

文章编号:1005-3085(2003)07-0053-10

## SARS 传播预测的数学模型

周义仓<sup>1</sup>, 唐 云<sup>2</sup>

(1-西安交通大学数学系,西安 710049; 2-清华大学数学科学系,北京 100084)

**摘 要:** SARS 的传播是 2003 年全国大学生数学建模竞赛的赛题之一,这是一个完全开放、国内外一直在探索的问题。同学们提交的论文中建立了许多模型,对 SARS 的传播和预测进行研究。本文对竞赛情况和需要探讨的问题进行了简单的总结。

**关键词:** SARS;传播;数学模型;预测

**分类号:** AMS(2000) 34B08

**中图分类号:** O241.81

**文献标识码:** A

SARS(Severe Acute Respiratory Syndrome,严重呼吸道传染病)是 21 世纪第一个在世界范围内传播的传染病。SARS 从 2002 年 11 月份开始在我国和世界范围内流行,到 2003 年 6 月 23 日为止,据世界卫生组织(WHO)报道,SARS 患者已经达到了 8459 人,其中 802 人死亡[1]。此时中国的 SARS 患者已经为 5326 人,其中 347 人死亡[2],这给人民生活和国民经济的发展带来了很大的影响。

SARS 是由一种冠状病毒引起的传染性很强的呼吸道传染病,它主要通过近距离空气飞沫以及接触病人呼吸道分泌物和密切接触进行传播,也可能通过病人飞沫污染物(如通过手、衣物、食物、水或环境等途径)传播。SARS 潜伏期一般为 2—11 天。SARS 患者的主要症状有:发热头痛、全身酸痛和不适、乏力,部分病人在早期也会有轻度的呼吸道症状(如咳嗽、咽痛等)[3,4]。

在全球抗击 SARS 的过程中,人们对 SARS 传播和发展趋势进行了大量的研究。如 Donnelly 等人对 SARS 的潜伏期、死亡率进行了估计[5],Lipsitch 和 Riley 等人研究了 SARS 传播中的再生数和传播趋势[6,7]。Chowell 等人通过建立 SEIJR 模型,研究了 SARS 在加拿大、香港和新加坡的传播情况[8]。国内许多科研院所也有许多工作是利用模型对 SARS 传播的分析与预测作出了定量的研究,对 SARS 的防控起到了积极作用。已公开发表的论文有杨方廷、陈吉荣等人关于北京 SARS 传播过程的仿真、参数和数据处理[9,10],方兆本等 SARS 流行规律的建模及预报[11],周义仓等建立的描述我国 SARS 传播与控制的离散 SEQIR 模型[12],以及王稳地等人对北京 SARS 模型的模拟等[13]。

在 SARS 传播的过程中,我国政府和世界卫生组织每天发布疫情信息,这些数据为我们的建模和分析提供了良好的基础。

## 1 命题的想法

我国大学生数学建模竞赛已经进行了十几年,吸引了越来越多学生的兴趣,学校和社会逐步认识到竞赛对教学改革、对学生能力培养的重要作用。扩大受益面、提高竞赛水平是今后大学生数学建模竞赛的主要目的。选择社会关注的热点问题、增加赛题的灵活性和开放性是提高竞赛质量一个重要的途径。

SARS 的流行是一场突如其来的灾难,它波及到了我国绝大部分省、市、区。从4月起,北京一度成了 SARS 流行最为严重的一个城市,4月下旬每天的 SARS 病人和疑似病人大幅度增加,北京周围的一些省区和全国也出现了较多的感染情况,对我国的人民生活和经济发展造成了严重的影响。从4月下旬起,全国各地都行动起来展开了一场抗击 SARS 的攻坚战。到6月下旬每天的新增病例下降到零,世界卫生组织就解除了对北京的旅游警告,我国人民的日常生活和经济发展逐步恢复正常。数学建模对传染病传播过程的描述、分析、预报和控制能起到积极的作用。为了唤起社会对数学作用的认识,全国大学生数学建模竞赛组委会专家组就把 SARS 的传播这场密切联系到国民生活的重大事件提炼成赛题,作为今年的 A、C 题。

我们知道许多院校在竞赛培训过程中非常关注 SARS 的建模和预测,已把它作为校内竞赛题或训练题。这表明大家在自觉地应用数学建模的理论和方法来分析和解决社会急需探讨的问题,数学的应用意识在不断的加强和深化,这是多年来数学建模活动效果的反映。同时也给 SARS 建模的命题带来了困难,考虑到这一点,在“SARS 的传播”赛题中,除了建模、预测及分析对经济的影响外,增加了评论一篇早期发表的论文,及给报社写短文部分。

在命题讨论过程中我们也知道已经有不少 SARS 研究的资料,特别是在网上可以搜索到许多关于 SARS 建模和预测的文章,有可能导致过多的引用,甚至抄袭,给评判带来困难。但从另一方面看,SARS 建模和预测是一个十分复杂的问题,从传播机理、建模原则,到数据的收集、整理、模型的建立和模型参数的确定,都有一系列问题需要继续研究。网上和刊物中大量丰富的文献资料可以使学生更多、更深入地了解 SARS 的传播过程和建模预测方法,学习真正怎样将数学和计算机技术用来研究与解决人们所关注的实际问题。

竞赛题目的第一问是提供了一篇北京大学在早期对 SARS 进行建模和分析预测的文章,让参加竞赛的同学进行评价,目的是希望让学生了解别人对 SARS 传播建模和预测的处理方法。该模型在早期有它的应用价值,在 SARS 流行结束后再对它作出分析和评价,有利于学生找到更加实用的模型和方法。题目的第二问是提供了北京市4月20日到6月12日已确诊的 SARS 累计病例数、现有的疑似 SARS 病例数、累计死亡人数和累计治愈出院人数,希望学生建立起自己的模型,以对北京等地 SARS 的感染情况进行研究,定量地描述,并分析控制措施对 SARS 传播的影响。特别是训练学生学习利用已给的数据确定模型中的参数,进行分析、计算和比较。题目的第三问让同学收集 SARS 对经济某个方面影响的数据,建立相应的数学模型并进行预测,由于担心学生资源的限制,题目中还提供了北京市从1997年1月到2003年8月接待的海外游客人数作为参考。题目的第四问是让学生给当地报刊写一篇通俗短文,说明建立传染病数学模型的重要性,当然也希望能将自己模型和预测结果简要地介绍给民众,使更多的人认识到数学建模和预测工作的重要意义。

把“SARS 的传播”作为赛题也是希望学生尽可能地发挥自己的特长,收集尽可能多的数据,在网络和刊物中搜索尽可能多的参考文献,建立数学模型解决问题。该题目没有(也

不可能有)固定的模型和标准答案,这也促使各赛区的阅卷人员也深入钻研、思考问题,了解国内外 SARS 建模的研究现状,从学生的答卷中挑选出有特点的优秀论文。这道赛题后来全国组委会只给了评阅的要点供给各赛区参考,并且着重对 A 题的第二、第三问提出一些要求,包括以下几个方面:

1) 学生答卷中应包含对传播机理和传播状况的叙述(如:传播途径、传播方式、潜伏期和传播地区等),并且给出建模原理、方法、思路或框图。

2) 模型中的人口至少有 3 类:易感者、患者和恢复(与死亡)者,仔细一些的可以再加入潜伏者、隔离者、疑似病人、确诊病人、治愈者和留观者等,要弄清楚他们之间的关系。

3) 模型还应包含对于传染率、治愈率和死亡率等重要概念的清晰表述。模型分析和计算中要给出上述参数的估计方法和估计值,还可包括平均治愈天数、隔离率和潜伏期等。

4) 模型的结果应该提供预测值(用数量或曲线来说明高峰期和持续时间)和隔离措施的效果(包括提前和推迟控制时间的影响,隔离人数多少的影响及遗漏病人的影响)。对于结果的分析应包括误差分析,及模型与方法的通用性分析(模型除北京外,是否还用于其它地区,如外地、香港、全国和全球等)。

5) 关于对经济的影响,要求收集某方面的数据(除旅游业外,还有如餐饮业、航运等),预测没有 SARS 时的变化趋势;与实际统计数据比较,用差距说明影响的大小;预测该领域恢复到正常发展水平的时间。

全国组委会还提供了一些网址和关于 SARS 传播模型的论文报告作为评阅时参考,包括北大金融数学系的模型、清华力学系的模型、超级传播模型、多伦多与香港的报告和模型、台湾模型、Anderson 模型、Murray 模型、Dye 模型、加拿大多伦多病例分析及疾病传播模型等。

## 2 解答综述

一些参赛队对上面要点中提到的几个方面都给出了很好的想法。具体情况如下:

### 1) 对已有模型的评价

对题目之中给出的“SARS 疫情分析及对北京疫情走势的预测”一文,每个队都对文章中的建模思路、模型、参数和预测进行了客观的评价。概括起来文章的优点有:模型简单明了地反映了疾病的传播过程,抓住了 SARS 传播过程中两个主要特征:传染期  $L$  和传染率  $K$ ,  $L$  是指平均每个病人可以直接感染他人的时间,  $K$  是指平均每病人每天可传染的人数,对其给出了估计方法。对北京、广东与香港的疫情进行了分析比较,预测值与实际统计较接近;模型特别简单、计算量小,容易理解和使用;模型的灵活性在于适当挑选初始值  $N_0$  和参数  $K$ 、 $L$  就可以描述不同地区、不同控制措施下的 SARS 传播情况。模型的不足在于对如何确定初始值  $N_0$  和参数  $K$  与  $L$  缺乏一般的原则或算法,这种指数变化的趋势作为长期预测不合理。另外还需要指出的是不少论文还对模型进行了推导、计算和改进,例如分段确定参数或使用随时间变化的  $K$ ,这些使得模型的理论基础更加完备,实用性更加广泛。

### 2) SARS 建模和预测

大部分答卷都在叙述了 SARS 传播机理的基础上,作出类似于下面这些基本合理的假设:①单位时间内感染的人数与现有的感染者成比例;②单位时间内治愈人数与现有感染者成比例;③单位时间内死亡人数与现有的感染者成比例;④ SARS 患者治愈恢复后不再被感染;⑤各类人口的自然死亡可以忽略;⑥忽略迁移的影响。这些比例系数可以是常数、时间的函数、时间和各类

人口的函数、或分几段取常数。当然,还可以根据需要做其它假设。建立模型一般是利用房室(compartment)结构,将总人口分为易感者  $S$ 、患者  $I$ 、恢复者  $R$ ,再仔细一些的还有潜伏者  $E$ 、隔离者  $Q$ 、疑似病人  $P$  和确诊病人  $J$  等类型。叙述或作出各类人口之间流动的示意图,并根据传染病建模的一般原理建立起如  $SIR$ 、 $SEIR$ 、 $SEQPIJR$  等类模型。这些模型基本思路相同,差异在于人口分类的多少,关键在于参数的确定。例如最简单的  $SIR$  模型为

$$\begin{aligned}\frac{dS(t)}{dt} &= -\beta S(t)I(t), & S(t_0) &= S_0, \\ \frac{dI(t)}{dt} &= \beta S(t)I(t) - (\delta + \gamma)I(t), & I(t_0) &= I_0, \\ \frac{dR(t)}{dt} &= \gamma I(t), & R(t_0) &= R_0.\end{aligned}\quad (1)$$

在模型(1)中, $S(t)$  是  $t$  时刻易感者的数量,它等于总人口减去患者和恢复者的数量, $I(t)$  是  $t$  时刻患者的数量, $R(t)$  是  $t$  时刻恢复者的数量, $\beta S(t)$  是单位时间内每个患者感染的人数, $\gamma$  是患者的恢复率, $\delta$  是患者的死亡率。由于隔离等控制措施的不断加强和治疗情况的变化, $\beta$ 、 $\gamma$ 、 $\delta$  也是随时间而变化的;另外,由于易感者的数量特别大,可以近似看作常数,且将常数合并到  $\beta$  中去。在实际应用中,我们最关心的是感染者数量的变化。取时间单位为天,将模型中的第 2 个方程离散化得递推关系为

$$I(t+1) = I(t) + \beta(t)I(t) - (\delta(t) + \gamma(t))I(t), \quad I(0) > 0. \quad (2)$$

在离散化的模型(2)中, $\beta(t)$  的含义是每天每个 SARS 感染者传染的人数,是一个十分重要的参数,其确定的原则是:当天新增 SARS 病人人数除以当天 SARS 感染者人数,再进行曲线拟合即可。 $\gamma(t)$  和  $\delta(t)$  是 SARS 患者每天治愈和死亡所占的比例,可以一起确定,其方法是当天 SARS 感染治愈和死亡人数除以当天 SARS 感染人数,再进行曲线拟合即可。例如,利用卫生部公布的 4 月 20 日至 5 月 15 日全国的数据[14] 进行计算,可以得到  $\beta(t)$  随时间变化的关系如图 1 中折线所示,用指数曲线  $\beta(t) = ae^{-bt}$  对其进行回归拟合

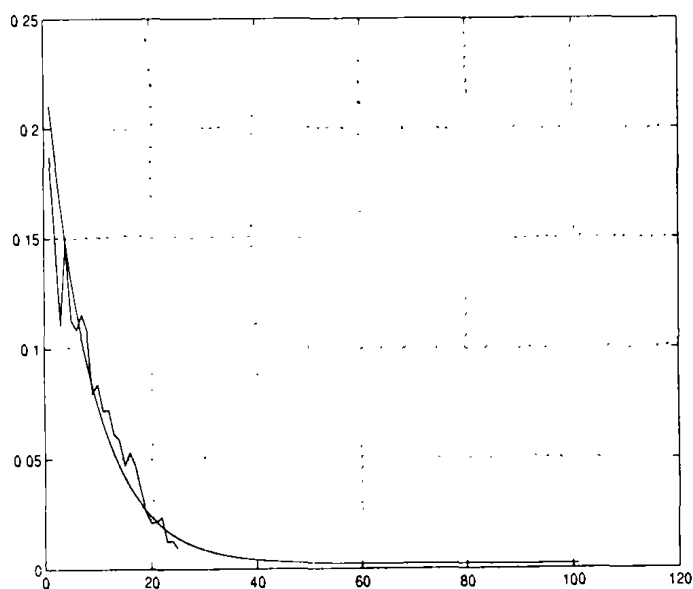


图 1 每天每个 SARS 病人平均感染的人数,其中光滑曲线表示指数拟合的  $\beta(t)$  曲线,而折线表示实际计算结果

得到  $\beta(t)$  的表达式,其曲线如图 1 中的光滑曲线所示。同理得到  $\delta(t) + \gamma(t)$  的表达式。将这些函数代入(2) 进行递推计算得每天的 SARS 感染者人数(见图 2)。

从图 2 中可以看出,这个非常简单的模型、参数确定方法所预测的结果与实际的统计值比较一致。为了进一步检验模型和参数确定方法的合理性,我们分别利用北京、山西、内蒙、广州等地的数据替换全国的数据,进行同样的计算、预测和对比,发现结果都比较符合。注意:在这样的简单预测中仅用了 25 天的数据对模型中的参数进行估计和曲线拟合,从 5 月

15 日以后的预测值比实际统计值小,这是由于我国政府不断地加强隔离控制措施和改善治疗效果,这些因素在 5 月 25 日之前的数据中反映的不够。在实际应用中不断增加和更新数据进行短期的预测,预测的误差就会减少。

在竞赛答卷中,有许多不同的模型和方法来描述 SARS 的传播过程。其中最多的是上述房室结构,包括有隔离的 SQIR 模型,有潜伏(也有将潜伏作为时滞处理)的 SEIR 模型,考虑 SARS 诊断困难和我国实际做法的有疑似的 SPIR 模型,有区分自由流动病人和收治病人的 SIJR 模型,还有考

虑医护人员感染的模型,涉及更多房室的模型等。从模型的形式看有常微分方程模型、偏微分方程模型、差分递推模型、控制反馈模型、小世界网络模型、神经网络模型、信息传输系统模型、基于复杂网络的概率演化模型、时间序列模型、混沌时间序列模型、自回归模型、非齐次泊松过程预测模型、Bezier 曲线模型、二阶 Logistic 回归模型和 Monte Carlo 法模拟模型等。在这些模型中所涉及的因素包括媒体宣传对心理和行为的影响、公共卫生水平、控制力度、平均气温、地区差异、人口流动、职业分布、年龄分布、医院内外的传播和超级传播者等。考虑到 SARS 病毒与一般传染病不同的是有着相当高的突变率,有的队还建立起 SARS 病毒的分子进化模型,对“毒王”现象做出合理的解释。

模型中的参数一般分为常量参数和变量参数,常量参数利用国家公布的数据或平均值确定,变量参数通过计算和回归的方法确定。优秀答卷都给出了北京市 SARS 传播的持续时间、高峰期、控制措施的影响等,也用所建立的模型对全国、山西、内蒙、香港和加拿大等地区的传播情况进行分析预测,与实际数据对比显示模型的适应情况。

在这些答卷中,不少参赛队把 SARS 的传播分为控制前和控制后两个阶段,分别进行建模、参数确定和预测,这样的处理方法对于研究控制措施实施迟早对 SARS 传播的影响十分方便。有一些队在传染率中引入一个因子,通过这个因子大小的调节,就可以分析控制措施的强弱对 SARS 传播的影响。考虑到心理压力对于免疫力下降的影响,有的在赛期还利用 BBC 网上聊天系统对人们在 SARS 流行期间的心里恐慌程度进行调查,将结果引入模型参数中。在分析控制措施效果的基础上,还有一些队提出了自己的建议,如提高整个人口的抵抗能力,加强对与感染接触人员的监控等。在描述人口流动对 SARS 的影响时,有些队把实际区域抽象成二维网络,每个格子规定其地域特性,人口按一定的密度分布在不同的格子内,将人按年龄分为 4 类,不同年龄人群的活动范围不一样,规定他们的移动规则,再根据 SARS 的特点定义传播规则,根据控制措施定义隔离规则,再对北京市的 SARS 传播情况进行模拟和预测。

### 3) 对经济的影响分析

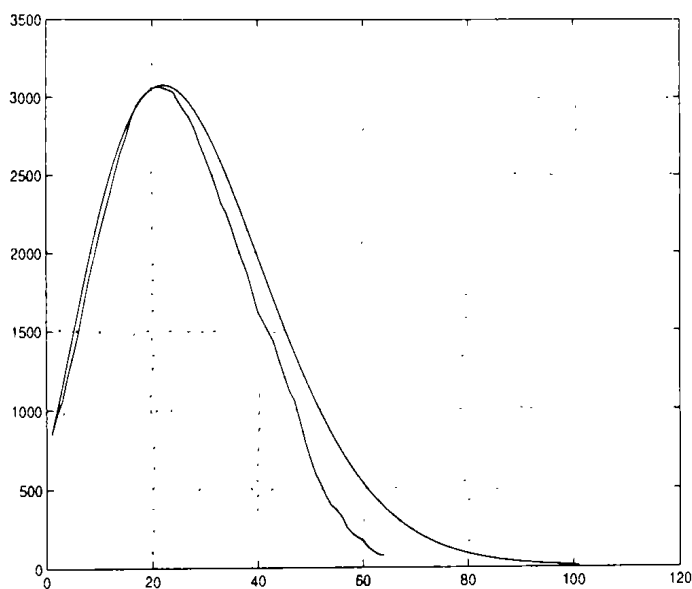


图 2 每天 SARS 感染者随时间变化的关系。

上面的光滑实线表示预测值,下面的曲线表示统计值。

SARS 对我国经济的影响是多方面的,在总体上对我国经济的正常运行有负面作用,但它对国民经济不同部门的影响不同。SARS 的流行阻碍我国的旅游、交通运输和外商投资等方面的发展,另一方面,它促进了医药、汽车的销售和电子商务的发展。题目中给出的北京市海外游客的数据是一个典型的例子,绝大部分参赛队就是对这些数据进行分析,首先观察到在 SARS 流行前游客数量是随季节变化且不断增加的。结合这个特点建立模型预测不受 SARS 影响时北京市的外来旅游人数,再与 2003 年的实际数据比较,给出了同期减少的量,给出游客数量恢复的期望时间。在这一部分用到的模型或方法包括曲线回归模型、曲面拟合模型、灰色系统的 GM(1,1) 预测模型、ARIMA 模型、三次指数平滑法、平均趋势预测法及消费乘数模型预测损失法等。

在这一部分也有少数参赛队收集到了其它数据分析 SARS 对经济发展的影响。有参赛队收集了我国民航系统从 1985 年以来的运营数据,建立模型对民航业的 3 大指标(总周转量,旅客运输量,货邮运输量)进行预测,通过无 SARS 情况下民航运营情况的预测和与 2003 年的分析对比,估计出了民航受 SARS 影响的经济损失,并对民航业的恢复期进行了预测。有参赛队在分析非典对经济的影响时着重讨论了受非典影响的旅游百货行业和医药行业的效益变化,考虑到股市作为国民经济的晴雨表,能很好地反映出各行各业经济效益的变化趋势,他们在每个行业中各选取了三十家品级较高的上市公司,找出它们业绩的变化特征,以此来反映整个行业的效益变化趋势,利用灰色预测模型对这两个行业的效益做出了预测;同时考虑到非典的影响,又对此模型进行了修正,加入了信心指数。利用修正后的模型,他们分别预测了 2003 年旅游百货行业和医药行业上半年业绩,与实际数据吻合较好,用此模型进一步分析得出,在十一长假期间,旅游百货行业会有强劲的反弹。也有参赛队考虑心理状态等影响旅游的因素,更关心旅游市场的恢复期预测。他们的出发点是大多数人的出游的决策受安全、收入、费用、时间等多因素的影响,不同的因素对于不同的人来说具有不同的影响力,各个因素的权重不同。但当象 SARS 这种对人身安全威胁极大的因素出现时,安全几乎成了人们考虑是否出游的决定因素。也就是说影响人们出游的各个因素的权重发生了很大的变化。而当 SARS 得到控制以后,虽然安全系数大大升高,但是安全顾虑仍然会持续一段时间,还是人们考虑出游的重要因素,直到人们对 SARS 的顾虑完全消除。因而建立模型时,采用了层次分析和数据拟合相结合的方法。首先通过层次分析法和一些合理的假设,归纳出一个带有参数的多项式。再根据该多项式的形式,结合已知数据表中部分合理的数据,作出了一条拟合曲线,由此曲线得到了对未来几个月游客数量的预测。

#### 4) 给报社写短文

给报刊写短文就是要让群众认识到传染病的危害,认识到数学模型在研究传染病流行规律方面的重要作用。

传染病历来是危害人类健康的大敌,历史上传染病一次次的流行给人类生存和国计民生带来了巨大的灾难。20 世纪是人类征服传染病取得最辉煌成果的时期:肆虐了近千年的天花终于被消灭了;麻风病、脊髓灰质炎被彻底消灭的日子也为期不远了;白喉、麻疹、百日咳、破伤风等病已在许多国家得到遏制;多种抗生素的问世,使一度给人类造成巨大灾难的“瘟疫”不再危害人间。然而,人类要征服传染病,道路依然曲折漫长。随着国际贸易和交往的发展、生态环境的变化以及病原体和传播媒介抗药性的增强,原来已灭绝或被控制的许多传染病(如性病、结核病)等再次抬头并且不断蔓延。一些新近出现的传染病也来势凶猛(如

艾滋病、SARS)。人类将长期面临着传染病的严峻威胁,对传染病发病机理、传染规律和防治策略研究的重要性日益突出,且已成为当今世界需要迫切解决的一个重大问题。正如世界卫生组织 WHO 发表的报告所述,传染病依然是人类的第一杀手。

疾病的传播问题是有一定规律,可以预测。早在 1760 年 D. Bernoulli 就曾用数学研究过天花的传播;1906 年 Hamer 为了理解麻疹的反复流行,构造并研究了一个离散时间模型;1911 年公共卫生医生 Ross 博士利用微分方程模型对疟疾在蚊子与人群之间传播的动态行为进行了研究,其结果表明,如果将蚊子的数量减少到一个临界值以下,那么疟疾的流行将会得以控制。Ross 的这项研究使他获得了 Nobel 医学奖。1926 年 Kermack 与 McKendrick 研究了 1665—1666 年黑死病在伦敦的流行规律以及 1906 年瘟疫在孟买的流行规律。传染病动力学的建模与研究于 20 世纪叶开始蓬勃地发展[15,16]。

传染病传播的数学模型可以用来描述疾病的发展变化过程,有效、经济、迅速、明确地给出疾病发生的数量规律,为人们的决策提供数量依据;并对各种检测、防治、治疗及控制措施进行评价和比较,以控制疾病的流行。2001 年伦敦大学帝国学院的流行病学家 Roy Anderson 等还为英国的口蹄疫爆发建立起数学模型,提供了大量的客观证据,促使英国首相托尼·布莱尔领导的政府下令大规模屠宰牲畜,阻止了疾病的蔓延,这就是一个成功的例子。

这次 A、C 题的参赛队大多能写出一篇较好的短文,说明数学建模在研究传染病流行规律方面的重要作用。简述自己的建模方法和所得到的结论,提出了对 SARS 传播进行控制的很好的建议。

给报刊写通俗短文还应考虑到读者面,应该用读者可接受的语言来说明建立传染病数学模型的重要性。这方面,有些参赛队做得是比较成功的,本文的附录就是一个范例。

这里还应当指出,传染病的数学模型和预测只能为有关部门提供决策的数量参考,由于统计数据和传播机理等方面的限制,不可能预测得很精确,只能预测传播发展的趋势,尤其是在一个传染病传播的初期更是如此,不应将数学模型预测和分析的效果过分夸大。

### 3 对 SARS 建模和预测进一步的思考

SARS 在我国和世界范围内的流行已经结束,但人们仍然关心还会有下一次的 SARS 流行,别的传染病出现应该怎样应对,有关 SARS 和新发传染病的研究一直在继续。有一点是肯定的,除 SARS 外,今后还会有其它类型传染病在不同的范围内流行,所以,公共卫生体系需要不断加强,对传染病的研究工作不能停止。这次竞赛通过对 SARS 的建模、预测及其经济影响的分析,可以使得我们对 SARS 和其它传染病的传播机理有更深入的了解,发现问题、积累经验。这些参赛学生中有一部分以后会在医疗卫生部门或政府机关工作,数学建模的训练和竞赛经历会使得他们在今后的工作中发挥作用,对防治工作起到帮助。

这次竞赛中经过三天的奋战,同学们经受了锻炼,对传染病建模有了初步的体会,取得了良好的成绩。但在答卷中也存在下面几个问题。

1) 不少队在确定模型的参数时几乎用上了赛题中给出的所有数据,然后再代入模型中作预测,指出预测结果与实际统计值很接近,几乎没有误差。这反映出他们没有真正体会传染病建模和预测的思想。实际中最需要的是在一种疾病传播初期收集到少量数据的基础上建模预测以后的发展情况,当一个疾病流行结束后,所有数据都有了,预测就失去价值。所以应当只用一部分数据进行预测,而用其余数据作检验。

2) 在疾病的传播过程中,对流行状况影响很大的是再生数,即一个病人在患病期间平均感染的人数,这在流行病传播中是个十分重要的参数。要将一个传染病控制住,实际上就是要通过各种措施使得再生数小于1。答卷中提到再生数概念和对它作出估计的很少,也没有分析我国的控制措施对再生数的影响。

3) 阅读的资料、收集的数据不足。许多队仅对北京的 SARS 传播情况进行了建模预测,没有涉及其它地区,经济的影响也是就题目之中给出的旅游数据进行了分析,没有收集其它数据,没有看到 SARS 对某些行业带来的促进作用。

4) 尽管全国组委会特别要求参赛同学签名保证不抄袭别人的成果。如果引用公开的资料(包括网上查到的资料),必须在正文和参考文献中明确列出。但仍有一些答卷明显地违背了他们的保证。全国组委会已经取消了这些队的全国评奖资格。

在这次数学建模竞赛过程中,同学们也遇到了不少困难,如传播机理和过程不是十分清楚,在确定参数时能得到的数据不足,缺乏患者被感染和感染给别人、确诊与治愈等的个案资料,而潜伏期的感染情况、超级传播者的传播链、流动人口的数据及流动人口中的感染情况等数据也明显不足。对于 SARS 及类似传染病的建模分析还有许多问题值得探讨,包括:

i) 如何从模型中能反映出对疾病的传播规律有重要影响的、但不能完全量化的因素或不易确定的因素。如全体人口都是易感者,但真正与 SARS 病人接触的仅是极少数;对潜伏期、就诊或隔离前以及治疗期的接触与传染率的估计;对于不同级别(如学校、单位和居民区半封闭状态)的隔离措施的描述和效果的评价;各个省、区之间人口流动的影响;少病例的省市和地区内疾病传播情况的描述;医院内部感染情况的描述;SARS 向农村、牧区扩散情况的描述;隐性感染者对疾病传播的影响等。

ii) 探讨在建立模型时如何对总人口进行分组以及对随机现象做出描述;如何根据流行病学的特征,以及年龄、行为、地域分布和流动等因素,研究影响传染病预测控制机制的人口分组分层问题;鉴于患病者传染力之间的差异、受体易感程度的差异、人口流动以及重疫区人口向周边地区和农村、牧区扩散的随机性、各地区间发病情况与预防控制强度的差异、统计数据的误差等,研究具有诸多随机因素的建模问题。

iii) SARS 及突发性传染病的数据挖掘问题。对于像“SARS”这类新的突发恶性传染病,在流行初期由于病因不清和经验不足,难以作出合理判断。我们需要研究如何从人类以往积累的大量传染病流行资料中提取类似的相关数据和信息进行模式识别,并建立实时自适应学习机制,用以不断完善数学模型。我们需要研究从仅有的少量数据和以前积累的相关传染病资料中提取信息,建立初步模型,估计疾病传播的流行病学参数和流行模式,探讨基于少量数据的理论和方法,发展一种对突发恶性传染病的快速反应决策技术。我们应当进一步完善已有的工作,在建模的基础上开发出一套实用的预测软件。

#### 参考文献:

- [1] WHO. <http://www.who.int/csr/sars/country/2003-06-23/en/>
- [2] Ministry of Health P. R. China. <http://www.moh.gov.cn/was40/detail?record=82&channelid=34385>
- [3] B. S. Kamps, & C. Hoffmann, SARS Reference - 05/2003, SARSreference.com, Flying Publisher,



May 2003

- [4] 孟继鸿,邱海波. SARS 基础与临床. 南京:东南大学出版社,2003
- [5] C A Donnelly et al., Epidemiological determinants of spread of causal agent of severe acute respiratory syndrome in Hong Kong, The Lancet, Published online May 7, 2003, <http://image.thelancet.com/extras/03arr4453web.pdf>
- [6] M Lipsitch, et al., Transmission dynamics and control of severe acute respiratory syndrome, Published online May 23, 2003;10. 1126/science. 1086616 (Science Express Reports)
- [7] S Riley, et al., Transmission dynamics of the etiological agent of SARS in Hong Kong: impact of public health interventions, Published online May 23, 2003;10. 1126/science. 1086478 (Science Express Reports)
- [8] G Chowell, P W Fenimore, M A Castillo Carsow, and C Castillo Chavez, SARS out break in Ontario, Hong Kong, and Singapore: the role of diagnosis and isolation as a control mechanism, Journal of Theoretical Biology, 2003(224), 1 - 8,
- [9] 杨方廷,侯立华,韩军等. 北京 SARS 疫情过程的仿真分析. 系统仿真学报, 2003(15:7), 991 - 994
- [10] 陈吉荣,杨方廷,战守义. 北京 SARS 仿真模型的参数和初始值的处理. 系统仿真学报, 2003(15:7), 995 - 998
- [11] 方兆本,李红星,杨建萍. 基于公开数据的 SARS 流行规律的建模及预报. 数理统计与管理, 2003(22:5), 48 - 57
- [12] Zhou Yicang, Ma Zhien, Fred Brauer. A Discrete Epidemic Model for SARS Transmission and Control in China, Mathematical and Computer Modeling, 2003
- [13] Wang Wendi, Shigui Ruan. Simulating SARS outbreak in Beijing with limited data, Journal of Theoretical Biology, 2004
- [14] Ministry of Health P R China. [http://www.moh.gov.cn/was40/outline?\\_page=4&channelid=3485](http://www.moh.gov.cn/was40/outline?_page=4&channelid=3485)
- [15] N T J Bailey. The Mathematical theory of infectious disease, 2nd ed., Hafner, New York, 1975
- [16] 马知恩,周义仓,王稳地,靳祯. 传染病动力学的数学建模与研究. 北京:科学出版社, 2004

## The Mathematical Modeling and Prediction of SARS Transmission

ZHOU Yi-cang<sup>1</sup>, TANG Yun<sup>2</sup>

(1 - Xi'an Jiaotong University, Xi'an 710049; 2 - Tsinghua University, Beijing 100084)

**Abstract:** SARS transmission is one of the contest problems for CUMCM - 2003. It is an open problem investigated in over the world. Several mathematical models to predict SARS transmission were established in the articles submitted by the student teams. A brief summary of the contest results is presented in this paper.

**附录:短文**(作者为中山大学 梁田贵,汪卉琴,刘琴)

### SARS 与数学模型

九月的北京,秋高气爽,大街上人来人往,热闹非凡。人们对于那段谈“非”色变的日子似乎已渐渐淡忘,只是偶尔在大街上看到一两个人戴着口罩,回想起那段日子,还会为自己当时莫名的恐慌而自嘲。

人们恐慌的心理,才是传染病最可怕的影响。

当流言满天飞,大家盲目抢购的时候,是因为我们不了解疫情的情况,所以恐慌。

当大家害怕得藏在家里,见到亲人朋友连握手都担心的时候,是因为我们对传染病的传播途径不了解,所以恐慌。

当政府说明已采取了许多措施,而大家看到医院每日收治病人仍然不断攀升的时候,是因为我们不了解传染病的发展规律,所以恐慌。

这一切的恐慌,都是因为我们不了解疫情而带来的。而要全面了解、控制疫情,除了相关的医学知识外,重要的问题之一是建立恰当的数学模型。

如果在疫情一开始,我们对传染病的常用数学模型有一定的认识,了解如何确定一种传染病的危害程度,何时进行控制,如何用科学的观点,详细的数据向民众解释,我们就不会因为盲目而恐慌。

如果我们了解基本传染病传播规律以及相应模型的决定因素,那么我们在生活中也可以通过控制这些因素以有效控制疾病,而不必恐慌。

如果我们知道传染病的发病曲线,我们也不会为新闻发布新的各种疫情数字而恐慌。

今天,在网上搜索一下“SARS 数学模型”,我们可以找到近千篇文章,十几种基本的传染病数学模型及其改进的模型,也许这些模型不一定很符合这次 SARS 传播的实际情况,也许我们对这些模型中的各种深奥的公式与定理不太了解。但是,模型得到的一些结论如:高峰期的到达时间,控制因素的效果等,至少可以让我们心安,不必盲目恐慌。而更精确的模型,以利于我们更好地制定相应的应急措施,有效的医疗保障系统等。

SARS 曾让我们惊慌失措,也让我们相信科学,认识与了解科学。拥有医学与数学等科学知识,我们应该有信心,流行病并不可怕。

-----  
(上接 52 页)

## The Mathematical Modeling and Analysis of SARS

WANG Yi - feng, TIAN Yi, YANG Qian

Advisor: SHANG Shou - ting

(Harbin Institute of Technology, Harbin 150001)

**Abstract:** In this paper, two mathematical models for the spreading of SARS before and after the control respectively were established based on the data displayed on the Internet before June 2003 and the classic SEIR model for infectious disease, in which people are defined into several groups reasonably. Each parameters was solved by reasonable estimation, method of curve fitting and method of statistical theory. Then the primary analysis is laid on the after-control-model, solving the equations by Runge-Kutta method, analyzing the specific situations in Peking, Inner Mongolia, Guangdong and Hongkong where the SARS was rampant. Finally, the evaluation was made about the rationality and practicability of the models, suggesting the method for further revising of the model.

**Keywords:** differential equation; Runge-Kutta method; method of curve fitting; statistical theory