

文章编号 1001-8166(2004)06-0925-06

SARS 传播时间过程的参数反演和趋势预测

韩卫国^{1,2}, 王劲峰¹, 刘旭华^{1,2}

(1. 中国科学院地理科学与资源研究所资源与环境信息系统国家重点实验室, 北京 100101;

2. 中国科学院研究生院, 北京 100039)

摘要 以公布的香港和北京 SARS 疫情数据为实例, 采用 SIR 模型对 SARS 传播的时间过程进行参数反演, 获取两地 SARS 高峰期、住院人数、移出系数等重要参数, 模型计算结果与实际数据基本相符, 通过参数反演很好地解释了 SARS 时间传播过程, 说明 SIR 模型可以用于 SARS 传播的数据拟合、趋势预测和过程模拟。

关键词 SIR 模型 模拟退火法 刚性问题 非线性极值

中图分类号 TP79 R181.8 **文献标识码** A

0 引言

SARS(严重急性呼吸系统综合症)作为一种冠状病毒感染的新颖传染性疾病, 它的爆发和蔓延极大地威胁到人类的身体健康和生命安全, 直接影响到社会稳定与经济发展。自世界卫生组织公布该传染病以来, 许多国家的公共卫生部门和科研院所对研究这种传染病的传播和感染规律、预防和控制措施、疫苗和特效药物等方面投入了大量的人力和物力, 已取得了可喜的成就。我国政府也启动了“863”计划、国家自然科学基金等多个 SARS 应急课题, 这些课题的研究对认识 SARS 传播与扩散的时空规律、分析 SARS 的发展态势、辅助传染病预防和控制决策起到了很好的帮助作用^[1-5]。

通过数学模型来描述 SARS 的传播过程、分析感染人数的变化规律、预测 SARS 的高峰期、评估政府措施效果等, 已有大量的文献见诸于各种科学杂志和互联网。这些研究方法大致分为 2 类: 采用确定性的微分方程或动力学模型^[1, 3, 9-11, 13, 14, 20, 21];

基于统计学、随机过程和时间序列分析对数据进

行统计建模^[7, 8, 12, 19]。

王铎等^[6]采用 SIR 模型的简化形式(见式(6))进行数据拟合和趋势预测, 取得了较好的结果, 但是该计算未能直接利用 SIR 模型进行计算, 无法获取模型的各个参数来进行传染病分析。若直接采用 SIR 模型进行计算, 存在着刚性微分方程求解和非线性极值问题, 需要采用新的方法来解决。本文主要采用 SIR 模型对 SARS 的时间传播过程进行参数反演和趋势预测, 将模拟退火法和刚性微分方程求解算法相结合用于 SIR 模型的参数反演, 解决了刚性微分方程求解和非线性极值问题, 并以北京和香港两地公布的 SARS 疫情数据, 采用上述方法进行参数反演, 计算结果说明通过 SIR 模型可以预测 SARS 的动态变化趋势, 并可用于 SARS 时间传播过程的情景模拟。

1 SIR 模型

SIR 模型由 Kermack 等在 1927 年提出, 该模型不是从医学的角度考虑各种传染病的病理知识, 而是按照一般的传染病传播机理建立, 并通过多次传

收稿日期: 2004-01-29, 修回日期: 2004-05-24.

基金项目: 国家高技术研究发展计划(863 计划)项目“SARS 流行病学资料的实时收集、分析和趋势预测”(编号: 2003AA208401), 国家自然科学基金项目“SARS 传播时空模型研究”(编号: 40341002)资助。

作者简介: 韩卫国(1976-) 男, 山西省夏县人, 博士研究生, 主要从事空间分析、时空数据挖掘、GIS 应用与开发方面研究。

E-mail: hanwg@reis.ac.cn

染病(如印度孟买 1905—1906 年发生的瘟疫)数据的验证。在模型中,将人群分为易感染者 S (Susceptible)、感染者 I (Infective) 和移出者 R (Removed) 3 类,传播过程如图 1。由于大多数传染病如天花、流感、麻疹等治愈后,病愈的人均有很强的免疫力,他们既非健康者,也不会再次患上该病,退出了传染系统^[15, 16, 22, 23],SARS 正好符合这类传染病的传染规律,设定 s 为易感人数, i 为 SARS 感染人数, r 为移出人数(包括治愈、死亡、隔离等),采用 SIR 模型进行 SARS 时间传播过程的参数反演和过程预测。

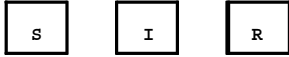


图 1 SIR 模型
Fig.1 SIR model

模型具体形式如下:

$$\frac{ds}{dt} = -kS(t)I(t) \quad (1)$$

$$\frac{di}{dt} = kS(t)I(t) - \mu I(t) \quad (2)$$

$$\frac{dr}{dt} = \mu I(t) \quad (3)$$

式中 $s(t)$ 、 $i(t)$ 和 $r(t)$ 分别表示 t 时刻易感人数、感染人数和移出人数,满足式(4):

$$s(t) + i(t) + r(t) = N \quad (4)$$

N 为该地区的总人数。 k 为感染者的日接触人数占总人数比率, μ 为日移出人数和感染人数比率。

设 s_0 为初始易感人数,假定 $r(0) = 0$,由式(1)、(3)和(4)得出,

$$\frac{dr}{dt} = \mu(N - r(t) - s_0 \exp(-\frac{r(t)}{s_0})) \quad (5)$$

其中, $\mu = 1/k$ 称为阈值,当 $s_0 > \mu$ 时,传染病开始流行,否则传染病不会流行。

对式(5)进行泰勒展开(假定 $r \ll N$),化简后生成常微分方程,对其求解,得到式(6):

$$\frac{dr}{dt} = \frac{1}{2s_0} \operatorname{sech}^2\left(\frac{r}{2}\right) \quad (6)$$

其中, $\mu = \frac{1}{2s_0} \frac{1}{\left(\frac{s_0 - 1}{s_0} + 2s_0(N - s_0)\right)^2}$,
 $= \operatorname{tanh}^{-1}\left(\frac{s_0 - 1}{s_0}\right)$

由式(3)和(5)得到 $i(t)$ 的表达式:

$$i(t) = N - r(t) - s_0 \exp(-\frac{r(t)}{s_0}) \quad (7)$$

2 方法

模拟退火法是近年提出的一种适合解大规模组

合优化问题的通用有效近似算法。该算法将固体在恒定温度下达到热平衡过程的模拟引入优化过程中来,是一种对 Metropolis 算法进行迭代的组合优化算法。该算法允许任意选取初始解和随机数序列,具有描述简单、运行高效、健壮、运用广泛和使用灵活等特点,适用于多种组合优化问题,同样也可用于非线性优化问题^[17]。

刚性微分方程的特点是方程右端的 Jacobi 矩阵具有广泛散布的特征值,没有一个解分量是不稳定的,但至少有一个解分量是非常稳定的,这样造成计算求得的解是缓慢变化的,但又存在迅速衰减的扰动,这种扰动的出现,给慢变解的数值计算带来很大困难^[18]。本文采用隐式 Runge-Kutta 微分方程求解方法,保证求解过程的稳定性和解的收敛性。

将模拟退火法和隐式 Runge-Kutta 刚性微分方程求解算法组合用于反演 SIR 模型参数,不仅有效克服了传统优化算法进行计算结果严重依赖于参数初值,搜索易陷入局部最优等缺陷,而且较大地提高了计算效率及收敛精度,解决了刚性微分方程求解和非线性极值问题,这种方法对于此类模型的参数反演具有普遍的适用性。求解步骤如下:

(1) 给定 r 和 i 的取值区间和初值。

(2) 代入式(5),采用隐式 Runge-Kutta 法计算该刚性微分方程的数值解。

(3) 计算数值解与实际数据的差值平方和,作为模拟退火法的目标函数值。

(4) 如果平方和满足指定条件,得到反演参数解,退出计算。否则,通过模拟退火法在取值区间内重新取值,若该值为边界值,转至步骤 5,否则转至步骤 2,进行下一次计算。

(5) 重新设定取值区间和初值,转至步骤 2,继续计算。最后,反演得到 SIR 模型参数解。

3 实例分析

3.1 北京数据

在 SIR 模型的参数反演计算中,采用北京市 4 月 20 日至 6 月 23 日公布的 SARS 传染病的各项数据,通过每日公布的累计确诊病例、累计死亡病例和累计治愈病例,计算得到当日住院治疗病例(当日住院病例 = 累计确诊病例 - 累计死亡病例 - 累计治愈病例),如图 2 所示:

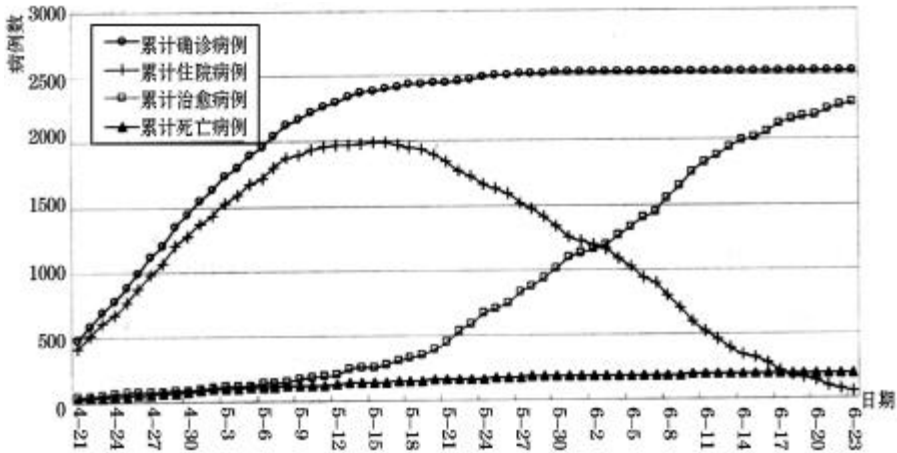


图 2 北京 SARS 疫情公布数据

Fig.2 SARS data released by Beijing health authority

经研究发现,部分与 SARS 病人没有接触的人自身携带冠状病毒的抗体,对 SARS 具有免疫力,以及存在部分携带冠状病毒但并不具有传染性的隐性感染者,所以 $R(t)$ 和 dR/dt 的数据无法根据公布数据确定。在计算中,取每日住院病例为 $I(t)$,对(5)式进行刚性微分方程组求解后,由(7)式得到 $I(t)$,

其中 N 取 2002 年北京市公布人口数 14 223 000(包括常住人口和流动人口),按照上节所述方法反演得到 SIR 模型的参数 β 和 γ ,其中 β 为 5.734136, γ 为 13 982 845.5。根据参数绘制出 $I(t)$ 的曲线,与实际住院病例数比较如图 3 所示。

由图 3 可以看出,北京 SARS 爆发于 2003 年 3

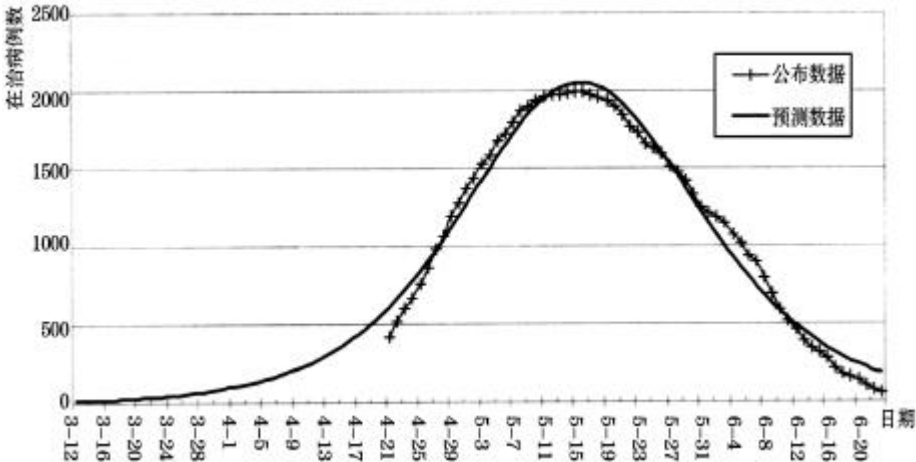


图 3 北京 SARS 疫情 SIR 模型结果与实际数据比较

Fig.3 Comparison of SIR model solution and actual SARS data of Beijing

月初,缓慢增长期为 3 月 24 日至 4 月 18 日,快速发展期为 4 月 19 日至 5 月 13 日,高峰期 5 月 20 日左右,下降期为 5 月 27 日至 6 月 22 日,消退期为 6 月 23 日至 7 月 20 日,全部历时为约 5 个月。比较

和 $S_0, \beta_0 > \gamma$,SARS 开始流行;当 $S(t) = S_0$ 时, $I(t)$ 达到最大值,接着开始下降。预测最大住院人数为 2 050 人左右,实际最大住院人数为 1 991 人。与实际数据相比,平均绝对误差为 30 人,平均相对误差

为 15%。将反演参数代入 SIR 模型计算的结果与 SARS 疫情的实际数据基本吻合,并可反推出 4 月 20 日前的理论在治病例数。

3.2 香港数据

在 SIR 模型进行香港 SARS 的参数反演计算时,采用香港政府公布的 3 月 17 日至 6 月 23 日的

疫情数据,采用同样的计算方法得到每日住院治疗病例如图 4 所示:

采取同样的方法反演参数,其中 N 取 2003 年香港政府公布的最新人口数 6 708 000,计算得到参数和 1 其中 1 为 5.309323, 为 6 602 185。根据参数绘制出 $x(t)$ 的曲线,与住院病例数比较如图 5 所示:

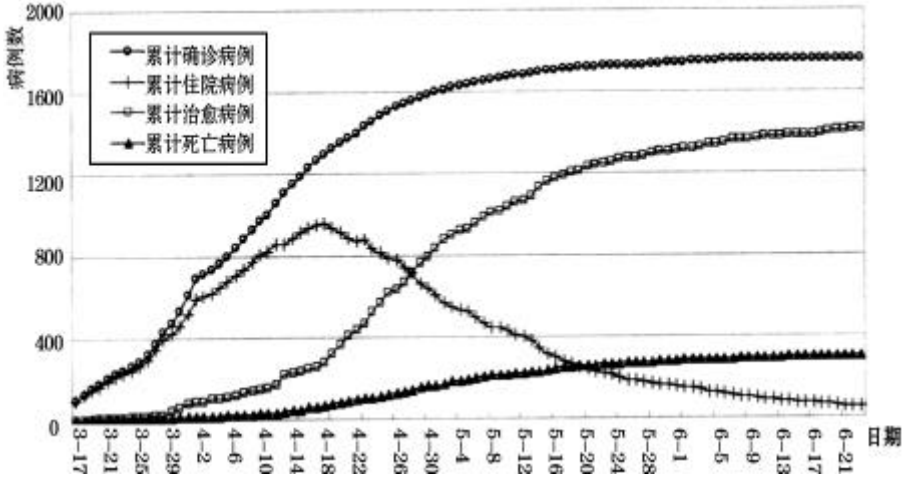


图 4 香港 SARS 疫情公布数据

Fig.4 SARS data released by Hong Kong health authority

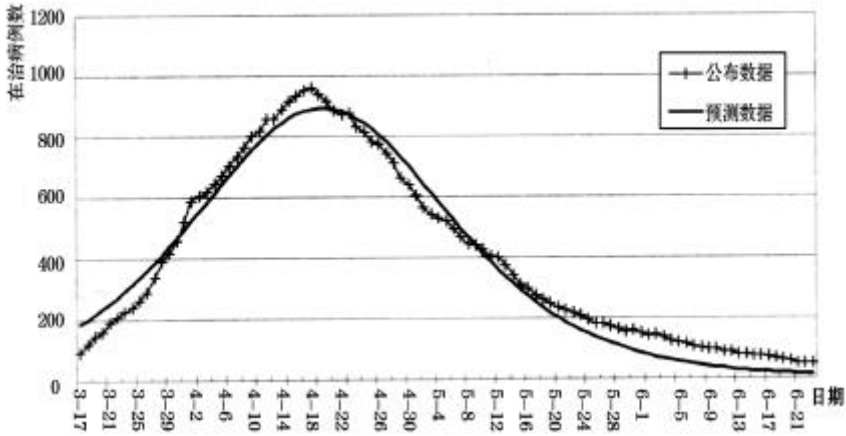


图 5 香港 SARS 疫情 SIR 模型结果与实际数据比较

Fig.5 Comparison of SIR model solution and actual SARS data of Hong Kong

由图 5 可以看出,香港 SARS 传播的周期比北京早 1 个月,爆发于 2003 年 2 月初,缓慢增长期为 2 月 10 日至 3 月 17 日,快速发展期为 3 月 18 日至 4

月 15 日,高峰期 4 月 20 日左右,下降期为 4 月 27 日至 5 月 19 日,消退期为 5 月 21 日至 6 月 20 日。同样和 s_0 的比较与北京有同样的规律。预测最大

香港 SARS 疫情数据来源于香港政府网站 <http://sc.info.gov.hk/gb/www.info.gov.hk/isd/news/cindex.htm>

住院人数为 910 人左右,实际最大住院人数为 960 人左右。与实际数据相比,平均绝对误差为 35 人,平均相对误差为 18%。

比较发现香港和北京的 R 相差不大,这与两地均为人群密集的国际大都市以及相近的医疗水平和设施相符合。同样比较 N ,两者都在 0.9833 左右,说明 SARS 在两地的传播基本遵循相似的规律。而香港的传播周期早于北京 1 个月,它的计算结果的参数可以作为北京防治 SARS 的参考,辅助预测北京的最大住院人数,确定开设专门治疗医院数。作者于 5 月 9 日曾采用当时香港疫情数据的反演参数做过北京 SARS 传播趋势的预测,结果与后来的发展趋势基本相同。

两地的计算结果说明 SIR 模型可以较好地解释 SARS 的动态变化趋势,通过参数反演很好地解释了 SARS 时间传播过程,SARS 的传播规律仍旧符合一般的传染病规律。

4 讨论与展望

SARS 结束后,北京和香港的卫生部门对两地的 SARS 确诊病例进行复查,发现有一部分病例属于误诊,但是详细的误诊信息(如误诊病例数、哪天误诊)无法获取到。如果剔除误诊病例的数据,采用 SIR 模型进行参数反演和过程预测,计算出来的参数会与现在的计算结果有所不同,但传播的趋势和时间过程不会改变。

总的来说,SIR 模型可以对 SARS 传播的时间过程进行参数反演和趋势预测,这些对认识 SARS 传播规律和指导控制策略的制定有着重要的意义。通过上述计算,可以看出减少 SARS 患者的接触人数,尤其是在发现首例患者后,对于控制传染病的传播和蔓延有着非常重要的作用。及时发现患者、严格的隔离、有效的防护以及切断传播途径等措施可以极大地降低传播率和感染率,有效地控制 SARS 和其他传染病传播。

在采用 SIR 模型或其他模型进行 SARS 传播过程的趋势预测时,对感染人数、移出人数等的理解不同,会造成模型的计算结果不同。SIR 模型在以下几个方面可以进一步改进和完善:

(1) 模型中没有把疾病流行期间的人员流动、严厉的隔离措施、输入性病例、疑似病例、隐性感染等因素考虑进来,不能准确预测模型中的其他未知量(如移出人数)模型需要进一步细化,考虑改进为分段形式。

(2) 通过模型计算,我们也发现缺乏详尽的、真实反映 SARS 实际情况的数据,无法合理确定部分参数的具体取值,模型的分析需要更准确的数据。

(3) SIR 模型只能反映传染病时间传播过程,不能反映空间传播过程,为此我们提出了传染病时空传播模型^[24],正在进行验证和改进。

致谢:感谢李小文院士、郑晓瑛、曾光、张霆教授的指导和支持。

参考文献(References):

- [1] Gong Jianhua(龚建华), Sun Zhanli(孙战利), Li Xiaowen(李小文), et al. Simulation and analysis of control of SARS[J]. Journal of Remote Sensing(遥感学报), 2003, 7(4): 260-265 (in Chinese).
- [2] Yan Shouyong(阎守崑), Liu Yalan(刘亚岚), Li Xiaowen(李小文), et al. SARS spatial transmission pattern and its application to design of disease surveillance and control information system[J]. Journal of Remote Sensing(遥感学报), 2003, 7(4): 266-272 (in Chinese).
- [3] Shi Yaolin(石耀霖). Dynamics stochastic model of SARS transmission[J]. Chinese Science Bulletin(科学通报), 2003, 48(7): 1 373-1 377 (in Chinese).
- [4] Meng Wei(孟伟), Luo Hong(罗宏). SARS transmission and environmental safety[J]. Research of Environmental Sciences(环境科学研究), 2003, 16(3): 1-2 (in Chinese).
- [5] Yang Zhicong(杨智聪), Du Lin(杜琳), Wang Ming(王鸣), et al. Analysis of effect of air pressure and air temperature on outbreak and epidemic of SARS[J]. Chinese Journal of Public Health(中国公共卫生), 2003, 19(9): 1 028-1 030 (in Chinese).
- [6] Wang Duo(王铎), Zhao Xiaofei(赵晓飞). Empirical analysis and forecasting of SARS epidemic situation[J]. Journal of Peking University(Health Sciences)(北京大学学报,医学版), 2003, 35(81): 72-74 (in Chinese).
- [7] Wang Zhengxing(王正行), Zhang Jianwei(张建玮), Tang Yinnan(唐毅南). A model analysis and evaluation of the SARS epidemic in Beijing[J]. Physics(物理), 2003, 32(5): 341-344 (in Chinese).
- [8] Ye Yanlin(叶沿林), Pang Danyang(庞丹阳), Liu Xunxu(刘循序). Analysis of the current SARS epidemic and the forecast for Beijing[J]. Physics(物理), 2003, 32(5): 345-347 (in Chinese).
- [9] Yang Fangting(杨方廷), Hou Lihua(侯立华), Han Jun(韩军), et al. Simulation analysis for the SARS epidemiological process in Beijing[J]. Journal of System Simulation(系统仿真学报), 2003, 15(7): 2 991-2 994 (in Chinese).
- [10] Chen Jiong(陈吉荣), Yang Fangting(杨方廷), Zhan Shouyi(战守义), et al. Processing on the parameters and initial values of SARS simulation model[J]. Journal of System Simulation(系统仿真学报), 2003, 15(7): 2 995-2 998 (in Chinese).

- [11] Huang Desheng(黄德生), Guan Peng(关鹏), Zhou Baosen(周宝森). Fitness of SIR model and discussion of epidemic characteristic of SARS based on logistic model[J]. Chinese Journal of Public Health(中国公共卫生), 2003, 19(6): T1-T2(in Chinese).
- [12] Fang Zhaoben(方兆本), Li Hongxing(李红星), Yang Jianping(杨建萍). Models and prediction of the SARS prevalence[J]. Application of Statistics and Management(数理统计与管理), 2003, 22(5): 48-53(in Chinese).
- [13] Li Zhonglai(李仲来), Cui Hongjian(崔恒建), Yang Hua(杨华), et al. SIM models and piecewise SIM models on SARS forecasting[J]. Journal of Remote Sensing(遥感学报), 2003, 7(5): 345-349(in Chinese).
- [14] Wu Kaichen(吴开琛), Wu Kaiyu(吴开昱), Chen Wenjiang(陈文江), et al. Mathematical model and prediction of epidemic trend of SARS[J]. China Tropical Medicine(中国热带医学), 2003, 3(4): 421-426(in Chinese).
- [15] Jiang Qiyuan(姜启源). Mathematical Models[M]. Beijing: Higher Education Press, 2001. 111-119(in Chinese).
- [16] Yan Gang(颜刚). Further study of SIR model[J]. Journal of First Military Medical University(第一军医大学学报), 2001, 21(2): 141-142(in Chinese).
- [17] Kang Lishan(康立山), Xie Yun(谢云), You Shiyong(尤矢勇), et al. Nonlinear Parallel Algorithm - Simulated Anneal Algorithm[M]. Beijing: Science Press, 2000. 22-55(in Chinese).
- [18] Xu Xuhai(徐绪海), Zhu Fangsheng(朱方生). Numerical Method of Stiff Differential Equation[M]. Hubei: Wuhan University Press, 1997. 16-34(in Chinese).
- [19] Donnelly Christl, Ghani Aza, Leung Gabriel, et al. Epidemiological determinants of spread of causal agent of severe acute respiratory syndrome in Hong Kong[EB/OL]. http://www.who.int/csr/don/20030304_sars_hk.pdf, 2003, 5.
- [20] Riley S, Fraser C, Donnelly C A, et al. Transmission dynamics of the etiologic agent of SARS in Hong Kong: Impact of public health interventions[J]. Science, 2003, 300: 1961-1966.
- [21] Lipsitch M, Cohen T, Cooper B, et al. Transmission dynamics and control of severe acute respiratory syndrome[J]. Science, 2003, 300: 1966-1970.
- [22] Kermack W O, McKendrick W G. A contribution to the mathematical theory of epidemics[J]. Proceedings of the Royal Society of London, Series A, 1927, 115: 700-721.
- [23] Bailey Norman T J. The Mathematical Theory of Infectious Diseases and its Applications(2nd)[M]. New York: Oxford University Press, 1975.
- [24] Wang J. Modeling epidemic in space and time[A]. In: Association of American Geographers Annual Meeting[C]. US, Los Angeles, 2002.

BACK ANALYZING PARAMETERS AND PREDICTING TREND OF SARS TRANSMISSION

HAN Wei-guo^{1,2}, WANG Jin-feng¹, LIU Xu-hua^{1,2}

(1. State Key Laboratory of Resources & Environmental Information System, Institute of Geographic Sciences & Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China;
2. Graduate School of Chinese Academy of Sciences, Beijing 100039, China)

Abstract: This paper uses SIR model to back-analyze the parameters of SARS transmission based on the data released by the health authorities of Beijing and Hong Kong, we get the important parameters such as the peak period, the hospitalized cases and the removed parameter. It can be seen that these parameters of the model allow for better understanding of the SARS transmission because the result fits the actual data approximately. As a result, SIR model could be used to fit data, predict trend and simulate process of SARS transmission.

Key words: SIR Model; Simulated Anneal Algorithm; Stiff Problem; Non-linear Extremum.