

· 新型冠状病毒肺炎疫情防控 ·

时间序列模型应用于新型冠状病毒肺炎 疫情预测效果比较研究

李忠奇 陶必林 占梦瑶 吴祝超 吴继周 王建明

南京医科大学公共卫生学院全球健康中心流行病学系 211166

通信作者:王建明, Email:jmwang@njmu.edu.cn

【摘要】目的 比较常见时间序列模型应用于新型冠状病毒肺炎(COVID-19)疫情预测的效果。**方法** 收集2020年4月1日至9月30日美国、印度和巴西3个国家COVID-19每日确诊病例数,分别建立差分自回归移动平均(ARIMA)模型和循环神经网络(RNN)模型,使用平均绝对百分比误差(MAPE)和均方根误差(RMSE)等指标,比较不同模型预测9月21-30日确诊病例的表现。**结果** 应用ARIMA模型预测美国、印度和巴西疫情的MAPE分别为13.18%、9.18%和17.30%,RMSE分别为6 542.32、8 069.50和3 954.59;应用RNN模型预测美国、印度和巴西疫情的MAPE分别为15.27%、7.23%和26.02%,RMSE分别为6 877.71、6 457.07和5 950.88。**结论** ARIMA和RNN模型的COVID-19预测效果存在地区差异,ARIMA模型的预测效果在美国和巴西较优,而RNN模型的预测效果在印度较优。

【关键词】 新型冠状病毒肺炎; 差分自回归移动平均模型; 循环神经网络模型; 预测

基金项目:国家重点研发计划(2017YFC0907000);国家自然科学基金(81973103);高校哲学社会科学重大研究项目(2020SJZDA096)

A comparative study of time series models in predicting COVID-19 cases

Li Zhongqi, Tao Bilin, Zhan Mengyao, Wu Zhuchao, Wu Jizhou, Wang Jianming

Department of Epidemiology, Center for Global Health, School of Public Health, Nanjing Medical University, Nanjing 211166, China

Corresponding author: Wang Jianming, Email: jmwang@njmu.edu.cn

【Abstract】 Objective To compare the performances of different time series models in predicting COVID-19 in different countries. **Methods** We collected the daily confirmed case numbers of COVID-19 in the USA, India, and Brazil from April 1 to September 30, 2020, and then constructed an autoregressive integrated moving average (ARIMA) model and a recurrent neural network (RNN) model, respectively. We applied the mean absolute percentage error (MAPE) and root mean square error (RMSE) to compare the performances of the two models in predicting the case numbers from September 21 to September 30, 2020. **Results** For the ARIMA models applied in the USA, India, and Brazil, the MAPEs were 13.18%, 9.18%, and 17.30%, respectively, and the RMSEs were 6 542.32, 8 069.50, and 3 954.59, respectively. For the RNN models applied in the USA, India, and Brazil, the MAPEs were 15.27%, 7.23% and 26.02%, respectively, and the RMSEs were 6 877.71, 6 457.07, and 5 950.88, respectively. **Conclusions** The performance of the prediction models varied with country. The ARIMA model had a better prediction performance for COVID-19 in the USA and Brazil, while the RNN model was more suitable in India.

【Key words】 COVID-19; Autoregressive integrated moving average model; Recurrent neural network model; Predicting

DOI: 10.3760/cma.j.cn112338-20201116-01333

收稿日期 2020-11-16 本文编辑 万玉立

引用本文:李忠奇,陶必林,占梦瑶,等.时间序列模型应用于新型冠状病毒肺炎疫情预测效果比较研究[J].中华流行病学杂志,2021,42(3):421-426. DOI: 10.3760/cma.j.cn112338-20201116-01333.



Fund programs: National Key Research and Development Program of China (2017YFC0907000); National Natural Science Foundation of China (81973103); Key Project of Philosophy and Social Science Research in Colleges and Universities in Jiangsu Province (2020SJZDA096)

2019 年年末新发现的新型冠状病毒肺炎 (COVID-19) 已造成全球大流行, 成为影响人类健康的重大公共卫生问题^[1]。准确预测 COVID-19 的发病趋势可以帮助政府部门合理分配医疗卫生资源, 及时调整防控策略, 有效遏制疫情。时间序列分析是根据系统观测获得的数据, 通过曲线拟合和参数估计来建立数学模型的一类方法。差分自回归移动平均 (autoregressive integrated moving average, ARIMA) 模型和循环神经网络 (recurrent neural network, RNN) 模型是两种常见的时间序列模型。其中, ARIMA 模型已被用于预测意大利、西班牙和法国等国家的 COVID-19 疫情, 并被证明具有良好的预测性能^[2], 但尚未有研究探索应用 RNN 模型预测 COVID-19 并比较其预测能力。因此, 本研究通过构建 ARIMA 模型和 RNN 模型, 旨在比较这两种模型应用于预测不同地区 COVID-19 疫情的表现。

资料与方法

1. 数据收集: 据 WHO 统计数据, 截至 2020 年 9 月 30 日, 美国、印度和巴西分别确诊 7 077 015、6 225 763 和 4 745 464 例 COVID-19, 是全球累计确诊病例数最多的 3 个国家^[1]。因此, 从 WHO 发布 COVID-19 疫情数据的网站 (<https://covid19.who.int/>) 收集上述 3 个国家 2020 年 4 月 1 日至 9 月 30 日期间每日确诊病例数, 用于构建时间序列模型并评估其预测性能。

2. ARIMA 模型的构建: ARIMA 模型是最经典、最成熟的时间序列模型, 该模型对数据的分布类型不做特殊要求, 已被广泛应用于结核病^[3]、手足口病^[4]和乙型肝炎^[5]等传染病的预测。ARIMA 模型可以表示为 ARIMA(p, d, q), 其中 p 表示自回归的阶数, d 表示差分的次数, q 表示移动平均的阶数。构建 ARIMA 模型包含 4 个步骤: ①通过时间序列图和单位根检验判断序列是否平稳, 若不平稳, 则通过差分的方式使其平稳; ②根据平稳序列的自相关系数图和偏自相关系数图初步确定 p 和 q 的取值, 以构建一些备选模型, 其中 p 的取值参考偏自

相关系数图, q 的取值参考自相关系数图; ③根据最小化贝叶斯信息准则 (Bayesian information criterion, BIC) 值的原則确定最佳模型; ④对模型的系数进行假设检验并检测其残差是否为白噪声, 如果模型的某项系数不显著或其残差被证明为非白噪声, 则应重新选择模型^[6]。确定最佳 ARIMA 模型后, 用其预测 9 月 21–30 日的 COVID-19 确诊病例数。

3. RNN 模型的构建: ARIMA 模型本质上是一种线性回归模型, 考虑到 COVID-19 的发病趋势可能是非线性, 因此, 构建了另一种非线性模型, 即 RNN 模型。RNN 模型是一种常见的神经网络模型, 除了具有强大的非线性拟合能力外, 还拥有记忆功能。神经网络模型通常由输入层、隐藏层和输出层组成, 传统的神经网络模型层与层之间完全连接, 但每一层中的神经元彼此之间没有连接, 而 RNN 模型则在隐藏层的神经元间添加连接, 其隐藏层的输入不仅包括输入层的输出, 还包括隐藏层先前的输出, 从而使其具有跨时间步长传递信息的能力, 特别适合用于时间序列分析^[7-8]。本研究将 2020 年 4 月 1 日至 9 月 30 日的每日确诊数据划分为训练集、测试集和预测集, 共构建了 10 个 RNN 模型 (RNN1~RNN10), RNN1 使用前 1 天的数据预测当日的数据, RNN2 使用前 2 天的数据预测当日的数据, 以此类推。在训练前对数据进行归一化处理, 使用公式: $X' = \frac{X - X_{min}}{X_{max} - X_{min}}$ 将所有数据转化到

区间 $[0, 1]$ 间。其中, X 是原始值, X_{max} 是原始值的最大值, X_{min} 是原始值的最小值, X' 是转化后的值。对每个 RNN 模型进行 3 次训练, 并使用测试集 (9 月 1–20 日的确诊数据) 比较预测性能, 以确定最佳的 RNN 模型。此外, 对于每个 RNN 模型, 将学习率分别设置为 0.05、0.1 和 0.2, 将隐藏层的维度分别设置为 3、5 和 10, 通过观察其在训练集上误差的下降情况确定合适的训练步数, 根据不同参数组合下模型在测试集上的表现确定最合适的参数。确定最佳 RNN 模型后用其预测 9 月 21–30 日的确诊病例数。

4. 模型预测性能比较: 采用平均绝对百分比误

差(mean absolute percentage error, MAPE)和均方根误差(root mean square error, RMSE)比较模型的预测性能^[9-10]。MAPE = $\frac{1}{n} \sum_{i=1}^n \frac{|X_i - \hat{X}_i|}{X_i} \times 100$, RMSE =

$\sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - \hat{X}_i)^2}$, 其中, X_i 是时刻 i 的实际值, \hat{X}_i 是模型在时刻 i 的预测值, n 为样本量。

5. 统计学分析: 使用 R 4.0.2 中的“forecast”“tseries”和“rnn”包构建 ARIMA 和 RNN 模型, 检验水平 $\alpha=0.05$, 以 $P<0.05$ 为差异有统计学意义。

结 果

1. ARIMA 模型: 2020 年 4 月 1 日至 9 月 20 日, 美国、巴西和印度的 COVID-19 每日确诊病例数见图 1。美国和巴西的日确诊病例数在 7 月底和 8 月初达到峰值, 而后略有下降, 而印度的日确诊病例数则一直波动性上升, 未出现明显的下降趋势。

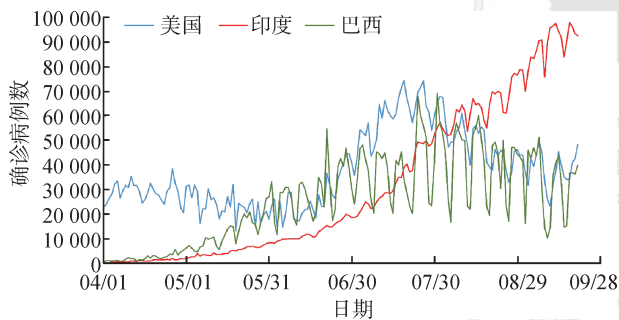


图 1 美国、印度和巴西 2020 年 4 月 1 日至 9 月 20 日的每日确诊病例数

对病例序列进行两次差分后经单位根检验, 结果显示, 序列已经平稳 ($P<0.01$)。平稳序列的自相关系数和偏自相关系数见图 2 A1~A3 和 B1~B3。对于美国和巴西, 初步确定 $p=1\sim 8, q=1$; 对于印度, 初步确定 $p=1\sim 8, q=1\sim 2$, 最终确定 ARIMA(6, 2, 1) 为最优 ARIMA 模型。对于美国, 尽管 ARIMA(6, 2, 1) 的 BIC 值略大于 ARIMA(7, 2, 1), 但考虑到模型复杂度和 p 的取值较高可能会导致过拟合问题, 因此, 仍选择 ARIMA(6, 2, 1) 为美国的最优 ARIMA 模型(表 1)。参数假设检验显示, 模型各项系数均有统计学意义 ($P<0.05$), 其残差序列的自相关性和偏自相关性已经基本消除, 可认为其为白噪声序列(图 2 C1~C3 和 D1~D3)。

使用 ARIMA(6, 2, 1) 预测 3 个国家 9 月 21-

30 日的确诊病例数, 结果见表 2。

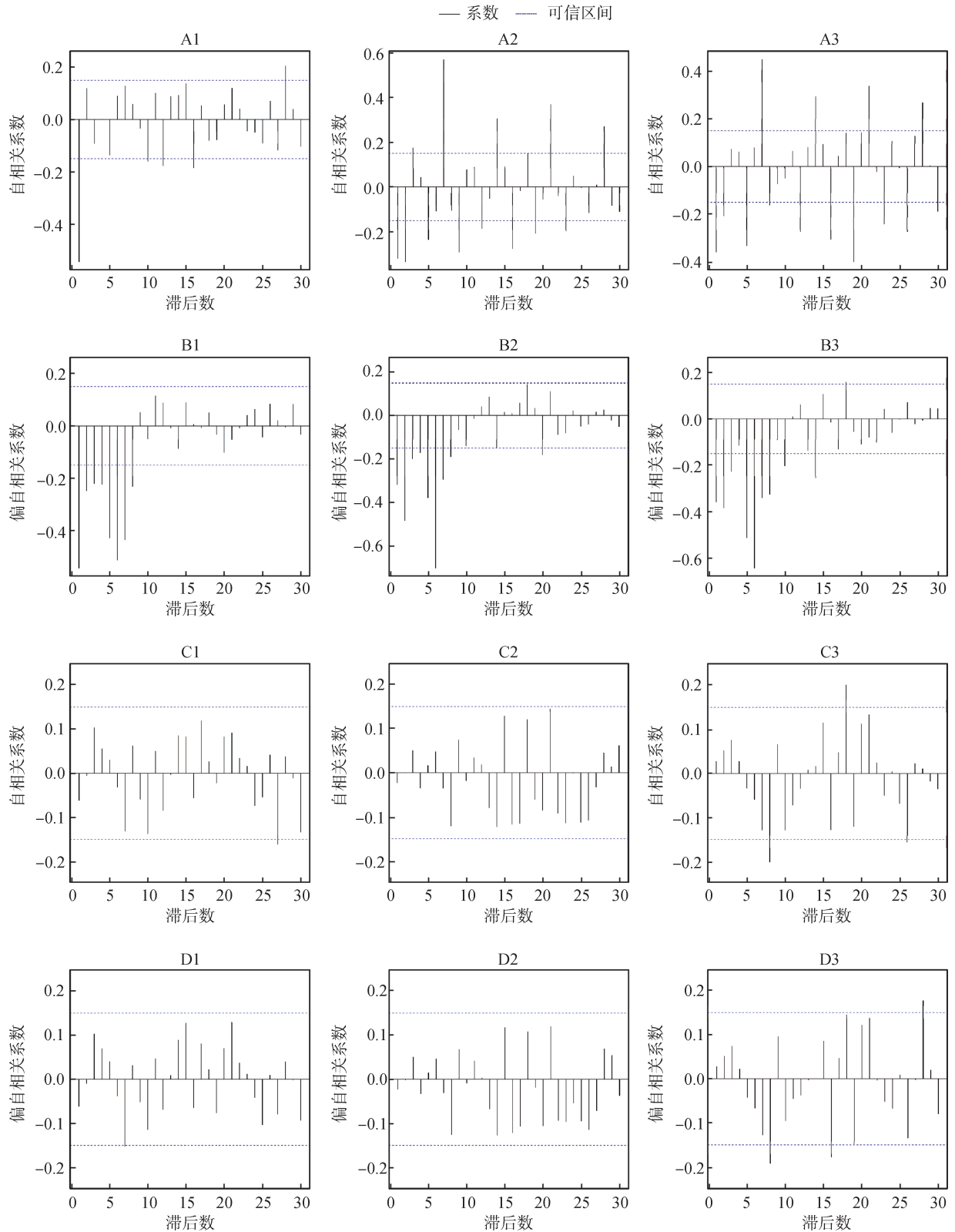
2. RNN 模型: 比较 3 个国家不同 RNN 模型在测试集上的表现, 结果见表 3。最终确定 RNN8、RNN1 和 RNN6 分别为美国、印度和巴西的最优 RNN 模型, 其在测试集上的 MAPE 值较小。考虑在 3 次训练中最小的测试集 MAPE 值, 因此, 选择第二次训练后的 RNN8、第三次训练后的 RNN1 和 RNN6 分别预测美国、印度和巴西 9 月 21-30 日的确诊病例数, 结果见表 2。

3. 模型预测性能比较: 使用两种模型分别预测 3 个国家 9 月 21-30 日共计 10 d 的病例数, 发现模型的预测性能在不同国家存在差异。在美国, ARIMA 模型有 6 d 优于 RNN 模型; 在印度, RNN 模型有 6 d 优于 ARIMA 模型; 在巴西, ARIMA 模型有 8 d 优于 RNN 模型。随后, 使用 MAPE 和 RMSE 两个指标比较两个模型在预测集上的总体表现(表 4)。结果显示, ARIMA 模型在美国和巴西的 MAPE 和 RMSE 值较小, 而 RNN 模型在印度的 MAPE 和 RMSE 值较小。因此, 总体而言, ARIMA 模型在美国和巴西的表现要优于 RNN 模型, 而 RNN 模型在印度的表现要优于 ARIMA 模型。

讨 论

构建合适的预测模型能够为政府合理配置医疗卫生资源和制定疾病防控策略提供参考。COVID-19 疫情发生后, 已有研究者采用时间序列模型来预测新发病例的变化趋势^[2, 11-12], 但是, 较少有研究比较不同模型在不同地区的预测效果。本研究构建了两种不同的时间序列模型, 并在全球累计确诊病例数最多的 3 个国家(美国、印度和巴西)进行评估, 结果发现, ARIMA 模型预测的准确性在美国和巴西较高, 而 RNN 模型预测的准确性在印度较高。

ARIMA 模型具有构建相对简单、对数据分布类型没有特殊要求的特点, 是使用最为广泛的时间序列模型。例如, Liu 等^[13]使用 ARIMA(0, 3, 1) 模型来预测 2009-2011 年中国肾综合征出血热的发病率, Moosazadeh 等^[14]构建 SARIMA(0, 1, 1)(0, 1, 1)₁₂ 模型预测伊朗 2015 年涂阳结核病的发病率。COVID-19 疫情发生以来, 一些学者使用传染病动力学模型来预测其流行趋势^[15-17]。然而, 这些动力学模型的构建过程比较复杂, 需要考虑较多的因素, 使用者需要具备较深的专业知识。因此, 一些



注:A:平稳序列的自相关系数图;B:平稳序列的偏自相关系数图;C:ARIMA模型残差的自相关系数图;D:ARIMA模型残差的偏自相关系数图;1:美国;2:印度;3:巴西

图2 自相关系数图和偏自相关系数图

表 1 美国、印度和巴西的备选 ARIMA 模型

国家	备选模型	BIC 值	系数均显著
美国	ARIMA (1,2,1)	3 475.10	否
	ARIMA (2,2,1)	3 479.41	否
	ARIMA (3,2,1)	3 473.19	否
	ARIMA (4,2,1)	3 461.30	是
	ARIMA (5,2,1)	3 428.73	是
	ARIMA (6,2,1)	3 409.66	是
	ARIMA (7,2,1)	3 406.67	是
	ARIMA (8,2,1)	3 417.96	否
印度	ARIMA (1,2,1)	3 250.78	否
	ARIMA (2,2,1)	3 231.62	否
	ARIMA (3,2,1)	3 235.52	否
	ARIMA (4,2,1)	3 225.27	否
	ARIMA (5,2,1)	3 192.53	是
	ARIMA (6,2,1)	3 132.30	是
	ARIMA (7,2,1)	3 137.04	否
	ARIMA (8,2,1)	3 141.67	否
	ARIMA (1,2,2)	3 239.87	否
	ARIMA (2,2,2)	3 214.24	是
	ARIMA (3,2,2)	3 234.40	否
	ARIMA (4,2,2)	3 235.09	否
	ARIMA (5,2,2)	3 176.01	否
	ARIMA (6,2,2)	3 136.96	否
	ARIMA (7,2,2)	3 142.09	否
	ARIMA (8,2,2)	3 144.88	否
巴西	ARIMA (1,2,1)	3 683.94	否
	ARIMA (2,2,1)	3 673.82	否
	ARIMA (3,2,1)	3 673.03	否
	ARIMA (4,2,1)	3 659.22	否
	ARIMA (5,2,1)	3 582.65	是
	ARIMA (6,2,1)	3 553.35	是
	ARIMA (7,2,1)	3 557.09	否
	ARIMA (8,2,1)	3 561.82	否

注: BIC: 贝叶斯信息准则

研究者开始尝试使用 ARIMA 模型来预测 COVID-19 的流行情况^[2,11-12]。本研究确定 ARIMA

(6,2,1)为美国、印度和巴西的最优 ARIMA 模型,该模型用于预测 3 个国家 9 月 21-30 日确诊病例数时的 MAPE 均<20%,具有较高的准确性。

尽管 ARIMA 模型在预测 COVID-19 时具有良好的表现,但是,考虑到 COVID-19 的发生往往受到多种因素的综合作用,如医疗卫生资源的配给、防控措施的实施和环境变化等,其发病趋势可能呈现出非线性特征。而 RNN 模型则可以弥补 ARIMA 模型只能处理线性特征的不足,该模型除了具有普通神经网络强大的非线性拟合能力外,还拥有记忆功能,不仅使用当前数据作为输入信息,还可利用其长期经验^[18]。由于训练集的构成会影响神经网络模型的预测性能,因此本研究分别使用前 1 天至前 10 天的数据作为训练集的输入,使用当日的数据作为训练集的输出,共构建了 10 个不同的 RNN 模型。考虑到 RNN 模型的构建需要人工设置一些参数,如学习率、隐藏层的维度和训练步数等,且在训练时其初始权重和阈值是随机分配的,即使训练集相同,其测试集的输出也不会完全相同,所以对具有不同参数组合的 RNN 模型进行了 3 次训练,通过比较其在测试集上的表现确定最佳的 RNN 模型。

通过比较研究,ARIMA 模型的预测性能在美国和巴西要优于 RNN 模型,而 RNN 模型的预测性能在印度要优于 ARIMA 模型。这种差异可能归因于两个原因,一是不同地区的发病趋势可能存在差异,二是两种模型建立时未考虑干预措施、诊断手段和报告质量等因素的影响。因此,在预测 COVID-19 疫情时,需针对不同地区的特点应选择不同的最优模型。

本研究存在局限性。首先,我们仅评估了两

表 2 ARIMA 和 RNN 模型预测美国、印度和巴西 9 月 21-30 日 COVID-19 确诊病例的结果

日期	美国			印度			巴西		
	实际值	ARIMA	RNN	实际值	ARIMA	RNN	实际值	ARIMA	RNN
9月21日	41 695	44 217	44 071	86 961	89 808	83 791	33 057	30 728	31 266
9月22日	36 766	41 177	39 732	75 083	84 666	81 510	16 389	16 338	18 833
9月23日	39 145	39 173	35 772	83 347	89 342	75 004	13 439	19 539	23 600
9月24日	49 176	40 544	39 195	86 508	95 945	79 798	33 536	31 225	26 952
9月25日	40 043	44 208	48 820	86 052	94 653	81 307	33 281	33 855	40 821
9月26日	41 254	48 007	45 227	85 362	91 820	81 099	32 817	35 635	37 228
9月27日	50 070	50 374	43 497	88 600	90 843	80 779	31 911	36 845	38 568
9月28日	48 958	48 527	45 384	82 170	88 102	82 219	28 378	28 224	27 246
9月29日	35 217	46 142	46 379	70 589	84 848	79 193	14 318	18 449	17 390
9月30日	32 688	44 734	41 302	80 472	88 658	71 808	13 155	20 817	21 341

表3 美国、印度和巴西的备选RNN模型

国家	模型	学习率	隐藏层的维度	训练步数	MAPE ^a (%)	MAPE ^b (%)	MAPE ^c (%)
美国	RNN1	0.05	10	200	15.14	15.09	15.12
	RNN2	0.2	3	100	17.58	17.65	17.62
	RNN3	0.1	3	200	18.18	18.96	18.53
	RNN4	0.1	3	250	18.14	17.54	17.88
	RNN5	0.1	5	200	14.96	15.17	15.04
	RNN6	0.2	3	120	12.19	12.48	12.07
	RNN7	0.1	3	300	12.14	12.35	12.12
	RNN8	0.1	5	400	11.68	11.64	11.80
	RNN9	0.05	5	700	12.15	11.96	12.45
	RNN10	0.2	5	150	13.47	13.91	13.46
印度	RNN1	0.05	10	250	10.09	10.09	9.79
	RNN2	0.1	10	150	11.69	10.92	11.20
	RNN3	0.05	10	300	12.26	12.63	12.42
	RNN4	0.05	10	300	12.49	12.26	12.42
	RNN5	0.05	10	300	12.97	12.80	13.05
	RNN6	0.2	10	80	13.20	12.95	12.97
	RNN7	0.1	10	150	13.93	13.59	13.75
	RNN8	0.2	10	100	13.71	13.71	13.43
	RNN9	0.1	10	250	15.71	15.17	15.53
	RNN10	0.05	10	400	15.75	15.32	15.85
巴西	RNN1	0.2	10	100	38.10	38.46	38.03
	RNN2	0.05	10	800	37.82	38.23	38.31
	RNN3	0.05	10	1 000	45.28	44.63	45.13
	RNN4	0.05	3	1 500	46.27	46.27	45.82
	RNN5	0.05	3	1 500	44.36	44.10	44.44
	RNN6	0.2	3	600	36.24	36.36	35.93
	RNN7	0.05	10	800	43.79	43.82	43.94
	RNN8	0.05	3	1 000	45.57	45.34	45.16
	RNN9	0.2	5	300	44.04	44.01	43.71
	RNN10	0.1	10	500	43.17	43.60	43.55

注:MAPE:平均绝对百分比误差;^a第一次训练;^b第二次训练;^c第三次训练

表4 ARIMA和RNN模型在3个国家预测性能的比较

国家	评价指标	模型	
		ARIMA	RNN
美国	MAPE (%)	13.18	15.27
	RMSE	6 542.32	6 877.71
印度	MAPE (%)	9.18	7.23
	RMSE	8 069.50	6 457.07
巴西	MAPE (%)	17.30	26.02
	RMSE	3 954.59	5 950.88

注:MAPE:平均绝对百分比误差;RMSE:均方根误差

种模型预测 COVID-19 的表现,其他种类的时间序列模型,如灰色模型和指数平滑模型等未被考虑。其次,无论是 ARIMA 模型还是 RNN 模型,本质上都属于短期预测模型,无法预测其长期趋势,因此需要不断地收集和更新数据以保证其预测性能。第三,本研究使用的 COVID-19 发病数据来源于 WHO 官方公布的结果,与各国实际的发病数据可能存在一定的差异。第四,相比传染病动力学模型等其他预测模型,时间序列模型无法体现干预措施、诊断手段和报告质量等因素对疫情的影响。

最后,COVID-19 具有冬春季高发的特点,本研究仅收集了 2020 年 4 月 1 日至 9 月 30 日的病例数据,无法体现其病例分布的周期性和季节性。

利益冲突 所有作者均声明不存在利益冲突

参 考 文 献

- [1] World Health Organization. WHO Coronavirus Disease (COVID-19) Dashboard[EB/OL]. [2020-11-09]. <https://covid19.who.int/>.
- [2] Ceylan Z. Estimation of COVID-19 prevalence in Italy, Spain, and France[J]. *Sci Total Environ*, 2020, 729: 138817. DOI:10.1016/j.scitotenv.2020.138817.
- [3] Li ZQ, Wang ZZ, Song H, et al. Application of a hybrid model in predicting the incidence of tuberculosis in a Chinese population[J]. *Infect Drug Resist*, 2019, 12: 1011-1020. DOI:10.2147/idr.s190418.
- [4] Liu WD, Bao CJ, Zhou YP, et al. Forecasting incidence of hand, foot and mouth disease using BP neural networks in Jiangsu province, China[J]. *BMC Infect Dis*, 2019, 19: 828. DOI:10.1186/s12879-019-4457-6.
- [5] Zheng YL, Zhang LP, Zhu XX, et al. A comparative study of two methods to predict the incidence of hepatitis B in Guangxi, China[J]. *PLoS One*, 2020, 15(6): e0234660. DOI: 10.1371/journal.pone.0234660.
- [6] Liu Q, Li ZQ, Ji Y, et al. Forecasting the seasonality and trend of pulmonary tuberculosis in Jiangsu Province of China using advanced statistical time-series analyses[J]. *Infect Drug Resist*, 2019, 12: 2311-2322. DOI: 10.2147/idr.s207809.
- [7] Chen YY, Cheng QQ, Cheng Y, et al. Applications of recurrent neural networks in environmental factor forecasting: a review[J]. *Neural Comput*, 2018, 30(11): 2855-2881. DOI:10.1162/neco_a_01134.
- [8] Kim ZM, Oh H, Kim HG, et al. Modeling long-term human activeness using recurrent neural networks for biometric data[J]. *BMC Med Inform Decis Mak*, 2017, 17 Suppl 1: 57. DOI:10.1186/s12911-017-0453-1.
- [9] Wang YW, Shen ZZ, Jiang Y. Comparison of ARIMA and GM (1, 1) models for prediction of hepatitis B in China[J]. *PLoS One*, 2018, 13(9): e0201987. DOI: 10.1371/journal.pone.0201987.
- [10] Akinbobola A, Omotosho JB. Predicting malaria occurrence in southwest and north central Nigeria using meteorological parameters[J]. *Int J Biometeorol*, 2013, 57(5): 721-728. DOI:10.1007/s00484-012-0599-6.
- [11] Ilie OD, Cojocariu RO, Ciobica A, et al. Forecasting the spreading of COVID-19 across nine countries from Europe, Asia, and the American continents using the ARIMA models[J]. *Microorganisms*, 2020, 8(8): 1158. DOI: 10.3390/microorganisms8081158.
- [12] Malki Z, Atlam ES, Ewis A, et al. ARIMA models for predicting the end of COVID-19 pandemic and the risk of second rebound[J/OL]. *Neural Comput Appl*, 2020. DOI: 10.1007/s00521-020-05434-0.
- [13] Liu QY, Liu XD, Jiang BF, et al. Forecasting incidence of hemorrhagic fever with renal syndrome in China using ARIMA model[J]. *BMC Infect Dis*, 2011, 11: 218. DOI: 10.1186/1471-2334-11-218.
- [14] Moosazadeh M, Khanjani N, Nasehi M, et al. Predicting the incidence of smear positive tuberculosis cases in Iran using time series analysis[J]. *Iran J Public Health*, 2015, 44(11): 1526-1534.
- [15] Wang K, Ding L, Yan Y, et al. Modelling the initial epidemic trends of COVID-19 in Italy, Spain, Germany, and France[J]. *PLoS One*, 2020, 15(11): e0241743. DOI: 10.1371/journal.pone.0241743.
- [16] IHME COVID-19 Forecasting Team. Modeling COVID-19 scenarios for the United States[J/OL]. *Nat Med*, 2020. DOI: 10.1038/s41591-020-1132-9.
- [17] 刘雅姝, 吴琪俊, 陆一涵, 等. 新型冠状病毒肺炎 (COVID-19) 传染病预测模型分析[J]. *公共卫生与预防医学*, 2020, 31(3): 10-13. DOI:10.3969/j.issn.1006-2483.2020.03.003.
- [17] Liu YS, Wu QJ, Lu YH, et al. Analysis of models of infectious disease dynamics of COVID-19[J]. *J Public Health Prev Med*, 2020, 31(3): 10-13. DOI: 10.3969/j.issn.1006-2483.2020.03.003.
- [18] Lecun Y, Bengio Y, Hinton G. Deep learning[J]. *Nature*, 2015, 521(7553): 436-444. DOI:10.1038/nature14539.