DOI:10.16781/j.0258-879x.2016.08.0969

·论　著·

基于乘积SARIMA模型的肺结核发病率预测

胡晓媛[1]，孙庆文[2]，王玲玲[3]，李　敏[1*]
1.第二军医大学海军医学系航海特殊损伤防护教研室，上海 200433
2.第二军医大学基础医学部数理教研室，上海 200433
3.解放军309医院全军结核病研究所，北京 100091

[摘要] **目的**　应用乘积季节自回归移动平均(seasonal autoregressive integrated moving average, SARIMA)模型对肺结核发病率进行预测研究，探讨其可行性并为肺结核病的防治工作提供科学依据。**方法**　应用EViews 7.0.0.1软件对我国2004年1月至2012年12月的肺结核逐月发病率建立乘积SARIMA模型并进行拟合，选取2013年1月至12月肺结核发病率数据评价模型的预测性能。**结果**　建立的SARIMA$(2,0,2)\times(0,1,1)_{12}$模型能较好地拟合既往时间段内肺结核的发病率，对2013年1月至12月肺结核发病率的预测与实际发病率趋势基本吻合，平均误差绝对值为0.416 992，平均误差绝对率为5.350 8%。**结论**　乘积SARIMA模型能较好地模拟和预测肺结核发病率在时间序列上的变动趋势，将其应用于肺结核发病预测是可行的，具有推广应用前景。

[关键词] 乘积季节ARIMA模型；肺结核；发病率；预测
[中图分类号] R 521　[文献标志码] A　[文章编号] 0258-879X(2016)08-0969-06

Multiplicative SARIMA model for prediction of pulmonary tuberculosis incidence

HU Xiao-yuan[1], SUN Qing-wen[2], WANG Ling-ling[3], LI Min[1*]
1. Department of Nautlical Injury Protection, Faculty of Naval Medicine, Second Military Medical University, Shanghai 200433, China
2. Department of Mathematics & Physics, College of Basic Medical Sciences, Second Military Medical University, Shanghai 200433, China
3. Institute for Tuberculosis Research, No. 309 Hospital of PLA, Beijing 100091, China

[**Abstract**] **Objective**　To examine the feasibility of using multiple seasonal autoregressive integrated moving average (SARIMA) model for predicting pulmonary tuberculosis (TB) incidence, so as to provide scientific evidence for the prevention and treatment of TB. **Methods**　EViews 7.0.0.1 software was used to create a SARIMA fit model for seasonal incidence of TB on a monthly basis from January 2004 to December 2012, and the predicting performance of the model was tested with TB data from January to December in 2013. **Results**　The established SARIMA $(2,0,2)\times(0,1,1)_{12}$ model could better fit with the previous TB incidence; and it basically well predicted the TB incidence of the 12 months of 2013, with the mean absolute error being 0.416 992 and the mean absolute error rate being 5.350 8%. **Conclusion**　The established multiplicative SARIMA model can better simulate and predict the trend of TB incidence with time, and it may have a future in predicting the incidence of TB.

[**Key words**] multiple seasonal ARIMA model; pulmonary tuberculosis; incidence; forecasting
[Acad J Sec Mil Med Univ, 2016, 37(8): 969-974]

世界卫生组织(WHO)于2015年10月28日发布的《2015年全球结核病报告》中指出，2014年结核病在全球范围夺去了150万人的生命，仍然是最严重的公共卫生威胁之一[1]；WHO估算我国2014年新发肺结核人数为93万，仅次于印度和印度尼西亚位，居全球第3位[1]。虽然我国近年来逐步形成了由点到面、由局部到整体、由基层到中央的网络监测系统，但由于我国地域广阔、监管机构复杂、监测资料不完善等因素，缺少有效的预测预警分析手段，未能有效掌控肺结核病的发病趋势，使得该病防治

[收稿日期] 2016-04-07　[接受日期] 2016-05-23
[基金项目] 中国博士后科学基金(2013M542491). Supported by China Postdoctoral Science Foundation (2013M542491).
[作者简介] 胡晓媛，博士. E-mail: huxiaoyuan1978@163.com
*通信作者 (Corresponding author). Tel: 021-81871120, E-mail: linlimin115@hotmail.com

工作处于被动的局面[2-3]。建立最优肺结核发病预测模型是正确预测肺结核发病水平、合理分配卫生资源和持续有效地开展肺结核病预警工作的重要前提。近年一些学者利用传染病监测数据建立自回归移动平均（autoregressive integrated moving average，ARIMA）预测疫情的发病率，取得了理想的预测效果[4-5]。本研究在时间序列分析基础上，采用乘积季节 ARIMA（seasonal ARIMA，SARIMA）模型对我国 2004 年 1 月至 2013 年 12 月的肺结核逐月发病率进行拟合和预测，以期证明逼近准确的肺结核发病率预测可以量化对该病防治的决策。

1　资料和方法

1.1　*资料来源*　原始数据来源于中华人民共和国国家卫生和计划生育委员会发布的 2004 年 1 月至 2013 年 12 月的全国法定传染病疫情概况以及国家统计局发布的 2004 年至 2013 年人口统计资料。

1.2　*研究内容*　通过研究发现我国肺结核月发病率序列的自相关系数明显有一个周期步长为 12 个月的强相关性，因此尝试采用时间序列分析法，对我国 2004 年 1 月至 2012 年 12 月的肺结核发病率原序列做周期步长为 12 个月的一阶季节差分变换，建立多个乘积 SARIMA$(p,d,q)\times(P,D,Q)_s$ 模型，对肺结核逐月发病率进行拟合，通过时间序列特征分析、模型识别、参数估计及检验、模型诊断筛选出最优预测模型，并选取 2013 年 1 月至 12 月肺结核发病率数据以评价模型的预测性能。

1.3　*研究方法*

1.3.1　*理论与模型介绍*　如果一个时间序列$\{Y_t=y_t-\mu, t=1,2,3,\cdots\}$满足如下模型[6]：

$$(1-L)^d(1-L^s)^D Y_t=\frac{\theta(L)\Theta(L)\varepsilon_t}{\phi(L)\Phi(L)}$$

则称其为季节周期 s 的乘积 SARIMA$(p, d, q)\times(P, D, Q)_s$ 模型。其中：

$$\phi(L)=1-\phi_1 L-\cdots-\phi_p L^P,$$
$$\theta(L)=1-\theta_1 L-\cdots-\theta_p L^P,$$
$$\Phi(L)=1-\Phi_1 L^s-\cdots-\Phi_p L^{Ps},$$
$$\Theta(L)=1-\Theta_1 L^s-\cdots-\Theta_p L^{Qs}$$

分别是 p 阶自回归过程 $AR(p)$ 的自回归算子、q 阶移动平均过程 $MA(q)$ 的移动平均算子、P 阶季节自回归算子、Q 阶季节移动平均算子。

1.3.2　*序列的平稳性检验、模型的识别与参数估计*　只有对于平稳时间序列，才可以使用自回归过程和移动平均过程。对于非平稳时间序列，可以尝试通过差分变换将其变为平稳时间序列。表现在上述模型中，就是差分的阶数 d、季节周期 s 和季节差分的阶数 D 的选择问题。序列的平稳性检验通常有以下 6 种：ADF 检验、DFGLS 检验、PP 检验、KPSS 检验、ERS 检验、NP 检验。至于自回归的阶数 p、季节自回归的阶数 P、移动平均的阶数 q 以及季节移动平均的阶数 Q 的选择，一般是通过观察样本的相关图和偏相关图，根据其特点，初步为模型定阶。模型的定阶可能需要反复尝试并根据模型检验和拟合的结果进行选择，并没有一个普遍适用的做法[7]。初步确定模型的阶数后，使用 EViews 7.0.0.1软件对模型参数进行估计。

1.3.3　*模型的检验与诊断*　检验的内容主要包括：一是模型参数估计时对参数的检验；二是对模型的残差序列进行白噪声检验，以确保残差序列中不再包含还可以改进模型估计的有用信息。

1.3.4　*模型的预测*　经过检验的模型，可以用于预测。对于 SARIMA 模型，预测一般分为样本内预测和样本外预测[8]。样本内预测又可以分为静态预测和动态预测。静态预测是指在预测时如果需要用到滞后期的数据，用真实数据代入；动态预测是指在预测时如果需要用到滞后期的数据，则用先前的预测数据代入，这相当于只用序列一开始的几个数据，预测其后所有的结果。因此，与静态预测相比，动态预测的效果通常要差。样本外预测，如果是多期预测，则只能用动态预测，如果是用静态预测，最多只能预测样本外的第一个时期。对模型预测结果的评估主要依据以下 4 个指标：误差均方根（RMSE）、平均绝对误差（MAE）、平均相对误差绝对值（MAPE）、Theil 不等系数（TIC）。

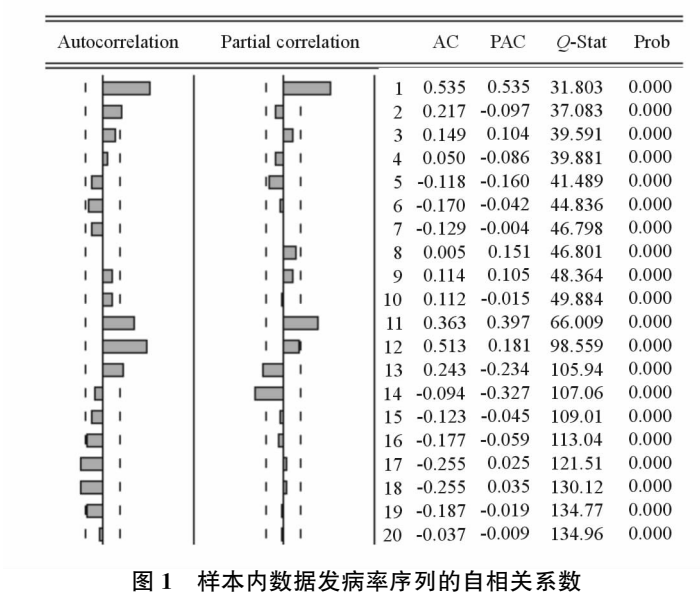
本研究将全部观测分为两部分，一部分作为样本内数据（2004 年 1 月至 2012 年 12 月的肺结核发病率），用于模型的参数估计、识别及诊断，另一部分用于样本外预测（2013 年 1 月至 12 月肺结核发病率），并与实际观测比较，以考察模型的实用性及预测性能。

1.4　*统计学处理*　应用 Excel 记录数据资料和模型预测结果；应用 EViews 7.0.0.1 进行 SARIMA 模型的参数估计、模型拟合及检验。

2　结　果

2.1　*时间序列特征分析*　从样本内月发病率时间

序列$\{IR_t, t=1,2,\cdots 108\}$的自相关系数(图 1)明显可以看出该序列存在周期步长为 12 个月的强相关性,所以本研究尝试对原序列做周期步长为 12 个月的一阶季节差分变换,使样本内数据发病率序列$D(IR,0,12)$趋于平稳(图 2)。季节差分后序列平稳性检验结果显示,在平稳性 1%水平下不能拒绝$D(IR,0,12)$是平稳序列。

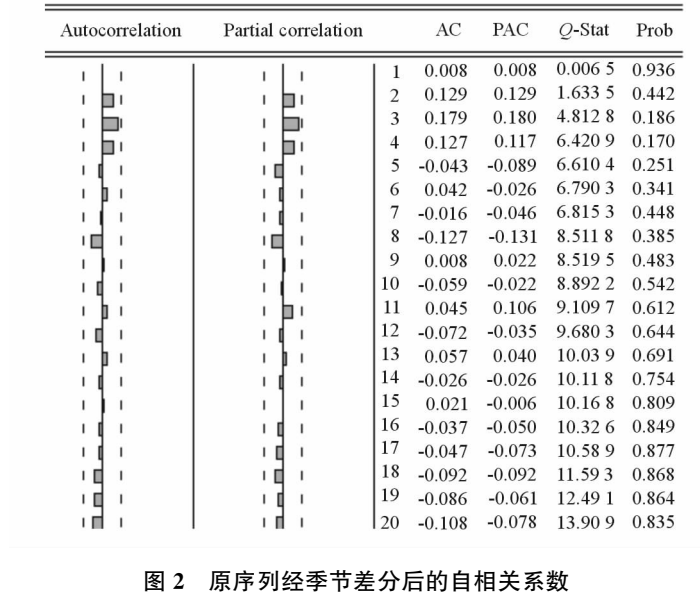


Autocorrelation	Partial correlation		AC	PAC	Q-Stat	Prob
		1	0.535	0.535	31.803	0.000
		2	0.217	-0.097	37.083	0.000
		3	0.149	0.104	39.591	0.000
		4	0.050	-0.086	39.881	0.000
		5	-0.118	-0.160	41.489	0.000
		6	-0.170	-0.042	44.836	0.000
		7	-0.129	-0.004	46.798	0.000
		8	0.005	0.151	46.801	0.000
		9	0.114	0.105	48.364	0.000
		10	0.112	-0.015	49.884	0.000
		11	0.363	0.397	66.009	0.000
		12	0.513	0.181	98.559	0.000
		13	0.243	-0.234	105.94	0.000
		14	-0.094	-0.327	107.06	0.000
		15	-0.123	-0.045	109.01	0.000
		16	-0.177	-0.059	113.04	0.000
		17	-0.255	0.025	121.51	0.000
		18	-0.255	0.035	130.12	0.000
		19	-0.187	-0.019	134.77	0.000
		20	-0.037	-0.009	134.96	0.000

图 1　样本内数据发病率序列的自相关系数

Fig 1　AC function of incidence rate sequence in the sample data

AC: Autocorrelation; PAC: Partial correlation; Prob: *Q* statistic value is greater than the probability of the *Q* value of the sample calculation



Autocorrelation	Partial correlation		AC	PAC	Q-Stat	Prob
		1	0.008	0.008	0.006 5	0.936
		2	0.129	0.129	1.633 5	0.442
		3	0.179	0.180	4.812 8	0.186
		4	0.127	0.117	6.420 9	0.170
		5	-0.043	-0.089	6.610 4	0.251
		6	0.042	-0.026	6.790 3	0.341
		7	-0.016	-0.046	6.815 3	0.448
		8	-0.127	-0.131	8.511 8	0.385
		9	0.008	0.022	8.519 5	0.483
		10	-0.059	-0.022	8.892 2	0.542
		11	0.045	0.106	9.109 7	0.612
		12	-0.072	-0.035	9.680 3	0.644
		13	0.057	0.040	10.03 9	0.691
		14	-0.026	-0.026	10.11 8	0.754
		15	0.021	-0.006	10.16 8	0.809
		16	-0.037	-0.050	10.32 6	0.849
		17	-0.047	-0.073	10.58 9	0.877
		18	-0.092	-0.092	11.59 3	0.868
		19	-0.086	-0.061	12.49 1	0.864
		20	-0.108	-0.078	13.90 9	0.835

图 2　原序列经季节差分后的自相关系数

Fig 2　AC function of original sequence with seasonal difference

AC: Autocorrelation; PAC: Partial correlation; Prob: *Q* statistic value is greater than the probability of the *Q* value of the sample calculation

2.2　模型识别　对平稳化处理后的发病率序列建立乘积 SARIMA$(p,d,q)\times(P,D,Q)_s$ 模型,根据样本相关图和偏相关图的特点反复试验,为模型定阶。

(1) 做周期步长为 12 个月的一阶季节差分变换,观察季节差分后的自相关系数图中 AC 值和 PAC 值(图 2),发现图形的时点基本都在可信区间内,因此根据季节差分后自相关系数图的特点,初步确定模型形式为 SARIMA$(p,0,q)\times(P,1,Q)_{12}$和 SARIMA$(p,1,q)\times(P,1,Q)_{12}$。

(2) 一般情况下,p、q 以及 P、Q 的识别较难,超过 2 阶的情况很少[9],因此,通过观察季节差分后的自相关系数图中 AC 和 PAC,初步确定 $P=2$,$q=1$[10]。p、q 及 P、Q 采取 0、1、2 组成的不同参数组合从低阶到高阶对模型进行反复调试和检验,确

定 3 个备选模型：SARIMA(2,0,2)×(0,1,1)$_{12}$；SARIMA(1,0,1)×(0,1,1)$_{12}$；SARIMA(0,1,1)×(0,1,1)$_{12}$。最后根据模型参数估计、残差序列的白噪声检验结果及拟合优度进行综合判断，筛选出最优预测模型。

2.3　*模型参数估计及检验*　(1) 用 EViews 7.0.0.1 分析样本内数据后实现备选模型的参数估计(表 1)，由此可见备选模型的参数估计均具有统计学意义。(2) 参数估计之后，对模型的残差序列进行白噪声检验，以确保残差序列中不再包含还可以改进模型估计的有用信息(表 2)。根据各滞后期 *Q*-Stat的 *P* 值，检验结果认为残差之间不存在相关性，即备选模型的残差序列是白噪声序列。(3) 备选模型拟合优度统计量比较显示，备选模型 SARIMA(2,0,2)×(0,1,1)$_{12}$ 的赤池信息准则(AIC)和 Schwarz 贝叶斯信息准则(SBC)值相对较小(表 3)，说明该模型与序列的拟合度最好。

表 1　备选模型参数估计及检验

Tab 1　Parameter estimation and test of alternative models

Variable	SARIMA(2,0,2)×(0,1,1)$_{12}$						SARIMA(1,0,1)×(0,1,1)$_{12}$				SARIMA(0,1,1)×(0,1,1)$_{12}$	
	C	AR(1)	AR(2)	MA(1)	MA(2)	SMA(12)	C	AR(1)	MA(1)	SMA(12)	MA(1)	SMA(12)
Coefficient	−0.266	−0.156	0.748	0.287	−0.713	−0.868	−0.232	0.860	−0.781	−0.839	−0.675	−0.865
Std. Error	0.046	0.057	0.064	0.090	0.093	0.031	0.055	0.039	0.082	0.047	0.072	0.030
t-Statistic	−5.789	−2.713	11.729	3.185	−7.706	−27.567	−4.237	22.147	−9.545	−17.839	−9.403	−29.161
Prob	0.000	0.008	0.000	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

SARIMA：Seasonal autoregressive integrated moving average；C：Constant；AR：Autoregressive parameters；MA：Moving average parameters；SMA：Seasonal moving average parameters

表 2　备选模型残差序列的白噪声检验

Tab 2　White noise test for residual series of alternative models

To lags	SARIMA(2,0,2)×(0,1,1)$_{12}$				SARIMA(1,0,1)×(0,1,1)$_{12}$				SARIMA(0,1,1)×(0,1,1)$_{12}$			
	AC	PAC	*Q*-Stat	Prob	AC	PAC	*Q*-Stat	Prob	AC	PAC	*Q*-Stat	Prob
6	0.042	−0.026	6.790	0.341	−0.014	−0.019	0.891	0.989	−0.041	−0.059	2.157	0.905
12	−0.072	−0.035	9.680	0.644	−0.038	−0.040	3.098	0.995	−0.009	−0.040	4.147	0.981
18	−0.092	−0.092	11.593	0.868	−0.072	−0.084	4.106	1.000	−0.038	−0.061	4.737	0.999
24	−0.255	−0.264	29.444	0.204	−0.105	−0.151	11.901	0.981	−0.078	−0.147	9.130	0.997

SARIMA：Seasonal autoregressive integrated moving average；AC：Autocorrelation；PAC：Partial correlation；Prob：*Q* statistic value is greater than the probability of the *Q* value of the sample calculation

表 3　备选模型拟合优度统计量

Tab 3　Goodness of fit statistics of alternative models

	SARIMA(2,0,2)×(0,1,1)$_{12}$	SARIMA(1,0,1)×(0,1,1)$_{12}$	SARIMA(0,1,1)×(0,1,1)$_{12}$
AIC	1.925 663	2.094 243	2.247 831
SBC	2.088 001	2.201 774	2.301 597

SARIMA：Seasonal autoregressive integrated moving average；AIC：Akaike information criterion；SBC：Schwarz bayesian information criterion

2.4　*模型诊断*　经过对备选模型的定阶、参数估计、残差序列白噪声检验及采用 AIC 和 SBC 衡量模型与序列的拟合度比较，最终选择我国肺结核发病率的预测模型为 SARIMA(2,0,2)×(0,1,1)$_{12}$，模型表达式为：

$$\nabla_{12} IR_t + 0.266\ 8 = \frac{(1-0.286\ 848L+0.712\ 964L^2)(1+0.868\ 312L^{12})\varepsilon_t}{(1+0.155\ 758L-0.748\ 1L^2)}$$

其中 IR_t 是发病率，∇_{12} 表示滞后期为 12 的一阶差分，$\nabla_{12} IR_t = IR_t - IR_{t-12}$，$L^s$ 是滞后算子，$L^s\varepsilon_t = \varepsilon_{t-s}$

2.5　*模型验证与预测*　模型验证：由图 3 可见，SARIMA(2,0,2)×(0,1,1)$_{12}$ 模型样本内静态预测、样本内动态预测与肺结核实际发病数的拟合程度较好，实际值均在预测区间(95%CI)以内，且预测趋势与实际趋势基本吻合，符合本次建模要求。

模型预测：利用建立的 SARIMA(2,0,2)×(0,1,1)$_{12}$ 模型，预测 2013 年 1 月至 12 月发病率。由图 4 可见，2013 年 1 月至 12 月发病率真实值均在该模型预测区间(95%CI)以内，且预测趋势与实际趋势基本吻合。由表 4 可见该模型预测值误差绝对值在0.036 673～1.239 393，误差绝对率在 0.408 454%～16.179%。平均误差绝对值为 0.416 992，平均误差绝对率为 5.350 8%，提示该模型具有较好的预测性能。

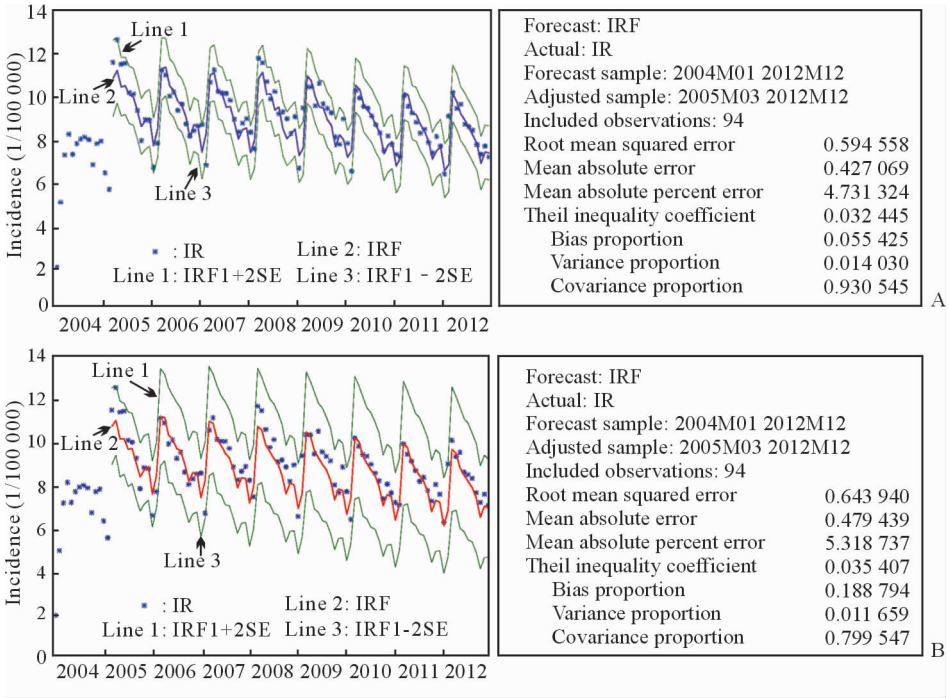


图 3　SARIMA(2,0,2)×(0,1,1)$_{12}$模型样本内静态预测拟合图(A)和动态预测拟合图(B)

Fig 3　The fitting chart of static forecast (A) and dynamic forecast (B) in the sample by SARIMA(2,0,2)×(0,1,1)$_{12}$ model

IR: Incidence rate; IRF: Incidence rate forecast value; SE: Standard error; SARIMA: Seasonal autoregressive integrated moving average

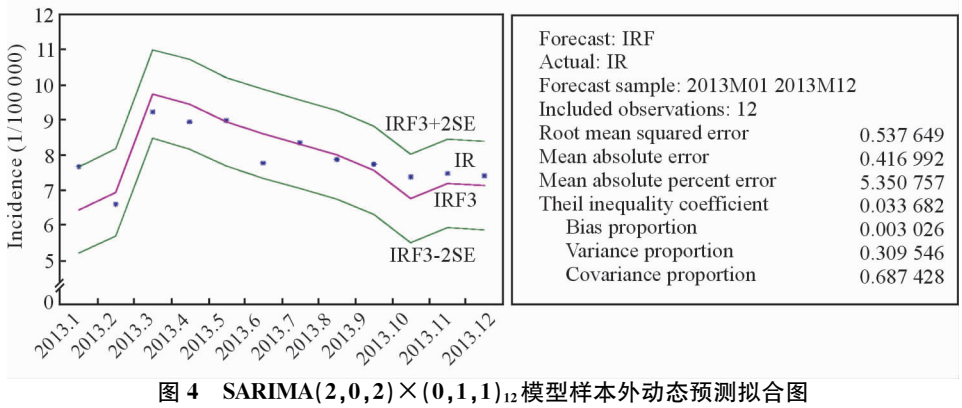


图 4　SARIMA(2,0,2)×(0,1,1)$_{12}$模型样本外动态预测拟合图

Fig 4　The fitting chart of dynamic forecast outside the sample by SARIMA(2,0,2)×(0,1,1)$_{12}$ model

IR: Incidence rate; IRF: Incidence rate forecast value; SE: Standard error; SARIMA: Seasonal autoregressive integrated moving average

表 4　SARIMA(2,0,2)×(0,1,1)$_{12}$模型对 2013 年肺结核发病率的预测结果

Tab 4　SARIMA(2,0,2)×(0,1,1)$_{12}$ model in prediction of pulmonary tuberculosis incidence of 2013

Month	Actual value	Predicted value	Absolute error value	Absolute error rate(%)	Month	Actual value	Predicted value	Absolute error value	Absolute error rate(%)
2013. 1	7. 660 503	6. 421 110	1. 239 393	16. 179 000	2013. 7	8. 343 451	8. 300 335	0. 043 116	0. 516 765
2013. 2	6. 583 426	6. 923 418	0. 339 990	5. 164 332	2013. 8	7. 867 232	7. 996 233	0. 129 000	1. 639 713
2013. 3	9. 226 659	9. 733 268	0. 506 610	5. 490 720	2013. 9	7. 732 597	7. 554 359	0. 178 238	2. 305 021
2013. 4	8. 943 868	9. 446 776	0. 502 910	5. 622 959	2013. 10	7. 367 717	6. 750 824	0. 616 893	8. 372 919
2013. 5	8. 978 482	8. 941 809	0. 036 673	0. 408 454	2013. 11	7. 466 268	7. 181 030	0. 285 238	3. 820 356
2013. 6	7. 756 629	8. 602 610	0. 845 980	10. 906 540	2013. 12	7. 399 391	7. 119 526	0. 279 865	3. 782 271

SARIMA: Seasonal autoregressive integrated moving average

3 讨　论

建立最优预测模型，提高预测准确性，对于肺结核病防治工作意义重大。在肺结核流行趋势的预测中，数学模型起着极其重要的作用。目前有多种数学模型被应用于传染病预测，如灰色预测模型、马尔科夫链预测模型[11]、回归模型[12]、神经网络模型(ANN)[13]、ARIMA 模型等[14]。其中 ARIMA 模型是对有时间性变动的序列提取季节趋势建立模型的方法，其精确度较高，是传染病时间序列预测模型中最重要的手段[15]。因此，对于肺结核发病率的预测，采用乘积 SARIMA 模型是较为理想的。

本研究基于样本内数据，通过时间序列特征分析、模型识别、模型参数估计及检验、模型诊断建立最优预测模型——乘积 SARIMA$(2,0,2)\times(0,1,1)_{12}$ 模型。预测结果显示 2013 年 1 月至 12 月发病率真实值均在该模型预测区间(95%CI)以内，且预测趋势与实际趋势基本吻合，平均误差绝对值为0.416 992，平均误差绝对率为 5.350 8%，提示该模型对肺结核发病趋势进行了较准确的跟踪和预测，能较好地模拟和预测肺结核发病率在时间序列上的变动趋势，可以为肺结核的控制和管理提供量化参考依据。

乘积 SARIMA 模型是在样本内数据发病率序列趋于平稳的基础上较好地拟合了历史数据，是精度较高的短期预测方法。当用于长期预测时，应根据监测数据调整模型的各项参数。若序列太短，则可靠性较差。在时间序列预测模型中，对于模型的定阶，没有固定的统一的规则可以严格遵循，因此，模型的定阶过程存在一定的主观性，需要摸索和试错。如果环境因素或其他影响发病率的外界因素发生较大的变化或突变，则原来时间序列的内部规律会受到冲击和破坏，预测模型的外推(样本外的预测)可能不再有效。

乘积 SARIMA 模型是从数据上反映肺结核发病率的发展趋势，可以为肺结核的防治工作提供数据参考依据，建议有关部门在制定具体肺结核防控策略时还应考虑其他综合因素的影响。

[参 考 文 献]

[1] World Health Organization. Global Tuberculosis Report 2015[R/OL]. [2016-07-02]. *http://www.who.int/tb/publications/global_report/gtbr15_executive_summary_zh.pdf?ua=1*.

[2] 杨　召，叶中辉，尤爱国，郭奕瑞，张肖肖，梁淑英，等. 乘积季节 ARIMA 模型在结核病发病率预测中应用[J]. 中国公共卫生，2013，29：469-472.

[3] 牛成虎，梅光辉，石　敏，高红韦. 我国肺结核发病率的发展动向及预测研究[J]. 现代生物医学进展，2009，9：561-564.

[4] HU W，TONG S，MENGERSEN K，CONNELL D. Weather variability and the incidence of cryptosporidiosis：comparison of time series poisson regression and SARIMA models[J]. Ann Epidemiol，2007，17：679-688.

[5] GOMEZ-ELIPE A，OTERO A，VAN HERP M，AGUIRRE-JAIME A. Forecasting malaria incidence based on monthly case reports and environmental factors in Karuzi，Burundi，1997-2003[J]. Malar J，2007，6：129.

[6] 赵国庆. 经济分析中的时间序列模型[M]. 天津：南开大学出版社，2012：17-22.

[7] CRYER J D，CHAN K S. 时间序列分析及应用[M]. 2 版. 潘红宇，译. 北京：机械工业出版社，2011：40-58.

[8] 易丹辉. 数据分析与 EViews 应用[M]. 北京：中国人民大学出版社，2008：141-148.

[9] 温　亮，徐德忠，林明和，夏结来，张治英，苏永强. 应用时间序列模型预测疟区疟疾发病率[J]. 第四军医大学学报，2004，25：507-510.

[10] 宇传华. SPSS 与统计分析[M]. 北京：电子工业出版社，2007：577-614.

[11] UYS P W，VAN HELDEN P D，HARGROVE J W. Tuberculosis reinfection rate as a proportion of total infection rate correlates with the logarithm of the incidence rate：a mathematical model[J]. J R Soc Interface，2009，6：11-15.

[12] SHILOVA M V，GLUMNAIA T V.[Prediction of the rate of tuberculosis mortality (calculation methodology)][J]. Probl Tuberk Bolezn Legk，2006(1)：22-28.

[13] HAMDY K E，YOSRY A A，FARAG I Y. Prediction of hourly and daily diffuse fraction using neural network，as compared to linear regression models[J]. Energy，2007，32：1513-1523.

[14] 胡晓媛，吴　娟，孙庆文，沙　琨，王玲玲，李　敏. ARIMA 模型与 GRNN 模型对肺结核发病率预测的对比研究[J]. 第二军医大学学报，2016，37：115-119.
HUI X Y，WU J，SUN Q W，SHA K，WANG L L，LI M. Comparative study on ARIMA model and GRNN model for predicting the incidence of tuberculosis[J]. Acad J Sec Mil Med Univ，2016，37：115-119.

[15] 孟　蕾，王玉明. ARIMA 模型在肺结核发病预测中的应用[J]. 中国卫生统计，2010，27：507-509.

[本文编辑]　魏学丽