

## 基于 SEAIRD-LSTM 混合模型对传染病的预测分析

李佳琴, 李虎飞

北方民族大学数学与信息科学学院 宁夏银川

**【摘要】** 在各类传染病交替爆发的背景下, 利用深度学习和统计相关理论对传染病的传播机制和流行趋势进行研究和探索是非常重要的。首先, 文章通过对 SEAIRD 模型与 LSTM 模型的预测结果进行线性回归, 提出 SEAIRD-LSTM 混合预测模型。然后, 对俄罗斯地区 COVID-19 传染病感染人数与死亡人数进行预测, 并将 SEAIRD-LSTM 混合预测模型与多项式回归、逻辑回归、SEIR 以及 LSTM 模型对比, 结果表明 SEAIRD-LSTM 混合预测模型在 RMSE、MAE、MAPE 和  $R^2$  评价指标下都取得了较好的预测效果。最后, 对德国和英国地区新冠数据进行预测, 通过真值与预测值的误差率对比, 验证了 SEAIRD 模型的可适性以及 SEAIRD-LSTM 混合预测模型的预测准确度。

**【关键词】** 传染病模型; SEAIRD 模型; LSTM 预测; 混合预测模型

**【基金项目】** 北方民族大学研究生创新项目 (YCX22103)

**【收稿日期】** 2023 年 3 月 21 日 **【出刊日期】** 2023 年 5 月 15 日 **【DOI】** 10.12208/j.aam.20230005

### Prediction analysis of infectious diseases based on SEAIRD-LSTM hybrid prediction model

Jiaqin Li, Hufei Li

School of Mathematics and Information Science, North Minzu University, Yinchuan, Ningxia

**【Abstract】** In the context of the outbreak of various infectious diseases, it is very important to use deep learning and statistical theories to study and explore the transmission mechanism and epidemic trend of infectious diseases. Firstly, the SEAIRD-LSTM hybrid prediction model is proposed by linear regression of SEAIRD model and LSTM model. Then, the number of COVID-19 infections and deaths in Russia was predicted, and the hybrid prediction model of SEAIRD-LSTM was compared with polynomial regression, logistic regression, SEIR and LSTM models. The results show that the SEAIRD-LSTM hybrid prediction model achieves good prediction effect under RMSE, MAE, MAPE and  $R^2$  evaluation indexes. Finally, the COVID-19 data in Germany and the UK were predicted, and the adaptability of the SEAIRD model and the prediction accuracy of the SEAIRD-LSTM hybrid prediction model were verified by comparing the error rate between the true value and the predicted value.

**【Keywords】** Infectious Disease Model; SEAIRD Model; LSTM Prediction; Hybrid Prediction Model

### 1 引言

从淋巴腺鼠疫、麻疹、非典、埃博拉等到现在人们所熟知的新型冠状病毒肺炎 (COVID-19)<sup>[1]</sup>, 传染病一直伴随着人类发展并且影响着人类的生存。COVID-19 是由严重急性呼吸综合征冠状病毒 (SARS-CoV-2) 所引起的传染病<sup>[2,3]</sup>, 自 2019 年底爆发至今, 仍在全球多个国家和地区肆虐。截至到 2023 年 3 月 16 日, 我国累计确诊病例有 9918 万例, 全球累计确诊病例高达 7.6 亿人次, 累计死亡病例高达 687 万人次。因此, 对传染病模型的研究具有十分重要的意义。

当前对传染病传播发展趋势的预测方法一般包括两种: 一种是针对数学的动态建模方法, 如 SIR、SEIR 等模型, 基本都是对传染病的数学模型<sup>[4]</sup>和传播动力学建立预测模型; 另一种是基于数据驱动的方法。鉴于常用的自回归综合移动平均及其衍生的时间序列分析方法无法较好的预测具有非线性传染率的传染病, 学

者们提出基于深度学习的一般预测方法, 如卷积神经网络<sup>[5]</sup>、基于注意力机制的递归多通道神经网络<sup>[6]</sup>、长短期记忆递归神经网络<sup>[7]</sup>等方法可以提高复杂高维的传染病数据特征来预测发展趋势<sup>[8]</sup>。Tiago 等<sup>[9]</sup>利用 EEMD-ARIMAX 组合模型和 ARIMAX 模型对巴西新冠肺炎病例数进行预测研究, 最终得到组合模型的预测性能优于单一模型的预测性能。何振欢等<sup>[10]</sup>构建了一种 EEMD-LSTM 混合模型, 应用禽霍乱的发病数据分析模型的预测性能, 为禽霍乱的有效防控以及该模型的其他应用研究提供理论依据。文献<sup>[11]</sup>提出并探讨了 ARIMA-LSTM-XGBoost 加权组合模型在预测肺结核发病趋势中的应用, 通过对比单一模型发现组合模型能够提升模型预测效果, 对肺结核发病预测具有应用价值, 也可为其它传染病预测提供新的方法和思路。由此可见, 组合预测模型可以利用模型的优势来弥补单一模型的缺陷, 从而达到提高预测精度和增加稳定性的效果, 且已被大多数学者证实有效<sup>[12]</sup>。

基于此, 文章以欧洲部分国家感染情况利用混合预测模型对 COVID-19 进行预测, 揭示传染病的传播规律, 有助于对已采取的防控措施效果进行评价, 并为下一步疫情防控提供决策建议。

### 2 SEAIRD-LSTM 混合模型建立

SEAIRD 传染病模型可以很好地捕捉疫情的历史变化, 但对疫情数据的波动原因无法解释, LSTM 模型能更准确的获取数据波动的信息。因此, 本文采用对 SEAIRD 传染病模型和 LSTM 深度学习模型的预测结果进行线性回归, 形成 SEAIRD-LSTM 混合预测模型。其中, SEAIRD-LSTM 模型由三部分组成:

(1) 对于 SEAIRD 模块, 通过和经典的传染病传播模型 SI、SIR 和 SEIR 等相比, 加入 COVID-19 传染病无症状感染者因素, 与此次疫情传播的特点也吻合。SEAIRD 模型如式 (2-1) 将人口分为 6 类, 分别是易感人群 (S)、暴露人群 (E)、无症状感染者 (A)、有症状感染人群 (I)、康复人群 (R)、死亡者 (D)。在爆发的疫情中假设没有感染过此传染病且迄今健康的人群就是易感人群, 密切接触者就是暴露人群, 无症状感染者表示被感染但是未表现出相关症状, 有症状感染即被感染并且表现出相关症状, 康复表示治愈或者康复后的病例, 其中的参数描述见表 1。

$$\begin{cases} \frac{dS}{dt} = -\frac{\alpha_I S(t)I(t)}{N} - \frac{\alpha_A S(t)A(t)}{N} - \frac{\alpha_E S(t)E(t)}{N} + \theta R(t), \\ \frac{dE}{dt} = \frac{\alpha_I S(t)I(t)}{N} + \frac{\alpha_A S(t)A(t)}{N} + \frac{\alpha_E S(t)E(t)}{N} - p\beta_A E(t) - (1-p)\beta_I E(t), \\ \frac{dA}{dt} = p\beta_A E(t) - \gamma_A A(t), \\ \frac{dI}{dt} = (1-p)\beta_I E(t) - \gamma_I I(t) - \mu I(t), \\ \frac{dR}{dt} = \gamma_A A(t) + \gamma_I I(t) - \theta R(t), \\ \frac{dD}{dt} = \mu I(t), \end{cases} \quad (1)$$

表 1 参数描述与取值

参数	描述	取值范围
$\alpha_I$	有症状感染的感染概率	(0, 1)
$\alpha_A$	无症状感染的感染概率	(0, 1)
$\beta_I$	暴露者向感染人群转变的概率	(0, 1)
$\beta_A$	暴露者向无症状感染者转变的概率	(0, 1)
$\gamma_I$	感染者恢复概率	(0, 1)
$\gamma_A$	无症状感染者恢复概率	(0, 1)
$\mu$	病死率	(0, 1)
$\theta$	恢复者再次成为易感者概率	(0, 1)
$p$	无症状感染者比例	0.58

(2) 对于 LSTM, 在递归神经网络 RNN 基础上通过输入门、输出门和遗忘门的特性对 RNN 存在的长依赖问题进行改善。本文的神经网络架构使用 LSTM 来捕捉观测值的动态时间行为。LSTM 模型本质上是 RNN 模型的一种改进, 它规避了标准 RNN 中梯度爆炸和梯度消失的问题, 另外简化了训练和参数的设计, 能够使得训练效果越好, 学习速度越快, 能够解决长期相依问题, 被广泛应用于对时间序列相关数据的预测中。每个 LSTM 的神经单元是由细胞状态、输入门、遗忘门、输出门组成。LSTM 的运算过程如下:

输入门: 输入门决定让多少新的信息加入到细胞状态中, 数学表达式为:

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i),$$

其中,  $\sigma$  表示 sigmoid 激活函数 (取值介于 0-1 之间),  $W_i$  为输入门的权重矩阵,  $b_i$  为输入门的偏置系数,  $x_t$  为当前细胞的输入,  $h_{t-1}$  为上一个细胞的输出。

遗忘门: 遗忘门的主要作用是用来决定当前状态需要丢弃之前的哪些信息, 数学表达式为:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f),$$

$$\tilde{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c),$$

其中,  $\sigma$  表示 sigmoid 激活函数,  $W_f$  为遗忘门的权重矩阵,  $b_f$  为遗忘门的偏置系数,  $W_c$  为记忆单元的权重矩阵,  $b_c$  为记忆单元的偏置系数,  $\tilde{C}_t$  表示由 tanh 层生成的新候选值。

输出门: 决定输出什么值, 一个输出到同层下一单元, 一个输出到下一层的单元上, 数学表达式为:

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o),$$

其中,  $\sigma$  表示 sigmoid 激活函数,  $W_o$  为输出门的权重矩阵,  $b_o$  为输出门的偏置系数。

长记性:

$$C_t = f_t C_{t-1} + i_t \tilde{C}_t$$

短记性:

$$h_t = o_t \tanh(c_t)$$

(3) 对 SEAIRD-LSTM 混合预测模型的线性回归, 虽然 SEAIRD 传染病模型对 COVID-19 疫情可以很好的拟合和预测, 但在建模时也有不足之处, 当对疫情进行预测时数据波动太大, 可能对预测结果造成大的误差, 而 SEAIRD 模型根本没有办法解释, 因此为了能更好的捕捉疫情的预测情况, 本文在 SEAIRD 模型预测基础上, 对 SEAIRD 模型和 LSTM 模型的预测结果和真实数据之间采用了线性回归对两个模型进行修正如图 1 所示。经过历史训练对两个模型的输出结果给定不一样的权重, 为了最终获得更精确的结果。

### 3 SEAIRD-LSTM 混合模型求解

回归分析是一种预测性的建模技术, 用于研究因变量和自变量之间的关系, 经常用于预测分析、时间序列模型, 并使用曲线拟合数据点来查找变量之间的因果关系。线性回归是回归问题之一, 线性回归假设目标值和线性相关的特征值满足多元线性方程, 通过构造损失函数, 我们可以求解参数和何时损失函数最小化。假设 SEAIRD 模型和 LSTM 模型的预测值和观测值之间是线性相关。使用 SEAIRD-LSTM 混合预测模型进行预测的主要步骤是:

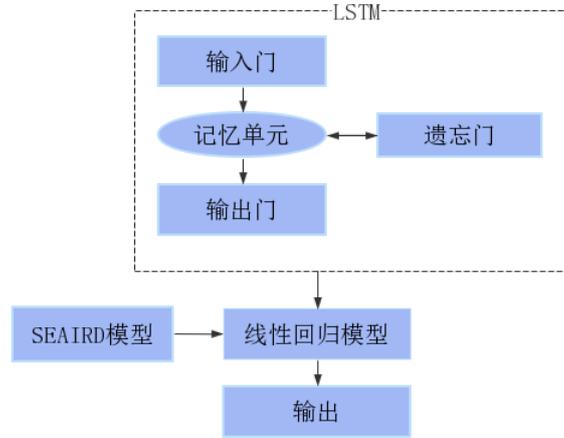


图 1 SEAIRD-LSTM 模型框架

步骤 1: 对疫情数据进行 SEAIRD 建模, 使用历史数据进行模型参数反演、优化、拟合和检验, 通过优化过后的模型得到预测值  $\hat{z}$ 。

步骤 2: 将疫情数据作为输入变量, 并对数据进行归一化处理, 设定时间窗口  $n$ , 构建一对多的 LSTM 预测模型, 通过训练模型和优化参数, 得到模型预测值  $\hat{e}$ 。

步骤 3: 通过线性回归计算 SEAIRD-LSTM 模型的预测值:

$$\hat{y} = \omega x + b,$$

其中,  $\hat{y}$  表示预测值, 需要通过一些已知的自变量  $x = [\hat{z}, \hat{e}]$  和对应的因变量  $y$ , 求解系数矩阵  $\omega$  和偏置单元  $b$ , 使得对于新变量  $x$ , 预测值  $\hat{y}$  和真实值  $y$  更加接近。为求解最佳参数定义损失函数:

$$L = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2,$$

即预测值与真实值的均方误差。要求解损失函数最小化时的  $\omega$  和  $b$ , 优化函数可以转化为:

$$(\omega^*, b^*) = \arg \min_{(\omega, b)} \sum_{i=1}^n (\omega x_i + b - y_i)^2$$

#### 4 模型评估方法

本文选用 4 种模型评价来指标掌握模型预测的准确程度, 分别为均方根误差 (RMSE)、平均绝对误差 (MAE)、平均绝对百分比误差 (MAPE)、和决定系数  $R^2$ , 其中均方根误差是预测值与实际值之间的近似程度, 平均绝对误差表示预测值与实际值之间绝对误差的平均值, 平均绝对百分比误差是绝对误差和实际值百分比的误差损失预期, 三种评价指标值越小表示模型效果越好, 而  $R^2$  是越接近于 1 效果越好。4 种评价指标表达式为:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}$$

$$MAPE = \frac{100\%}{N} \sum_{i=1}^N \frac{|y_i - \hat{y}_i|}{|y_i|}$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|$$

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2}$$

其中,  $\hat{y}_1, \dots, \hat{y}_N$  为预测结果,  $y_1, \dots, y_N$  为实际值,  $N$  为样本预测。

### 5 模型预测分析

#### 5.1 模型预测分析一

选取 2022 年 1 月 18 日至 4 月 18 日俄罗斯地区疫情数据, 利用优化算法得到传染病模型参数向量:

$$q_k = (0.0116, 0.052, 0.5362, 0.1623, 0.055, 0.2035)$$

利用 SEAIRD-LSTM 混合模型对 2022 年 4 月 19 日至 4 月 28 日的疫情数据进行预测。为了对本文提出的方法进行评价, 选取了几种主流的传染病预测算法进行比较, 包括多项式回归 (Polynomial)、Logistic 回归、SEIR 和 LSTM。如图 2 所示, 对感染人数进行预测时, Logistic、SEIR 模型整体预测偏差较大, 多项式回归预测相较于真实数据预测感染者人数前期偏低, LSTM 和 SEAIRD-LSTM 混合模型预测与真实感染人数较为接近。从对死亡人数的预测来看, 如图 3 所示, Logistic 回归的预测值整体偏低, SEIR 模型预测值整体偏大, 多项式回归、LSTM 和 SEAIRD-LSTM 模型与真实死亡人数最为接近。计算 5 种模型评价指标, 根据评价指标, 选择单一模型 LSTM 和混合模型 SEAIRD-LSTM 进行预测对比。可以发现, 结合传统动力学模型和考虑实际因素的 LSTM 的混合预测模型 SEAIRD-LSTM 能取得较好的结果。

模型评价指标如表 2 所示, 可见, 多项式回归效果较差, 模型可能过于注重对每一个训练数据的拟合, 容易造成过拟合现象。SEIR 和 Logistcs 模型预测效果也一般。结合图 2、3 和表 2, 可以看出相较于其他几种方法, LSTM 的预测精度分别为 MSE: 1338, MAE: 2814, MAPE: 0.13, R2: 0.92, 仅次于 SEAIRD-LSTM 混合模型。SEAIRD-LSTM 混合模型的预测精度有明显的优势, RMSE、MAE、MAPE 和  $R^2$  分别为 1037, 1095, 0.09 和 0.98。

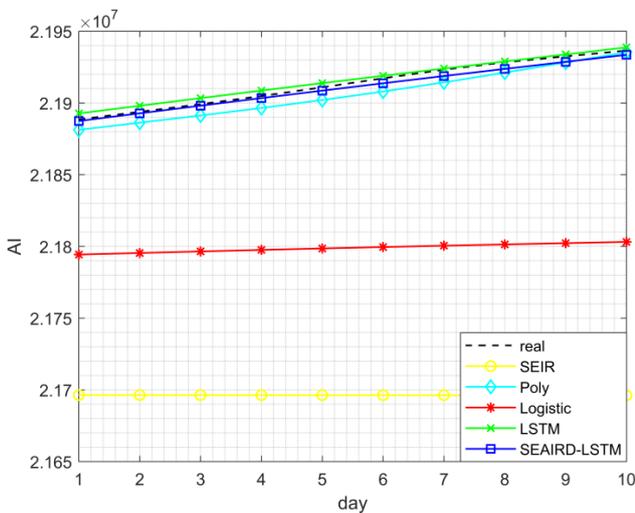


图 2 5 种方法预测感染人数与真实数据对比

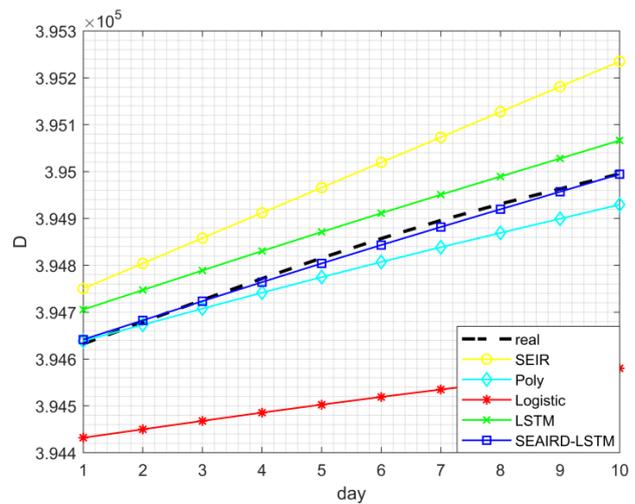


图 3 5 种方法预测死亡人数与真实数据对比

表 2 5 种预测模型评价指标

方法	RMSE	MAE	MAPE/%	R <sup>2</sup>
SEIR	1435	8127	5.83	0.53
Poly	2253	2650	2.13	0.75
Logistic	4682	3358	2.96	0.69
LSTM	1338	2814	0.13	0.92
SEAIRD-LSTM	1037	1095	0.09	0.98

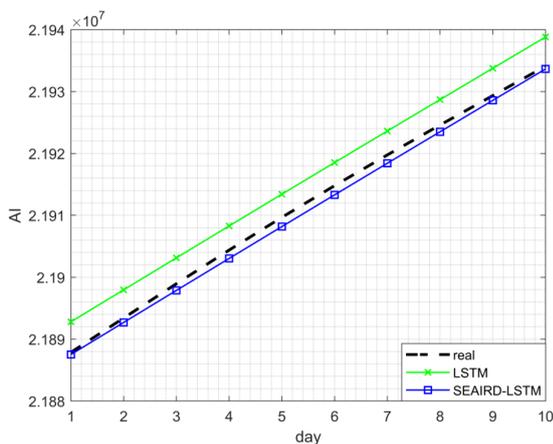


图 4 LSTM 与 SEAIRD-LSTM 混合模型预测与真实感染数据对比 (左图)

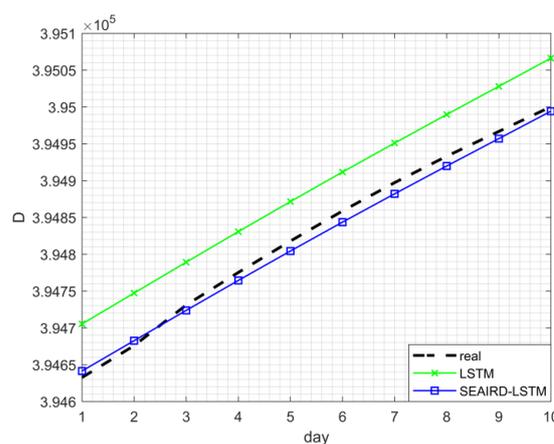


图 5 LSTM 与 SEAIRD-LSTM 混合模型预测与真实死亡数据对比 (右图)

通过对 LSTM 和 SEAIRD-LSTM 混合模型与真实数据对比, 可以看出两个模型预测趋势基本一致, 但 SEAIRD-LSTM 混合模型预测感染人数与真实数据更为接近, 如图 4 所示。图 5 对比两种预测方法与真实死亡人数, 可以看出 SEAIRD-LSTM 混合模型对死亡人数的预测更为逼近, 随着预测时间变长, 虽然 SEAIRD-LSTM 混合模型出现较小的误差, 但是整体预测效果还是良好的。同样, 从两种模型对感染人数与死亡人数的预测误差率图中也能看出, SEAIRD-LSTM 混合模型预测效果最佳, 两种模型误差率都远小于 0.01, 如图 6 和图 7。也可以说明, 在传染病模型参数参演的基础上, SEAIRD-LSTM 混合预测模型的预测效果良好。反之, 模型的参数反演效果也值得肯定。

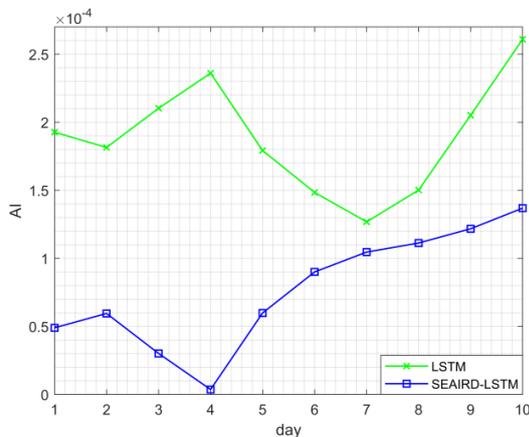


图 6 LSTM 与 SEAIRD-LSTM 混合模型预测感染人数误差率图

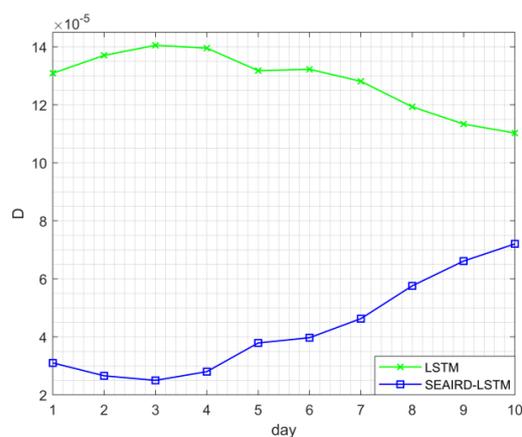


图 7 LSTM 与 SEAIRD-LSTM 混合模型预测死亡人数误差率图

### 5.2 模型预测分析二

为验证传染病模型的可适用性以及混合模型预测准确度, 进一步对德国地区疫情数据进行传染病参数反演及预测。选取 2022 年 1 月 15 日至 4 月 15 日数据进行参数反演, 得到模型参数向量:

$$q_k = (0.1244, 0.5571, 0.5362, 0.3766, 0.0669, 0.2376)$$

并利用 SEAIRD-LSTM 混合模型对 2022 年 4 月 16 日至 4 月 25 日疫情趋势进行预测。

对德国地区同样采用 5 种预测模型进行预测对比, 如图 8 所示, 对感染人数进行预测时, Logistic、SEIR 模型整体预测偏差较大, 多项式回归预测感染者人数相较于真实数据随着时间越长偏差越大, LSTM 和 SEAIRD-LSTM 混合模型预测与真实感染人数较为接近。从对死亡人数的预测来看, 如图 9 所示, Logistic 回归与 SEIR 模型的预测值整体偏低, 多项式回归预测值随着预测时间的变长偏差也越大, LSTM 和 SEAIRD-LSTM 模型与真实死亡人数最为接近。结合来看, 混合模型 SEAIRD-LSTM 在使用重构后的参数对模型进行预测, 发现多项式回归、Logistic 回归以及 SEIR 模型与真实数据相比预测结果偏差较大, 主要可能由于外界因素干扰造成的数据严重波动。而结合传统动力学模型和考虑实际因素的 LSTM 混合模型 SEAIRD-LSTM 能取得较好的结果。

通过对 LSTM 和 SEAIRD-LSTM 混合模型与真实数据对比, 可以看出两个模型预测趋势基本一致, 但 SEAIRD-LSTM 混合模型预测感染人数与真实数据更为接近, 如图 10 所示。图 11 对比两种预测方法与真实死亡人数, 可以看出 SEAIRD-LSTM 混合模型对死亡人数的预测更为逼近, 随着预测时间变长, 虽然 SEAIRD-LSTM 混合模型中间出现较小的误差, 但是整体预测效果还是良好的。同样, 从图 12 和图 13 两种模型对感染人数与死亡人数的预测误差率图中也能看出, SEAIRD-LSTM 混合模型预测效果最佳, 其误差率远小于 0.01。可以说明, 在传染病模型参数参演的基础上, 对德国地区疫情数据预测时 SEAIRD-LSTM 混合预测模型的预测效果良好。反之, 模型的参数反演效果也值得肯定。

### 5.3 模型预测分析三

另外, 为验证传染病模型的可适用性以及混合模型预测准确度, 对英国地区也进行传染病参数反演及预测。选取 2021 年 11 月 5 日至 2 月 5 日疫情数据, 利用优化算法得到模型参数向量:

$$q_k = (0.1027, 0.4013, 0.4549, 0.4981, 0.0218, 0.1958)$$

最后利用 SEAIRD-LSTM 混合模型对 2022 年 2 月 6 日至 2 月 15 日疫情趋势进行预测。

对英国地区亦采用 5 种预测模型进行预测对比, 如图 14 所示, 对感染人数进行预测时, Logistic、SEIR 模型整体预测偏差较大, 多项式回归预测感染者人数相较于真实数据随着时间越长偏差越大, LSTM 和 SEAIRD-LSTM 混合模型预测与真实感染人数较为接近。从对死亡人数的预测来看, 如图 15 所示, Logistic 回归与 SEIR 模型的预测值整体偏低, 多项式回归预测值随着预测时间的变长偏差也越大, LSTM 和 SEAIRD-LSTM 模型与真实死亡人数最为接近。

结合来看, 混合模型 SEAIRD-LSTM 在使用重构后的参数时对模型进行预测, 发现多项式回归、Logistic 回归以及 SEIR 模型与真实数据相比预测结果偏差都较大, 而结合传统动力学模型和考虑实际因素的 LSTM 混合模型 SEAIRD-LSTM 能取得较好的结果。

通过对 LSTM 和 SEAIRD-LSTM 混合模型与真实数据对比, 可以看出两个模型预测趋势基本一致, 尤其 SEAIRD-LSTM 混合模型预测感染人数与真实数据更为接近, 随着预测时间变长, 虽然 SEAIRD-LSTM 混合模型预测前后期出现较小的误差, 但是整体预测效果还是良好的, 如图 16 所示。图 17 对比两种预测方法与真实死亡人数, 可以看出 SEAIRD-LSTM 混合模型对死亡人数的预测更为逼近。同样, 从图 18 和图 19 两种模型对感染人数与死亡人数的预测误差率图中也能看出, SEAIRD-LSTM 混合模型预测效果最佳, 其误差率远小于 0.01。可以说明, 在传染病模型参数参演的基础上, 对英国地区疫情数据预测时 SEAIRD-LSTM 混合预测模型的预测效果良好。反之, 模型的参数反演效果也值得肯定。

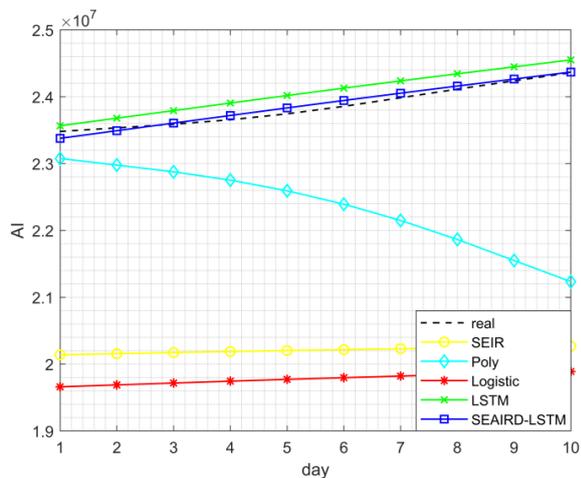


图 8 5 种方法预测感染人数与真实数据对比

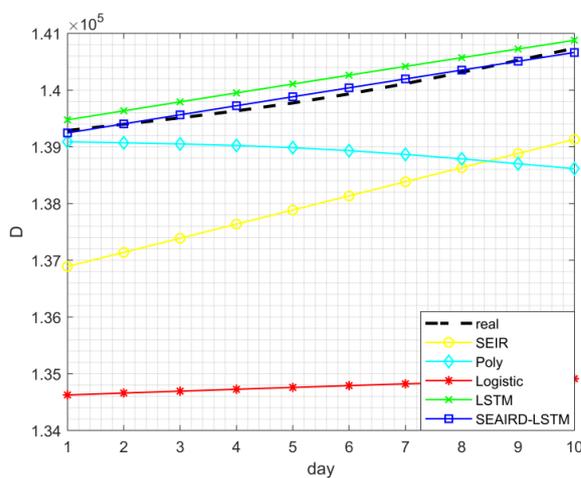


图 9 5 种方法预测死亡人数与真实数据对比

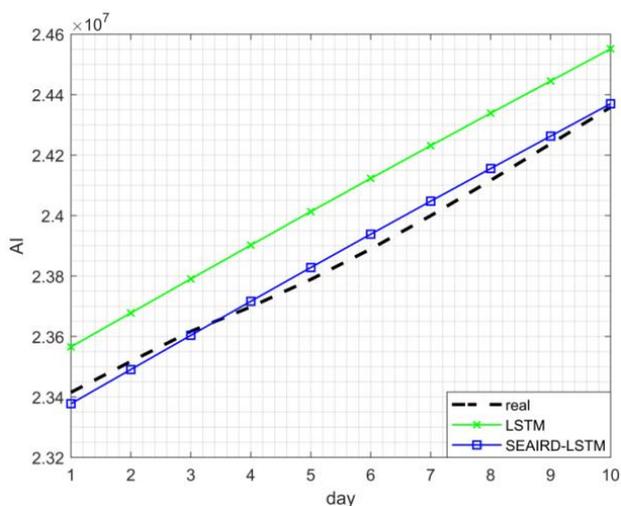


图 10 LSTM 与 SEAIRD-LSTM 混合模型预测与真实感染数据对比 (左图)

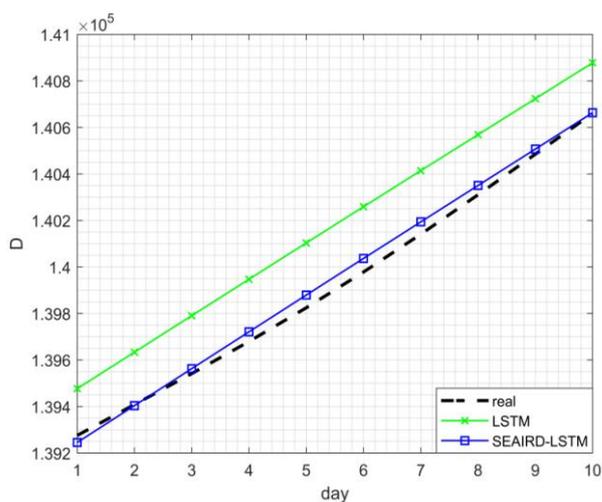


图 11 LSTM 与 SEAIRD-LSTM 混合模型预测与真实死亡数据对比 (右图)

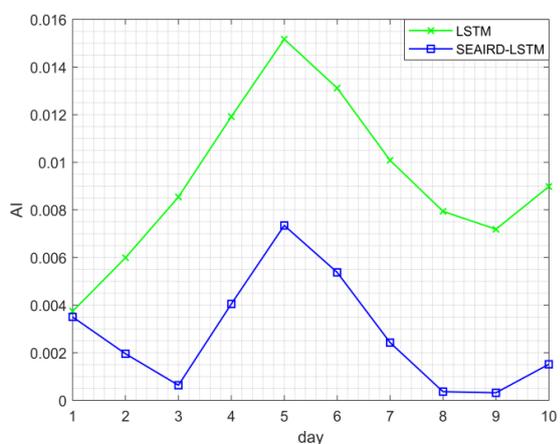


图 12 LSTM 与 SEAIRD-LSTM 混合模型预测感染人数误差率图 (左图)

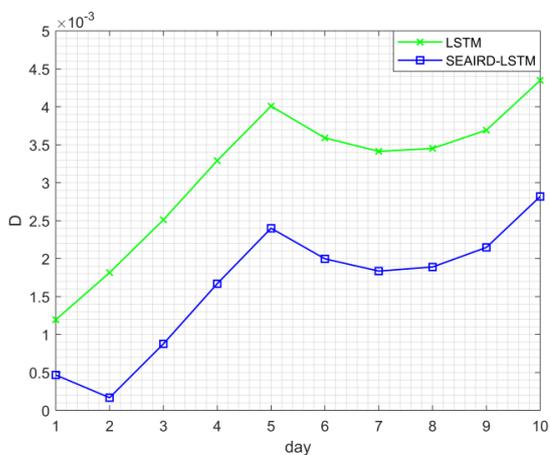


图 13 LSTM 与 SEAIRD-LSTM 混合模型预测死亡人数误差率图 (右图)

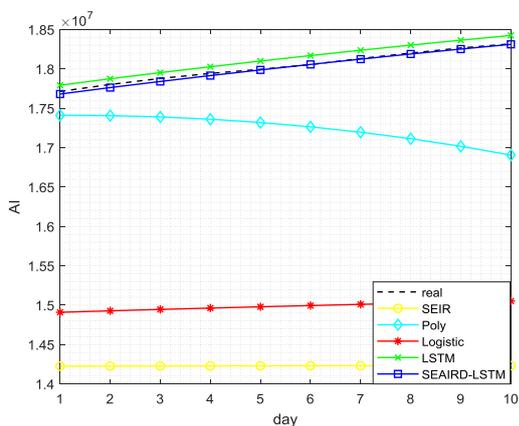


图 14 5 种方法预测感染人数与真实数据对比

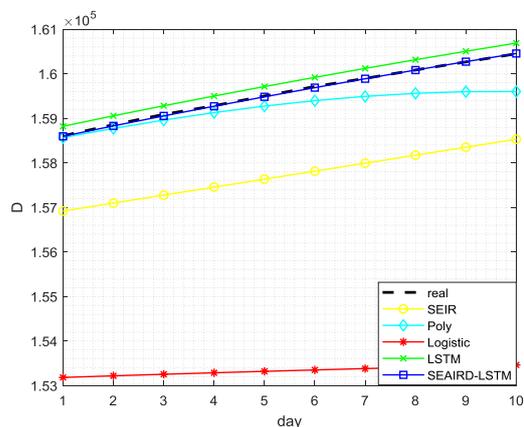


图 15 5 种方法预测死亡人数与真实数据对比

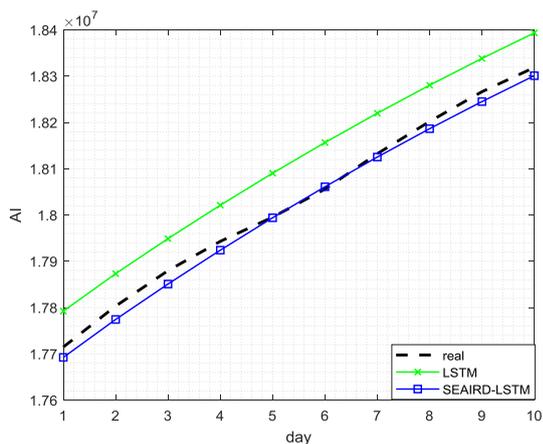


图 16 LSTM 与 SEAIRD-LSTM 混合模型预测与真实感染数据对比 (左图)

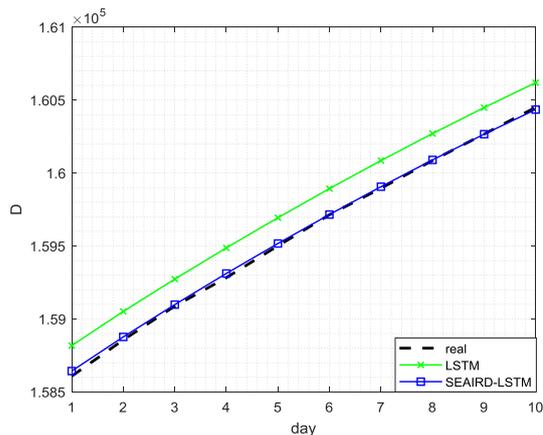


图 17 LSTM 与 SEAIRD-LSTM 混合模型预测与真实死亡数据对比 (右图)

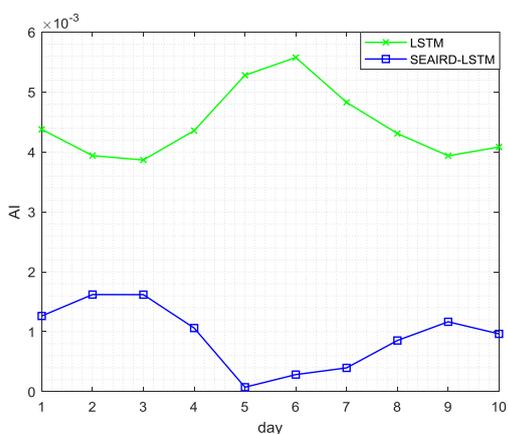


图 18 LSTM 与 SEAIRD-LSTM 混合模型预测感染人数误差率图 (左图)

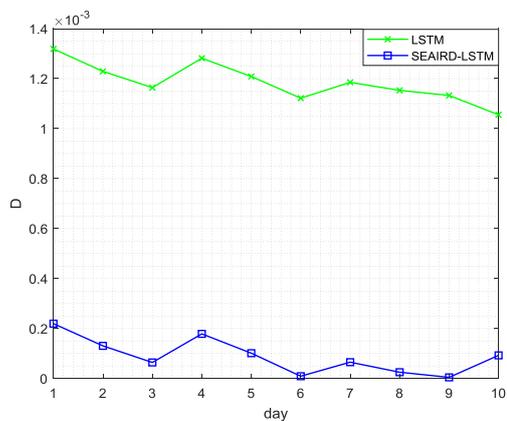


图 19 LSTM 与 SEAIRD-LSTM 混合模型预测死亡人数误差率图 (右图)

## 6 结论

本文提出一种 SEAIRD-LSTM 混合预测模型, 该模型方法集合了传统传染病预测方法和深度学习预测

方法。通过与多项式回归、Logistic 回归、SEIR 以及 LSTM 多种预测模型进行对比发现, 不管是对俄罗斯感染人数的预测还是对死亡人数的预测, 只有 LSTM 和 SEAIRD-LSTM 的预测值与实际数据误差较小。最后, 对比单一模型 LSTM 与混合模型 SEAIRD-LSTM, 可以发现 SEAIRD-LSTM 混合预测模型的预测精度和稳定性效果更好, 也验证了混合模型可以利用模型的优势来弥补单一模型的缺陷。

为了验证本文提出的混合模型预测的适用性, 将 SEAIRD-LSTM 应用于其它研究地区-德国、英国。对比多种模型预测结果, 可以发现利用 SEAIRD-LSTM 混合预测模型对德国和英国地区的感染人数和死亡人数预测与真实值对比误差都相对较小, 预测效果较好。LSTM 模型同样预测感染人数和死亡人数, 虽然整体预测趋势与 SEAIRD-LSTM 相同, 但预测精度和稳定性没有 SEAIRD-LSTM 混合模型效果好。因此本文提出的 SEAIRD-LSTM 混合预测模型有一定的可适用性, 可以推广到存在大量无症状感染者在疫情爆发区。

### 参考文献

- [1] Ciotti M, Ciccozzi M, Terrinoni A, et al. The COVID-19 pandemic[J]. *Critical reviews in clinical laboratory sciences*, 2020, 57(6): 365-388.
- [2] Lipsitch M, Swerdlow D L, Finelli L. Defining the epidemiology of Covid-19—studies needed[J]. *New England journal of medicine*, 2020, 382(13): 1194-1196.
- [3] Bulut C, Kato Y. Epidemiology of COVID-19[J]. *Turkish journal of medical sciences*, 2020, 50(9): 563-570.
- [4] Jiao J, Liu Z, Cai S. Dynamics of an SEIR model with infectivity in incubation period and homestead-isolation on the susceptible[J]. *Applied Mathematics Letters*, 2020, 107: 106442.
- [5] Fu B, Yang Y, Ma Y, et al. Attention-based recurrent multi-channel neural network for influenza epidemic prediction[C]. *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2018: 1245-1248.
- [6] Doni A R, Sasipraba T. LSTM-RNN Based Approach for Prediction of Dengue Cases in India[J]. *Ingénierie des Systèmes d'Information*, 2020, 25(3).
- [7] Derr T, Ma Y, Fan W, et al. Epidemic graph convolutional network[C]. *Proceedings of the 13th International Conference on Web Search and Data Mining*, 2020: 160-168.
- [8] 谢黎颖. 基于时空序列的传染病发展趋势预测[D]. 东华大学, 2021.
- [9] da Silva T T, Francisquini R, Nascimento M C V. Meteorological and human mobility data on predicting COVID-19 cases by a novel hybrid decomposition method with anomaly detection analysis: A case study in the capitals of Brazil[J]. *Expert Systems with Applications*, 2021, 182: 115190.
- [10] 何振欢, 肖建华. 基于 EEMD-LSTM 模型的禽流感预测研究[J]. *动物医学进展*, 2022, 43(11):34-38.
- [11] 赖晓莹, 钱俊. ARIMA-LSTM-XGBoost 加权组合模型在肺结核发病趋势预测的研究[J]. *现代预防医学*, 2021, 48(01):5-9.
- [12] 冯晨, 陈志德. 基于 XGBoost 和 LSTM 加权组合模型在销售预测的应用[J]. *计算机系统应用*, 2019, 28(10): 226-232.

版权声明: ©2023 作者与开放获取期刊研究中心(OAJRC)所有。本文章按照知识共享署名许可条款发表。

<http://creativecommons.org/licenses/by/4.0/>

