

新型冠状病毒肺炎疫情中后期的发展趋势分析

——基于意大利疫情数据

姚 艺,钱亚冠,陈晓霞,徐庆华,方科彬

(浙江科技学院 理学院,杭州 310023)

摘 要: 为分析新型冠状病毒肺炎疫情中后期的发展趋势,首先以意大利民事保护部门通报的疫情数据作为数据来源,利用3种模型根据实际数据进行拟合;其次利用SIR(susceptible-infective-removal,易感者、感染者、移出者)模型预测疫情的缓增期、最终累计确诊病例数和当日再生数 R_t ;最后分析物理隔离和医疗资源运作效率对后期发展趋势变化产生的影响。研究结果表明:1)意大利疫情的缓增期是从2020年9月29日开始,最终累计确诊病例数约达到26.89万人;2)自意大利“封城”后当日再生数 R_t 在不断减小,疫情得到有效控制;3)通过仿真试验发现,隔离措施对疫情大规模传播具有重要抑制作用,加强医院收治效率对感染人数峰值的快速下降起到了关键作用;4)意大利确诊病例数还会持续增加,很有可能因复工潮而引起疫情反弹。因此,中国仍需保持高度警惕,严格加强出入境管控,避免输入性病例流入。

关键词: logistic; LSTM; SIR; 意大利疫情; 统计预测

中图分类号: TP312.8; R563.1

文献标志码: A

文章编号: 1671-8798(2021)05-0369-09

Trend analysis of COVID-19 epidemic situation in middle to late stage

— A study based on epidemic data of Italy

YAO Yi, QIAN Yaguan, CHEN Xiaoxia, XU Qinghua, FANG Kebin

(School of Sciences, Zhejiang University of Science and Technology, Hangzhou 310023, Zhejiang, China)

Abstract: In order to analyze the development trend of COVID-19 epidemic situation in Italy in the middle to late period, the epidemic data reported by the Italian civil protection department served as the experimental data source, and three models were used to fit the actual data. Then, the SIR (susceptible to infective-removal) model was applied to predict the slow increase period of the epidemic, the final cumulative number of confirmed cases and the intraday regeneration number R_t . Finally, the effects of physical isolation and medical resource efficiency were analyzed on the trend change. The results show that: 1) The gradual increase of

收稿日期: 2020-11-05

基金项目: 国家自然科学基金项目(61902082);浙江省自然科学基金项目(LY17F020011);浙江省公益技术应用研究项目(LGG19F030001)

通信作者: 钱亚冠(1976—),男,浙江省嵊州人,副教授,博士,主要从事机器学习与大数据处理研究。E-mail: qianyg@yeah.net。

the epidemic in Italy started from 29 September 2020, and the cumulative number of confirmed cases reached approximately 268 900; 2) Since the “lockdown” in Italy, the intraday regeneration number R_t has been decreasing, and the epidemic has been effectively controlled; 3) Through the simulation experiment, it is found that the isolation measures have a significant restraining effect on the large-scale spread of the epidemic, and the improvement of hospital admission efficiency plays a key role in the rapid decline of the peak number of infected people; 4) The number of confirmed cases in Italy will continue to increase, and it is highly likely that the epidemic will rebound due to the resumption of work. Therefore, China still needs to remain on high alert and strictly strengthen entry and exit controls to prevent imported cases from entering the country.

Keywords: logistic; LSTM; SIR; epidemic situation in Italy; statistical prediction

新型冠状病毒肺炎(以下简称新冠肺炎)发生以来,对疫情未来趋势的预测一直是学术界关注的热点。针对此次疫情不少研究者对其展开探讨。Zhao 等^[1]基于 2020 年 1 月 10 日至 24 日的公开数据进行指数增长趋势曲线拟合,由此判断新型冠状病毒早期传播能力接近或略高于 SARS。倪顺江^[2]把复杂网络理论和流行病学研究相结合,建立了传染病动力学模型,试验发现传染病感染人数初期时以指数形式增长。这些研究都认为新冠肺炎初期传播具有指数增长的特点,对判断新冠肺炎初期传播的规模和特点具有重要意义。但传统的曲线拟合完全是基于数据趋势来进行预测,无法考虑传染病的传播速度、传播模式及各种防控措施的实施等动态信息,预测效果在初期过后并不可靠。张琳^[3]利用一般增长模型,分 3 个阶段拟合新冠肺炎确诊人数(初期,无障碍指数增长阶段;中期,次指数增长阶段;后期,次线性增长阶段)。而次指数增长方式的内在机制,主要有空间异质性、群聚感染及感染参数的时间异质性^[4]等几方面解释。在 SIR(susceptible-infective-removal)模型中,考虑空间异质性^[5]、传染率的时变性^[6]及异质混合模型^[7],都能够呈现出次指数的增长方式。Yang 等^[8]运用长短期记忆(long-short-term-memory, LSTM)模型预测中国疫情将在 2 月底达到高峰,并通过机器学习算法展示若取消湖北省的交通封闭措施,将导致湖北省在 3 月中旬出现第二次高峰。

2020 年 3 月 17 日,意大利新冠肺炎死亡人数超过武汉^[9],成为疫情严重国家。这引起了研究者的关注。陈晓平^[10]通过聚类分析方法研究各国新冠肺炎疫情发展的严重程度,将国外新冠肺炎疫情国家分为主要疫区、重点疫区、分散疫区 3 类,其中意大利属于主要疫区。刘波^[11]认为,意大利新冠肺炎疫情是全球化背景下的集体问题,这场疫情既考验欧盟处理公共突发事件的能力也给欧洲乃至世界产生了深刻而复杂的影响。

现有的新冠肺炎疫情预测方法包括曲线拟合(curve fitting)、传染病动力学模型(epidemic dynamics model)及人工智能算法(artificial-intelligence, AI)三大类^[12]。疫情预测的主流算法有改进的传染病动力学模型和 LSTM 模型^[13]。为了更好地预测意大利新冠肺炎各阶段的传播情况,本研究选用 logistic(逻辑回归)模型、SIR 模型和 LSTM 模型进行预测。通过对试验结果的分析 and 对比,提取各模型中有价值的预测信息,取长补短,为此次疫情提供更为全面细致的预测结果。

1 预测模型

1.1 logistic 模型

logistic^[14]模型是 Verhulst-Pear 在修正密度方程时提出的,常用于描述种群、传染病增长及商品销售量预测等方面。logistic 函数如下:

$$P(t) = \frac{K}{1 + \left(\frac{K}{P_0} - 1\right)e^{-rt}} \quad (1)$$

式(1)中: $P(t)$ 为 t 时刻的人口数; P_0 为初始容量; r 为人口增长率; K 为环境最大容纳量。

使用 logistic 模型预测的主要步骤如下:

1) 利用最小二乘法构造损失函数:

$$Q(K, P_0, r) = \sum_{t=1}^n (P_t - P(t))^2. \quad (2)$$

式(2): P_t 为 t 时刻实际确诊人数; $P(t)$ 为 t 时刻模型预测的确诊人数。

2) 将损失函数分别对模型参数 K, P_0, r 求导建立方程组, 即

$$\begin{cases} \frac{\partial Q(K, P_0, r)}{\partial K} = 0; \\ \frac{\partial Q(K, P_0, r)}{\partial P_0} = 0; \\ \frac{\partial Q(K, P_0, r)}{\partial r} = 0. \end{cases}$$

求出与真实确诊病例数据最接近的 logistic 函数:

$$P(t) = \frac{268\,999}{1 + \left(\frac{268\,999}{220\,260} - 1\right)e^{-0.01t}}.$$

由上述 logistic 函数定义可知此次疫情累计确诊病例人数最大会达到 268 999 人; 模型的初始容量 P_0 为 22 060 人, 人口增长率 r 为 0.01。

3) 利用 $P(t)$ 函数预测在测试集上的确诊病例数, 用平均绝对百分比误差^[15] (mean absolute percentage error, MAPE) 指标计算模型精确度。

$$e_{\text{MAPE}} = \frac{\sum_{i=1}^n \frac{|Y_i - y_i|}{y_i}}{n} \times 100\%. \quad (3)$$

式(3)中: Y_i 为预测值; y_i 为真实值; n 为预测点的个数。

MAPE 值大小与预测精度呈反比关系, MAPE 值越小, 预测越准确。一般情况下, 如果 MAPE 值小于 20%, 则模型拟合效果好。

4) 利用缓增期公式计算疫情缓增期的初始时间 $t = \frac{\ln a + 1}{b} \frac{317}{b} = 151(\text{d})$ (其中 $a = \frac{K}{P_0} - 1, b = r$), 即累计确诊病例数在第 151 d (2020 年 9 月 29 日) 后进入缓增期。

由于意大利于 2020 年 5 月 4 日实施解封, 导致实际累计病例数在后期仍持续向上增长, 而在 logistic 模型中, 人口数在后期达到饱和状态后会维持水平无法向上增长。因此, 本研究把数据分成 2020 年 2 月 21 日至 5 月 3 日和 5 月 4 日至 8 月 18 日分别进行拟合, 并在两段输入数据里划分训练集和测试集 (比例为 7 : 3)。

1.2 SIR 模型

SIR 模型是一种仓室模型^[16], 它将发病地区的所有人分为易感者 (Susceptible, S)、感染者 (Infective, I) 及移出者 (Recovered, R) 3 类, 并通过某一群体转移至另一群体的传播学机制来建立常微分方程组, 以揭示疫情传播的规律。

模型设置了如下假定: 1) 冠状病毒通过飞沫直接传播; 2) 总人口数是一个常数, 不发生变化, 任何时刻的 3 类人群总数不变; 3) 易感者为尚未感染的健康人群; 4) 感染者是已被感染、具有传染性的人群; 5) 感染者一段时间之后可以康复, 且成为移出者后对于该疾病具有免疫能力。

SIR 模型的常微分方程组形式^[17] 如下:

$$\begin{cases} \frac{dS}{dt} = -\beta IS; \\ \frac{dI}{dt} = \beta IS - \gamma I; \\ \frac{dR}{dt} = \gamma I; \\ S(t) + I(t) + R(t) = N. \end{cases} \quad (4)$$

式(4)中: β 为易感者被传染的平均概率; γ 为感染者痊愈的平均速率; t 为冠状病毒传播天数; N 为人口总数。

观察式(4)可以看出, $\frac{dS}{dt} < 0$ 始终成立,则 $S(t)$ 是关于 t 的单调递减函数且下界为0,故极限

$$\lim_{t \rightarrow \infty} S(t) = S_{\infty} \quad (5)$$

存在。式(4)的第2个方程意味着 $I(t)$ 的变化依赖 $S(t)$ 的大小。如果

$$\frac{dI}{dt} = I(t)[\beta S(t) - \gamma I] < 0, \quad (6)$$

那么当 $t \rightarrow \infty$ 时可以得到 $I_0 > I(t) \rightarrow 0$,此时疾病将消亡而不会暴发。根据基本再生数理论,定义:

$$R_0 = \frac{\beta S_0}{\gamma}. \quad (7)$$

基本再生数 R_0 ^[18]是刻画传染病是否会暴发的阈值。当 $R_0 < 1$ 时,疾病不会暴发,且随时间的推移逐渐消亡;当 $R_0 > 1$ 时,疾病会暴发一段时期,在达到 $I(t)$ 最高值后随时间的推移逐渐消亡。同时,我们定义相应的当日再生数为

$$R_t = \frac{\beta_t S(t)}{\gamma_t}. \quad (8)$$

式(8)中: $S(t)$ 为当日易感者的人群数量; β_t 为当日的平均传染率; γ_t 为当日的平均恢复系数。

使用SIR模型预测的主要步骤如下:

- 1) 将输入数据进行规范化处理,输入数据的区间为2020年2月21日至8月18日。
- 2) 利用最小二乘法构造损失函数:

$$L(\beta, \gamma) = a \sum_t (I(t) - I_t)^2 + \sum_t (R(t) - R_t)^2. \quad (9)$$

式(9)中: $I(t)$ 为真实感染者病例数; $R(t)$ 为真实抵抗者病例数; I_t 为模型感染者病例数; R_t 为模型抵抗者病例数。为了加强模型的拟合效果,损失函数中对确诊病例数的误差项进行了加权,取权重 $a=9$ 。

- 3) 将损失函数分别对模型参数 β, γ 建立方程组,即

$$\begin{cases} \frac{\partial L(\beta, \gamma)}{\partial \beta} = 0; \\ \frac{\partial L(\beta, \gamma)}{\partial \gamma} = 0. \end{cases}$$

以求出一个既满足SIR常微分方程关系又与真实确诊病例数据最接近的SIR曲线,最后求得的SIR模型常微分关系如下:

$$\begin{cases} \frac{dS}{dt} = -0.000\ 001\ 29IS; \\ \frac{dE}{dt} = 0.000\ 001\ 29IS - 0.031\ 408\ 08I; \\ \frac{dR}{dt} = 0.031\ 408\ 08I. \end{cases}$$

此时,易感者的平均被传染率 β_t 为0.000 001 29;感染者痊愈的平均速率 γ_t 为0.031 408 08。

- 4) 利用求得的SIR模型常微分关系预测确诊病例数据,用MAPE指标验证模型的准确性。

- 5) 根据前面得到的当日再生数公式,求得 $R_t = \frac{\beta_t S(t)}{\gamma_t} \approx 0.002\ 840\ 04 \ll 1$,说明此时疫情已经得到

控制。

1.3 LSTM模型

LSTM模型^[19]能够对长时间跨度的时间序列保持良好的记忆,该模型在时间序列的预测问题上有着突出的表现。LSTM模型原理是充分利用过去若干天数的数据信息预测未来若干天数的数据。使用LSTM模型进行确诊病例数预测的主要步骤如下:

- 1) 读取疫情数据,将输入数据进行规范化处理并划分为训练集和测试集,划分时序窗口,设置窗口长度 $L=3$,学习率 $\alpha=0.005$,迭代偏置为 $0.000\ 01$,最大迭代次数 $T=2\ 000$;
- 2) 将输入数据经过输入层的所有节点传递到隐藏层,隐藏层结合 Sigmoid 激活函数将变量映射到 $[0,1]$ 之间以决定信息是否向下层传递;
- 3) 运用 tanh 函数计算当前状态 x_t 、上一刻 LSTM 的隐藏层状态 h_{t-1} 和记忆单元状态 c_{t-1} 等参数值经过遗忘门叠加处理形成的新记忆单元状态值 c_t ;
- 4) 利用 Sigmoid 函数获得初始输出,并利用 tanh 函数对数据缩放得到预测输出 h_t ;
- 5) 计算损失函数 $E=\frac{1}{2}(p_{t+L}-x_{t+L})^2$;
- 6) 权重和偏置更新,基于损失函数 E 使用 Adam 梯度优化器更新 W 和 b ,当更新前后差值的绝对值小于 ϵ 时结束循环;
- 7) 输入测试数据集,验证模型准确性;
- 8) 运用训练后的 LSTM 模型对未来确诊人数进行预测,即用历史确诊人数 $x_t, x_{t+1}, \dots, x_{t+L-1}$ 预测下一时刻的确诊人数 p_{t+L} 。

2 试验结果分析

2.1 试验数据

试验数据来源于意大利民事保护部门的疫情通报(<http://www.salute.gov.it/portale/home.html>),时间范围为 2020 年 2 月 21 日至 8 月 18 日,包括现存确诊人数、累计确诊人数、死亡人数和治愈人数 4 个指标,数据质量良好,无缺失值。其中,logistic 模型使用的数据为 2020 年 2 月 21 日至 8 月 18 日的累计确诊人数;SIR 模型使用的数据为 2020 年 2 月 21 日至 8 月 18 日的现存确诊人数和治愈人数;LSTM 模型使用的数据为 2020 年 2 月 21 日至 8 月 18 日的累计确诊人数和现存确诊人数。

2.2 模型精度比较

为了避免试验的偶然性,我们进行了重复试验,取均值作为该模型的 MAPE 指标值,训练结果见表 1。

从试验结果中可以看到 3 种模型的 MAPE 值均远小于 20%,即 3 种模型预测结果均有效。其中,LSTM 模型的 MAPE 值最小,为 0.367 5%。3 种模型得到的 MAPE 指标值均非常小,其原因可能是分母实际确诊病例的数值过大。模型 SIR 的 MAPE 指标值与其他两种模型相差 20 倍之

多。导致这一差异的原因可能是 SIR 模型无法将突发的“解封”措施考虑在内,使得预测效果在后期并不理想。而 Logistic 模型则通过分段预测的方法大大提高了其预测精度。

通过试验发现,Logistic、SIR 和 LSTM 模型在预测中均各有优势与不足。因此,研究者需要综合考虑,提取模型各自最有价值的预测信息以得到更为全面的预测结果,3 种模型的优缺点对比见表 2。

表 2 3 种模型的优缺点对比

Table 2 Comparison of advantages and disadvantages of three models

模型	优点	缺点
logistic	1) 可以将疫情预测曲线明确划分成三个阶段:渐增期、快增期、缓增期,即可以计算疫情缓增期的开始时间 2) 可以预测疫情最终的累计确诊人数	预测精度一般
SIR	1) 可以预测当日再生数 2) 可以用于模拟仿真	预测精度低
LSTM	模型预测精度高	不适于长期预测

表 1 MAPE 值比较

Table 1 Comparison of MAPE values

算法	MAPE/%
LSTM	0.367 5
logistic	0.489 5
SIR	11.418 8

2.3 缓增期的模型预测

由 logistic 模型预测得到的确诊病例数结果如图 1 所示。红色的虚线为模型感染病例数的预测值，模型预测区间为 2020 年 2 月 21 日至 12 月 15 日，蓝色的实线为截至 2020 年 8 月 18 日官方公布的实际数据。

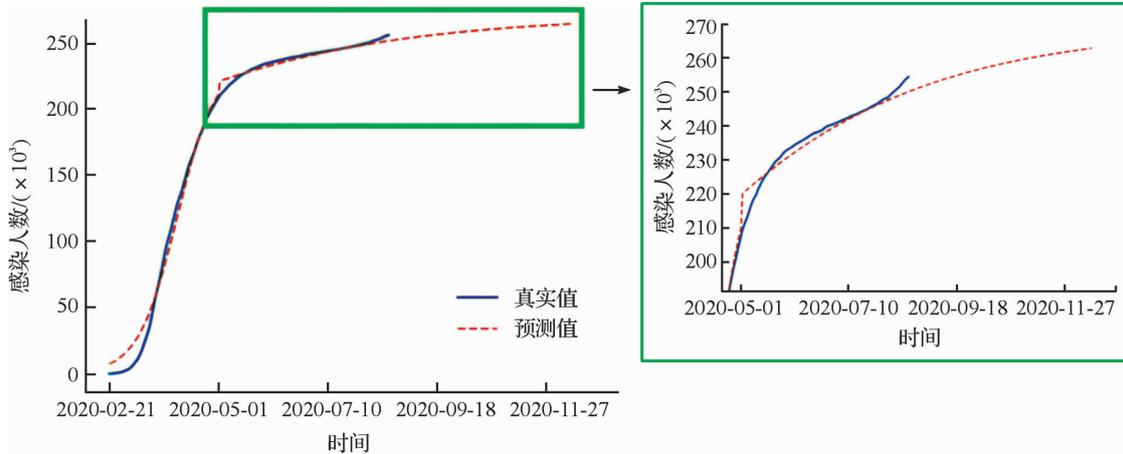


图 1 logistic 模型累计确诊病例趋势预测

Fig. 1 Trend prediction of cumulative confirmed cases by Logistic model

根据上文试验分析模型预测在 2020 年 9 月 29 日后疫情进入缓增期，随后确诊病例数继续向上增长，最终的累计确诊人数 K 会达到 268 999 人。从图 1 中可以看出，模型的预测曲线与实际数据的总体增长趋势一致，在疫情后期两者的偏差随着时间节点的推移而逐渐变大。

2.4 当日再生数的模型预测

由 SIR 模型预测得到的确诊病例数结果如图 2 所示。在 SIR 模型中用 I 类人群拟合实际确诊病例数。 $I(t)$ 、 $R(t)$ 为真实的确诊人数和移出者人数； S_t 、 I_t 、 R_t 分别为模型预测的易感者人数、感染者人数和移出者人数。从图 2 中可以看出，模型预测在 2020 年 4 月 21 日出现拐点，现存感染人数的峰值可达到约 124 554 人(略高于真实数据)。但在 2020 年 7 月 10 日后，模型预测曲线开始偏离真实数据，可见，在疫情后期 SIR 模型的预测效果不佳。

以 2020 年 2 月 11 日至 3 月 8 日为第一个时间段，每增加一天构成一个新的时间段，用 SIR 模型求解其当日再生数 R_t ，得到了当日感染再生数随时间推移的演化趋势图，预测区间为 2020 年 3 月 8 日至 7 月 10 日，如图 3 所示。

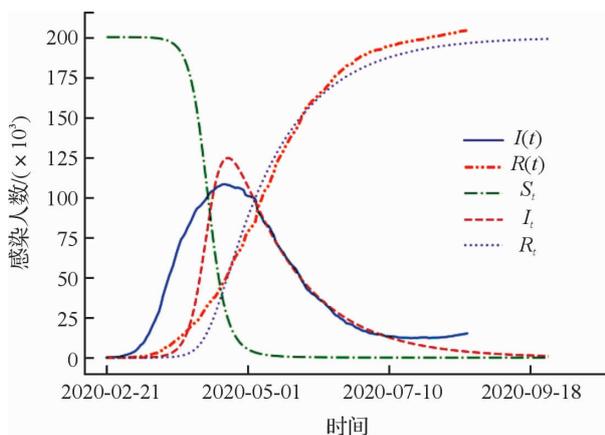


图 2 SIR 模型现存确诊病例趋势预测

Fig. 2 Trend prediction of existing confirmed cases by SIR model

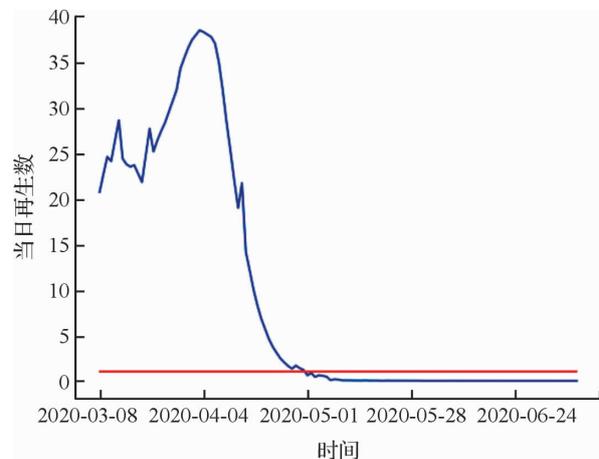


图 3 当日再生数随时间的演化情况

Fig. 3 Change of intraday regeneration number over time

从图3中可以发现,在2020年3月8日意大利全国“封城”之后当日基本再生数上下不稳定地波动,总体向上增长。在2020年4月4日时当日再生数达到峰值,之后开始快速下降,隔离措施产生了作用,且当日再生数在2020年5月1日至7月10日一直小于1,疫情得到有效控制。

2.5 隔离措施对疫情传播的影响

从2020年3月8日起,意大利正式签署“封城令”,开始实施欧洲史上规模最大的封城措施。为验证物理隔离对阻断新冠肺炎传播是否有效果。本研究通过提高传染率来模拟意大利不采取封城措施的情形,并与SIR模型预测到的感染者人数放在一起做对比分析,结果如图4所示。

图4中, $I(t)$ 为官方提供的确诊人数; I_t 为模型预测的感染者人数。SIR模型参数 β 为易感者被感染的平均概率。采取隔离措施相当于降低传染率 β ,两者高度相关。将模型中的参数 β 提高到原来的2倍来模拟没有采取“封城”措施时感染者病例数随时间变化的曲线 $I_t(2\beta)$ 。

通过分析我们发现,目前的防控措施对疫情起到了很好的抑制作用。如果不采取相应隔离措施,现存感染人数会比目前高出30000人左右,也就是说,如果没有封锁措施,将导致感染人数峰值更高且更快达到峰值,造成更大的生命财产损失以及社会恐慌。

2.6 集中收治对感染人数峰值的影响

在新冠肺炎传播的高峰期,医疗人员和物资的供应是一个关键问题。医院的收治效率与SIR模型中的恢复系数 γ 呈正相关。本研究通过降低恢复系数来模拟医院收治效率不高的情况,结果如图5所示。 $I(t)$ 为官方提供的确诊病例数; I_t 为模型预测的感染者病例数; $I_t(\gamma/2)$ 为模型中的恢复系数 γ 减少到原来一半时的感染者病例数。

从图5中可以看出,集中收治在疫情高峰期可以避免约3万人被感染,在疫情后期现存病例数也明显低于未采取集中收治的状况。综上,集中收治对感染者病例数的迅速回落发挥了关键作用。

2.7 意大利疫情后期的增长趋势

由LSTM模型预测得到的确诊病例数结果如图6~7所示,模型预测的区间为2020年2月11日至8月24日。红色的虚线为模型在训练集上对累计和现存感染人数的预测值,绿色的虚线为模型在测试集上对累计和现存感染人数的预测值,蓝色的实线为截至2020年8月18日官方公布的数据。

从图7中可以看出,在2020年4月21日现存确诊病例数达到峰值后开始迅速下降,隔离措施初见成效。但疫情好转的同时也给意大利经济带来重创,为此意大利在2020年5月4日被迫实行解封措施。从图7可以看出,确诊病例数在达到低点后又开始继续上升且增长速率越来越快,进入二次增长阶段。如果意大利政府在这种情况下放任不管、不采取任何防控措施,势必会导致疫情二次爆发。

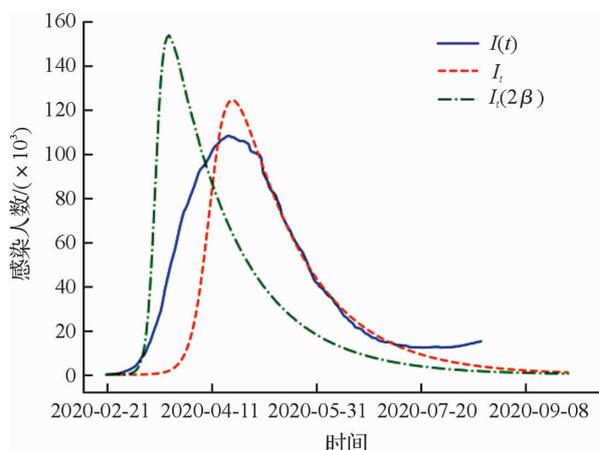


图4 防控隔离措施对疫情的影响

Fig. 4 Impact of physical isolation measures on epidemic

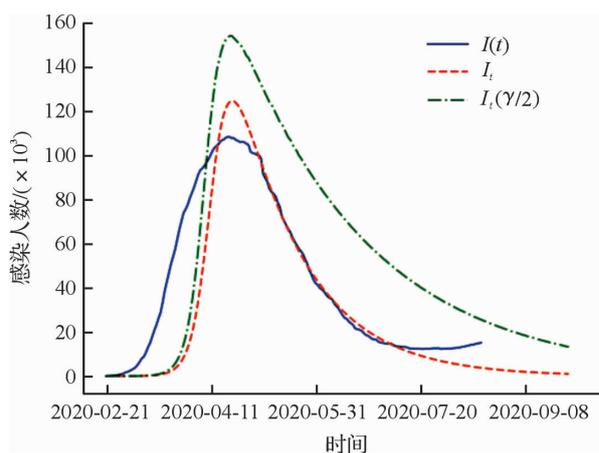


图5 医院收治效率对疫情的影响

Fig. 5 Impact of hospital admission efficiency on epidemic

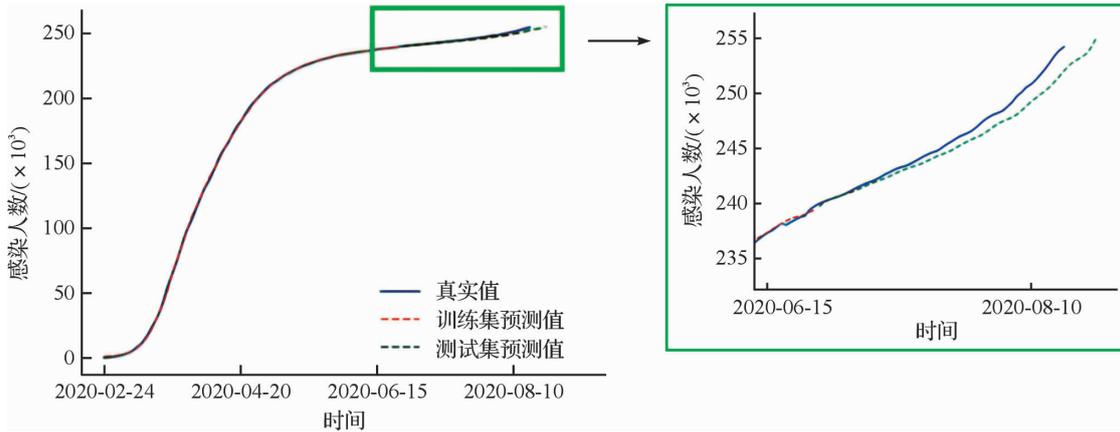


图 6 LSTM 模型累计确诊病例趋势预测

Fig. 6 Trend prediction of cumulative confirmed cases by LSTM model

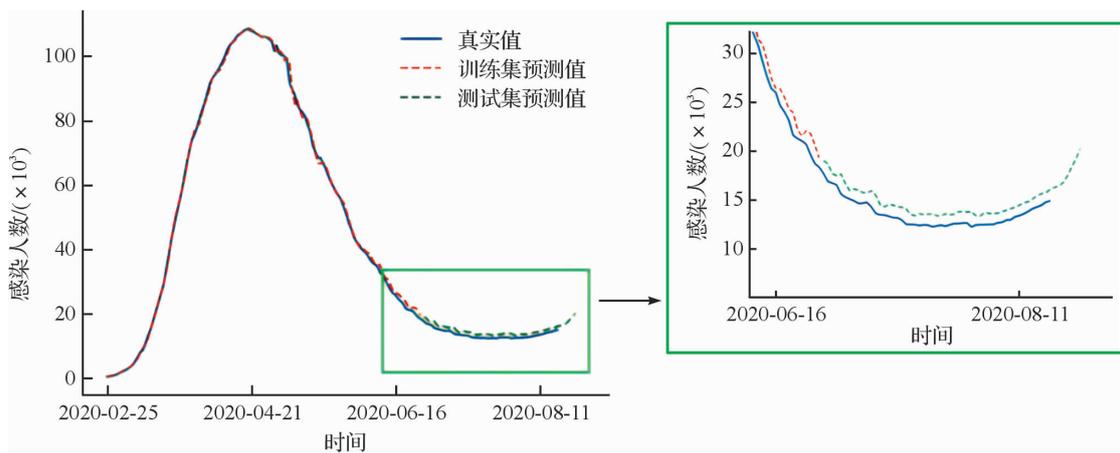


图 7 LSTM 模型现存确诊病例趋势预测

Fig. 7 Trend prediction of existing confirmed cases by LSTM model

3 结 论

本研究根据意大利“封城”和“解封”两个重要时间节点,将疫情传播划分为初期(指数上升阶段)、中期(次指数增长阶段)和后期(二次增长阶段)3阶段,并利用 logistic 回归模型、SIR 模型和 LSTM 模型,对意大利疫情中后期传播过程中感染者的发展趋势进行建模和分析。通过试验验证,拟合结果与实际数据高度吻合,从而可为疫情趋势分析提供有效的数据分析支持。由此得出如下结论:

1) 意大利的缓增期出现于 2020 年 9 月 29 日左右,预计最终累计确诊病例数将达到 26.89 万人。当日再生数 R_t 在不断减小,这与政府的防控措施及民众的居家防护有直接关系。

2) 根据模型预测结果,在疫情后期意大利确诊人数会进入二次增长阶段,建议意大利推迟全部企业复工复产及学校开学的时间。目前,意大利疫情控制还不稳定,甚至海外一些国家不仅仅是面对新冠疫情,例如“西尼罗病毒”和“沙门氏菌病毒”席卷美国,新加坡受到“登革热”疫情的困扰。为了防止这些传染病流入中国,我们还需加强出入境管理。

参考文献:

- [1] ZHAO S, MUSA S S, LIN Q, et al. Estimating the unreported number of novel coronavirus (2019-nCoV) cases in China in the first half of January 2020: a data-driven modelling analysis of the early outbreak[J]. Journal of Clinical Medicine, 2020, 9(2): 388.
- [2] 倪顺江. 基于复杂网络理论的传染病动力学建模与研究[D]. 北京:清华大学, 2009.

- [3] 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.
- [4] SZENDROI B, CSÁNYI G. Polynomial epidemics and clustering in contact networks[J]. *Proceedings of the Royal Society B: Biological Sciences*, 2004, 271(supplement 5): S364.
- [5] CHOWELL G, SATTENSPIEL L, BANSAL S, et al. Mathematical models to characterize early epidemic growth: a review[J]. *Physics of Life Reviews*, 2016(18): 66.
- [6] DRAKE J M, KAUL R R B, ALEXANDER L W, et al. Ebola cases and health system demand in Liberia[J]. *PLoS Biology*, 2015, 13(1): e1002056.
- [7] GRENFELL B T, BJÄRNSTAD O N, FINKENSTÄDT B F. Dynamics of measles epidemics: scaling noise, determinism, and predictability with the TSIR model[J]. *Ecological Monographs*, 2002, 72(2): 185.
- [8] 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.
- [9] 陶短房. 意大利疫情: 为何死亡率这么高? [N]. 中国经营报, 2020-03-16(E03).
- [10] 郑丽晖, 陈建宝, 陈晓平. 新型冠状病毒(COVID-19)传播的统计建模与风险预测[J]. 应用数学学报, 2020, 43(2): 324.
- [11] 刘波. 意大利疫情数字的背后[J]. 人民论坛, 2020(10): 20.
- [12] 黄丽红, 魏永越, 沈思鹏, 等. 常见新型冠状病毒肺炎疫情预测方法及其评价[J]. 中国卫生统计, 2020, 37(3): 322.
- [13] 宁晴, 鲍泓, 徐成. 新冠病毒疫情预测模型研究方法评述[C]//第二十四届网络新技术与应用年会论文集. 北京: 中国计算机用户协会网络应用分会, 2020.
- [14] 百度百科. Logistic 模型[EB/OL]. [2020-09-02]. <https://baike.baidu.com/item/Logistic%E6%A8%A1%E5%9E%8B>.
- [15] 贾俊平, 何晓群, 金勇进. 统计学[M]. 5版. 北京: 中国人民大学出版社, 2012: 334.
- [16] KERMACK W O, MCKENDRICK A G. A contribution to the mathematical theory of epidemics[J]. *Proceedings of the Royal Society of London Series A*, 1927, 115(772): 700.
- [17] 喻孜, 张贵清, 刘庆珍, 等. 基于时变参数-SIR模型的 COVID-19 疫情评估和预测[J]. 电子科技大学学报, 2020, 49(3): 357.
- [18] 崔玉美, 陈姗姗, 傅新楚. 几类传染病模型中基本再生数的计算[J]. 复杂系统与复杂性科学, 2017, 14(4): 14.
- [19] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. *Neural Computation*, 1997, 9(8): 1735.