年 2 月 24 日

摘要

本工作建立了传染病动力学 SARD 新型模型，在模型函数定义和若干假设下，得到模型中四个函数 S、A、R、D 的解析表达式，解决了传染病广泛采用的 SIR 模型，疫情三个函数 S、I、R 没有解析解的难题。SIR 模型的感染者 I 在 SARD 模型中用 A、R、D 计算得到。

选用湖北外其他地区 COVID-19 疫情 5 天数据，用本 SARD 模型计算得到，从 1 月 30 日到至今 20 多天期间 A、R、I 三类病例数与通报符合良好，验证证明 SARD 模型能定量准确分析传染病 COVID-19 疫情。

湖北外其他地区 SARD 模型计算 COVID-19 疫情要点如下。

- (1) 现有确诊最大病例数日期区，即拐点区在 2 月 9 到 11 日期间，与通报的相同。
现有确诊最大数 9150 人（取 3 位有效数字），与通报数差 40 人。
- (2) 定量预测，累计感染者总数 1 万 3 千人，偏差 400 人，相对偏差 3%。
- (3) 定量预测，疫情在 3 月 11 日走向尾声，那时现有确诊人数 5 百人左右，很多地级市现有确诊人数将清零。
- (4) 基本传染数 $R_0=3.3$ （1 月 24 日），偏差 0.1，选用传染周期 14 天。

定量预测，全国累计感染者总数 8 万 1 千人，偏差 2 千 4 百人，相对偏差 3%。疫情在 3 月 22 日走向尾声，那时现有确诊人数 5 百人左右。

1、前言

新型冠状病毒 2019 年底武汉出现，疫情 2020 年 1 月开始爆发以来，中国和其他国家很多研究机构 and 高等院校的流行病、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。

传染病建模计算早在 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，尝试建模计算数据与病例统计数据拟合更接近。

本研究工作的目标是寻找疫情关键函数的解析表达式。参考前人的模型，本文作者提出传染病动力学 SARD 新型模型，并用于 COVID-19 流行病疫情定量分析和预测。

2、传染病动力学 SARD 新型模型

本研究工作建立了传染病动力学 SARD 新型模型，模型中 S、A、R、D 分别称为易感者，累计感染者，累计痊愈者，累计死亡者。本模型定义 R 等于痊愈率 γ 乘 A，D 等于死亡率 δ 乘 A，易感者 S 等于 N 减去 A，N 是累计感染者 A 的总数。模型还大胆但合理假设，痊愈率 γ

是逻辑回归函数，死亡率 δ 是线性函数。

推导得到 SARD 模型中两个函数 S、A 一阶微分方程。求解 A 的一阶微分方程，得到 A 是逻辑回归函数。拟合疫情通报数据 A，得到逻辑回归函数 A 的两个参数，其中一个 N，是累计感染者总数 N。利用公式 $S=N-A$ ，这样四个函数 S、A、R、D 都找到解析表达式。

本工作建立的 SARD 模型模型，解决了传染病广泛采用的 SIR 模型中疫情三个函数 S、I、R 没有解析解的难题。感染者 I 在 SARD 模型中用公式 $I=A-R-D$ 计算得到。

每日新增感染者 B 等于当日 A 减去昨天的。其他一些参数也可以计算得到。如计算感染者的每天基本传染数（日期的函数）。本文中采用数学函数常用标记 R_0 （日期），区别其他文章中代表早期的基本传染数 R_0 。

模型 SARD 中各类名称 A、R、D 和 I 与中国卫生健康委员会报告中的病例名称有一一对应关系。

流行病学统计学和公众通常把现有确诊数最大日作为拐点日。通报数据和模型计算显示，现有确诊数最大日的前后一天，现有确诊人数百分比差别很小，再考虑到流行病学统计学的涨落性质，所以这里把现有感病者数 I 最大日的前后共 3 天作为现有感病者数最大日期区，对应拐点区。另外，传染病统计学习惯称为拐点与数学定义不同，数学中函数的拐点是函数二阶导数为零的点。

3、建模计算湖北外其他地区 COVID-19 疫情的详情及分析和讨论

本文选用 2 月 13 到 17 日的 5 天疫情数，采用 SARD 模型计算用于分析及预测疫情。

湖北外其他地区，从图 1 中看出，模型计算得到的累计确诊病例数 A，从 1 月 30 日到至今 20 多天期间与通报符合良好。

模型计算定量预测，累计感染者总数 N 为 1 万 3 千人，偏差 400 人，相对误差 3%。基本传染数 $R_0=4.2$ （1 月 24 日），偏差 0.1，选用传染周期 14 天。

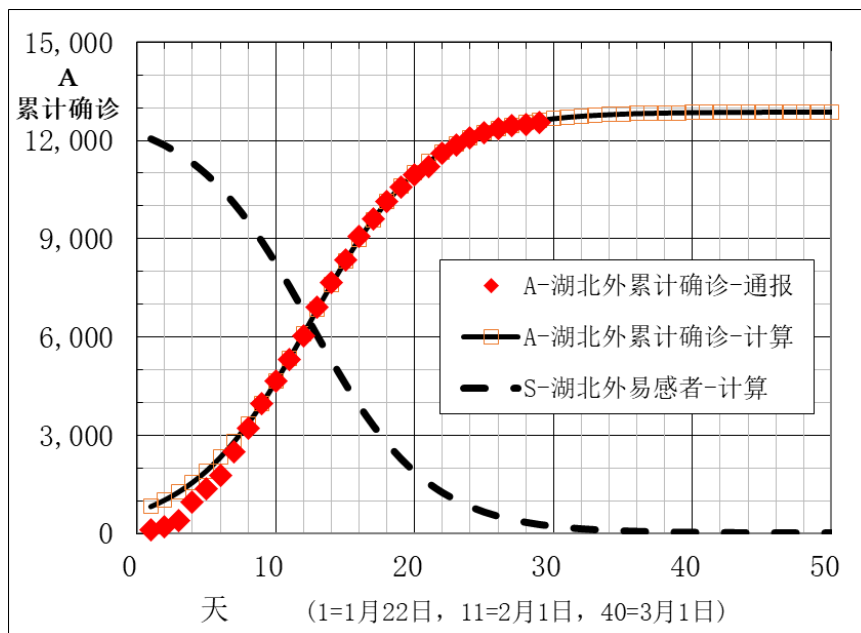


图 1 湖北外其他地区通报累计确诊 A 和模型计算得到的 A 以及易感者 S

易感者 S 用公式 $S = N - A$ 计算得到。管控传染病疫情最关键的是隔离，以便尽量减少舱室中人数 N。这次 COVID-19 疫情数据看，武汉“封城”后，武汉外其他地区隔离比较到位，

城市封到小区，农村封到村，各类感染者、接触者分类医治隔离，外来人员在家隔离等。

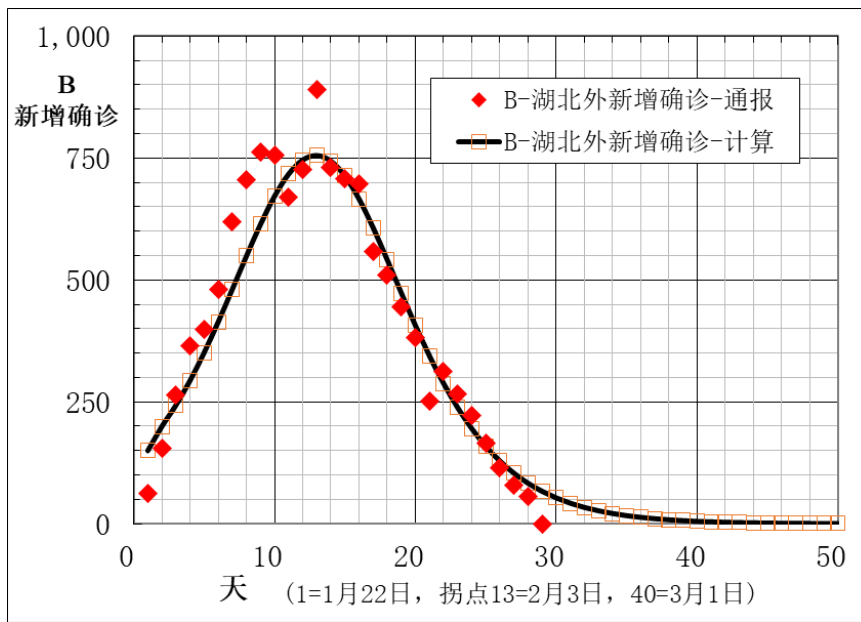


图 2 湖北外其他地区每日新增确诊数 B 通报和模型计算数据（拐点区在 2 月 2 到 4 日）

图 2 中看出，通报新增确诊数 B 涨落比较大，这是由于 B 是累计确诊 A 的相邻一天差，数值比较小，最大病例小于 1000。另外，根据误差理论，B 的偏差是 A 的两倍。尽管如此，通报数大多数还是分布在模型计算数据曲线的两侧附近。

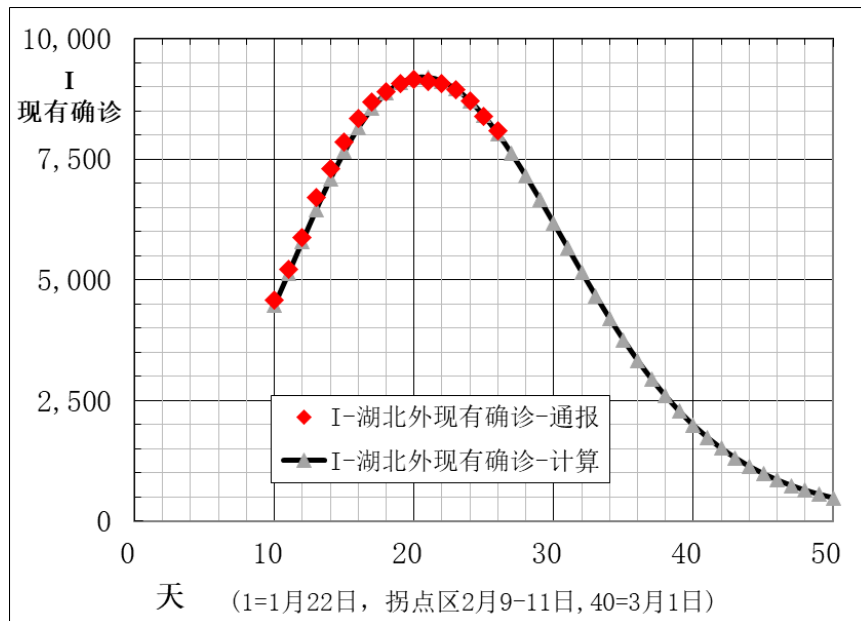


图 3 湖北外其他地区现有确诊数 I 的通报和模型计算数据（拐点区在 2 月 9 到 11 日）

湖北外其他地区，图 3 显示，模型计算得到的现有确诊数 I，与通报数一致性非常好。

模型计算得到的现有确诊最大数日期区，即拐点区，在 2 月 9 到 11 日期间，与通报的相同。模型计算最大现有确诊人数 9150 人（取 3 位有效数字），与通报数差仅 40 人。模型计算预期，疫情在 3 月 10 日前后走向尾声，即湖北外全国所有地区，那时现有确诊人数合计 5 百人以上，很多地级市现有确诊人数将清零。

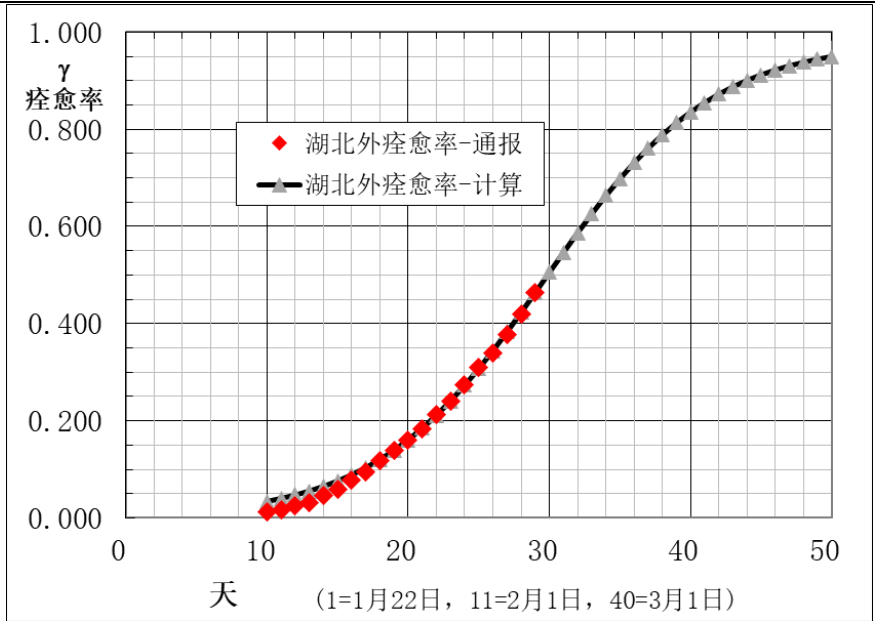


图 4 湖北外其他地区痊愈率通报和模型计算数据

湖北外其他地区，采用逻辑回归函数拟合痊愈率 γ ($\gamma=R/A$)，与通报数痊愈率 γ 符合良好，见图 4。表明本文作者提出的传染病动力学 SARD 新型模型，假设痊愈率 γ 也是逻辑回归函数非常合理。

图 5 给计算得到累计痊愈数 R 与通报数符合良好。累计痊愈数 R 是两个逻辑回归函数 A 和痊愈率 γ 乘积，分母中包含 3 个指数函数项。

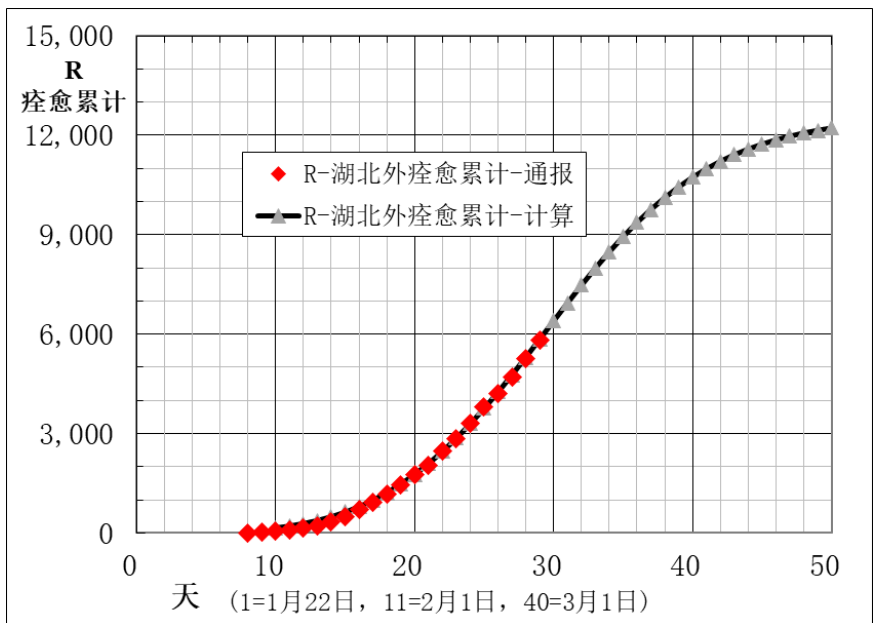


图 5 湖北外其他地区累计痊愈数 R 通报和模型计算数据

4、建模计算湖北地区 COVID-19 疫情的详情及分析和讨论

本文选用 2 月 13 到 17 日的 5 天疫情数，采用 SARD 模型计算用于分析及预测疫情。

从图 6 中看出，采用 SARD 模型计算得到的累计确诊数 A 与疫情通报的符合良好，除了在 2 月 5 日到 12 日的 8 天期间外，这表明在这期间通报给出累计感染数不可靠。因而图 7 中新增确诊数 B，图 8 中现有确诊数 I 也不可靠。

模型计算湖北地区，定量预测累计感染者总数 N 为 6 万 8 千人。基本传染数 $R_0=4.2$ (1 月 24 日)，偏差 0.1，选用传染周期 14 天。

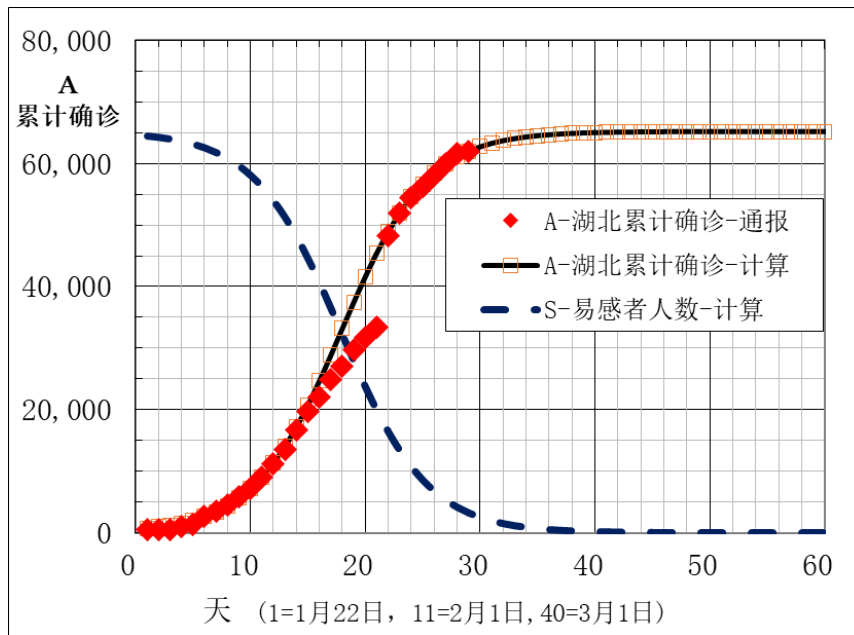


图 6 湖北累计确诊数 A 通报和模型计算得到的 A 以及易感者 S

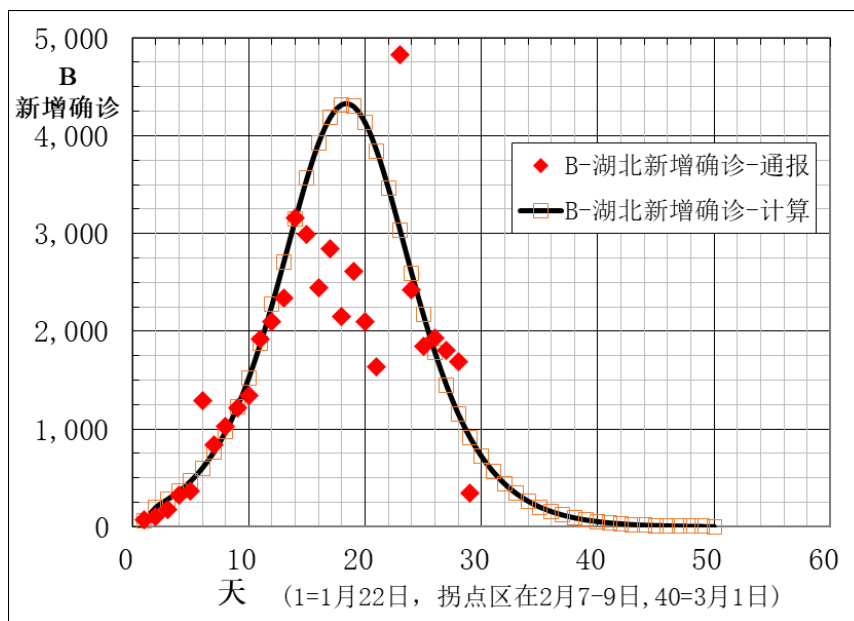


图 7 湖北每日新增确诊数 B 通报和模型计算数据

湖北地区，图 8 显示，模型计算现有确诊数 I 与通报数，在 2 月 12 日以后一致性非常好，在 2 月 12 日前偏差大。这表明在 2 月 12 日前湖北地区通报数据可能有问题。

模型计算得到，现有确诊最大数日期区，即拐点区在 2 月 16 到 18 日期间，与通报的相同，比湖北外其他地区晚 7 天。现有确诊最大人数 50,100 人（取 3 位有效数字），与通报数差 200 人。模型计算，预期疫情在 3 月 22 日走向尾声，比湖北外其他地区晚 11 天，那时现有确诊人数 5 百人左右。

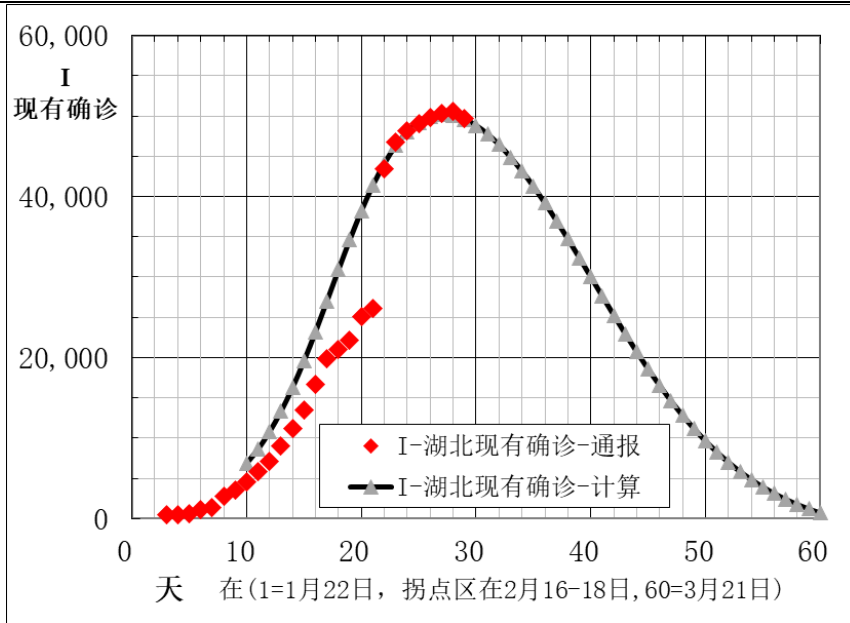


图 8 湖北地区现有确诊数 I 通报和模型计算数据 (预测拐点区 2 月 16 到 18 日)

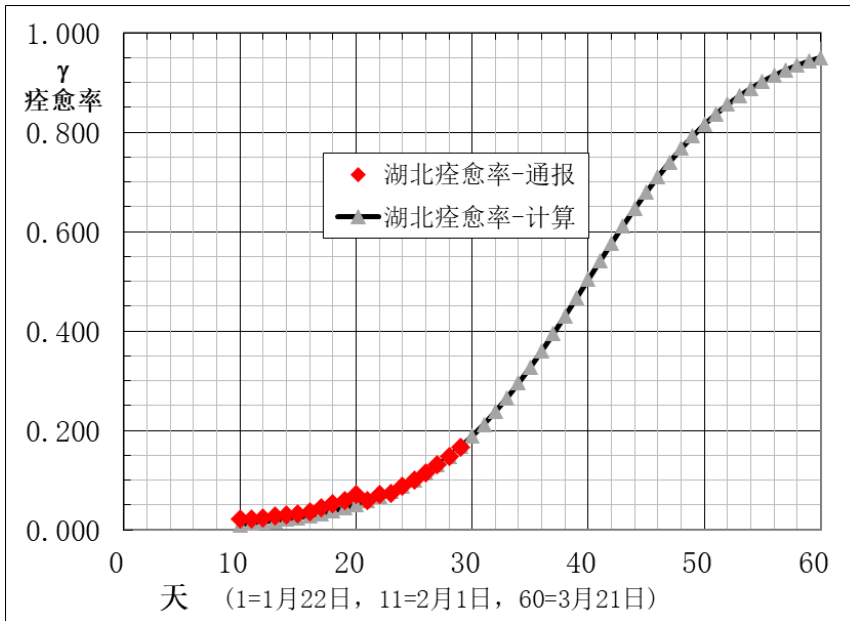


图 9 湖北痊愈率通报和模型计算数据

图 9 看出, 模型计算痊愈率 γ 在 1 月 31 日到 2 月 10 日期间比通报的偏低, 之后几乎相同, 再一次表明模型假设痊愈率 γ 也是逻辑回归函数合理。

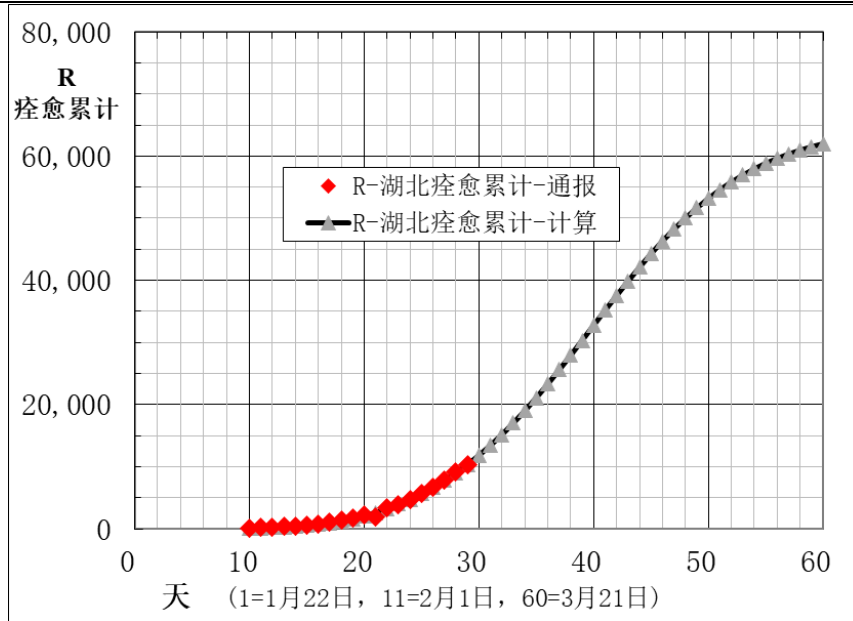


图 10 湖北痊愈累计数 R 通报和模型计算数据

5、结论

本研究建立了传染病动力学 SARD 新型模型。本模型定义 R 等于痊愈率 γ 乘 A，D 等于死亡率 δ 乘 A。模型还大胆但合理假设，痊愈率 γ 是逻辑回归函数，死亡率 δ 是线性函数。

求解一阶微分方程组，得到累计感染者 A 是逻辑回归函数，拟合疫情通报数据，得到逻辑回归函数 A 的两个参数，其中一个 N，是累计感染者总数 N。易感者 S 等于累计感染者 A 的总数 N 减去 A ($S=N-A$)，这样四个函数 S、A、R、D 都找到解析表达式。解决了传染病广泛采用的 SIR 模型，疫情三个函数 S、I、R 没有解析解的难题。现有感染者 I 在 SARD 模型中用 A、R、D 计算得到。

选用湖北外其他地区 COVID-19 疫情 5 天数据，用本 SARD 模型计算得到，从 1 月 30 日到至今 20 多天期间 A、R、I 三类病例数与通报符合良好，验证证明 SARD 模型能定量准确分析传染病 COVID-19 疫情。

SARD 模型计算定量预测 COVID-19 疫情，湖北外其他地区，累计感染者总数 1 万 3 千人，疫情在 3 月 11 日走向尾声，那时现有确诊人数 5 百人左右，很多地级市现有确诊人数将清零。全国累计感染者总数 8 万 1 千人，疫情在 3 月 22 日走向尾声，那时现有确诊人数 5 百人左右。

感谢

本文中 COVID-19 疫情大部分通报数据由天津陈革正高工汇总提供，并参考了斯坦福大学 Michael Levitt 教授发表在网上的这次疫情模型计算论文中疫情早期数据，这里一併感谢。

后记

正文题外话作为后记。

本文作者 2010 年 2 月 17 日写的，传染病动力学 SARD 新型模型，用于 COVID-19 疫情分析和预测一文，应《海纳智库》主编，孙君泓教授邀请，前几天已经发表在《海纳智库》及网站上。

这几天，作者重新进行了数值计算，模型计算与通报数据更加一致，因此决定再写一篇，在《海纳智库》以及网站上发表。这第二篇相关文章，基本按照科学期刊的格式编写，比较严谨，对一些数据给出了误差。内容也进行了大量修改，表达更准确。考虑同一内容第二篇相关文章，为了完整性复制了第一篇的部分内容，但文中图采用重新数值计算得到的数据。在网上发表时，把微分方程，逻辑回归函数表达式，复杂的数值计算过程略去。

作者手头有关资料缺乏，还没有阅读过在学术刊物上发表的传染病建模计算的论文，所以，若有人发现了类似的传染病模型论文，烦请告知发表本论文的网站，或《海纳智库》。发现论文内容问题，也请相关学者、专家及其网友提交给网站。另外，医学领域的学者模型数值计算有困难，作者可以协助计算，并提供计算得到的数据。

本文作者建议，传染病学和其他专业学者，可以采用本 SARD 模型进一步深化研究，写出可以在 Nature, Science, 以及中国和国际医学类等学术刊物上发表的高质量论文。提醒大家，到时不要忘记感谢本文作者章其初，SARD 模型提出者，以及发表的网站及《海纳智库》。

《海纳智库》及网站是志愿者编辑和维护的。本文在网上发表前，主编孙君泓教授，花费了大量时间做修改，这里对孙君泓教授和网站本文的编辑一併感谢。



章其初博士照片