新冠肺炎疫情传播的一般增长模型拟合与预测

张　琳*

(北京邮电大学理学院　北京 海淀区　100876)

【摘要】采用一般增长模型拟合了 2020 年 1 月 15 日−2 月 15 日全国新型冠状病毒肺炎的累计确诊人数，及 2020 年 1 月 23 日−2 月 15 日的全国累计疑似人数和全国累计密切接触人数，模型与国家卫生健康委员会公布的数据吻合。根据一般模型得出的方程形式，分 3 个阶段拟合了确诊人数，发现确诊人数在经历了初期 (2020 年 1 月 15 日−1 月 27 日) 的无障碍指数增长，中期 (2020 年 1 月 27 日−2 月 6 日) 的次指数增长后，已在 2020 年 2 月 6 日进入了次线性增长阶段。拟合结果适时反映了疫情防控工作的进展，同时为疫情的发展趋势提供预测和参考。

关　键　词　指数；　拟合；　新冠肺炎疫情；　非线性拟合；　次指数；　次线性

中图分类号　TP391　　**文献标志码**　A　　doi:10.12178/1001-0548.2020037

Fitness of the Generalized Growth to the COVID-19 Data

ZHANG Lin*

(School of Science, Beijing University of Posts and Telecommunications　Haidian Beijing　100876)

Abstract　A generalized growth model is applied to fit the time series of cumulative confirmed cases between Jan. 15 to Feb. 15, 2020. Moreover, the same formula is also applied to the time series of cumulative susceptive cases and cumulative close contact cases from Jan. 23 to Feb. 15, 2020. The model tallies with data published by the National Health Commission. The sub-exponential and sub-linear growth reflect the time heterogeneity during the transmission of COVID-19, which provide the reference to the prediction of the growth trend of the transmission.

Key words　exponential growth;　fitting;　novel corona virus;　nonlinear fitness;　sub-exponential growth;　sub-linear growth

2020 年 1 月 31 日，世界卫生组织宣布，新冠肺炎疫情构成全球突发公共卫生事件。在病毒传播过程中，挖掘疫情发展数据的内在规律，结合合理的理论模型进行预测，能够为疫情防控提供有价值的参考。

本文基于病毒传播初期过后的次指数增长现象[1]，在指数增长微分方程的基础上，采用了一般形式的微分方程，并根据该一般方程的显式解，对疫情发展数据进行拟合。本文拟合的疫情发展数据来自国家卫生健康委员会的公开数据，包括 2020 年 1 月 15 日−2 月 15 日全国累计确诊病例数、以及 2020 年 1 月 23 日−2 月 15 日全国累计疑似病例数和全国累计密切接触人数。非线性拟合的结果与已发布数据高度吻合。在非线性拟合方程的基础上，本文给出了此后 10 日 (2020 年 2 月 16 日−2 月 25 日) 的趋势预测，为疫情防控提供参考。

1　模型与方法

理论上，在病毒传播的初期，因资源不受限制，其增长方式呈指数函数，满足：

$$y'(t) = ry(t) \tag{1}$$

式中，$y(t) = y(0)\mathrm{e}^{rt}$，$y(0)$为初始感染人数，$r$为增长速率。然而，在很多已经爆发的传染病中，相比于指数增长，次指数的增长趋势在疫情传播初期更为常见。如 HIV/AIDS[2-4]、埃博拉[5] 和手足口[6] 等实证数据，都呈现了次指数的增长方式。

次指数增长方式的内在机制，主要有如下几方面解释：空间异质性、群聚感染[6-7]、以及感染参数的时间异质性[4] 等。在 SIR 模型中，考虑空间异质性[4, 7-9]、传染率的时变性[6-7, 9-12] 以及异质混合模

收稿日期：2020－02－09；修回日期：2020－02－15；网络首发日期：2020－02－21
基金项目：国家自然科学基金 (61671005)
作者简介：张琳 (1983－)，女，博士，主要从事随机过程、传播建模及数据分析等方面的研究. Email: zhanglin2011@bupt.edu.cn

型[13-14]，都能够呈现出次指数的增长方式。考虑到2020年1月23日武汉开始“封城”，加之全国各省市防疫工作的陆续开展，新冠肺炎传播在时间、空间上异质性明显，且发病多呈现集聚感染的情况，都与次指数增长方式产生的机制吻合。

下文的拟合中，在次指数模型的基础上，拓展模型参数范围，使其不仅适用于疫情发展初期的指数增长和次指数增长，还适用于疫情发展到当前阶段，增长速度放慢情形的拟合，拟合结果与国家卫生健康委员会每日发布的累计确诊人数、累计疑似人数和累计密切接触人数的公布数据高度契合。

本文在文献 [2, 7] 给出的病毒传播的一般增长 (generalized-growth) 模型基础上，对尺度参数进行扩展。首先，一般增长模型在式 (1) 的指数增长微分方程基础上，引入可调参数$p \in [0,1]$，称为增长率的负加速度参数。一般增长模型刻画的累计感染人数$y(t)$满足：

$$y'(t) = ry(t)^p \tag{2}$$

式中，$r > 0$表示增长率；$p \in [0,1]$为增长率的负加速度参数，当$p = 0$时，累计感染人数呈线性增长；当$p = 1$时，方程退化为式 (1)。

微分方程 (2) 的精确解为：

$$y(t) = \left(r(1-p)t + y(0)^{1-p}\right)^{\frac{1}{1-p}} \tag{3}$$

式中，$y(0)$为初值；当$0 < p < 1$时，$y(t)$呈次指数增长模式，且p越接近于1，$y(t)$越接近指数增长；参数p的范围扩展为$p \leqslant 1$，当$p < 0$时，$y(t)$将呈次线性增长。下文对累计确诊人数、累计疑似人数及累计密切接触人数的拟合方程均为式 (3)。

新冠肺炎疫情的发展经历了不同的增长模式。本文将确诊数据的拟合分为三个阶段：无障碍指数增长、次指数增长及次线性增长阶段，反映了疫情防控工作的阶段进展，为新冠肺炎疫情的趋势预测提供参考。

2 拟合结果

本文数据采集自国家卫生健康委员会每日疫情通报。数据采用每天更新的前一天0～24时全国病例更新数据，采集数据包括全国确诊累计病例数、全国疑似累计病例数和全国密切接触累计人数，数据截至2020年2月15日24时。

以式 (3) 对2020年1月15日−2月15日全国新冠肺炎的累计确诊人数、2020年1月23日−2月15日的累计疑似人数和累计密切接触人数，进行非线性拟合，拟合参数包括r、p和$y(0)$。

2.1 累计确诊人数

参考疫情发展的关键时间节点，全国新冠肺炎累计确诊病例数的拟合分三个阶段，第一阶段为2020年1月15日−1月27日，本阶段的疫情确诊人数拟合参数为：

$$r = 0.198 \quad p = 1.098 \quad y(0) = 43.2$$

从拟合参数看，p约为1，这一阶段新冠肺炎确诊人数呈指数增长，其原因是疫情发展初期，病毒广泛传播。

第二阶段为2020年1月27日−2月6日。本阶段拟合参数为：

$$r = 12.6 \quad p = 0.559 \quad y(t_0) = 3167$$

此阶段尺度参数$p < 1$，增长呈现次指数的幂函数增长。经过上一阶段疫情的迅猛发展，人们防护意识增强，减少出行，主动隔离，疫情传播得到一定程度的抑制，同时诊断能力有限。

第三阶段为2020年2月6日至本文数据截止时间2020年2月15日。除去临床确诊病例数，本阶段的拟合参数为：

$$r = 9.6 \times 10^9 \quad p = -1.434 \quad y(t_1) = 2.8 \times 10^4$$

这一阶段，$p < 0$，式 (3) 的增速小于线性。“封城”14日后，第一个隔离期结束，火神山医院启用，全国各地医护人员驰援疫情发源地，疫情得到有效控制。在这一阶段，控制成效显著。

2020年2月12日−2月15日发布了含临床诊断的确诊病例数。采用这4日的数据，得到如下拟合参数：

$$r = 4.2 \times 10^6 \quad p = -0.659 \quad y(t_2) = 5.7 \times 10^4$$

尽管4天的数据量少，仍能为疫情发展趋势提供参考。实证数据与拟合曲线如图1所示。

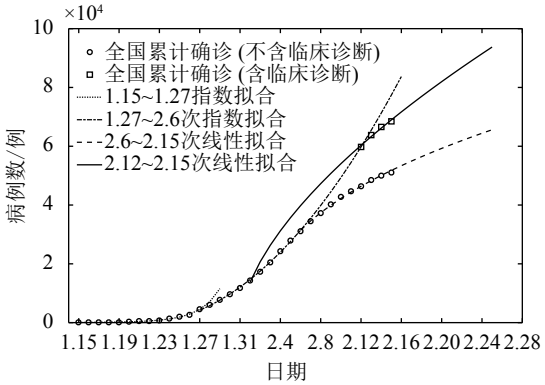

图1 新冠肺炎全国累计确诊人数及其三阶段拟合曲线

在图 1 中，2020 年 1 月 27 日−2 月 6 日的次指数拟合曲线 (图中点划线)，其预测值与 2 月 12 日公布的含临床确诊病例数据非常接近；2 月 12 日−2 月 15 日含临床诊断的确诊病例数据的次线性拟合曲线 (图中实线) 向前追溯 11 日，与 2 月 1 日公布的确诊病例数据接近。尽管这两条曲线无法得知 2 月 12 日之前含临床诊断的确诊数据的演化方式，仍能为认识该阶段疫情发展提供参考。

2020 年 2 月 12 日−2 月 15 日含临床诊断的累计确诊病例数呈次线性增长，拟合参数为$p = -0.659$。结合下文累计疑似病例数与累计密切接触人数的增长趋势，次线性增长方式将在一定时间段内持续，其次线性的参数p将进一步减小，增速会放缓。疫情目前处于胶着对峙阶段。

2.2　累计疑似人数与累计密切接触人数

新冠肺炎累计疑似病例数拟合的分为两个阶段。第一阶段为 2020 年 1 月 23 日−2 月 3 日，拟合参数为：

$$r = 75.8 \quad p = 0.409 \quad y(0) = -0.004\,3$$

这一阶段为次指数增长，$p = 0.409$。武汉“封城”后，人们对防控疫情的重视程度提升，疫情得到了一定程度的控制，疫情发展呈现次指数。

第二阶段为 2020 年 2 月 3 日至本文数据截止时间 2 月 15 日。拟合参数为：

$$r = 9.4\times10^{8} \quad p = -1.126 \quad y(t_1) = 3.6\times10^{4}$$

疑似人数随着医疗资源的配置得到快速诊断，疑似人数次线性指数$p = -1.126$，累计疑似人数增长呈次线性，且次线性程度较大。实证数据与拟合曲线如图 2 所示。

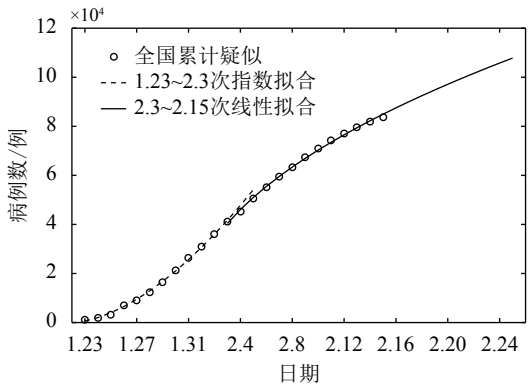

图 2　新冠肺炎全国累计疑似人数及其两阶段曲线拟合

新冠肺炎疫情累计密切接触人数的拟合与疑似人数类似。同样分为两个阶段，第一阶段为 2020 年 1 月 23 日−2 月 3 日，拟合参数为：

$$r = 97.1 \quad p = 0.472 \quad y(0) = 1\,531$$

第二阶段为 2020 年 2 月 3 日−2 月 15 日，拟合参数为：

$$r = 1.4\times10^{8} \quad p = -0.666 \quad y(t_1) = 1.8\times10^{5}$$

从两阶段的参数看，疑似与接触在次指数增长阶段的参数较为相似，但在次线性阶段的参数反映了累计密切接触人数的下降幅度并没有预期的快，其原因或与确诊后的追踪滞后有关。累计密切接触人数的实证数据与拟合曲线如图 3 所示。

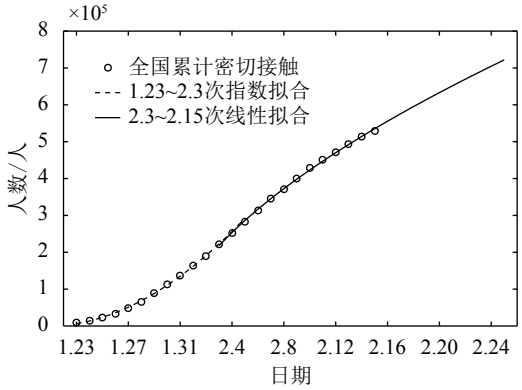

图 3　新冠肺炎全国累计密切接触人数及其两阶段曲线拟合

3　结 束 语

本文借助一般增长模型，对新冠肺炎疫情的确诊、疑似和密切接触人数进行分阶段非线性拟合。拟合结合疫情发展过程，进行合理分段，拟合结果与实证数据高度吻合，能够为疫情的防控效果以及发展趋势提供参考。

指数、次指数、次线性函数的增长方式均由式 (3) 给出，其背后的产生机制可从非时齐随机过程和随机环境中的随机过程的理论中得到启示。新冠疫情的发展过程可用年龄依赖的分枝过程刻画，其中年龄用来刻画每位患者从感染到确诊并隔离住院的这段时间，在此期间，新冠肺炎病毒携带者具有传染性，并能够感染给更多与之接触的人，可建模为该患者的“后代”。结合新冠肺炎疫情传播的实际情况，该年龄依赖分枝过程的年龄分布、后代分布随疫情防控的进展，均呈现非时齐的特点。文献 [15] 也指出，疾病传播过程可由非时齐分枝过程相对应。尽管关于非时齐年龄依赖分枝过程已有相关研究[16-17]，但与新冠肺炎疫情发展一致的非时齐模型并未得到深入研究，这也将是接下来的研究

工作，借助非时齐的年龄依赖分枝过程，从仿真和理论两方面，揭示新冠疫情传播的内在机制。

致谢 感谢北京师范大学新闻传播学院吴晔教授的交流与讨论。

参 考 文 献

[1] 杰弗里·韦斯特. 规模[M]. 北京: 中信出版社, 2018. WEST G. Scale[M]. Beijing: China Citic Press, 2018.

[2] CHOWELL G, VIBOUD C, SIMONSEN L, et al. Characterizing the reproduction number of epidemics with early sub-exponential growth dynamics[J]. Journal of Royal Society Interface, 2018, 13: 20160659.

[3] COLGATE S A, STANLEY E A, HYMAN J M, et al. Risk behavior-based model of the cubic growth of acquired immunodeficiency syndrome in the United States[J]. Proceedings of the National Academy of Sciences of the United States of America, 1989, 86: 4793-4797.

[4] SZENDROI B, CSANYI G. Polynomial epidemics and clustering in contact networks[J]. Proceedings of the Royal Society B, 2004, 271: S364-S366.

[5] MAY R M, ANDERSON R M. Transmission dynamics of HIV infection[J]. Nature, 1987, 326: 137-142.

[6] CHOWELL G, VIBOUD C, HYMAN J M, et al. The Western Africa Ebola virus disease epidemic exhibits both global exponential and local polynomial growth rates[J]. PLoS Currents, 2015, 7: 1-10.

[7] VIBOUD C, SIMONSEN L, CHOWELL G. A generalized-growth model to characterize the early ascending phase of infectious disease outbreaks[J]. Epidemics, 2015, 15: 27-37.

[8] KISKOWSKI M, CHOWELL G. Modeling household and community transmission of Ebola virus disease: Epidemic growth, spatial dynamics and insights for epidemic control[J]. Virulence, 2015, 7: 63-73.

[9] CHOWELL G, SATTENSPIEL L, BANSAL S, et al. Mathematical models to characterize early epidemic growth: A review[J]. Physics of Life Review, 2016, 18: 66-97.

[10] CHOWELL G, HENGARTNER N W, CASTILLO-CHAVEZ C, et al. The basic reproductive number of Ebola and the effects of public health measures: The cases of Congo and Uganda[J]. Journal of Theoretical Biology, 2004, 229: 119-126.

[11] BARBAROSSA M V, DENES A, KISS G, et al. Transmission dynamics and final epidemic size of Ebola virus disease outbreaks with varying interventions[J]. PLoS One, 2015, 10: e0131398.

[12] DRAKE J M, KAUL R B, ALEXANDER L W, et al. Ebola cases and health system demand in Liberia[J]. PLoS Biology, 2015, 13: e1002056.

[13] FINKENSTA D T, GRENFELL B T. Time series modelling of childhood diseases: A dynamical systems approach[J]. Journal of Applied Statistics, 2000, 49: 187-205.

[14] GRENFELL B T, BJORNSTAD O N, FINKENSTADT B F. Dynamics of measles epidemics: Scaling noise, determinism, and predictability with the TSIR model[J]. Ecological Monographs, 2002, 72: 185-202.

[15] CLANCY D, O'NEILL P. Approximation of epidemics by inhomogeneous birth-and-death processes[J]. Stochastic Processes and their Applications, 1998, 73: 233-245.

[16] AGRESTI A. On the extinction times of varying and random environment branching processes[J]. Journal of Applied Probability, 1975, 12: 1, 39-46.

[17] BANSAYE V, CABALLERO M E, MELEARD S. Scaling limits of population and evolution processes in random environment[EB/OL]. (2018-07-09). http//arxiv.org/abs/1802.02362.

编 辑 蒋 晓