

考虑人群流动数据的 COVID-19 传播模型

邓春燕, 郭强, 傅家旗

(上海理工大学 管理学院, 上海 200093)

摘要: 新冠疫情随着人群的流动和互相接触在各地快速传播和扩散。通过提出一种考虑地区间人群流动和地区内人群流动状态的疫情传播模型(简称 D-SEIR 模型), 综合考虑人群的流动情况、易感人群和潜伏人群的数量变化情况, 从而预测确诊人群的变化。经过深圳市和山东省的实证数据实验结果表明, D-SEIR 模型能够准确拟合 COVID-19 传播过程中的易感人群和确诊人群的演化趋势。疫情未稳定前, 模型对深圳市易感人群和确诊人群变化趋势的拟合优度 r^2 可达 0.9661, 0.8474, 山东省的拟合优度 r^2 可达 0.8576, 0.8795。疫情稳定后, 模型对深圳市易感人群和确诊人群变化趋势的拟合优度 r^2 可达 0.9720, 0.8180, 山东省的拟合优度 r^2 可达 0.8615, 0.9122。模型有助于发现 COVID-19 在人群流动下的传播规律和特性, 为复工复产背景下不同城市和地区的风险预警提供决策依据。

关键词: COVID-19; D-SEIR 模型; 人群流动

中图分类号: TP 391 **文献标志码:** A

COVID-19 propagation model considering crowd flow data

DENG Chunyan, GUO Qiang, FU Jiaqi

(Business School, University of Shanghai for Science and Technology, Shanghai 200093, China)

Abstract: The new crown epidemic spreads and spreads rapidly in various regions along with the movement of people and mutual contact. An epidemic spread model (D-SEIR model for short) that considers the population flow among regions and within regions was proposed. Through comprehensive consideration of the flow of people, and the changes in the number of susceptible and latent people, the changes in the confirmed population can be predicted. The experimental results of empirical data in Shenzhen and Shandong show that the proposed D-SEIR model can accurately fit the evolutionary trends of susceptible and diagnosed populations during the transmission of COVID-19. Prior to the change in the epidemic situation, the goodness-of-fit r^2 of the change trend of the susceptible and the confirmed population in Shenzhen reaches 0.9661 and 0.8474, The fitness r^2 of Shandong province reaches 0.8576 and 0.8795. With the stabilization of the epidemic situation, the goodness-of-fit r^2 of Shenzhen reaches 0.9720 and 0.8180. The fitness r^2 of Shandong province reaches 0.8615 and 0.9122.

收稿日期: 2020-06-09

基金项目: 国家自然科学基金资助项目 (71771152, 617773248); 国家社科基金重大项目 (18ZDA088)

第一作者: 邓春燕 (1997-), 女, 硕士研究生. 研究方向: 数据挖掘. E-mail: dcy02130213@163.com

通信作者: 郭强 (1975-), 女, 教授. 研究方向: 网络科学、商务智能、科学知识图谱.

E-mail: qiang.guo@usst.edu.cn

The model of this paper is helpful to understand the propagation rules and characteristics of COVID-19 under the flow of people, and provides decision-making basis for early risk warning of different cities and regions under the background of resumption of working and production.

Keywords: COVID-19; D-SEIR model; crowd movement

对流行性传染病的实证数据进行建模分析，有助于深入研究传染病的传播规律，发现传染病的传播弱点和趋势，为采取有效的疫情防控手段、保持疫情防控成果提供重要科学支持^[1-3]。新型冠状病毒肺炎 (COVID-19) 作为新型流行性传染病，由于大家缺乏对其传播规律和特性的了解，2020 年初在各地迅速蔓延。根据国家卫生健康委员会官方网站公布的数据，全国累计确诊人数在 2020 年 2 月 17 日达到最高峰，随后开始下降，疫情得到转变和稳定^[4]。国家采取的区域封闭式管理和治疗措施在疫情管控上卓有成效，目前疫情也已经稳定，复工复产皆有序进行^[5]。根据疫情前期爆发、中期管控下降、后期稳定的特点，全面研究 COVID-19 的传播特性、传播弱点，为实施精准防控、防止少数潜伏或者确诊病例有可能带来的疫情反弹、保持疫情防控成果、实施疫情防控工作提供科学依据。

在 COVID-19 突然爆发期间，关于 COVID-19 传播的研究也有很多，为国家管理疫情提供了科学支持。除了预测 COVID-19 持续时间^[5-7]、估算高峰时期感染人数^[8-9]、分析隔离对疫情管控的作用^[10-11]等，研究工作还包括基于实时的疫情感染数据，利用经典传染病 SI 和 SEIR 等模型来分析 COVID-19 的传播机制^[12-13]。比如：利用经典传染病模型中的病例指数增长假设分析人群流动对 COVID-19 传播的影响^[14-15]；基于 SI 模型构建复杂人流网络的 COVID-19 传播模型，在未考虑传染病潜伏过程下分析不同省市的疫情传播特点^[16]；利用疫情的动态变化数据建立 SEIR 模型分析区域隔离的有效性^[17]；通过构建分数阶 SEIR 模型描述疫情的动态传播，利用微分方程残差幂级数求解来描绘感染过程，但是忽略区域人数流动总数的变化^[18]；基于 SEIR 模型建立考虑公共交通工具内部因素的疫情传播模型，分不同站点阶段刻画疫情扩散机制^[18]；考虑离散时间内输入患者对特定地区的影响和风险改进 SEIQR 模型^[20]；还有以传染病模型验证感染患者在发病症状不明显

时期和未隔离时的扩散传染率^[21]。这些研究成果都让人们对于 COVID-19 有了更多的认识，了解到 COVID-19 的传播机制和隔离的重要性，为国家实施疫情管控工作提供了帮助。

通过已有的研究可以发现 SI 或者 SEIR 模型都能有效分析 COVID-19 的传播过程，但这些工作并未全面考虑疫情由严重到稳定这一过程中 COVID-19 传播效力的不同，人群流动和疫情稳定时的管控会使 COVID-19 的传播规律发生变化。因此，本文尝试根据我国由严控至逐步复工复产 (2020 年 1 月 19 日—2020 年 3 月 6 日) 期间不同的 COVID-19 传播程度，构建考虑人群流动情况下，分析易感人群和潜伏人群的变化对确诊人群影响的疫情传播 D-SEIR 模型，并以实证数据进行验证拟合，利用传播模型更加准确分析 COVID-19 的传播规律。本文所有研究数据均来自于 2020 年 1 月 19 日—2020 年 3 月 6 日这一时间段。

1 COVID-19 传播模型

经典 SEIR 模型常适用于具有潜伏期的传染病传播研究，通过把人群划分为易感人群、潜伏人群、确诊人群、治愈人群来构建传播微分方程，并假设整体人数不变来求解微分方程中的传染率参数，再代入传播微分方程以此刻画传染病的传播过程^[22-23]。其中：易感人群是指未感染传染病的健康人群；潜伏人群是指已经被传染病感染但未有发病症状的人群；确诊人群是指具有传染病发病症状并诊疗确诊的人群；治愈人群则是指经过隔离治疗治愈出院的患者人群。

由于 COVID-19 具有一定潜伏期^[24]，本文同样将 COVID-19 传播中的人群分为：易感人群 (S)、潜伏人群 (E)、确诊人群 (I)、治愈人群 (R)。再构建考虑地区间人口流动和地区内部人口流动情况的疫情传播模型 (简称 D-SEIR 模型)，综合考虑人口的流动数据、易感人群和潜伏人群的数量变化来预测染病人群的变化。构建模型的过程中为使

模型更加有效, 假设流动人群皆为易感人群和潜伏人群, 确诊人群已经被隔离收治。另外由于疫情在由严重转至稳定期间的传播规律不同, 而2月17日全国确诊人数出现下降趋势, 疫情开始稳定, 本文将构建的 D-SEIR 模型分为两部分: 第一部分是对疫情初始至稳定前的 COVID-19 传播模型进行构建; 第二部分是对构建的 COVID-19 传播模型基于疫情稳定后的情况进行调整, 让 D-SEIR 模型更加符合真实情况。

1.1 考虑人群流动的 D-SEIR 模型

基于国家进行区域隔离的政策和疫情处于春节特殊时期的特点, 本文将各地区构建为一个半封闭系统。半封闭系统是指该地区在特殊时期内没有人员流出, 只有人员流入。因为疫情恰逢春节, 有很多出门旅游和在外长期务工的人员回到他们所在家乡, 各地区就有大量的人员流入, 而后疫情爆发, 政府采取隔离措施限制各地区人员的流出, 所以将行政区域构建为一个半封闭系统。而这部分回到自身所属地区的人群可能是易感状态也可能是潜伏状态, 他们对自己的感染情况并不知晓。基于此绘制出各地区人群流动的 D-SEIR 模型示意图, 如图 1 所示。图中: 下标 q 表示本地区, p 表示其他地区; S_p , E_p 分别表示从其他地区 p 流入的易感人群和潜伏人群; λ 表示不同人群之间转化的调节参数; ε 表示不同人群之间转化的误差常数。

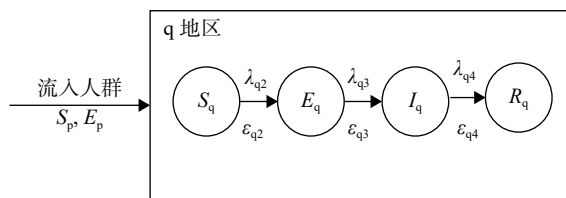


图 1 D-SEIR 模型示意图

Fig.1 Schematic diagram of D-SEIR model

从图 1 可以看出, 易感人群 S_p 和潜伏人群 E_p 流入到 q 地区, 那么 q 地区原本的易感人群和潜伏人群数量会增多, 潜伏人群增多后转化为确诊人群, 确诊人群的数量也会发生变化。基于此, 定义 q 地区 t 时刻的易感人群数为 $S_q(t)$, 潜伏人群数为 $E_q(t)$, 确诊人群数为 $I_q(t)$, 治愈人群数为 $R_q(t)$, 流入 q 地区的易感人群数为 $S_p(t)$, 潜伏人群数为 $E_p(t)$, 构建考虑人口流动的 D-SEIR 模型, 模型动力学方程如式(1)所示。

$$\begin{cases} S_q(t) = S_p(t-1) + [S_q(t-1) - \Delta E_q(t)] + \Delta S_q(t) \\ E_q(t) = E_p(t-1) + [E_q(t-1) - \Delta I_q(t)] + \Delta E_q(t) \\ I_q(t) = I_q(t-1) - \Delta R_q(t) + \Delta I_q(t) \\ R_q(t) = R_q(t-1) + \Delta R_q(t) \end{cases} \quad (1)$$

式中: $\Delta S_q(t)$ 为 t 时刻易感人群数量相较于 $t-1$ 时刻的变化量; $\Delta E_q(t)$ 为 t 时刻潜伏人群数量相较于 $t-1$ 时刻的变化量; $\Delta I_q(t)$ 为 t 时刻确诊人群数量相较于 $t-1$ 时刻的变化量; $\Delta R_q(t)$ 则为 t 时刻治愈人群数量相较于 $t-1$ 时刻的变化量。因为各地区的流入人群只有易感人群 S_p 和潜伏人群 E_p , 所以流入人群的 $S_p(t)$ 和 $E_p(t)$ 关联到模型 (1) 中的 $S_q(t)$, $E_q(t)$ 部分。

由于 COVID-19 在疫情期间不断传播扩散, 通过把病毒传染给易感人群, 易感人群的数量减少并转化为潜伏人群, 影响潜伏人群时刻变化量 $\Delta E_q(t)$ 。然后, 潜伏人群一旦有发病症状, 经过确诊变为确诊人群, 影响确诊人群时刻变化量 $\Delta I_q(t)$, 而确诊人群经过隔离治疗会转变为治愈人群, 影响治愈人群时刻变化量 $\Delta R_q(t)$ 。因此, 为准确反映各地区每天疫情的变化情况, t 时刻易感人群 $\Delta S_q(t)$ 主要由前一时刻的易感人群 $S_q(t-1)$ 决定, 潜伏人群 $\Delta E_q(t)$ 主要由前一时刻的易感人群 $S_q(t-1)$ 决定, 确诊人群 $\Delta I_q(t)$ 主要由前一时刻的潜伏人群 $E_q(t-1)$ 决定, 治愈人群 $\Delta R_q(t)$ 由前一时刻的确诊人群 $I_q(t-1)$ 决定。据此, 本文根据式(2)来确定易感人群的时刻变化量 $\Delta S_q(t)$ 、潜伏人群的时刻变化量 $\Delta E_q(t)$ 、确诊人群的时刻变化量 $\Delta I_q(t)$ 和治愈人群的时刻变化量 $\Delta R_q(t)$ 。

$$\begin{cases} \Delta S_q(t) = \lambda_{q1} S_q(t-1) + \varepsilon_{q1} \\ \Delta E_q(t) = \lambda_{q21} S_q(t-1)^2 + \lambda_{q22} S_q(t-1) + \varepsilon_{q2} \\ \Delta I_q(t) = \lambda_{q3} E_q(t-1) + \varepsilon_{q3} \\ \Delta R_q(t) = \lambda_{q4} I_q(t-1) + \varepsilon_{q4} \end{cases} \quad (2)$$

式中: λ_q 为调节常数; ε_q 为误差常数。

另外在模型进行实证数据拟合前, 先把 q 地区的病例患者进行时刻状态分类统计。通过一名患者的发病时间和治愈时间, 判断具体时刻患者的身体状态。在发病时间之前包括易感状态和潜伏状态, 在发病时间和治愈时间之内为感染状态, 治愈后则为治愈状态。由于国家卫健委发布的新冠病毒防控中宣布新冠病毒的潜伏期为 1~14 d, 通常为 3~7 d^[4], 为了让模型对疫情中 COVID-19 潜伏扩散的情况预估复现更为准确, 在分析过程中包含大部分患者的真实染病潜伏情况。本文以

发病时间的前 7 d 内为界限判定患者的潜伏状态,发病前 7 d 外为易感状态。从而统计出 q 地区每一时刻的易感人群、潜伏人群、确诊人群、治愈人群数量,再计算 q 地区每时刻处于不同状态的总人群数量 $S_q(t)$, $E_q(t)$, $I_q(t)$, $R_q(t)$, 以及不同状态下的人群数量时刻变化值 $\Delta S_q(t)$, $\Delta E_q(t)$, $\Delta I_q(t)$, $\Delta R_q(t)$ 。基于上述数据分类统计,分别代入式(2)进行拟合求解,将拟合参数结果代入便可求出模型的拟合变化值。以拟合求出的各时刻人群变化量为结果,各时刻人群变化量的真实数据为标准,作图对比 2 月 17 日前的 COVID-19 的传播感染过程。

1.2 疫情稳定后的 D-SEIR 模型调整

由于疫情确诊人数在 2 月 17 日后开始下降,疫情开始稳定,稳定后的感染人数传播效力、健康人员管控力度、治愈效果和前期相比都发生了变化。构建的初始 D-SEIR 模型是以疫情严重时段的 COVID-19 为基础建立的,用此模型求解结果直接预测稳定后的感染情况会有偏差。因此,将 D-SEIR 模型中加入调节参数,使模型对疫情后期的预测拟合更加符合实际情况。

在加入调节参数过程中,考虑到疫情后期的易感人群、潜伏人群、确诊人群、治愈人群变化趋势逐渐稳定,感染人数减少,治愈人数增多,疫情整体呈现好转趋势,将调节参数 γ 针对易感人群、潜伏人群、确诊人群、治愈人群的不同时刻总数进行调整,通过对易感人群、潜伏人群、确诊人群、治愈人群的二次影响反映疫情稳定后的 $\Delta S_q(t)$, $\Delta E_q(t)$, $\Delta I_q(t)$ 和 $\Delta R_q(t)$ 变化趋势。结合式(2)对 D-SEIR 模型进行参数修正,将参数修正指标 γ_q 依次加入式(2)中,修正后的 D-SEIR 预测模型如式(3)所示。

$$\begin{cases} \Delta S_q(t) = (\lambda_{q1} + \gamma_{q1})S_q(t-1) + \varepsilon_{q1} \\ \Delta E_q(t) = (\lambda_{q21} + \gamma_{q2})S_q(t-1)^2 + (\lambda_{q22} + \gamma_{q2})S_q(t-1) + \varepsilon_{q2} \\ \Delta I_q(t) = (\lambda_{q3} + \gamma_{q3})E_q(t-1) + \varepsilon_{q3} \\ \Delta R_q(t) = (\lambda_{q4} + \gamma_{q4})I_q(t-1) + \varepsilon_{q4} \end{cases} \quad (3)$$

根据调整后的 D-SEIR 模型式(3),便可计算出疫情严控至复工复产后的确诊人群变化情况 $\Delta I_q(t)$ 和易感人群变化情况 $\Delta S_q(t)$, 以此分析模型调整后对于各地区疫情描绘的结果,并将预测趋势与实证数据的变化趋势进行对比,分析 COVID-19 传染病在疫情后期的传播特性。

2 D-SEIR 模型评价标准

为了判断 D-SEIR 模型对新冠肺炎 S , E , I , R 人群时刻变化量的拟合准确度,本文以拟合优度 r^2 为评价标准。 r^2 经常被用来分析线性回归模型拟合结果的好坏^[25],取值在 0~1 之间,值越靠近于 1 表明模型对目标变量的刻画能力越强,模型拟合效果也越好。因此,本文以拟合优度 r^2 来判断 D-SEIR 模型对 S , E , I , R 人群时刻变化量的刻画程度,计算公式如式(4)所示。

$$r^2 = 1 - \frac{\sum (y - \hat{y})^2}{\sum (y - \bar{y})^2} \quad (4)$$

式中: y 为实证数据不同人群的时刻变化量 $\Delta S_q(t)$, $\Delta E_q(t)$, $\Delta I_q(t)$, $\Delta R_q(t)$; $\bar{y}$ 为对实证数据不同人群时刻变化量均值的计算; $\hat{y}$ 则是通过式(2)拟合出的不同人群的时刻变化量。

3 实证数据分析

3.1 数据简介

考虑到地区感染情况的代表性和疫情期间人群流动的特点,首先以市级地区为例,本文选取人群流动多、传播情况复杂的一线城市深圳市为研究对象;然后以省份地区为例,选取沿海省份疫情感染程度在全国居中的山东省为研究对象。最终,本文以这两个地区的 COVID-19 感染情况进行实验分析。其中,深圳市 419 名病例,山东省 558 名病例,数据包括每日病例报告时间、发病时间、发病原因、居住地、活动区域和治愈时间,数据统计时段为省市严控至复工复产的 2020 年 1 月 19 日—2020 年 3 月 6 日。然后通过病例的居住地和活动区域,区分感染患者的地理区域,整理成省市流入人群和省市本地人群两大类,代入 D-SEIR 模型中。经过统计,详细数据如表 1 所示。

其中,流入人群感染数是居住于外省市或者于外省市旅游,并在 2020 年 1 月 19 日—2020 年

表 1 实验统计数据
Tab.1 Design parameters

区域	统计时间	流入人群感染数	本地人群感染数
深圳	2020-01-19—2020-03-06	161	242
山东	2020-01-21—2020-03-01	266	292

3月6日来到统计地区的确诊人群总数。本地人群感染数是本地居住人群, 并且未离开过本地区的被感染人群总数, 本地人群主要通过与其他潜伏或者确诊人群接触而患病。

然后进行时刻状态人群分类统计, 根据发病时间、诊疗出院时间、潜伏时间7 d来判断病例的时刻归属状态, 即易感人群、潜伏人群、确诊人群和治愈人群中的一类, 从而统计每一时刻的 $S_q(t)$, $E_q(t)$, $I_q(t)$, $R_q(t)$ 数。同时, 为了防止真实数据在以 1 d 为单位统计时可能出现延迟, 减少模型对 48 d 不同人群拟合带来的误差, 本文实验确定时刻单位为 2 d。基于疫情由严重至稳定的这一过程, 将 2 月 17 日之前作为模型初始数据拟合, 2 月 17 日后作为模型调整数据拟合。

3.2 D-SEIR 模型拟合结果

以深圳市为例, 将统计得到的时刻 S , E , I , R 数量和时刻变化量代入式(2), 进行 COVID-19 传播的 D-SEIR 传播模型拟合。求解得到 $\lambda_{q1} = -0.249\ 2$, $\varepsilon_{q1} = -0.639\ 7$, $\lambda_{q21} = 0.004\ 723$, $\lambda_{q22} = -0.789\ 6$, $\varepsilon_{q2} = -0.33$, $\lambda_{q3} = 0.411\ 5$, $\varepsilon_{q3} = -14.05$, $\lambda_{q4} = 0.061\ 57$, $\varepsilon_{q4} = -3.525$ 。拟合结果如图 2 所示,

再根据式(3)求得拟合优度 r^2 , 在图 2 中进行标注。

从图 2 可看出 $\Delta S_q(t)$, $\Delta E_q(t)$, $\Delta I_q(t)$ 的拟合结果(红色圆圈曲线)和实际数据(黑色方块曲线)贴合很近。在图 2(a)中, 从黑色方块曲线呈现的变化趋势来看, 可以将 ΔS_q 的变化分为两个阶段。第 1 阶段是从 1 月 21 日—2 月 10 日, ΔS_q 曲线由负值开始上升直至 0, 对应于图中 2 月 10 日这个节点。因为 $\Delta S_q(t)$ 为 t 时刻易感人群相较于 $t-1$ 时刻的变化量, 此时易感人群不断被感染, $\Delta S_q(t)$ 小于 $\Delta S_q(t-1)$, 所以第 1 阶段为负值。第 2 阶段是从 2 月 10 日—2 月 17 日, ΔS_q 的值始终为 0。因为在 419 名人员中易感人群已经被全部感染, 没有健康人员的存在, 所以时刻变化量 ΔS_q 始终为 0。从红色圆圈曲线来看, 拟合结果对应了 ΔS_q 的两阶段变化趋势, 在 2 月 10 日前曲线上升, 2 月 10 日后一直为 0, 对于易感人群各时刻的变化量 $\Delta S_q(t)$ 拟合结果很准确, ΔS_q 的拟合优度 r^2 值达到 0.966 1, 与 1 很相近, 表明拟合效果优。在图 2(b)中, 黑色方块曲线的变化趋势也可分为 2 阶段: 第 1 阶段为 1 月 21 日—1 月 25 日, 曲线 ΔE_q 迅速下降,

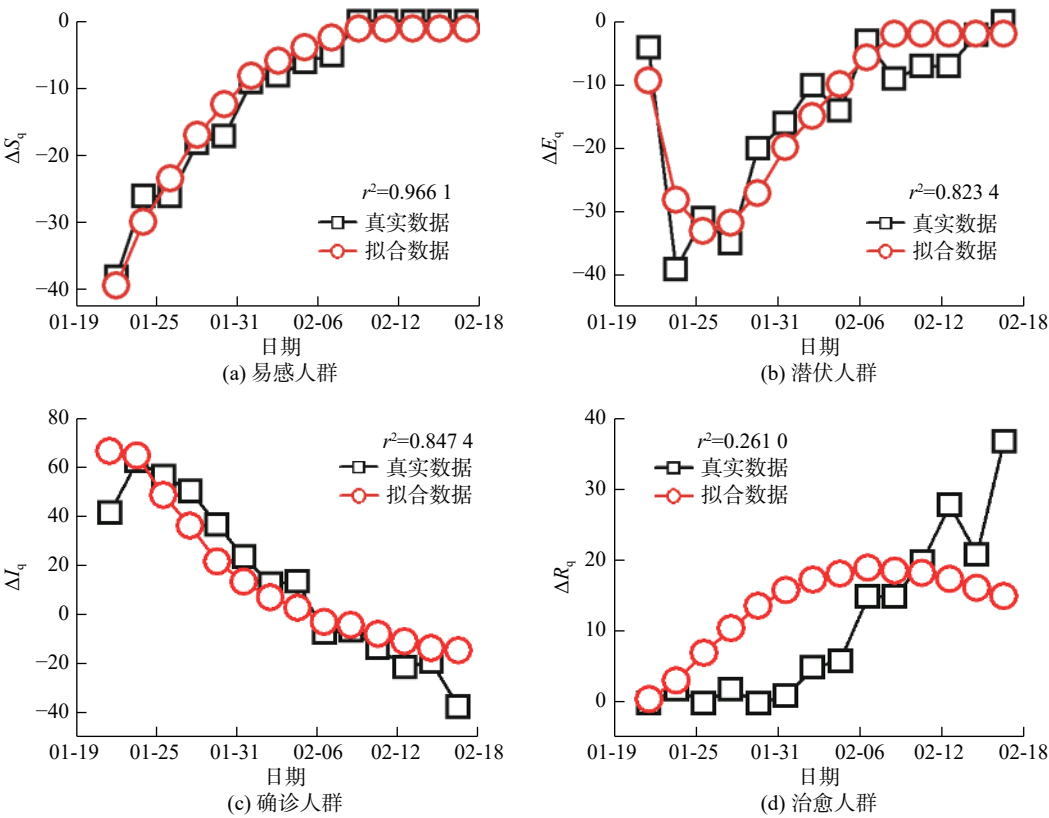


图 2 深圳市 D-SEIR 模型拟合图
Fig.2 Fitting diagram of Shenzhen D-SEIR model

并从正值变为负值，其中曲线最高节点1月21日为正值，此时疫情刚爆发，人群还在大范围流动，造成了病毒传播潜伏人群的增多，而后随着隔离，病毒传播被控制，潜伏人群也减少。第2阶段为1月25日—2月17日， ΔE_q 开始呈现上升趋势，并向0逐渐靠近，潜伏人群最终逐渐减少并最后稳定下来，但还是有少数未发现的潜伏病例在波动。红色圆圈曲线在 ΔE_q 的拟合上也很准确刻画潜伏人群变化量的走向，复现了1月25日前下降、后上升并稳定在0的趋势，拟合优度 r^2 值达到0.8234，与1相近，模型拟合效果较准确。

在图2(c)中， ΔI_q 的黑色方块曲线则一直下降，但在1月21日—2月4日之间为正值，也就是在这期间每天感染人数的变化量都为正值，说明感染人数一直在增加，这也对应疫情刚爆发时感染人数迅速增多的实际情况。而在2月4日—2月17日之间， ΔI_q 则下降为负值并且绝对值越来越大，说明疫情管控和治疗开始有效，感染人数在不断减少并且每天减少的变化量越来越大。 ΔI_q 拟合数据的红色圆圈曲线也从1月21日开始迅速下降，在2月4日后变为负值，与真实数据的变化趋势相近，拟合优度 r^2 值达到0.8474，与1也相近，拟合效果较好。从图2(d)中来看， ΔR_q 的黑色方块曲线趋势则可分为2阶段：第1阶段为1月21日—2月4日，疫情前期治疗手段还不明确，治愈人数很少，治愈人群时刻变化量 ΔR_q 都在0的周围波动。第2阶段则为2月4日—2月17日， ΔR_q 出现明显的上升趋势，治疗手段的采取都开始起效， ΔR_q 为正值并大幅度上升。对比来看， ΔR_q 的红色圆圈拟合结果曲线则是先上升后下降的趋势，与实际数据贴合的结果则不是很好，拟合优度 r^2 值为0.2610，结果与1较远。分析原因可能是治疗强度的外在因素，影响了感染患者单靠自身治愈的结果，使得 ΔR_q 不能单由 $I_q(t-1)$ 准确拟合，因此 r^2 的结果与1相差较远。

再以 COVID-19 在山东省的传播为例分析其传播规律，将详细数据处理得到的 S, E, I, R 人群时刻变化量，代入式(2)，进行 D-SEIR 传播模型拟合。求解得到 $\lambda_{q1} = -0.1781, \varepsilon_{q1} = -3.222, \lambda_{q21} = 0.001062, \lambda_{q22} = -0.2936, \varepsilon_{q2} = -4.811, \lambda_{q3} = 0.4791, \varepsilon_{q3} = -34.56, \lambda_{q4} = 0.07373, \varepsilon_{q4} = 0.444$ 。拟合结果和拟合优度 r^2 如图3所示。

由图3可看出， $\Delta S_q, \Delta E_q, \Delta I_q, \Delta R_q$ 的拟合

趋势(红色圆圈曲线)和实际数据(黑色方块曲线)贴合都很近。在图3(a)中，易感人群变化量 ΔS_q 的黑色方块曲线一直呈现上升趋势，将它分为2阶段。第1阶段为1月24日—2月9日， ΔS_q 曲线由负值开始上升直至到0，对应于图中的2月9日节点。第2阶段为2月9日—2月17日， ΔS_q 曲线在0附近小范围波动，此阶段疫情管控和治疗到位，新增被感染的易感人群也逐渐减少，停留在0附近。从 ΔS_q 的模型拟合结果(红色圆圈曲线)来看，拟合结果也刻画了易感人群变化量的时间趋势，从1月24日开始曲线稳定上升，与黑色方块曲线趋势对应，小范围波动有偏离，拟合优度 r^2 值为0.8567，与1较近，基本通过模型复现了 ΔS_q 的走势。由图3(b)可发现， ΔE_q 的黑色方块曲线的趋势并不稳定，两个数据点之间波动较大，但整体还是先下降后上升，也可分为2阶段：第1阶段为1月21日—2月3日，曲线趋势下降，从正值变为负值，1月21日为正值，潜伏人群增多后为负值，潜伏人群逐渐减少。第2阶段为2月3日—2月17日， ΔE_q 曲线呈现上升趋势并最后停留在0左右，病毒传播的潜伏人群基本已得到管控。红色圆圈曲线在 ΔE_q 的拟合上则描绘了潜伏人群变化量的走向，1月21日—2月3日下降、2月3日—2月17日上升并稳定在0，拟合优度 r^2 值为0.4104，虽然趋势得到复现，但因为真实数据的波动使得 r^2 并不太佳。

图3(c)中， ΔI_q 的黑色方块曲线趋势也可分为2阶段：第1阶段为1月24日—2月7日，感染人数变化量 ΔI_q 为正值并逐渐下降，也就是说感染人数在逐渐增加但是增加的人数逐渐减少；第2阶段为2月7日—2月17日之间， ΔI_q 开始下降为负值且绝对值越来越大，感染人数逐渐减少并且减少的趋势越来越快。 ΔI_q 模型拟合数据(红色圆圈曲线)也显示在1月24日—2月7日为正值下降，2月7日后则为负值下降，与真实数据的变化趋势相近，拟合优度 r^2 值也达到0.8795，拟合效果不错。图3(d)中， ΔR_q 的黑色方块曲线趋势也可分为2阶段：第1阶段为1月24日—2月1日，曲线平缓值为0左右，治愈人数尚少；第2阶段则为2月1日—2月17日， ΔR_q 出现明显的上升趋势，治愈人数逐渐增多， ΔR_q 为正值并大幅度上升。对比来看， ΔR_q 的红色圆圈拟合结果曲线则是先上升后下降，与实际数据贴合的结果不是很好，拟合

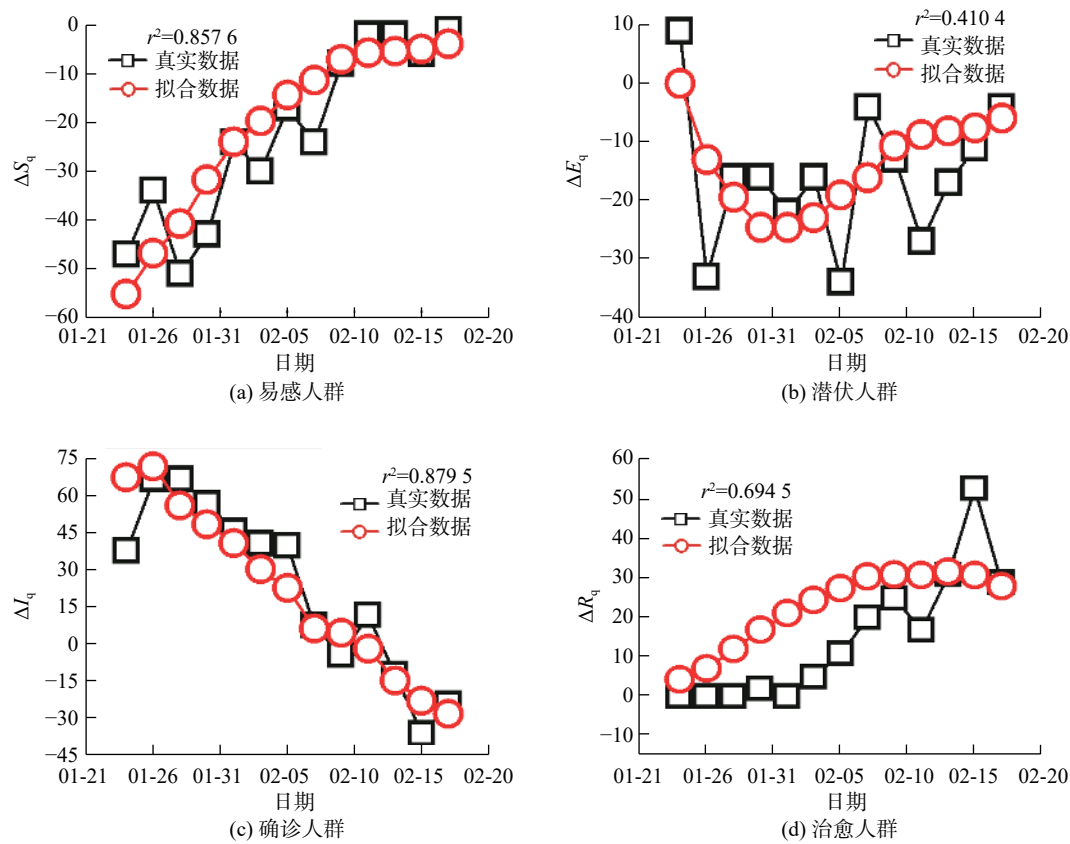


图 3 山东省 D-SEIR 模型拟合图

Fig.3 Fitting diagram of Shandong D-SEIR model

优度 r^2 值为 0.6945, 但是与真实数据的趋势还是有差别, 贴合不好, 因此 ΔR_q 由 $I_q(t-1)$ 来刻画趋势显得还不太够。

3.3 疫情稳定后 D-SEIR 模型调整结果

基于疫情严重时期建立了 D-SEIR 传播模型和式 (3) 的调整模型, 本文将其运用于疫情后期 COVID-19 传播拟合预测。由于疫情各个时间段的感染人数不同, 传播效力也不一样, 尤其是在疫情后期感染人数迅速减少, 隔离和复工复产也有影响因素, 因此通过式 (3) 调节参数的求解使模型更加符合 COVID-19 的整体传播趋势, 调节参数结果如表 2 所示。再将表 2 结果代入式 (3) 预测 I_q , ΔS_q , ΔI_q , 结果如图 4 所示。同时通过式 (3) 求得拟合优度 r^2 , 如图 4 所示。

从图 4 中可以发现, 调整后的模型对于深圳市和山东省的 I_q , ΔS_q , ΔI_q 的描绘曲线趋势都很贴合, 真实数据(黑色方块曲线)与拟合数据(红色圆圈曲线)的变化趋势一致。以深圳市为例, 图 4(a)中 I_q 的黑色方块曲线与红色圆圈曲线趋势可分为 3 个阶段。第 1 阶段 1 月 19 日—2 月 5 日 I_q 迅速上升,

表 2 调节参数结果表

Tab.2 Adjustment parameter results

区域	γ_{q1}	γ_{q2}	γ_{q3}	γ_{q4}
深圳	0.3628	0.3951	6.5294	0.0349
山东	1.0930	0.3653	-1.0403	0.0688

COVID-19 扩散感染情况越来越严重, 确诊人群数量一直上升。第 2 阶段 2 月 5 日—2 月 17 日, I_q 则出现下降趋势, 确诊人数随着控制和治疗逐渐减少, 疫情扩散得到控制。第 3 阶段 2 月 17 日—3 月 6 日, I_q 下降并趋于平稳, 在 2 月 25 日节点开始一直稳定, 再无新增和减少, 疫情已经完全得到控制, 剩下部分人群是在接受治疗的确诊人群。拟合预测红色圆圈曲线在 I_q 上面的第一、二阶段趋势高度一致, 除去第 3 阶段的平稳有偏离。图 4(b)中从 ΔS_q 的黑色曲线变化趋势来看, 红色圆圈曲线在 2 月 9 日节点前稳定上升, 2 月 9 日节点后也稳定在 0 左右, 复现了真实数据 ΔS_q 黑色方块曲线的变化趋势, r^2 则比不加调节参数时高, 为 0.9712, 无调节参数时图 2(a)中为 0.9661。再

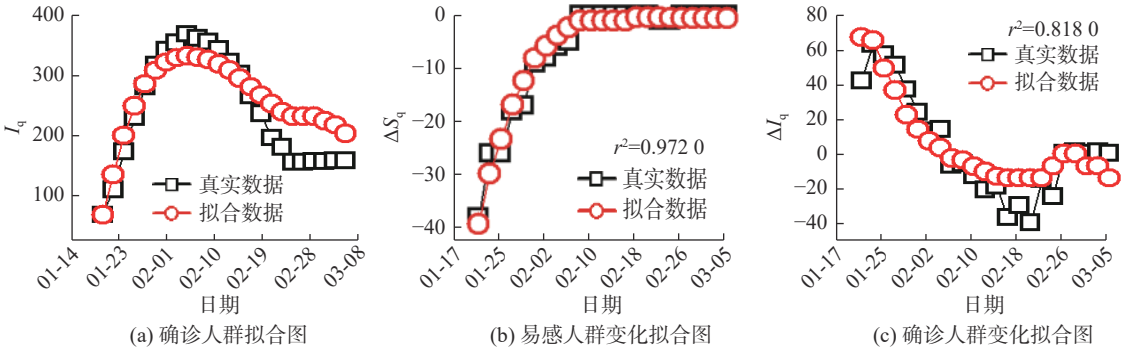


图 4 深圳市 D-SEIR 修正模型拟合图

Fig.4 Fitting diagram of Shenzhen D-SEIR modified model

观察图 4(c) 中 ΔI_q 的变化趋势, 黑色方块曲线与红色圆圈曲线皆呈现先下降后小幅上升的趋势, 也是由正值逐渐下降为负值, 确诊人群的数量由不断增加变为不断减少。在 2 月 17 日—3 月 6 日阶段, 因为黑色方块曲线 ΔI_q 由最低点上升后平稳, 确诊人群数量在减少, 调整过程为了平稳上升过程中的点波动, 所以在这阶段曲线稍许偏离, 对比图 2(c) 中 ΔI_q 的拟合结果, r^2 虽有一些下降, 由 0.8474 变为 0.8180, 但是在稳定后的确诊人群的演化趋势上更贴合。

再观察 D-SEIR 模型在疫情稳定后的调整模型对山东省的传播拟合结果, 如图 5 所示。从图 5(a) 中可发现中 I_q 的黑色方块曲线与红色圆圈曲线也共同呈现先上升后下降 2 阶段趋势, 第 1 阶段 1 月 21 日—2 月 9 日, 感染人数不断上升, 并在 2 月 9 日节点达到最高峰, 山东省此时感染情况最严重。第 2 阶段 2 月 9 日—3 月 3 日, 感染人群数量开始迅速下降, 不过没有趋于平稳的趋势, 但山东此时疫情也已经得到基本控制, 修正后的模型曲线红色圆圈曲线在这期间也与黑色方块曲线十分贴近, 复现了感染人群数量变化走势。图 5(b)

中则对模型修正后的易感人群变化量 ΔS_q 进行描绘, 根据曲线趋势也可分为 2 阶段, 第 1 阶段 1 月 23 日—2 月 17 日, ΔS_q 为负值并一直上升, 直至为 0, 易感人群被感染的变化量一直在减少, 直至疫情稳定后在无易感人员被感染, 变化量为 0。第 2 阶段 2 月 7 日—3 月 1 日, ΔS_q 一直在 0 左右波动, 被感染的易感人员在疫情稳定后就更少, 只是少数被感染。从两阶段来看模型修正的拟合结果(红色圆圈曲线)与真实数据(黑色方块曲线)很贴近, 拟合优度 r^2 达到 0.8615, 比图 3(a) 中的拟合优度 r^2 值 0.8576 有提高, 描绘结果更为准确。而图 5(c) 中确诊人群变化量 ΔI_q 在整体下降, 也可分为 2 阶段, 第 1 阶段 1 月 23 日—2 月 7 日, ΔI_q 为正值并逐渐下降, 确诊人群的增加变化量逐渐减少。第 2 阶段 2 月 7 日—3 月 1 日, ΔI_q 为负值并后期逐渐稳定, 感染人数在下降并最后稳定在接受治疗的感染人群总数上, 无新增感染人员。两阶段调整后的红色圆圈拟合趋势也更贴合黑色方块曲线的走向, 拟合优度 r^2 为 0.9122, 比图 3(c) 中未调整前的 r^2 值 0.8795 提高, 刻画确诊人群变化情况更为准确。

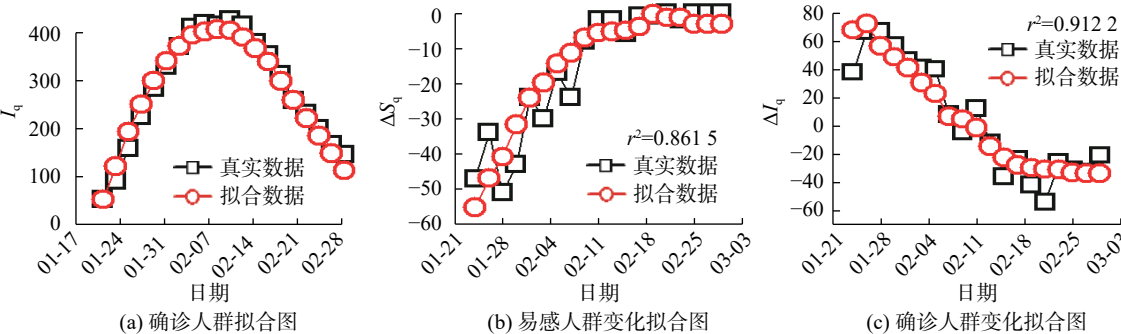


图 5 山东省 D-SEIR 修正模型拟合图

Fig.5 Fitting diagram of Shandong D-SEIR modified model

4 结束语

本文根据严控至逐步复工复产期间不同阶段的 COVID-19 感染情况,提出一种考虑地区之间人口流动和地区内部人口流动状态的疫情传播 D-SEIR 模型。通过综合考虑在人群流动之下,构建 D-SEIR 模型利用易感人群被感染的变化量和潜伏人群的变化量,拟合预测疫情在严重至稳定期间的易感人群和确诊人群的演变趋势。经过深圳市和山东省的实证数据实验结果表明,提出的 D-SEIR 模型能够准确预测 COVID-19 传播过程中的易感人群和确诊人群演化趋势。根据疫情未转变稳定前的传播情况,深圳市对易感人群变化趋势和确诊人群变化趋势拟合优度 r^2 可达 0.9661 和 0.8474,山东省的拟合优度 r^2 可达 0.8576 和 0.8795。当疫情转变稳定后,深圳市对疫情整体期间的易感人群变化趋势和确诊人群变化趋势拟合优度 r^2 可达 0.9720 和 0.8180,山东省的拟合优度 r^2 则可达 0.8615 和 0.9122。由实验结果可发现本文模型对于 COVID-19 在人群流动下的传播规律刻画准确,可为复工复产背景下不同城市和地区的风险预警提供决策依据。但同时本文的模型还可以更加完善,比如对于治愈人群和确诊人群的关系、治疗因素对疫情传播的抑制作用、流动人群中确诊人群的存在对疫情扩散的影响等,这些工作将在以后的研究中再作更细致的分析。

参考文献:

- [1] GIRE S K, GOBA A, ANDERSEN K G, et al. Genomic surveillance elucidates Ebola virus origin and transmission during the 2014 outbreak[J]. *Science*, 2014, 345(6202): 1369–1372.
- [2] ZHANG Z K, LIU C, ZHAN X X, et al. Dynamics of information diffusion and its applications on complex networks[J]. *Physics Reports*, 2016, 651: 1–34.
- [3] 胡兆龙,刘建国,任卓明.基于节点度信息的自愿免疫模型研究[J]. *物理学报*, 2013, 62(21): 218901.
- [4] 李盈科,赵时,楼一均,等.新型冠状病毒肺炎的流行病学参数与模型[J]. *物理学报*, 2020, 69(9): 21–30.
- [5] ANDERSON R M, HEESTERBEEK H, KLINKENBERG D, et al. How will country-based mitigation measures influence the course of the COVID-19 epidemic?[J]. *The Lancet*, 2020, 395(10228): 931–934.
- [6] ROOSA K, LEE Y, LUO R, et al. Real-time forecasts of the COVID-19 epidemic in China from February 5th to

February 24th, 2020[J]. *Infectious Disease Modelling*, 2020, 5: 256–263.

- [7] CHRISTOS K. A time series-based statistical approach for outbreak spread forecasting: Application of COVID-19 in Greece[J]. *Expert Systems with Applications*, 2021(166): 114077–114086.
- [8] KUCHARSKI A J, RUSSELL T W, DIAMOND C, et al. Early dynamics of transmission and control of COVID-19: a mathematical modelling study[J]. *The Lancet: Infectious Diseases*, 2020, 20(5): 553–558.
- [9] YOU C, DENG Y H, HU W J, et al. Estimation of the time-varying reproduction number of COVID-19 outbreak in China[J]. *International Journal of Hygiene and Environmental Health*, 2020, 228: 113555.
- [10] WU J T, LEUNG K, LEUNG G M. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study[J]. *The Lancet*, 2020, 395(10225): 689–697.
- [11] CHEN S M, YANG J T, YANG W Z, et al. COVID-19 control in China during mass population movements at New Year[J]. *The Lancet*, 2020, 395(10226): 764–766.
- [12] TANG B, WANG X, LI Q, et al. Estimation of the transmission risk of 2019-nCoV and its implication for public health interventions[J/OL]. *SSRN Electronic Journal*, (2020-01-27). <https://ssrn.com/abstract=3525558>.
- [13] 程晓敏,李芊璘,陆家海.从 2019 新型冠状病毒肺炎疫情谈新发传染病防控策略[J]. *传染病信息*, 2020, 33(1): 22–26.
- [14] LI W, YUE L. Correlation analysis of the population emigration rate in Wuhan and the trend of novel pneumonia in Hubei province[J/OL]. *SSRN Electronic Journal*, 2020(2020-09-02). <https://ssrn.com/abstract=3537086>.
- [15] DU Z W, WANG L, CAUCHEMEZ S, et al. Risk for transportation of coronavirus disease from Wuhan to other cities in China[J]. *Emerging Infectious Diseases*, 2020, 26(5): 1049–1052.
- [16] 傅家旗,刘敏,邓春燕,等.复杂人流网络下的 COVID-19 传播模型[J]. *电子科技大学学报*, 2020, 49(3): 383–391.
- [17] YANG Z F, ZENG Z Q, 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): 165–174.
- [18] 李琳娜,王欢,黄琼丹,等.分数阶 SEIR 传染病模型的残差幂级数解法[J]. *数学的实践与认识*, 2019, 49(15): 306–317.

- (编辑:石 瑛)

- (编辑: 丁红艺)