

· 预防医学研究 ·

## 基于SARIMA-ERNN组合模型预测我国细菌性痢疾发病率

张生奎,王镇德,杨 荔,张乐言,王永斌,袁聚祥

华北理工大学公共卫生学院,河北 唐山 063210

**[摘要]** 目的:探讨季节自回归滑动平均混合(SARIM)-Elman神经网络(ERNN)组合模型在我国细菌性痢疾发病率预测中的利用价值。方法:使用2005年1月—2016年12月我国细菌性痢疾的月发病率资料作为训练集分别建立SARIMA模型、ERNN模型以及SARIMA-ERNN组合模型,2017年我国细菌性痢疾的月发病率资料作为测试集用于模型间的效果评价。结果:SARIMA模型拟合及预测的平均相对误差(average relative error, MRE)、平均误差率(average error rate, MER)、均方根误差(root mean squared predict error, RMSE)和平均绝对误差(mean absolute error, MAE)分别为5.661 37、0.061 81、0.001 45、0.000 94和5.596 40、0.051 77、0.004 54、0.000 34;ERNN模型拟合及预测的MRE、MER、RMSE和MAE分别为5.348 57、0.056 05、0.017 08、0.000 79和5.544 30、0.044 55、0.000 36、0.000 30;SARIMA-ERNN组合模型拟合及预测的MRE、MER、RMSE和MAE分别为4.942 52、0.047 33、0.001 15、0.000 72和4.251 30、0.044 19、0.000 38、0.000 29。结论:与ERNN模型和SARIMA模型相比,SARIMA-ERNN组合模型具有较高的有效性和合理性,可以用于我国细菌性痢疾的短期预测和早期预警。

**[关键词]** 预测模型;细菌性痢疾;发病率

**[中图分类号]** R516.4

**[文献标志码]** A

**[文章编号]** 1007-4368(2019)06-925-07

**doi:**10.7655/NYDXBNS20190628

## Forecasting the incidence of bacillary dysentery in china based on SARIMA - ERNN combination model

Zhang Shengkui, Wang Zhende, Yang Li, Zhang Leyan, Wang Yongbin, Yuan Juxiang

School of Public Health, North China University of Science and Technology, Tangshan 063210, China

**[Abstract]** **Objective:** To explore the application of SARIMA-ERNN combination model in predicting the incidence of bacterial dysentery in China. **Methods:** Using the monthly incidence data of bacterial dysentery in China from January 2005 to December 2016 as the training set, the SARIMA model, the ERNN model, and the SARIMA-ERNN combination model were established respectively. The monthly incidence of bacterial dysentery in China in 2017 which was used to test the efficacy of the above models was used as a test set. **Results:** The MRE, MER, RMSE and MAE fitted and forecasted by SARIMA model were 5.661 37, 0.061 81, 0.001 45, 0.000 94 and 5.596 40, 0.051 77, 0.004 54, 0.000 34 respectively. The MRE, MER, RMSE and MAE fitted and forecasted by ERNN model fit and predicted were 5.348 57, 0.056 05, 0.017 08, 0.000 79 and 5.544 30, 0.044 55, 0.000 36, 0.000 30 respectively; The MRE, MER, RMSE and MAE fitted and forecasted by SARIMA-ERNN combination model were 4.924 25, 0.047 33, 0.001 15, 0.000 72 and 4.251 30, 0.044 19, 0.000 38, 0.000 29 respectively. **Conclusion:** The SARIMA-ERNN combination model has high validity and reasonability that can be used for short-term forecasting and early warning of bacterial dysentery in China.

**[Key words]** predictive model; bacterial dysentery; incidence

[J Nanjing Med Univ, 2019, 39(06):925-931]

细菌性痢疾(bacillary dysentery)是由志贺氏菌感染引起的一种主要通过粪-口途径传播的严重肠

**[基金项目]** 河北省重点职业病防治技术研究(13277709D)

\*通信作者(Corresponding author), E-mail: yuanjx@ncst.edu.cn

道传染病。在我国,细菌性痢疾常年散发,病例主要集中在夏季和秋季。目前我国细菌性痢疾仍有较高的发病率,严重威胁着人群健康<sup>[1-4]</sup>。因此,需要一种建模方法,可以根据历史时间序列数据为决

策者提供未来细菌性痢疾发病率的早期估计,有助于及早采取相应的预防措施,以期最大限度地降低患者的发病率、死亡率和不良健康结局。

传统的回归模型需要获取预测对象相关的影响因素的详细数据,而时间序列可以通过捕捉历史数据和时间变量的关系对疾病的未来趋势进行定量的描述。自回归滑动平均混合(autoregressive integrated moving average, ARIMA)模型是预测时间序列数据的常用统计方法<sup>[5-7]</sup>。ARIMA模型要求预测对象为平稳随机序列,但是实践中随着时间和环境的变化,大量医学现象总呈现出某种长期趋势,此时需要通过数据变换或者差分等方法对数据进行预处理。当疾病具有季节性趋势时经常使用季节自回归滑动平均混合(seasonal autoregressive integrated moving average, SARIMA)模型<sup>[8]</sup>。然而,传染病的发病受很多因素的影响,其变化往往还包含有复杂的非线性趋势,因此单一的线性模型并不适用。

与ARIMA模型不同,人工神经网络模型中引入了非线性函数,可以有效捕捉数据中的非线性关系,已被广泛用于复杂事件的预测实践中。应用较多的人工神经网络有BP神经网络(back propagation neural networks, BPNN)、RBF神经网络(radial basis function neural networks, RBFNN)和Elman神经网络(Elman neural networks, ERNN)等<sup>[9-11]</sup>。除了BPNN模型已被用于甲肝的预测外<sup>[12]</sup>,很少有研究使用RBFNN模型和ERNN模型来预测疾病。

单一模型由于其各自的适用条件,都存在信息源有限的问题,导致这些单一模型在应用中受到一定的限制。而组合模型可以充分有效地提取复杂的时间序列信息,克服单一模型适用条件有限的障碍,有助于改善模型的预测效果<sup>[13]</sup>。因此本研究通过比较SARIMA-ERNN组合模型与ERNN模型和SARIMA模型的预测效果,以验证该组合模型在预测我国细菌性痢疾时的适用性和可行性。

## 1 资料和方法

### 1.1 资料

细菌性痢疾2005—2017年的发病资料从中国疾病预防控制中心发布的法定传染病疫情概况表中获取。

### 1.2 模型构建

#### 1.2.1 SARIMA模型<sup>[9,14]</sup>

ARIMA模型的表达式为:  $Y_t = \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \dots + \varphi_p Y_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q}$

$Y_t$ 是时间序列在 $t$ 期的观测值, $p$ 是自回归模型的阶数, $q$ 是滑动平均模型的阶数, $e$ 是误差或偏差。SARIMA模型作为ARIMA模型的延伸,考虑了序列的季节性趋势。SARIMA模型通常称表示为SARIMA( $p, d, q$ ) $\times$ ( $P, D, Q$ ) $S$ ,式中 $p, d, q$ 和 $P, D, Q$ 分别是非季节性模型和季节性模型中自回归过程、差分过程、滑动平均过程的阶数, $S$ 表示周期。

#### 1.2.2 ERNN模型<sup>[15-16]</sup>

ERNN是一种非线性动态神经网络。其承接层可以将隐含层的输出值和上一输入层的输入值一并作为隐含层的输入,实现了网络内信号沿着向前和向后的方向同时传递,使得系统具有较强的动态映射能力。其表达式为:

$$\begin{aligned} y(k) &= g(\omega_3 x(k)) \\ x(k) &= f(\omega_2 x(k) + \omega_2(u(k-1))) \\ x_c(k) &= x(k-1) \end{aligned}$$

式中 $u(k-1)$ 和 $y(k)$ 分别是时间 $k$ 时网络的输入值和输出值, $x_c(k)$ 和 $x(k)$ 分别是承接层和隐含层的节点, $\omega_1, \omega_2, \omega_3$ 为对应各层间的权重矩阵。

#### 1.2.3 SARIMA-ERNN组合模型

SARIMA模型和ERNN模型具有不同的理论基础,两个模型的结合可以在一定程度上克服单一模型信息源应用不广泛的问题,取长补短,有助于提高预测的精确性和准确性。本研究首先使用SARIMA模型建立细菌性痢疾发病率的预测序列,该模型拟合结果与实际值之间存在高度的关联性,因此,本研究使用细菌性痢疾发病率序列的SARIMA模型拟合值作为ERNN模型的输入值,细菌性痢疾的实际发病率数据作为ERNN模型的输出值,借助newelm函数构建优化的ERNN模型。

#### 1.2.4 模型预测效果比较

对于SARIMA模型,使用Akaike信息准则(Akaike information criterion, AIC)和Schwarz贝叶斯准则(Schwarz Bayesian criterion, SBC)作为选择最佳模型的标准。然而在神经网络中,没有类似于AIC和SBC的模型选择标准。通常,将建模样本分为两部分:用于训练样本的训练集和用于评价效果的验证集,应选择训练集得到的值和验证集得到的值与实际观察值之间的偏差最小的模型为最优模型。本研究使用经验方法中广泛使用的测量偏差的指标:平均绝对误差(mean absolute error, MAE)、均方根误差(root mean squared predict error, RMSE)、平均误差率(average error rate, MER)和平均相对误差(average relative error, MRE)。

1.3 统计学方法

运用 SPSS13.0 建立 SARIMA 模型,用 MatlabR2014a 建立 ERNN 模型和 SARIMA-ERNN 组合模型,检验水准 $\alpha=0.05$ 。

2 结果

2.1 SARIMA 模型结果

根据细菌性痢疾发病率原始序列图以及单位根检验结果( $ADF=-2.729, P=0.072$ ),提示我国细菌性痢疾发病率序列为非平稳时间序列,再结合我国细菌性痢疾的发病周期,对原始序列进行 $D=1$ 周期为12的季节差分使序列转化为平稳序列。结合平稳序列和残差序列的自相关函数(autocorrelation function, ACF)图、偏自相关函数(partial autocorrelation function, PACF)图以及白噪声检验图确定合适的预测模型,然后根据 likelihood 值、AIC 值和 SBC 值选取最优模型。经过反复比较,得到最优的模型是  $ARIMA(1, 0, (11)) \times (0, 1, 1)_{12}$ , 具体方程为  $(1-B)(1-B^{12})x_t = (1 + 0.321B)(1 - 0.468B^{12})\varepsilon_t / (1 - 0.627B)$ 。 $AR1=0.627 (t=-9.469, P < 0.001)$ ,  $MA1=-0.321 (t=-3.748, P < 0.001)$ ,  $SMA1=0.468 (t=5.268, P < 0.001)$ ,  $AIC=-1440.424$ ,  $SBC=-1398.847$ ,  $\text{Log likelihood}=734.212$ 。最后,模型的 Ljung-Box Q 检验 ( $P > 0.05$ )和残差诊断(图1)的结果表明残差序列是“白噪声”。利用建立好的模型预测2017年细菌性痢疾发病率(表1第3列)。

2.2 ERNN 模型结果

由于细菌性痢疾发病率变化的周期为12个月,因此本研究中 ERNN 模型输入层和输出层神经元个

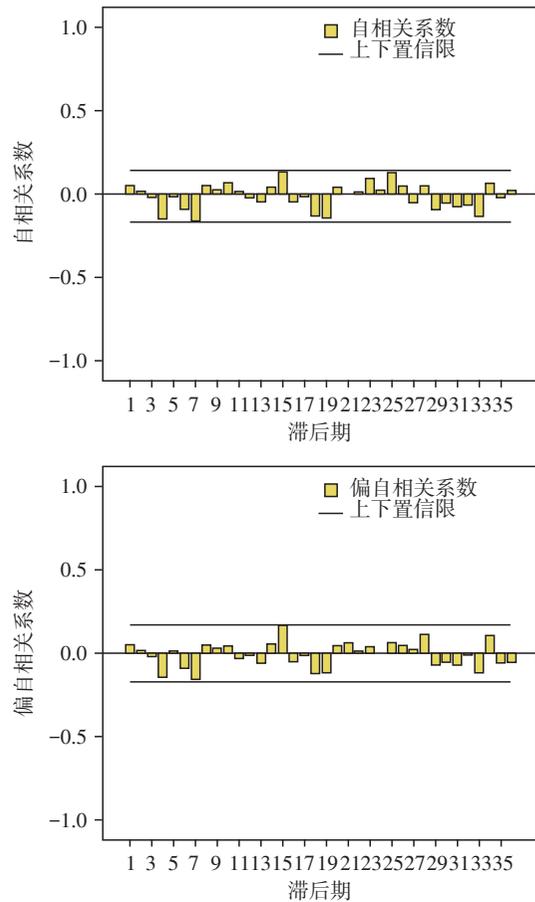


图1 SARIMA(1,0,(11))x(0,1,1)<sub>12</sub>残差序列的 ACF 和 PACF 图

Figure 1 The ACF and PACF plot of residuals from the SARIMA(1,0,(11))x(0,1,1)<sub>12</sub> model

数分别为12和1,根据经验公式并进行相应扩展,最终确定在区间4~21之间寻找最佳隐含层神经元个数。首先使用 Matlab 自带神经网络工具箱里面的

表1 2017年细菌性痢疾发病率(1/10万)3种模型的预测验证值

Table 1 Verification value of incidence of bacillary dysentery predicted by three models in 2017(1/100 000)

月份 (月)	实际值	SARIMA 模型		ERNN 模型		SARIMA-ERNN 组合模型	
		预测值	误差率(%)	预测值	误差率(%)	预测值	误差率(%)
1	0.003 7	0.003 6	2.702 7	0.003 1	16.216 2	0.003 5	5.405 4
2	0.003 7	0.003 3	10.810 8	0.003 2	13.513 5	0.003 5	5.405 4
3	0.004 2	0.004 5	7.142 9	0.004 4	4.761 9	0.004 4	4.761 9
4	0.005 1	0.005 3	3.921 6	0.005 1	0.000 0	0.005 1	0.000 0
5	0.007 8	0.007 4	5.128 2	0.007 2	7.692 3	0.007 5	3.846 2
6	0.010 1	0.009 9	1.980 2	0.009 7	3.960 4	0.011 0	8.910 9
7	0.011 7	0.011 9	1.709 4	0.011 7	0.000 0	0.011 6	0.854 7
8	0.011 0	0.011 6	5.454 5	0.011 4	3.636 4	0.011 6	5.454 5
9	0.007 9	0.009 1	15.189 9	0.008 2	3.797 5	0.008 4	6.329 1
10	0.006 1	0.006 1	0.000 0	0.005 9	3.278 7	0.005 8	4.918 0
11	0.004 4	0.004 2	4.545 5	0.004 1	6.818 2	0.004 3	2.272 7
12	0.003 5	0.003 8	8.571 4	0.003 6	2.857 1	0.003 6	2.857 1

dividerand 随机函数将 2005—2016 年我国细菌性痢疾发病率分为训练集(80%)和验证集(20%),然后使用循环控制语句循环 100 次寻找最佳的隐含节点。训练目标误差为 0.005,训练步数为 20 000。经过 20 000 次训练,当隐含层神经元个数为 4~21 时,网络的逼近误差依次为:0.013 2,0.009 8,0.019 3,0.005 5,0.009 0,0.024 3,0.009 0,0.005 8,0.004 9,0.022 5,0.011 7,0.032 4,0.009 1,0.012 3,0.012 7,0.005 5,0.012 3,0.017 1。可见当隐含层神经元为 12 时网络训练误差最小,模型训练性能图见图 2。因此选择 12-12-1 作为本研究最优的 ERNN 模型,残差的 ACF 图见图 3。

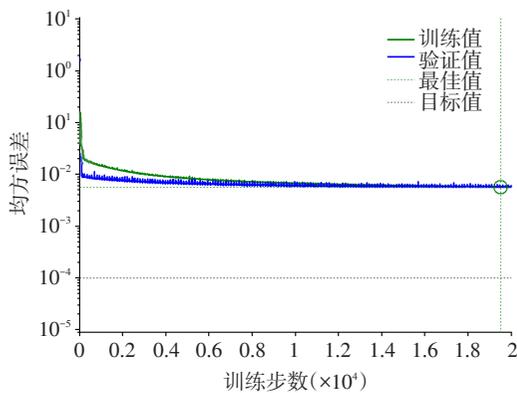


图2 ERNN 模型的训练与验证性能图

Figure 2 The training and validation performance for ERNN model

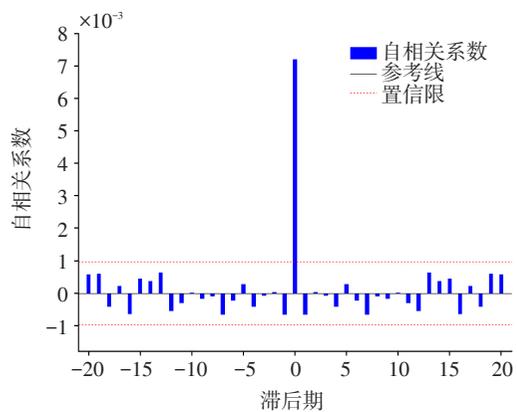


图3 ERNN 模型残差的 ACF 图

Figure 3 The ACF of residuals from ERNN model

### 2.3 SARIMA-ERNN 组合模型结果

对细菌性痢疾原始序列进行 D=1 的季节差分,以 2005 年 1 月—2016 年 12 月 SARIMA 模型的拟合值作为 ERNN 模型的输入值,以对应月份的细菌性痢疾实际发病率作为 ERNN 模型的输出值。网络输入层和输出层神经元个数均为 1,根据经验,为防止

过拟合的现象出现,隐含层神经元个数采用试凑法在 4~21 之间寻找,网络训练误差结果见表 2。根据训练结果最终选择隐含层神经元个数为 20,用确定好的网络构建 SARIMA-ERNN 组合模型,然后以 2017 年 ARIMA 模型的预测结果作为输入值,使用该组合模型预测 2017 年细菌性痢疾的发病率(表 1 第 5 列),残差的 ACF 图见图 4。

表 2 不同隐含层神经个数对应的网络训练误差  
Table 2 Network training error corresponding to the number of different hidden layer nerves

神经元个数	RES	MSE
4	0.014 1	0.012 4
5	0.004 6	0.006 2
6	0.017 6	0.014 6
7	0.012 3	0.011 5
8	0.009 9	0.009 8
9	0.010 2	0.006 3
10	0.012 3	0.014 3
11	0.005 3	0.005 7
12	0.004 6	0.004 8
13	0.009 4	0.005 3
14	0.004 6	0.004 6
15	0.005 0	0.003 4
16	0.017 5	0.017 5
17	0.009 2	0.015 5
18	0.010 4	0.009 7
19	0.012 7	0.009 1
20	0.004 2	0.003 1
21	0.009 9	0.010 9

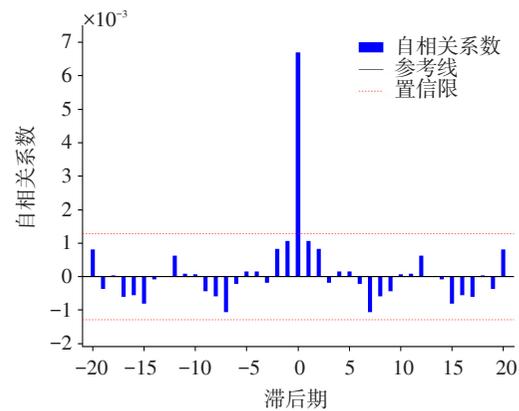


图4 SARIMA-ERNN 组合模型残差的 ACF 图

Figure 4 The ACF of residuals from SARIMA - ERNN model

### 2.4 模型比较

在拟合效果比较中 SARIMA-ERNN 组合模型的

MRE、MER、RMSE 和 MAE 都是 3 种方法中最小的;在预测效果比较中,除 ERNN 模型的 RMSE 是最小的外,MRE、MER 和 MAE 的最小值都表现在 SARIMA-ERNN 组合模型中。整体而言 SARIMA-ERNN 组合模型最优,其次为 ERNN 模型(表 3)。从 3 种模型拟合以及预测我国细菌性痢疾发病率的曲线图(图 5)可以看出,3 种模型的拟合结果都较为接近实际情况。

### 3 讨论

通过分析我国 2005—2017 年细菌性痢疾的发病序列,可以发现我国细菌性痢疾的发病率总体上

呈逐年下降趋势(年平均下降速度为 11.7%),每年的发病多集中在 6—10 月,7—8 月达到高峰,且随着时间的推移这种下降趋势逐渐趋缓。提示我国细菌性痢疾的发病率数据序列中含有线性成分和非线性成分,既具有长期趋势又具有季节性,呈现出复杂的动态性特点。随着我国医疗保健服务水平的不断提高和抗生素的应用,自上个世纪 90 年代以来我国细菌性痢疾的发病率和病死率大幅度下降<sup>[17]</sup>。国内外相关研究表明随着抗生素的广泛应用,近年来临床上频繁出现多重耐药志贺菌并迅速在不同地区间传播开来<sup>[18-19]</sup>。细菌性痢疾的主要易感人群为老人和儿童,而我国人口老龄化易感人群在总人口

表 3 三种方法对细菌性痢疾发病率的拟合及预测效果比较

Table 3 Comparison on incidence effects of bacillary dysentery fitted and predicted by three models

指标	拟合效果			预测效果		
	SARIMA 模型	ERNN 模型	SARIMA-ERNN 组合模型	SARIMA 模型	ERNN 模型	SARIMA-ERNN 组合模型
MRE (%)	5.661 37	5.348 57	4.942 52	5.596 40	5.544 30	4.251 30
MER	0.061 81	0.056 05	0.047 33	0.051 77	0.044 55	0.044 19
RMSE	0.001 45	0.017 08	0.001 15	0.004 54	0.000 36	0.000 38
MAE	0.000 94	0.000 79	0.000 72	0.000 34	0.000 30	0.000 29

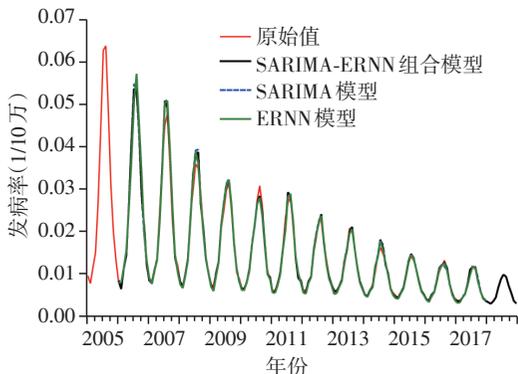


图 5 三种模型拟合及预测的细菌性痢疾发病率曲线

Figure 5 Comparison on curve of incidence of bacillary dysentery fitted and predicted by three models

中的比重逐渐增大,因此我国细菌性痢疾发病率下降趋势的趋缓可能与细菌耐药性和人口老龄化有关。以往一项研究分析了 1970—2012 年北京地区的气象因素与细菌性痢疾之间的关系,结果显示气温和降雨量与细菌性痢疾的病例数正相关<sup>[17]</sup>,青岛地区洪涝灾害对短期肠道传染病影响的研究也发现了类似的关系<sup>[20]</sup>。每年 6—10 月是我国大部分地区的雨季时间,这在一定程度上解释了我国细菌性痢疾季节性趋势的原因。

已有学者提出使用复杂的模型综合利用相关

的流行病学信息来预测传染病的发病率,比如 Cvjetanovic 等<sup>[21]</sup>在预测伤寒发病率时提出的机械模型(mechanistic models),为了解释疾病的过程,机械模型需要相关的人群流行病学信息来确定必要的传播参数以建立解释模型。然而,在中国和其他一些发展中国家,目前的公共卫生监测系统并未收集详细的基本流行病学信息,因为这些信息通常很难准确获得。相反,时间序列相对简单,需要较少的流行病学信息,作为一种有效的非解释性方法可以基于历史数据预测未来的流行趋势。因此,通过时间序列建立一个合适的细菌性痢疾发病率预测模型,及早掌握其发病趋势具有重要的现实意义。ARIMA 模型作为传统的线性模型,经过差分过程可用于分析呈季节性趋势的时间序列,已广泛用于传染病的预测。然而,实际上在许多时间序列事件中线性假设并不成立,且传染病的发病又受多种因素的影响,这就使得单独使用 ARIMA 模型预测传染病发病率受限。相反,人工神经网络时间序列模型通过非线性函数捕捉历史信息,借助灵活的非线性函数映射功能,人工神经网络可以有效地提取数据中的非线性关系,且可以以任意期望精度逼近任何连续可测函数。此类模型中的一种前馈神经网络 BPNN 已被用于传染病的预测。ERNN 作为一种典型的循

神经网络是在BPNN的基础上形成的,具有BPNN模型的优点,同时兼有BPNN模型没有的映射动态特征的功能。而组合模型具有各模型的特性,可以克服单一模型信息源不广泛的问题,能更好地捕捉数据的潜在联系,提高预测性能。然而目前预测我国传染病发病率的时间序列模型多为传统的SARIMA模型<sup>[22-24]</sup>,目前尚未见到使用SARIMA-ERNN组合模型预测我国细菌性痢疾发病率的相关研究。因此本研究建立同时具备SARIMA模型和ERNN模型优点的SARIMA-ERNN组合模型对此进行预测。

本研究结果发现,就模型的整体表现而言,SARIMA-ERNN组合模型最优,ERNN模型次之。一般情况下当模型拟合及预测的MRE不超过20%时就可以认为该模型为合格模型,若MRE小于10%则认为模型具有较高的准确度和精确度,理想状态下模型的MRE应不超过5%<sup>[25]</sup>。本研究显示SARIMA模型和ERNN模型在拟合和预测方面都达到了一定的预测准确度和精度,然而只有SARIMA-ERNN组合模型达到了理想水平,可见组合模型能更加自然而准确地捕捉到我国细菌性痢疾序列中复杂的动态规律,具有推广意义。与本研究结果一致,在预测某些传染病发病趋势时,单一的人工神经网络模型在建模和预测方面的性能都优于传统的ARIMA模型<sup>[9,26]</sup>。有学者比较了ARIMA模型与广义回归神经网络(GRNN)模型的ARIMA-GRNN组合模型与单一的ARIMA模型,ARIMA-BPNN组合模型与ARIMA模型,以及ARIMA模型与非线性自回归(NAR)模型的ARIMA-NAR组合模型与单一的ARIMA模型,在预测特定传染病时的性能,结果一致显示组合模型明显优于单一模型<sup>[27-30]</sup>。目前关于不同组合模型间的性能比较的研究较少,总的来说当理论基础不同的模型相组合有利于提高模型的预测效果<sup>[31]</sup>。

虽然组合模型和人工神经网络模型的预测性能在相关研究中都表现出一定的优越性,但SARIMA模型具有明显的建模优势且可以通过SAS、SPSS、Stata等主流统计软件轻松实现,所以SARIMA模型仍有较大的利用空间。但应该注意,首先,不同模型的组合过程并不是任意的,在选择组合模型时,不仅要考虑模型的预测精度,还要考虑原始数据的波动特征,备选的组合模型最好是基于不同的理论基础且具有相对较高的预测精度;其次,本研究结果仅基于我国细菌性痢疾的发病率资料,组合模型的性能在其他疾病预测中的作用有待于进一步验证;最后,因为基于SARIMA的组合模型是一个

短期预测模型,所以为了保持组合模型的精度和稳定性,需要通过添加新数据进行更新。

综上所述,SARIMA-ERNN组合模型在预测我国细菌性痢疾发病率时在一定程度上表现出了较高的精确度与准确度,可以用于我国细菌性痢疾的短期预测和早期预警。将来的研究可以探索构建基于其他理论基础的组合模型,并在其他传染病中进行类似的研究,以期推动在实践中引入更加先进的预测技术。

#### [参考文献]

- [1] Zhang H, Si Y, Wang X, et al. Environmental drivers and predicted risk of bacillary dysentery in Southwest China [J]. *Int J Environ Res Public Health*, 2017, 14(7): 782
- [2] Duan R, Liang J, Zhang J, et al. Prevalence of *Yersinia enterocolitica* bioserotype 3/O:3 among children with diarrhea, China, 2010-2015 [J]. *Emerg Infect Dis*, 2017, 23(9): 1502-1509
- [3] Xu C, Li Y, Wang J, et al. Spatial-temporal detection of risk factors for bacillary dysentery in Beijing, Tianjin and Hebei, China [J]. *BMC Public Health*, 2017, 17(1): 743
- [4] Xu C, Xiao G, Wang J, et al. Spatiotemporal risk of bacillary dysentery and sensitivity to meteorological factors in Hunan Province, China [J]. *Int J Environ Res Public Health*, 2017, 15(1): 47
- [5] Moosazadeh M, Nasehi M, Bahrampour A, et al. Forecasting tuberculosis incidence in Iran using box-jenkins models [J]. *Iran Red Crescent Med J*, 2014, 16(5): e11779
- [6] Kumar V, Mangal A. Forecasting malaria cases using climatic factors in Delhi, India: a time series analysis [J]. 2014, 2014: 482851
- [7] Yousefzadeh-Chabok S, Ranjbar-Taklimie F, Malekpouri R, et al. A time series model for assessing the trend and forecasting the road traffic accident mortality [J]. *Arch Trauma Res*, 2016, 5(3): e36570
- [8] Liu S, Chen J, Wang J, et al. Predicting the outbreak of hand, foot, and mouth disease in Nanjing, China: a time-series model based on weather variability [J]. *Int J Biometeorol*, 2018, 62(4): 565-574
- [9] Zhang X, Liu Y, Yang M, et al. Comparative study of four time series methods in forecasting typhoid fever incidence in China [J]. *PLoS One*, 2013, 8(5): e63116
- [10] Vickram AS, Kamini AR, Das R, et al. Validation of artificial neural network models for predicting biochemical markers associated with male infertility [J]. *Systems Biology in Reproductive Medicine*, 2016, 62(4): 258
- [11] Zhang G, Xia JJ, Liebschner M, et al. Improved Rubin-Bodner model for the prediction of soft tissue deformations

- [J]. *Medical Engineering & Physics*, 2016, 38 (11) : 1369-1375
- [12] Guan P, Huang DS, Zhou BS. Forecasting model for the incidence of hepatitis A based on artificial neural network [J]. *World J Gastroenterol*, 2004, 10(24):3579-3582
- [13] Yan W, Xu Y, Yang X, et al. A hybrid model for short-term bacillary dysentery prediction in Yichang City, China [J]. *Jpn J Infect Dis*, 2010, 63(4):264-270
- [14] Liu L, Luan RS, Yin F, et al. Predicting the incidence of hand, foot and mouth disease in Sichuan province, China using the ARIMA model [J]. *Epidemiol Infect*, 2016, 144 (1):144-151
- [15] Wang J, Wang J, Fang W, et al. Financial time series prediction using elman recurrent random neural networks [J]. *Comput Intell Neurosci*, 2016, 2016:4742515
- [16] Christodoulos C, Michalakelis C, Varoutas D. On the combination of exponential smoothing and diffusion forecasts: an application to broadband diffusion in the OECD area [J]. *Technol Forecast Soc Change*, 2011, 78(1):163-170
- [17] Yan L, Wang H, Zhang X, et al. Impact of meteorological factors on the incidence of bacillary dysentery in Beijing, China: a time series analysis (1970-2012) [J]. *PLoS One*, 2017, 12(8):e0182937
- [18] Song Q, Lin W, Gao H, et al. Changing antimicrobial resistance patterns and trends of Shigella isolates in Ningbo, Mid-East China, 2005-2013 [J]. *Int J Antimicrob Agents*, 2015, 45(5):559-560
- [19] Baker K S, Dallman T J, Behar A, et al. Travel- and community-based transmission of multidrug-resistant shigella sonnei lineage among international orthodox jewish communities [J]. *Emerg Infect Dis*, 2016, 22(9):1545-1553
- [20] Zhang F, Liu Z, Gao L, et al. Short-term impacts of floods on enteric infectious disease in Qingdao, China, 2005 - 2011 [J]. *Epidemiol Infect*, 2016, 144(15):3278-3287
- [21] Cvjetanovic B, Grab B, Uemura K. Epidemiological model of typhoid fever and its use in the planning and evaluation of antityphoid immunization and sanitation programmes [J]. *Bull World Health Organ*, 1971, 45(1):53-75
- [22] Adeola AM, Botai JO, Rautenbach H, et al. Climatic variables and malaria morbidity in mutale local municipality, South Africa: a 19-year data analysis [J]. 2017, 14 (11) : 1360
- [23] Xu Q, Li R, Liu Y, et al. Forecasting the incidence of mumps in zibo city based on a SARIMA model [J]. *Int J Environ Res Public Health*, 2017, 14 (8) : 925
- [24] Liu S, Chen J, Wang J, et al. Predicting the outbreak of hand, foot, and mouth disease in Nanjing, China: a time-series model based on weather variability [J]. *Int J Environ Res Public Health*, 2018, 62(4):565-574
- [25] Zhou L, Yu L, Wang Y, et al. A hybrid model for predicting the prevalence of schistosomiasis in humans of Qianjiang City, China [J]. *PLoS One*, 2014, 9(8):e104875
- [26] Wei W, Jiang J, Liang H, et al. Application of a combined model with autoregressive integrated moving average (ARIMA) and generalized regression neural network (GRNN) in forecasting hepatitis incidence in Heng County, China [J]. *PLoS One*, 2016, 11(6):e0156768
- [27] Wang H, Tian CW, Wang WM, et al. Time-series analysis of tuberculosis from 2005 to 2017 in China [J]. *Epidemiol Infect*, 2018:1-5
- [28] Zhang G, Huang S, Duan Q, et al. Application of a hybrid model for predicting the incidence of tuberculosis in Hubei, China [J]. *PLoS One*, 2013, 8(11):e80969
- [29] Ren H, Li J, Yuan ZA, et al. The development of a combined mathematical model to forecast the incidence of hepatitis E in Shanghai, China [J]. *BMC Infect Dis*, 2013, 13:421
- [30] Wang K W, Deng C, Li J P, et al. Hybrid methodology for tuberculosis incidence time - series forecasting based on ARIMA and a NAR neural network [J]. *Epidemiol Infect*, 2017, 145(6):1118-1129
- [31] Wu W, Guo J, An S, et al. Comparison of two hybrid models for forecasting the incidence of hemorrhagic fever with renal syndrome in Jiangsu Province, China [J]. *PLoS One*, 2015, 10(8):e0135492

[收稿日期] 2018-05-23