基于生态模型的新型冠状病毒肺炎疫情分析

杨安吉，颜忠诚*

(首都师范大学生命科学学院，北京　100048)

摘要:为了研究和解释新型冠状病毒肺炎(COVID-19)疫情的变化趋势,本研究利用生态增长模型对实际数据进行拟合分析.收集COVID-19疫情的相关数据,构建确诊患者例数、治愈患者例数和死亡患者例数3项数值的logistic模型,并预测世界6个区域内的疫情发展趋势.截至2020年4月18日,logistic增长曲线的拟合数值与实际观测数据差异较小,且曲线的决定系数(R^2)均>0.997 0.全国COVID-19疫情的累计确诊患者例数、累计治愈患者例数和累计死亡患者例数变化趋势基本符合logistic增长曲线;全球COVID-19疫情分析显示,世界各地区的感染、暴发有时间差异,通过对现有数据进行曲线拟合,预测不同区域在具有不同拐点时的最终确诊患者例数.本研究证实logistic曲线可以模拟COVID-19疫情的变化趋势,也可以对疫情做出粗略预测,为未来的疫情防治工作提供一些帮助.

关键词:新型冠状病毒肺炎;生态增长模型;变化趋势;回归分析;logistic曲线

中图分类号:Q141　　**DOI**:10.19789/j.1004-9398.2021.02.007

0　引　言

2019年12月,在湖北武汉出现了多例新型冠状病毒肺炎病例,2020年1月12日世界卫生组织(WHO)将其命名为2019-nCoV病毒,并于2月11日将新型冠状病毒肺炎命名为COVID-19(corona virus disease 2019)[1].截至2020年4月18日,COVID-19的大规模传播已造成全国82 735例确诊患者和4 632例死亡患者[2],在全国COVID-19疫情暴发前期,相关研究多围绕在流行病学(35.70%)、诊断治疗(24.60%)、病毒研究(13.50%)等方面[3],而对COVID-19疫情发展趋势的研究分析相对较少.趋势分析是通过研究相关指标的变化,初步预测发展趋势的方法,可以为疫情的防控工作提供有效的帮助.回归分析法是一种利用统计原理对数据进行数学处理,并确定因变量与自变量的相关关系,建立回归方程加以外推,用于预测因变量变化的分析方法[4].

目前针对COVID-19疫情发展的探究,有研究者采用一般增长[5]、二阶导数[6]、指数平滑[7]和离散随机[8]等模型进行分析,此类模型在疫情的一定阶段内拟合程度较高,适用于反映患者数量的短期变化,也有研究者基于SIR[9-10]或SEIR[11-12]及其复合模型[13-15]、人工神经网络[16-18]等方法对疫情的传播机理进行了解释.然而从早期研究认定潜伏期中位数的3.0 d[19],到后来的5.1 d[20],都证实了2019-nCoV病毒具有明显的潜伏特征,SIR模型没有考虑潜伏期的问题.SEIR模型可以加入潜伏期的参数,根据国家卫生健康委员会公布的数据[21],只能通过假设接触率、传染率等重要参数来分析疫情的发展趋势,由SEIR模型改进的复合模型也具有同样局限性,并且模型的相关参数在不同地区有较大的差异.网络动力学模型可以纳入更多参数来提升拟合精度,但是各国在疫情初期检测力度普遍较低,数据更新周期较长,可用于训练模型的数据与实际情况差异较大,限制了模型预测的可靠性.

本研究通过拟合患者比例和增长率的变化来分析疫情趋势,将复杂的外界影响因素视为患者数量增长的环境压力,以拟合生态增长曲线的方式解读2019-nCoV病毒的感染患者例数变化,分析全国COVID-19疫情现状并预估全球发展趋势.

1　数据与方法

1.1　数据来源

本文针对近期出现的COVID-19疫情进行数据分析,其中全国31个省(自治区、直辖市)和新疆生产建设兵团(未包含中国香港、澳门特别行政区及

收稿日期:2020-04-07

* 通信作者:yanzhongcheng@cnu.edu.cn

台湾地区)的原始数据均来自中华人民共和国卫生健康委员会每日发布的疫情消息[21],数据采集时间为2020年1月10日—4月18日,共计100 d,采用了网站报道中的每日新增确诊病例、死亡病例和治愈出院患者例数,以及累计确诊病例、死亡病例和治愈出院患者例数共6项数据,计算每日确诊增长率、治愈增长率和死亡增长率3项数值辅助展开分析.

全球COVID-19疫情的原始数据来自WHO每日发布的情况报告[22],数据采集时间为2020年1月20日—4月28日,共计100 d,采用了报道中世界各个国家或地区的累计确诊患者例数.根据WHO的地理划分,结合疫情各国输入途径和最早确诊时间,将世界主要分为6个内部交流密集的区域,分别是除去中东地区的亚洲及太平洋国家组成的亚太地区,欧洲地区,阿拉伯半岛、东北非和西亚构成的中东地区,非洲地区,美国和加拿大为主的北美地区,中美洲和南美洲组成的拉美地区等.

1.2 曲线拟合方法

曲线拟合是指选择适当的曲线类型来拟合观测数据,并用拟合的曲线方程分析变量之间的关系.本文将每日疫情数据制成散点图,采用单个变化量对时间的响应进行线性或非线性回归,通过比对不同模型的拟合程度,选取拟合程度高的模型进行迭代计算,取得拟合方程的系数,并绘制回归曲线.数据均采用Microsoft Excel 2019进行整理与分析.

结合S、R^2以及残差分析对拟合程度进行判断,其中S描述了数据值与拟合值的距离,表示评估模型描述响应值的程度,其值越低,模型描述响应的程度越高;R^2是度量拟合优度的统计量,介于0~1,也称拟合的确定系数,R^2越高可表示拟合优度越高;残差分析可以通过回归数值与残差的散点图来检验曲线的拟合程度,在拟合正常的残差图中,散点分布在以原点为中心的带状区域上,带状区域的宽度越狭窄,说明模型拟合的精度越高.

1.3 logistic增长曲线

增长曲线被用来描述物种迁入到新生态系统后的数量变化,物种的初始数量一般小于环境最大容纳量[23],又因为自然生态系统中的资源有限,所以物种的增长函数满足logistic方程式:

$$\frac{\mathrm{d}N}{\mathrm{d}t}=rN\left(\frac{k-N}{k}\right), \tag{1}$$

式中N为物种数量,t为时间,r为增长速度,k为环境最大容纳量.

在分析COVID-19疫情变化时,可以将患者看作在限定环境中数量增长的物种,实际中病毒的传播也存在人为管控或自然阻力,因此,2019-nCoV病毒的感染者例数不会无限制的增长,在一定空间范围内应符合logistic增长曲线,其积分方程式为

$$N=\frac{k}{1+e^{a-rt}}, \tag{2}$$

式中k可以通过实际数值、曲线拟合和多点法预估等方式获得,本次分析通过曲线拟合法求得k值;a表示初始患者数在k中的所占比例;r表示病例数在无干扰下的增长速度.a、r可通过进行线性回归统计得到[24],方程式为

$$\ln\left(\frac{k-N}{N}\right)=a-rt. \tag{3}$$

通过4个参数的拟合算法对logistic增长曲线进行处理,得到

$$y=d+\frac{e-d}{1+\left(\frac{x}{c}\right)^{b}}, \tag{4}$$

式中d为增长曲线的最大值,e为曲线最小值,c为曲线拐点即增长速度最大时所对应的x值,一般在曲线中为$\frac{k}{2}$值所对应的x值,b为其对应的曲线斜率.

根据实际情况可代入数据转化为

$$N=k+\frac{N_0-k}{1+\left(\frac{x}{T}\right)^{p}}, \tag{5}$$

式中N为患者例数,k为环境最大容纳量,也就是预估的患者总例数,N_0为初始患者例数,x为时间,T为增长速度最大时所对应的时间,p是该点对应的曲线斜率.

2 结果与分析

2.1 累计确诊病例回归分析

全国COVID-19疫情确诊病例变化趋势及回归分析如图1所示.数据显示,2020年1月10—15日确诊患者例数无增长,一直保持在41例,1月16日后确诊患者例数逐渐上涨,由于前期COVID-19疫情患者基数较小,所以增长率较高且波动较大,曾于1月18日达95.16%的增长率峰值.增长率曲线显示出后因患者数量不断增多,每日增长数量上升但增长率下降,实际中于2月4和12日2次达到的确诊患者例数增长峰值,此阶段内每日增量维持在较高水平.2月14日后增长率开始持续下降,3月7日后每

日增长率维持在0.05%以下,增长数量大体呈每日减少的状态.因此,确诊患者例数的变化表现出了前期增速较缓慢、中期增速加快、后期增速又趋于缓慢的大趋势.

基于2020年4月13日的数据,对去除境外输入情况的累计确诊病例数进行回归分析,得到logistic模型的R^2高于线性、多次、指数、对数和有理多项式等模型,累计确诊病例曲线的拟合结果是N=80 788.3+(41−80 788.3)/(1+exp(((x−30.279 7)/4.416 4)),k=80 788.3,S=1 415.66,R^2=0.998 3.通过对$\ln\left(\frac{k-N}{N}\right)$进行线性回归分析,$a$=6.230 1,$r$=0.226 1.将3个参数($k$、$a$和$r$)代入logistic方程的积分式中进行回归分析,得到logistic增长曲线N=80 788.0/(1+exp(6.230 1−0.226 1(x−2.720 2))),k=80 788.0,S=1 406.56,R^2=0.998 3.注:x表示从1月10日开始统计的时间,因为缺乏1月10日前的疫情数据,所以对logistic方程中的x值采用了修正.

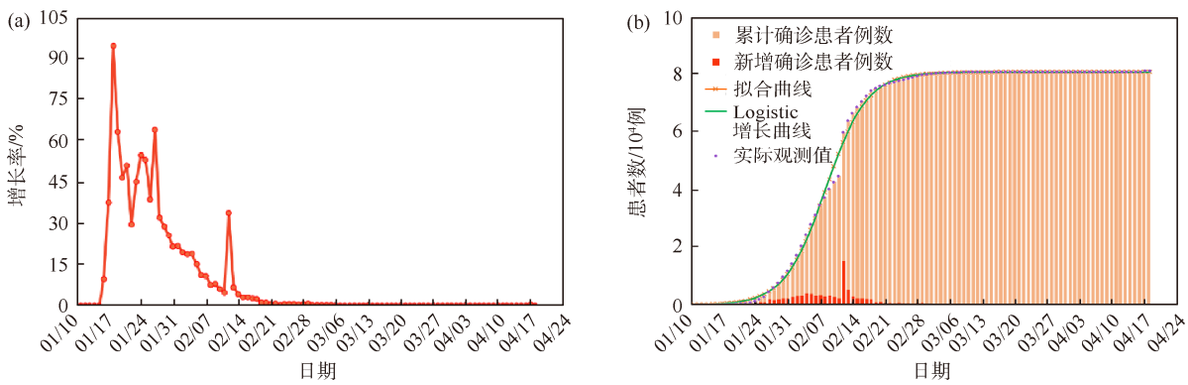


图1　全国COVID-19疫情确诊病例数据
(a)增长率;(b)回归分析

数据拟合结果表明S较小和R^2较高,通过对logistic增长曲线进行赋值,可认为全国COVID-19疫情的确诊病例变化趋势较符合logistic增长曲线.模型的$\frac{k}{2}$出现在自有统计数据后的30 d,对应实际日期是2020年2月8日,处在2月4—12日的2个峰值之间,增速峰值每天约为4 300例.考虑到使用新诊断方法后病例数的大幅增长,2月12日附近出现了单日较大的变动,可以推测2月8—11日的实际患病例数可能要高于当日统计的确诊例数,这与logistic增长曲线的预测值更为接近,因此,可以用logistic模型解释COVID-19疫情在全国31省(自治区、直辖市)和新疆生产建设兵团的总变化趋势.

2020年4月13日及其前后各5 d内确诊病例的实际数值和logistic增长曲线拟合值如表1所示.拟合值与去除境外输入的实际值总体较为接近,数据来源中3月4日前的每日新增病例包含了境外输入患者例数,模型内也将此类引起的增长纳入k值的预估,因此,k会稍高于国内实际增长的确诊患者.3月份后因境外输入病例持续增多并在国内造成传染,导致边境口岸、发达城市等地区检测出新的确诊患者,全国本土确诊病例的变化变为偶发性增加,4月16日湖北省核增确诊病例数后,实际值与拟合值出现了较大差异.

表1　累计确诊病例数据对比　　单位:例

日期	实际观测值(境外输入)	拟合值	差值
04/08	81 865(1 103)	80 786	24
04/09	81 907(1 141)	80 787	21
04/10	81 953(1 183)	80 787	17
04/11	82 052(1 280)	80 787	15
04/12	82 160(1 378)	80 787	5
04/13	82 249(1 464)	80 788	3
04/14	82 295(1 500)	80 788	7
04/15	82 341(1 534)	80 788	19
04/16	82 367(1 549)	80 788	30
04/17	82 719(1 566)	80 788	365
04/18	82 735(1 575)	80 788	372

2.2　累计治愈病例回归分析

全国COVID-19疫情治愈病例变化趋势及回归分析如图2所示.增长率曲线显示,前期因为病例基数小、治愈患者少的前提下,增长率变化波动较大.自2月份起,治愈病例数开始出现每日增长率下降,

但数量稳固上升的情况;在2月底,随着新确诊病例数和剩余患者例数的大量减少,每日新增的治愈例数也较之前有所减少.总体来说,累计治愈例数的增加也表现出了前期增速慢,中期增速快,后期增速慢的变化趋势.

基于2020年4月13日的数据,对去除境外输入情况的累计治愈病例数进行回归分析,曲线拟合结果是N=76 394.2+(2−76 394.2)/(1+exp((x-50.219 4)/6.998 7)),k=76 394.2,S=928.13,R^2=0.999 2.通过对$\ln\left(\frac{k-N}{N}\right)$进行线性回归分析,得到$a$=9.583 2,$r$=0.162 5.将参数$k$、$a$和$r$带入logistic方程的积分式中,得到logistic增长曲线N=76 394.0/(1+exp(9.583 2−0.162 5(x+8.803 1))),k=76 394.0,S=1 529.14,R^2=0.997 8,可知全国累计治愈例数变化基本符合logistic增长曲线.

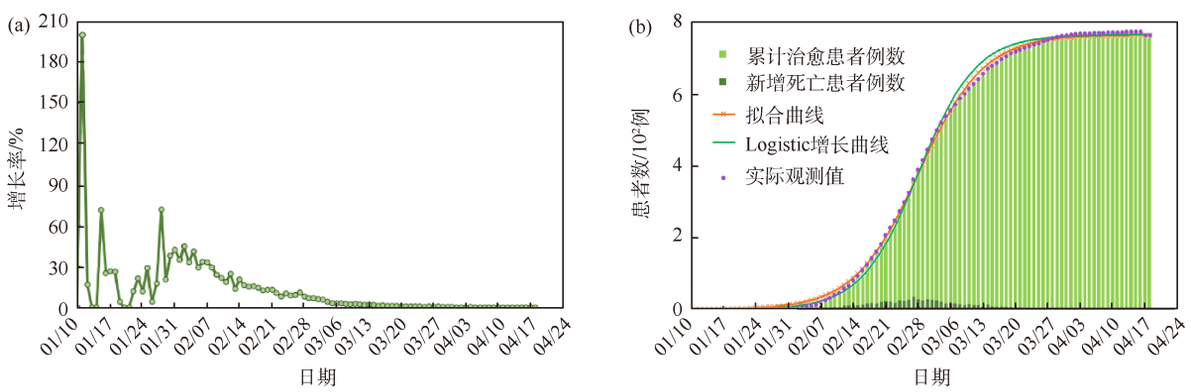


图2　全国COVID-19疫情累计治愈病例数据

(a)增长率;(b)回归分析

在每日补充数据进行回归分析的过程中,显示数据的增多会使参数k上升,a、r下降,S下降,R^2上升,让全国累计治愈例数的变化更加符合logistic增长曲线.考虑到现实中患者从确诊到痊愈需要一定的时间,累计治愈病例数的增长有一定滞后性,模型的$\frac{k}{2}$出现在自有统计数据后的50 d,对应实际中的2020年2月28日,与现实治愈病例的增长拐点2月27日差异较小.

2020年4月13日及其前后各5 d内治愈病例的实际数值和logistic增长曲线拟合值如表2所示.数据来源中3月28日后开始公布境外输入患者的治愈情况,因此,有部分境外输入的治愈病例被纳入k预估,3月30日—4月16日实际值和拟合值差距逐渐加大,4月17日国家卫生健康委员会重新核算内地治愈病例数后,观测值与本研究长期拟合数值的差异迅速缩小.

表2　累计治愈病例数据对比　　单位:例

日期	实际观测值(境外输入)	拟合值	差值
04/08	77 370(374)	76 276	720
04/09	77 455(408)	76 294	753
04/10	77 525(449)	76 309	767
04/11	77 575(481)	76 322	772
04/12	77 663(511)	76 333	819
04/13	77 738(559)	76 342	837
04/14	77 816(596)	76 350	870
04/15	77 892(636)	76 356	900
04/16	77 944(670)	76 362	912
04/17	77 029(709)	76 367	47
04/18	77 062(728)	76 371	37

2.3　累计死亡病例回归分析

全国COVID-19疫情死亡病例变化趋势及回归分析如图3所示.增长率曲线显示死亡病例具有前期波动较大、增长率较高的特点,湖北省因为疫情严重,统计困难,先后于2020年2月13日核减108例、4月17日核增1 290例,导致死亡病例增长率在疫情中后期出现2次较大波动.累计死亡病例的总体增长趋势为:1月10日—21日增长不明显;1月22日后随确诊病例、重症患者例数的大幅度增多而增加;2月25日后因为新增病例、正在就医患者例数的普遍减少,增长速度开始明显下降.全国COVID-19疫情死亡病例显示出了logistic增长曲线的变化趋势.

基于数据来源中4月13日的累计死亡病例进行回归分析,曲线拟合结果是N=3 345.6+(1−3 345.6)/(1+exp((x−37.596 4)/6.811 7)),k=3 345.6,S=56.52,

R^2=0.998 3.通过对$\ln\left(\frac{k-N}{N}\right)$进行线性回归分析,得到结果$a$=6.500 9,$r$=0.145 6,再将$k$、$a$和$r$代入logistic方程的积分式中,得到logistic增长曲线N=3 346.0/(1+exp(6.500 9−0.145 6(x+7.052 4))),k=3 346.0,S=56.24,R^2=0.998 3,通过对比拟合优度,判断此次全国COVID-19疫情累计死亡病例数的变化趋势基本符合logistic增长曲线.

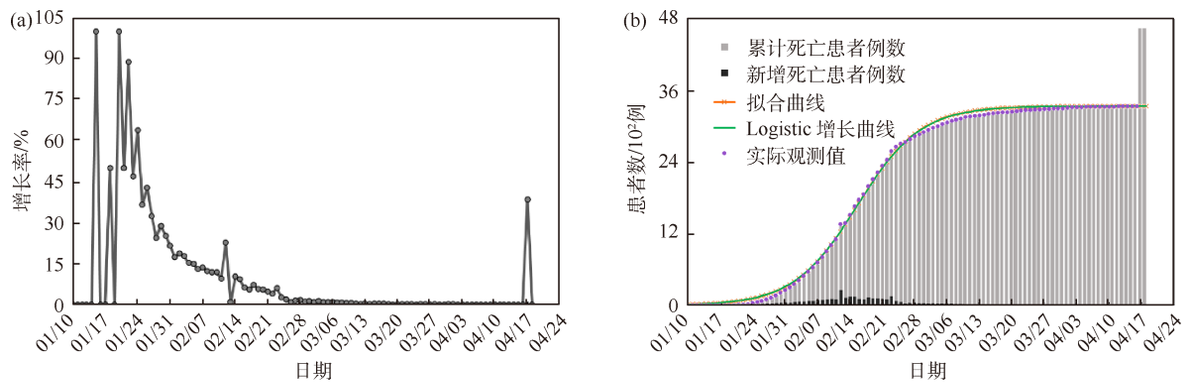


图3　全国COVID-19疫情累计死亡病例数据
(a)增长率;(b)回归分析

2020年4月13日前累计死亡病例的模型拟合值与实际观测值差异较小,增长曲线的S小,R^2高,模型的$\frac{k}{2}$出现在统计数据后的37 d,将2月12日的死亡例数核减后,实际中增长峰值出现在2月14日,与logistic增长曲线中拟合结果基本一致.4月17日的核增通报中指出死亡病例的迟报、漏报和误报现象主要出现在2月20日之前,根据2月20日后的死亡病例增长状况判断疫情前期的实际死亡病例数比曲线拟合值更高,实际累计死亡病例的变化趋势应该同样符合logistic增长曲线.

2020年4月13日及其前后各5 d内死亡病例的实际数值和logistic增长曲线拟合值如表3所示.4月13—18日全国实际增加死亡病例1例,与logistic增长曲线拟合结果一致,医疗水平的提高促使全国COVID-19确诊患者治愈出院的比例上升,4月份后死亡病例数增长缓慢,基本表现为间隔性增加,其中大部分为湖北患者,推测与该地区在4月1日还有400例未治愈的重症患者有关.

表3　累计死亡病例数据对比　　单位:例

日期	实际观测值	拟合值	差值
04/08	3 335	3 344	9
04/09	3 336	3 345	9
04/10	3 339	3 345	6
04/11	3 339	3 345	6
04/12	3 341	3 345	4
04/13	3 341	3 345	4
04/14	3 342	3 345	3
04/15	3 342	3 345	3
04/16	3 342	3 345	3
04/17	4 632	3 346	1 286
04/18	4 632	3 346	1 286

2.4　全球COVID-19疫情预测

全球6个区域COVID-19疫情累计确诊例数变化如图4所示.亚太地区于2020年2月中旬开始出现增长,时间略早于其他地区;欧洲、北美和中东地区于2月下旬开始暴发;拉美地区出现病例相对较晚但增长速度快,5月后有可能出现疫情暴发.2020年4月28日的全球数据拟合曲线显示各区域确诊病例数在5—6月内仍然会有较大增长,北美可能在5月中旬附近超过欧洲成为全球确诊病例最多的区域,非洲、拉美地区因为数据较少,拟合确诊病例数的变化幅度较小.

对全国累计确诊病例的logistic增长曲线进行赋值,得知从预测值>1到增速最大、模型的$\frac{k}{2}$所对应的日期,共经历的时间约60 d,实际中不同国家的疫情增长有时间差异,导致区域达到增速峰值的时间超过60 d,因此,分别对6个区域的累计确诊例数进行回归分析,将方程式(5)中的T设定为70、80、90和100 d,得到基于2020年4月28日的数据(表4),6个区域在不同增长拐点下的拟合确诊病例数,亚太、中东和非洲地区平均每日新增的病例数低于欧洲、北美和拉美地区,因此,logistic增长曲线

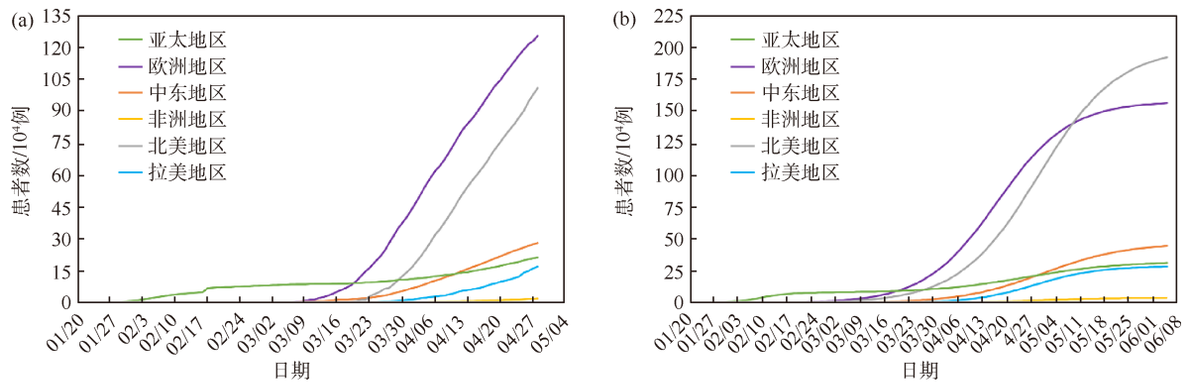


图4　全球6个区域COVID-19疫情确诊病例变化趋势

(a)实际确诊例数;(b)logistic增长曲线拟合例数

表4　基于logistic增长曲线预测不同时间节点下不同区域的累计确诊患者数

地区	70 d		80 d		90 d		100 d	
	患者数/例	日期[a]	患者数/例	日期[a]	患者数/例	日期[a]	患者数/例	日期[a]
亚太地区	170 320	03/29	199 908	04/08	256 872	04/18	343 208	04/28
欧洲地区	1 117 149	04/03	1 620 214	04/13	2 336 365	04/23	3 252 287	05/03
中东地区	255 701	04/07	387 476	04/17	583 284	04/27	859 379	05/07
非洲地区	60 265	05/05	91 021	05/15	135 952	05/25	209 778	06/04
北美地区	686 993	04/02	977 973	04/12	1 573 910	04/22	2 457 276	05/02
拉美地区	520 279	05/06	904 875	05/16	1 609 323	05/26	2 945 492	06/05

注:[a]2020年拟定增速峰值对应日期.

的拟合值较低.欧洲和中东地区的增速峰值设为80~90 d时,已有数据和logistic增长曲线拟合数值的差异较小,亚太和北美地区的增速峰值设为90~100 d时,实际值和拟合值的差异较小,非洲和拉美地区因为数据少,改变T后差异不明显.据此推测在区域内各国不出现疫情复发、突然暴发的条件下,亚太、欧洲、中东和北美地区的最终确诊例数可能为(25~35)万、(162~234)万、(38~59)万和(157~246)万例.

3　讨　论

本文基于logistic模型对全国和全球6个区域的COVID-19疫情进行了分析,全国COVID-19疫情的累计确诊例数、累计治愈例数和累计死亡例数变化趋势基本符合logistic增长曲线.尽管每日数据中存在人为统计原因造成的波动,logistic模型仍较好地拟合了全国疫情的发展趋势,因为logistic增长曲线不同于解释病毒传播机理的模型,是一类描述数据变化的模型,拟合数值时疫情传播过程中时间、空间和感染能力等差异对参数的影响较小,参数可以通过处理数据直接获取,模型在达到增速峰值后,具有较好的发展预测能力,说明使用生态模型来分析疫情变化有一定的可行性.

通过模型预估出整个增长周期为120 d左右,考虑到缺少1月10日之前的数据,实际的疫情可能会相较于拟合数值有所延长,如果能补充早期的发病数据,将可以提升logistic模型的精确性.本研究数据来源选择了我国31个省(自治区、直辖市)和新疆生产建设兵团,是因为这些地区受到中国政府的统一调控较强,可以视为一个限制因素类似的整体环境,从国家卫生健康委员会发布的消息,港、澳、台地区的增长周期要比内地晚近2周的时间,故未将其计入此次的分析数据中,模型拟合也证实全国的COVID-19疫情变化趋势较符合logistic增长曲线.据此推测使用回归分析法研究疫情走势需注意其适用条件,在一定地理范围内,且外界压力变化相似的情况下,logistic模型的拟合程度会更高.

2019-nCoV病毒的潜伏期较长,大众起初对疫情的认识不够透彻,部分国家的早期检测例数很少,在暴发前有将近20 d的低或无增长确诊病例的阶段,而实际感染者数量可能已经高于当时的报

道,按照logistic曲线方式出现了增长,只是随着近期开展的大面积检疫工作才被显现出来,导致确诊病例数的急剧增长.因此,在预测全球COVID-19疫情形势时,暂时不能将世界看作一个整体,因为各地区的增长情况表现出了较强的异质性,本研究认为在数据过少、全球拐点来临前,分区进行回归分析,可以较好地预测COVID-19疫情在全球的变化.

3月份后中国境外输入患者逐渐增多,造成了内地病例数再次增长,同时全球COVID-19疫情出现了更多影响确诊患者数量的因素,logistic增长曲线无法拟合出人员流动、区域二次增长等状况引发的病例数变化,数据大幅度调整对拟合的影响较大,并且增大现有数据与拟定拐点对应日期的距离,将导致拟合值的不确定性增强,因此,拟合结果只能表示该区域最终确诊病例数的大致范围,表现出模型受已知数据影响较大的局限性.但利用此方法可以通过补充数据而使拟合数值更加精确,在得知拐点后可预测疫情大致的持续时间,以生态学的角度来估测可能被感染的例数,为今后的防疫工作提供思路和参考.

4 结束语

本文通过生态模型分析COVID-19的发展趋势,为研究疫情变化提供了新的思路.全国病例数的变化趋势基本符合logistic增长曲线,随着数据的不断积累会修正曲线系数,能有效提高logistic模型的拟合程度.全国的疫情变化已趋于稳定,世界其他地区所处状态各不相同,建议下一步可以在全球疫情的模型构建方面开展相关研究,为防控工作提供更多的参考.抗击疫情需要全世界人民的携手努力,希望在各国家和地区的疾病预防与控制部门和医疗单位合理、规范、安全和透明的调查与救治下,此次疫情能早日结束.

参 考 文 献

[1] World Health Organization. WHO Director-General's remarks at the media briefing on 2019-nCoV on 11 February 2020 [EB/OL].(2020-02-11)[2020-04-04].https://www.who.int/dg/speeches/detail/who-director-general-s-remarks-at-the-media-briefing-on-2019-ncov-on-11-february-2020.

[2] 中华人民共和国卫生健康委员会.截至4月18日24时新型冠状病毒肺炎疫情最新情况[EB/OL].(2020-04-19)[2020-04-22].https://www.nhc.gov.cn/xcs/yqtb/202004/2d391a171acc4624a50a1188c8de7361.shtml.

[3] 巩玥,廖青云,于倩倩,等.冠状病毒研究态势分析[J].中国生物工程杂志,2020,40(1/2):21-37.

[4] 盛骤,谢式千,潘承毅.概率论与数理统计[M].4版.北京:高等教育出版社,2010:254-259.

[5] 张琳.新冠肺炎疫情传播的一般增长模型拟合与预测[J].电子科技大学学报,2020,49(3):345-348.

[6] CHEN X G,YU B. First two months of the 2019 Coronavirus Disease (COVID-19) epidemic in China:real-time surveillance and evaluation with a second derivative model[J]. Global Health Research and Policy,2020,5(5):18-48.

[7] 王旭艳,喻勇,胡樱,等.基于指数平滑模型的湖北省新冠肺炎疫情预测分析[J].公共卫生与预防医学,2020,31(1):1-4.

[8] 石耀霖,程惠红,黄禄渊,等.用离散随机模型研究湖北新冠肺炎COVID-19流行病动力学特征[J].中国科学院大学学报,2020,37(2):145-154.

[9] 李承倬,武文韬,潘振宇,等.基于SIR模型和基本再生数的浙江省新型冠状病毒肺炎防控效果分析[J].浙江医学,2020,42(4):311-314.

[10] VOLPERT V,BANERJEE M,PETROVSKII S. On a quarantine model of coronavirus infection and data analysis[J]. Mathematical Modelling of Natural Phenomena,2020,15:24.

[11] IWATA K,MIYAKOSHI C. A simulation on potential secondary spread of novel coronavirus in an exported country using a stochastic epidemic SEIR model[J]. Journals of Clinical Medicine,2020,9(4):944.

[12] 耿辉,徐安定,王晓艳,等.基于SEIR模型分析相关干预措施在新型冠状病毒肺炎疫情中的作用[J].暨南大学学报(自然科学与医学版),2020,41(2):175-180.

[13] LIN Q,ZHAO S,GAO D,et al. A conceptual model for the coronavirus disease 2019(COVID-19) outbreak in Wuhan,China with individual reaction and governmental action[J]. International Journal of Infectious Diseases,2020,93:211-216.

[14] CHEN T,RUI J,WANG Q,et al. A mathematical model for simulating the phase-based transmissibility of a novel corona-

virus[J]. Infectious Diseases Poverty,2020,9(1):18-25.
[15] 邵年,钟敏,程晋,等.基于FUDAN-CCDC模型对新冠肺炎的建模和确诊例数的预测[J].数学建模及其应用,2020,9(1):29-32.
[16] SANTOSH K C. AI-driven tools for coronavirus outbreak:need of active learning and cross-population train/test models on multitudinal/multimodal data[J]. Journal of Medical Systems,2020,44(5):1-5.
[17] Al-NAJJAR H,Al-ROUSAN N. A classifier prediction model to predict the status of Coronavirus COVID-19 patients in South Korea[J]. European Review for Medical and Pharmacological Sciences,2020,24(6):3400-3403.
[18] GARIRA W. The research and development process for multiscale models of infectious disease systems[J]. PLoS Computational Biology,2020,16(4):e1007734.
[19] GUAN W J,NI Z Y,HU Y,et al. Clinical characteristics of 2019 novel Coronavirus infection in China[J/OL].medRxiv preprint,2020.(2020-02-09)[2020-04-02].https://www.medrxiv.org/content/10.1101/2020.02.06.20020974v1.
[20] LAUER S A,GRANTZ K H,BI Q,et al. The incubation period of coronavirus disease 2019(COVID-19) from publicly reported confirmed cases:estimation and application[J]. Annals of Internal Medicine,2020,172(9):577-582.
[21] 中华人民共和国卫生健康委员会.疫情通报[EB/OL].(2020-04-19)[2020-04-22].https://www.nhc.gov.cn/xcs/yqtb/list_gzbd.shtml.
[22] World Health Organization. Coronavirus disease (COVID-2019) situation reports[EB/OL].(2020-04-29)[2020-04-30].https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/.
[23] 陆卫,王友保.应用逻辑斯蒂映射研究生态平衡[J].安庆师范学院学报(自然科学版),2003(1):53-55.
[24] 万昌秀,梁中宇.逻辑斯谛曲线的一种拟合方法[J].生态学报,1983(3):288-296.

Analysis of COVID-19 epidemic based on ecological model

YANG Anji, YAN Zhongcheng

(College of Life Sciences, Capital Normal University, Beijing 100048)

Abstract: The trend of novel coronavirus pneumonia (COVID-19) epidemic was studied and explained by fitting the actual data with the ecological growth model in this paper. This study collected the data of COVID-19 epidemic situation and constructed the logistic model of three numerical values: the number of confirmed patients, cured patients and dead patients. The logistic growth curve was used to predict the trend of epidemic situation in six regions of the world. As of April 18, 2020, the fitting value of logistic model are slightly different from the actual observation data, and the determination coefficient (R^2) of the curves are greater than 0.997 0. The trend of the cumulative number of confirmed cases, cured cases and dead cases of COVID-19 in China basically conform to the logistic growth curve. The analysis of global epidemic shows that the time of infection and outbreak are different in the regions. The number of confirmed cases in six regions with different inflexion points are inferred according to the curve fitting of existing data. This study confirmed that the logistic growth curve can simulate the trend of COVID-19, it can also make a rough forecast of the epidemic situation to provide data for the epidemic prevention and control work.

Keywords: COVID-19; ecological growth model; tendency; regression analysis; logistic growth curve

(责任编辑:王 媛)