

ARIMA 乘积季节模型在惠州市手足口病疫情预测中的应用

谢中勇, 陈伟明, 廖康荣 (广东省惠州市第一妇幼保健院, 广东惠州 516000)

摘要:目的 探讨 ARIMA 乘积季节模型在惠州市手足口病疫情预测中作用。方法 利用 2008–2018 年惠州市手足口病月发病率, 建立 ARIMA 模型; 用该模型预测 2019 年 1–6 月发病率, 并以实际发病率评价该模型。结果 模型 ARIMA(2, 0, 1)(1, 1, 0)₁₂ 为最优模型, BIC=6.087, Ljung-Box=20.195, $P=0.124$, 模型预测值在实际发病率 95% 可信区间范围内。结论 ARIMA 模型可较好地拟合并预测手足口病月发病率。

关键词: ARIMA 模型; 手足口病; 发病率

中图分类号: R 725.1; R 181.2⁺2

文献标志码: A

文章编号: 2096-3610(2021)02-0147-04

Application of multiple seasonal ARIMA model in epidemic prediction of hand-foot-mouth disease in Huizhou city

XIE Zhong-yong, CHEN Wei-ming, LIAO Kang-rong (Huizhou First Maternal and Child Health Hospital, Huizhou 516000, China)

Abstract: Objective To investigate the role of multiple seasonal ARIMA model in epidemic prediction of hand-foot-mouth disease (HFMD) in Huizhou city. Methods ARIMA model was established using monthly incidence of HFMD in Huizhou from 2008 to 2018. The incidence of HFMD between January and June 2019 was predicted by ARIMA model and testified by actual incidence. Results ARIMA model (2, 0, 1)×(1, 1, 0)₁₂ was optimal, with BIC=6.087, Ljung-Box=20.195, and $P=0.124$. The predictive value of the model was within 95% confidence interval of actual incidence. Conclusion ARIMA model is better fitting and can predict the incidence of HFMD in Huizhou city.

Key words: ARIMA model; hand-foot-mouth disease; incidence

2008 年始手足口病(HFMD)疫情在安徽省阜阳市发生,随后在全国各地均有流行, HFMD 已成为当今社会关注的重要公共卫生问题^[1], 2008 年 5 月起 HFMD 纳入丙类传染病^[2], 给我国儿童带来严重的健康威胁和疾病负担。近年来,越来越多的数学模型在传染病早期预测中得到广泛应用,当前传染病预警模型种类较多,包括有传染病动力学模型、神经网络模型、空间统计学等^[3-5],不同的模型有其各自的特点及适用范围。1970 年提出了非平稳自回归移动平均求和模型^[6](ARIMA)及建模使用方法,广泛应用于公共卫生领域。本研究基于 ARIMA 乘积季节模型对 2019 年 1–6 月惠州市手足口病发病率的趋势进行预测。

1 资料和方法

1.1 一般资料

本研究中所有手足口病数据来自惠州市“传染病

报告信息管理系统”子系统,经漏报调查等回顾性调查进行质量控制。人口数据来自《惠州市统计年鉴》。

1.2 模型基本原理

据惠州市 2008 年 1 月至 2018 年 12 月手足口病月发病数建立时间序列。手足口病月发病数时间序列属于季节性时间序列,故采用乘积季节模型,即 ARIMA(p, d, q)×(P, D, Q)_s。其中 d 为平稳化过程中差分的阶数, p, q 为自回归和移动平均阶数。P, Q 为季节性自回归和移动平均阶数, D 为季节差分阶数, s 为季节周期。

1.3 统计学处理

利用 SPSS 25.0 统计软件,通过 ARIMA 模型分析方法建立预测模型,将 2019 年 1–6 月实际发病率与模型测算出的预测发病率与相对误差(相对误差=|实际值-预测值|/实际值)进行比较,评价模型预测效果。

2 结果

2.1 序列平稳化、平稳性检验

绘制 2008 年 1 月–2018 年 12 月惠州市手足口病发病率时间序列图(图 1),该序列为非平稳序列,需利

基金项目: 广东省医学科研基金项目(No.B2019060, No.A2021 175)

收稿日期: 2020-05-29; **修订日期:** 2020-10-15

作者简介: 谢中勇(1981–),男,硕士,副主任医师

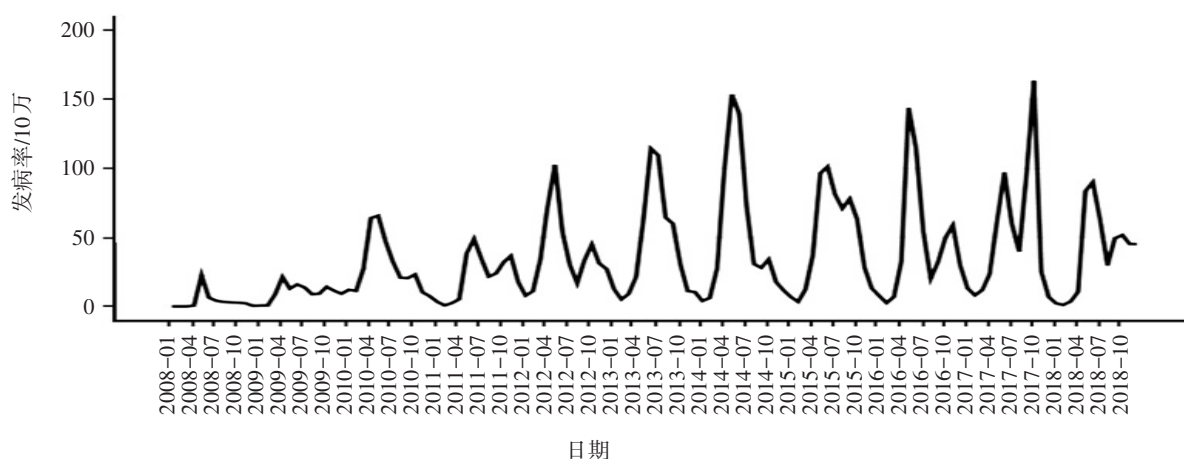


图1 惠州市手足口病发病率序列图

用季节差分,使序列平稳。由图1可见惠州市手足口病发病有明显的季节性,惠州市手足口病全年各月均有病例报告,以12个月为流行周期。每个流行周期出现2个流行高峰,大高峰出现在5~7月,小高峰出现在9~11月。差分后的时间序列自相关和偏自相关函数未见截尾和拖尾现象(图2、3),差分后的时间序列接近平稳,可以构建ARIMA乘积季节模型。

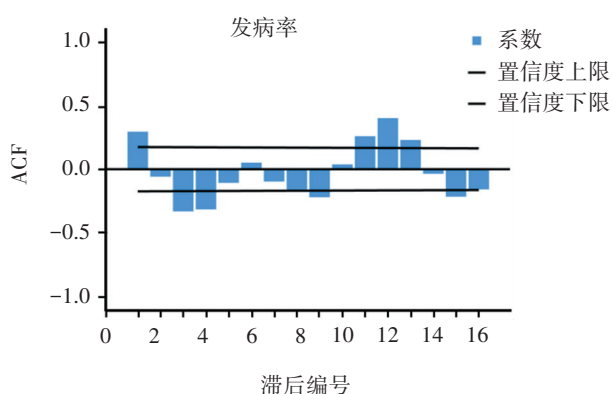


图2 原始序列经过一次季节差分后自相关系数图

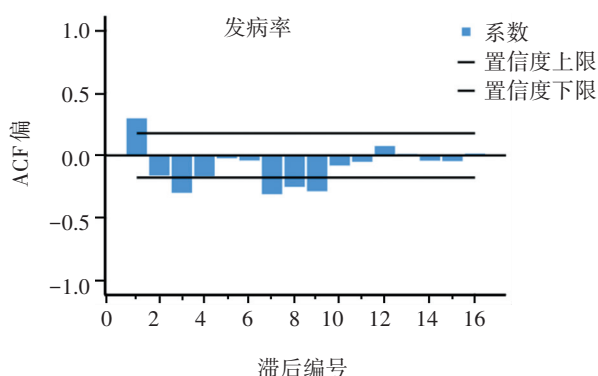


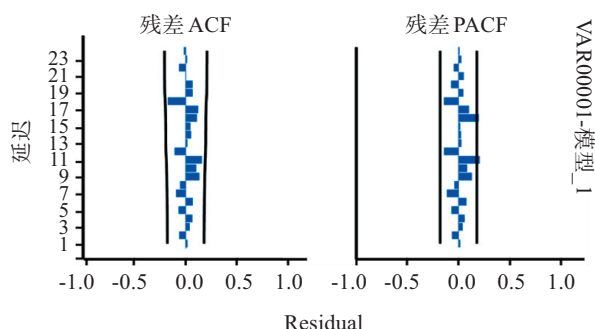
图3 原始序列经过一次季节差分后偏相关系数图

2.2 模型识别、模型参数估计和检验

根据差分后序列自相关函数图和偏自相关函数图,构建合适的ARIMA模型。本研究利用非线性最小二乘法估计模型参数,模型的拟合优度采用标准化的贝叶斯准则比较,标准化BIC值最小,并结合残差是否为白噪声、BIC值的大小、Ljung-Box Q统计量 $P>0.05$ 的模型为最优。为模型初步识别和定阶,需确定 p, d, q 和 P, D, Q 的值。惠州市手足口病月发病率序列通过一阶季节性差分后,自相关系数自二阶后截尾,偏自相关系数截尾,表明该模型应为AR(2)过程,且 $d=0, D=1, p=2, q=0, s=12$ 。根据序列季节化特征和平稳化处理过程, $d=0, D=1$ 。根据自相关函数图和偏自相关函数图, $p=1, q=1$ 。文献报道^[7],季节模型 P, Q 值较难判断,但参数 P, Q 很少超过2阶,模型形式初步判别为ARIMA($p, 0, q$)X($P, 1, Q$),通过比较模型参数、结合模型的拟合优度、残差以及系数间的相关性进行估计,采用Ljung-Box方法检验残差白噪声,非白噪声模型排除(表1)。经试验,模型ARIMA(2,0,1)(1,1,0)₁₂标准化BIC值(6.087)最小,平稳 $R^2=0.743$,Ljung-Box=20.195, $P=0.124$,残差序列的自相关系数及偏相关系数均在95%置信区间内,该模型被选为最优模型(图4)。

表1 备选ARIMA模型残差检验及拟合优度比较

序号	模型	LB		BIC	R^2
		统计量	P值		
1	ARIMA(2,0,1)(0,1,2) ₁₂	17.157	0.192	6.127	0.745
2	ARIMA(2,0,1)(1,1,0) ₁₂	20.195	0.124	6.087	0.743
3	ARIMA(2,0,1)(1,1,1) ₁₂	17.364	0.183	6.105	0.750
4	ARIMA(2,0,2)(0,1,1) ₁₂	19.302	0.114	6.178	0.732

图4 ARIMA(2,0,1)(1,1,0)₁₂模型自相关和偏相关系数图

2.3 评价模型预测效果

比较惠州市2019年1-6月的实际发病率与预测发病率相对误差,验证模型预测效果(图5,表2)。本研究针对惠州市手足口病预测数据与实际发病情况区间估计一致,在预测值95%CI内。惠州市2019年1-6月手足口病月发病率预测最小相对误差为

9.95%,手足口病发病的预测值同实际值之间存在一定差异,但包括发病高峰在内的基本趋势拟合良好,提示预测结果可信。

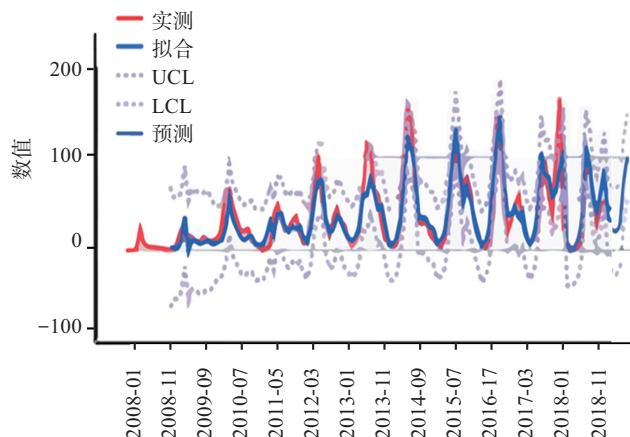


图5 2008-2018年手足口病月发病率预测图

表2 ARIMA模型实际值与预测值比较(/10万)

	1月	2月	3月	4月	5月	6月
2019年预测发病率	29.24	18.11	17.26	26.34	74.86	101.8
95%CI	-8.80~67.28	-27.47~63.7	-29.15~63.67	-20.07~72.75	28.43~121.29	55.36~148.25
2019年实际发病率	22.77	8.11	15.07	51.99	94.45	113.05
相对误差/%	22.418	125.30	14.53	49.37	20.74	9.95

3 讨论

动态监测是评估 HFMD 疫情变化的有效方法,并可提供科学依据以制定精准的防控政策^[8]。惠州市各级医疗卫生部门加强疫情监测、学校晨检等多措并举的措施,做到防治 HFMD 的关口前移^[9]。目前,及时有效地预测、预警 HFMD 发病趋势,是预防控制工作的重点和难点^[10]。

本研究运用 ARIMA 模型对 HFMD 进行短期预测,把相关因素的效应关联到时间变量中,克服了相关因素的相互影响^[11],适用于 HFMD 具有明显的季节特征变化的预测^[12]。本研究利用 2008 年 1 月至 2018 年 12 月 11 年共 132 月惠州市 HFMD 月发病率资料作为时间序列,通过拟合预测及效果评估等步骤,建立了 ARIMA(2,0,1)(1,1,0)₁₂ 为最优模型,从区间预测看,本预测模型相对误差最低为 6 月份的 9.95%,HFMD 发病的预测值同实际值之间存在一定差异,但 2019 年 1-6 月 HFMD 实际发病率基本在预测模型的 95% CI 内,与武汉、宁波、江门市等地^[13-15]应用 ARIMA 模型预测当地 HFMD 发病率的结果相似,其中以江门市预测模型的平均相对误差较小,可能与

各地人口、社会因素等不同有关^[16]。因此,ARIMA 模型能够在短期内、实际发病趋势无较大波动时,对惠州市 HFMD 发病率进行了较好的拟合预测,为 HFMD 疫情变化提供科学依据^[17]。在 HFMD 实际防控工作中,HFMD 疫情受气候因素、人口流动等诸多因素影响^[18],因此有学者根据当地的情况,采取组合模型对 HFMD 发病率进行预测^[19],减少混杂因素干扰,提高预测准确率,使其具有更高的指导价值。

参考文献:

- [1] YANG F, REN L L, XIONG Z H, et al. Enterovirus 71 outbreak in the people's republic of China in 2008[J]. J Clin Microbiol, 2009, 47(7): 2351-2352.
- [2] 佚名. 卫生部关于将手足口病纳入法定传染病管理的通知[J]. 首都公共卫生, 2008, 2(4): 145.
- [3] 吴家园, 赖天文, 刘华锋, 等. 广东省湛江市新型冠状病毒肺炎流行趋势的初步预测[J]. 广东医科大学学报, 2020, 38(2): 148-152.
- [4] LIU L, LUAN R S, 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.
- [5] LI Y, HUANG M, PENG L. A multi-group model for estima-

- ting the transmission rate of hand, foot and mouth disease in mainland China[J]. *Math Biosci Eng*, 2019, 16(4): 2305-2321.
- [6] BOXG E P, JENKINS G M. Time-Series analysis: forecasting and control[M]. San Francisco: Holden Day, 2007: 1-70.
- [7] 李晓松, 冯子健. 传染病时空聚集性探测与预测预警方法[M]. 北京: 高等教育出版社, 2014: 77-85.
- [8] 王丽萍, 曹务春. 实施传染病监测是预防控制传染病的有效途径[J]. *中华流行病学杂志*, 2017, 38(4): 417-418.
- [9] 李媛, 张振, 路滢, 等. 2005-2017年深圳市手足口病聚集性疫情流行特征及病原学分析[J]. *热带医学杂志*, 2019, 19(5): 660-663.
- [10] 潘浩, 胡家瑜, 吴寰宇, 等. GM (1, 1) 灰色模型和 ARIMA 模型在上海市手足口病发病率预测应用中的比较研究[J]. *中华疾病控制杂志*, 2011, 15(5): 445-448.
- [11] 孟凡东, 吴迪, 隋承光. 2004-2015 年中国狂犬病发病数据 ARIMA 乘积季节模型的建立及预测[J]. *中国卫生统计*, 2016, 33(3): 389-391, 395.
- [12] 陈玲, 程丽君, 赵向军. 恶性肿瘤住院量与住院费用的 ARIMA 乘积季节模型预测研究[J]. *中国卫生统计*, 2017, 34(4): 554-557.
- [13] 孙霞霞, 葛锦荣, 李巧方, 等. ARIMA 模型在宁波市北仑区手足口病发病率预测中的应用[J]. *现代预防医学*, 2018, 45(4): 582-586.
- [14] SATO R C. Disease management with ARIMA model in time series[J]. *Einstein(Sao Paulo)*, 2013, 11(1): 128-131.
- [15] 黄国, 朱宇平, 黄焕莺. 季节性 ARIMA 模型在江门市手足口病疫情预测中的应用[J]. *中国卫生统计*, 2019, 36(1): 65-67.
- [16] WANG C, CAO K, ZHANG Y, et al. Different effects of meteorological factors on hand, foot and mouth disease in various climates: a spatial panel data model analysis[J]. *BMC Infect Dis*, 2016(16): 233.
- [17] 汪鹏, 陈邦华, 孔德广, 等. 求和自回归移动平均模型在武汉市手足口病疫情预测预警中的应用[J]. *公共卫生与预防医学*, 2015, 26(1): 26-29.
- [18] 时振东, 王华义, 王加坤. 手足口病及聚集发病与气温关系[J]. *中国公共卫生*, 2014, 30(12): 1586-1588.
- [19] 王永斌, 许春杰, 尹素凤, 等. 中国手足口病发病率 ARIMA、RBF 及 ARIMA-RBF 组合模型拟合及预测效果比较[J]. *中国公共卫生*, 2017, 33(5): 760-763.

~~~~~  
(上接第 131 页)

- [2] MCCARTHY M M, PICK E, KLUGER Y, et al. HSP90 as a marker of progression in melanoma[J]. *Ann Oncol*, 2008, 19(3): 590-594.
- [3] TAS F, BILGIN E, ERTURK K, et al. clinical significance of circulating serum cellular heat shock protein 90(HSP90) level in patients with cutaneous malignant melanoma[J]. *Asian Pac J Cancer Prev*, 2017, 18(3): 599-601.
- [4] STRICKLER A G, VASQUEZ J G, YATES N, et al. Potential diagnostic significance of HSP90, ACS/TMS1, and L-plastin in the identification of melanoma[J]. *Melanoma Res*, 2014, 24(6): 535-544.
- [5] HOLMES J L, SHARP S Y, HOBBS S, et al. Silencing of HSP90 cochaperone AHA1 expression decreases client protein activation and increases cellular sensitivity to the HSP90 inhibitor 17-allylamino-17-demethoxygeldanamycin[J]. *Cancer Res*, 2008, 68(4): 1188-1197.
- [6] GUO Q C, SHEN J N, JIN S, et al. Comparative proteomic analysis of human osteosarcoma and SV40-immortalized normal osteoblastic cell lines[J]. *Acta Pharmacol Sin*, 2007, 28(6): 850-858.
- [7] SHAO J, WANG L, ZHONG C, et al. AHS1 regulates proliferation, apoptosis, migration, and invasion of osteosarcoma[J]. *Biomed Pharmacother*, 2016, 77: 45-51.
- [8] 谢汶蓉, 梁桂丽. AHA1 蛋白对恶性黑色素瘤细胞增殖的影响[J]. *广东医科大学学报*, 2018, 36(1): 32-35.
- [9] RASTRELLI M, TROPEA S, ROSSI C R, et al. Melanoma: epidemiology, risk factors, pathogenesis, diagnosis and classification[J]. *In Vivo*, 2014, 28(6): 1005-1011.
- [10] BECKER B, MULTHOFF G, FARKAS B, et al. Induction of Hsp90 protein expression in malignant melanomas and melanoma metastases[J]. *Exp Dermatol*, 2004, 13(1): 27-32.
- [11] XU W, BEEBE K, CHAVEZ J D, et al. Hsp90 middle domain phosphorylation initiates a complex conformational program to recruit the ATPase-stimulating cochaperone Aha1[J]. *Nat Commun*, 2019, 10(1): 2574.
- [12] MCDOWELL C L, BRYAN S R, OBERMANN W M. Expression of Hsp90 chaperone[corrected] proteins in human tumor tissue[J]. *Int J Biol Macromol*, 2009, 45(3): 310-314.
- [13] WORRALL C, SULEYMANOVA N, CRUDDEN C, et al. Unbalancing p53/Mdm2/IGF-1R axis by Mdm2 activation restrains the IGF-1-dependent invasive phenotype of skin melanoma[J]. *Oncogene*, 2017, 36(23): 3274-3286.