基于SI模型和SEIR模型的

新型冠状病毒疫情走向趋势的预测

摘要

自1月中旬以来，我国爆发了新型冠状病毒的传染病感染，造成了严重的经济损失和巨大的社会负担。对于疫情爆发的趋势是如何，疫情多久结束，我们可以利用相应的数学模型进行预测推演，这就是数学建模中常用的传染病模型。本文运用了传染病模型中比较常见的两种模型，那就是SI模型和SEIR模型。

注：本文使用的模型为理想化模型结果只供参考，现实情况更加复杂。

一、SI模型和SEIR模型的介绍

SI模型，即“易感者-感染者”模型，它的基本假设就是在人群中只存在两种人，那就是只有易感者（Susceptible）和感染者（Infectious）。通俗的理解来说，在此模型中，所有人都是易感者，所有确诊过新冠病毒的人都是感染者。SI模型的数学表达式如下：

其中、分别为易感者和感染者随时间变化的函数。为传染率，与每人每天接触人数和病人接触后传染疾病的概率有关。为参与本传染病模型的总人数。SI模型的解析解如下：

SEIR模型相比于SI模型，多了两种人，那就是潜伏者（Exposed）和康复者（Recovered）。在此模型中，感染者的定义与上一个模型有所不同，康复者和潜伏者在此模型中不算作感染者，并且康复者共同包含了死者和治愈者。SEIR模型的数学表达式如下：

其中为潜伏期的传染率，一般的SEIR模型这个值为0，但是此次新型冠状病毒在潜伏期也有一定的传染性，因此加上了这个值。、、、分别为易感者、潜伏者、感染者和康复者随时间变化的函数。为潜伏者变为感染者的转化率，与每天潜伏者发病的平均概率有关。为康复率，与平均住院天数有关。此模型没有解析解，只有数值解。

二、模型的建立与求解

要搭建上述模型需要对上述模型的相关参数进行求解，以拟合现有的数据,从卫健委官网上查询到如下数据。

表 1 卫健委统计数据

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 日期 | 累计确诊 | 累计治愈 | 累计死亡 | 感染者 |
| 1月20日 | 291 |  |  | 291 |
| 1月21日 | 440 |  | 9 | 431 |
| 1月22日 | 571 |  | 17 | 554 |
| 1月23日 | 830 | 34 | 25 | 771 |
| 1月24日 | 1287 | 38 | 41 | 1208 |
| 1月25日 | 1975 | 49 | 56 | 1870 |
| 1月26日 | 2744 | 51 | 80 | 2613 |
| 1月27日 | 4515 | 60 | 106 | 4349 |
| 1月28日 | 5974 | 103 | 132 | 5739 |
| 1月29日 | 7711 | 124 | 170 | 7417 |
| 1月30日 | 9692 | 171 | 213 | 9308 |
| 1月31日 | 11791 | 243 | 259 | 11289 |
| 2月1日 | 14380 | 328 | 304 | 13748 |
| 2月2日 | 17205 | 475 | 361 | 16369 |
| 2月3日 | 20438 | 632 | 425 | 19381 |
| 2月4日 | 24324 | 892 | 490 | 22942 |
| 2月5日 | 28018 | 1153 | 563 | 26302 |
| 2月6日 | 31161 | 1540 | 636 | 28985 |
| 2月7日 | 34546 | 2050 | 722 | 31774 |
| 2月8日 | 37198 | 2649 | 811 | 33738 |
| 2月9日 | 40171 | 3281 | 908 | 35982 |

通过上述数据进行拟合，即可得到各个参数的数值。

2.1 SI模型的建立与求解

从表 1中提取第一列和第二列数据运用MATLAB进行回归分析即可求取出公式内的参数，在套用公式即可对疫情发展进行预测。求得感染者人数随时间变化公式如下：

运用此公式即可进行预测，其预测结果和已有统计结果如下：

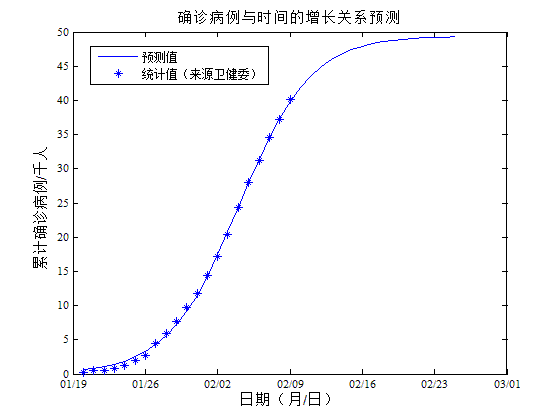


图 1 确诊病例随时间增长关系预测

从图中不难看出预测趋势与实际值差别不大，而且照此趋势发展在二月中下旬，累计确诊病例将会达到峰值，疫情将会得到控制，大约最后累计确诊数为5万人左右。

2.2 SEIR模型的建立与求解

由于SEIR模型没有解析解，因此无法列出其具体的表达式，只能通过数值计算对其进行拟合。由于进行数值计算，因此要将模型改写为如下形式：

将其继续改写为矩阵乘法形式便于MATLAB进行数值计算，

对该模型进行迭代，即可求出其相关参数，

在求得的参数中值得注意的是潜伏期传染率的值为0，因此可以大致推断出潜伏期传染性较小。

根据所求得的参数带入迭代模型得到如下结果：

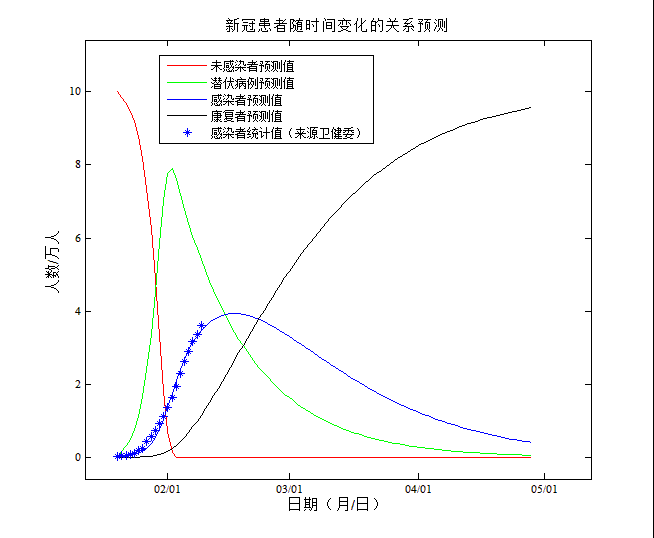


图 2 新冠患者随时间变化关系预测

根据图中不难看出感染者预测值将会在2月中下旬到达峰值，因此疫情将会在2月中下旬得到控制，总共波及人口大约为10万人左右。

三、模型的验证

在上述模型中只运用到了2月9日截止的数据，因此图 1、图 2中，实际值的散点图只到2月9日。为验证模型将2月10日的数据绘制出来，如下所示。

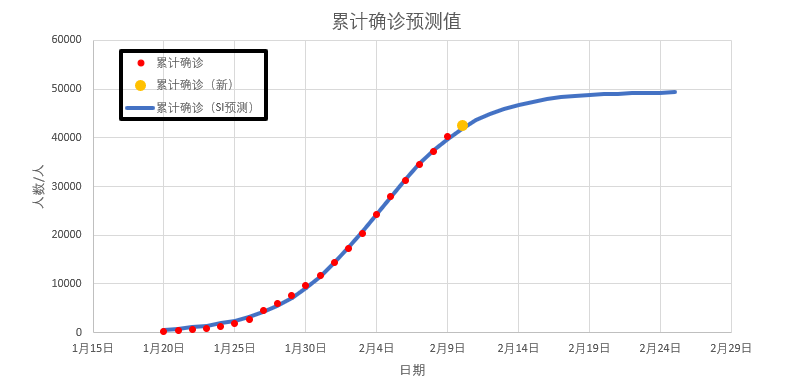


图 3 累计确诊对比

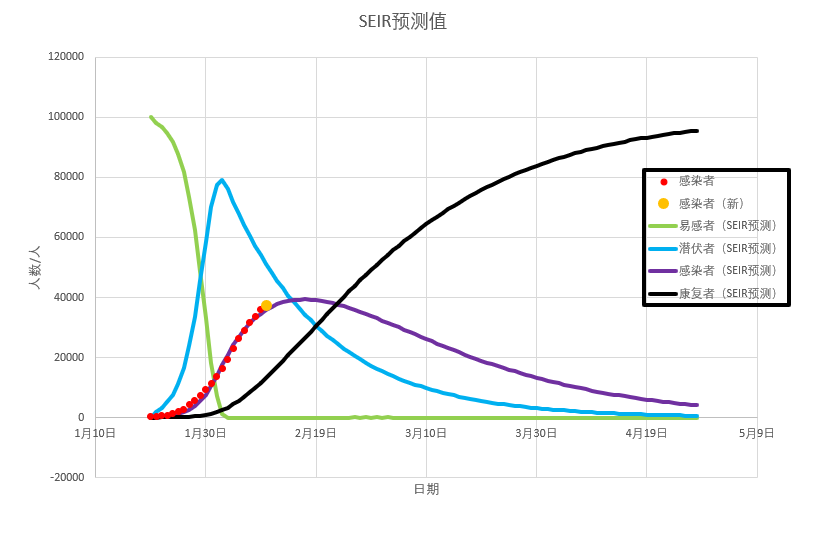


图 4 预测趋势对比

从上述两图不难看出，预测的基本趋势正确。

经过计算SI、SEIR的2月10日的误差分别为1.77%、4.54%，误差小于10%，SEIR模型的误差大于SI模型，这有可能是因为SI有具体的公式而SEIR则没有具体的公式。

四、结论

注：本文使用的模型为理想化模型结果只供参考，现实情况更加复杂。

经过数学模型的推演，得出如下结论：（1）本次新型冠状病毒确诊病例将会在2月中下旬达到顶峰，之后疫情将会得到基本控制；（2）潜伏期传染性应该比较低；（3）波及到的人口大约5到10万人左右。