

# 新冠疫情下六大洲发展趋势对比分析

乔姝婧      段成凤      苏思宁      张楠

2020/04/24

## 二. 文献及模型理论

## 2.1 参考文献

关于 2019nCoV 新冠肺炎的建模 ( ) — 基于 Logistic 模型的疫情预测。采用 Logistic 模型与 SEIR 带潜伏期传染模型对疫情进行分析与预测，发现在模型选取的方面，Logistic 大约能预计到的拐点的日期与患病人数的峰值，无法预测出疫情的全过程，同时在  $r$  与  $k$  值得优化问题上仍有待改进。

李慧聪和李金仙 ( 2020 ) 关于河南省新型冠状病毒 ( COVID-19 ) 疫情分析，建立非自治 SIAR 模型，利用 Matlab 对河南省的累计病例数进行数据拟合，刻画了疾病流行时间、规模、高峰时间等流行病学特征。

朱仁杰和唐仕浩 ( 2020 ) 基于改进 SIR 模型的新型冠状病毒肺炎疫情预测及朱仁杰和唐仕浩 ( 2020 ) 基于改进 SIR 模型的新型冠状病毒肺炎疫情预测及防控对疫情发展的影响，发现改进的 SIR 模型对 COVID-19 疫情发展趋势的分析基本可靠；除英国和美国以外，其他 5 个国家的疫情现阶段已经得到初步控制，而英国和美国还需要加强防控，以减少疫情带来的损失。

## 2.2 研究模型

- 阻滞增长模型
- SEIR 模型

## 2.3 阻滞增长模型

阻滞增长模型，又叫逻辑斯蒂（Logistic）模型，它是皮埃尔·弗朗索瓦·韦吕勒在 1844 或 1845 年在研究它与人口增长的关系时命名的。Logistic 模型是考虑到自然资源、环境条件等因素对人口增长的阻滞作用，对指数增长模型的基本假设进行修改后得到的。

## 2.3 阻滞增长模型

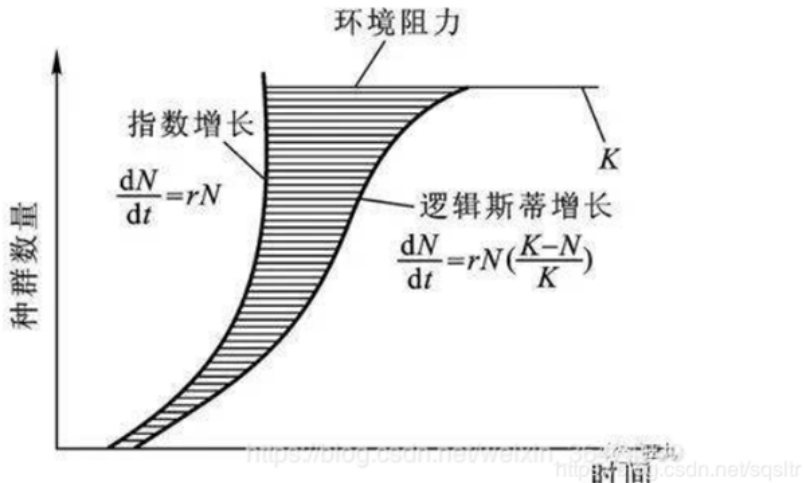


Figure 1: 阻滞增长模型示意图.

## 2.3 阻滞增长模型

Logistic 方程，即常微分方程：

$$\frac{dP}{dt} = \frac{rP * (K - P)}{K}$$

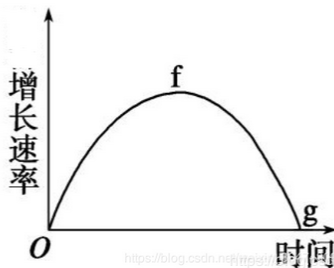
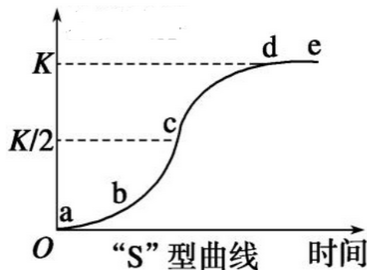
求解得 Logistic 函数：

$$P(t) = \frac{KP_0 e^{rt}}{K + P_0(e^{rt} - 1)}$$

其中  $P_0$  为初始值， $r$  衡量曲线变化快慢， $t$  为时间。 $dP/dt$  是种群增长率（单位时间个体数量的改变）， $P$  是种群的大小（个体的数量）， $K$  是可能出现的最大种群数（上渐近线）或承载力。

## 2.3 阻滞增长模型

广义 Logistic 曲线又称 S 形曲线。起初阶段大致是指数增长；然后随着开始变得饱和，增加变慢；最后，达到成熟时增加停止。



»»



## 2.3 阻滞增长模型

Logistic 曲线通常分为 5 个时期：

- 开始期，由于种群个体数很少，密度增长缓慢，又称潜伏期。
- 加速期，随个体数增加，密度增长加快。
- 转折期，当个体数达到饱和密度一半 ( $K/2$ )，密度增长最快。
- 减速期，个体数超过密度一半 ( $K/2$ ) 后，增长变慢。
- 饱和期，种群个体数达到  $K$  值而饱和。

## 2.3 阻滞增长模型

意义：

当  $P > K$  时，Logistic 系数是负值，种群数量下降

当  $P < K$  时，Logistic 系数是正值，种群数量上升

当  $P = K$  时，Logistic 系数等于零，种群数量不变

( $P$  表示种群大小， $K$  表示环境容纳量或种群的稳定平衡密度)

2003 年非典患者预测，部分学者利用逻辑斯蒂增长模型进行预测，并且准确率很高，所以我们也尝试利用 Logistic 增长模型预测全球的疫情发展趋势。

### 三. 描述性统计分析

## 描述性统计分析

截止 2020 年 4 月 16 日，全球新冠肺炎累计确诊人数 216.52 万，累计治愈人数 55.12 万，累计死亡人数 14.47 万，死亡率约为 6.69%。

表 1 六大洲治愈率与死亡率

	亚洲	欧洲	北美洲	南美洲	大洋洲	非洲
累计确诊	345846	1005111	719591	67753	7955	18918
累计治愈	166473	281060	68780	25932	4624	4375
累计死亡	12506	92692	35434	3064	77	970
治愈率(%)	48.14	27.96	9.56	38.27	58.13	23.13
死亡率(%)	3.62	9.22	4.92	4.52	0.97	5.13

»»

## 3.1 累计确诊

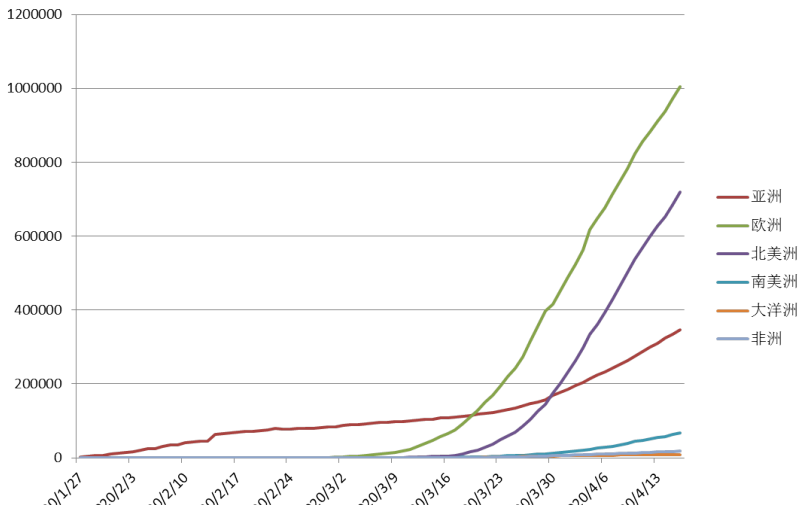
### 3.1.1 各洲首例确诊时间

洲	时间	首例确诊国家
欧洲	1 月 24 日	法国
美洲	1 月 21 日	美国
亚洲（不含中国）	1 月 20 日	日本、韩国、泰国
大洋洲	1 月 25 日	澳大利亚
非洲	2 月 14 日	埃及

Figure 2: 各洲首例确诊时间.

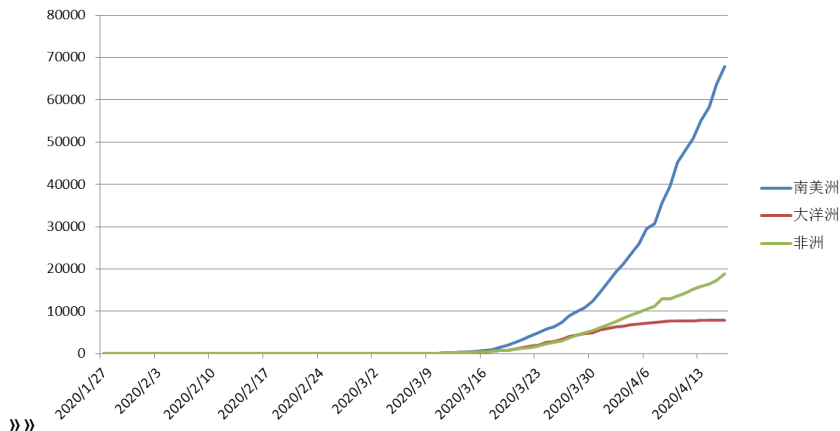
## 3.1 累计确诊

### 3.1.2 六大洲累计确诊人数



## 3.1 累计确诊

将亚洲、欧洲、北美洲去掉，只观察南美洲、大洋洲及非洲的情况。从图 2 中看出，南美洲跟非洲正处在呈指数增长的时期，而大洋洲的累计确诊人数已经放缓。



## 3.1 累计确诊

### 3.1.3 堆积面积图与百分比堆积面积图

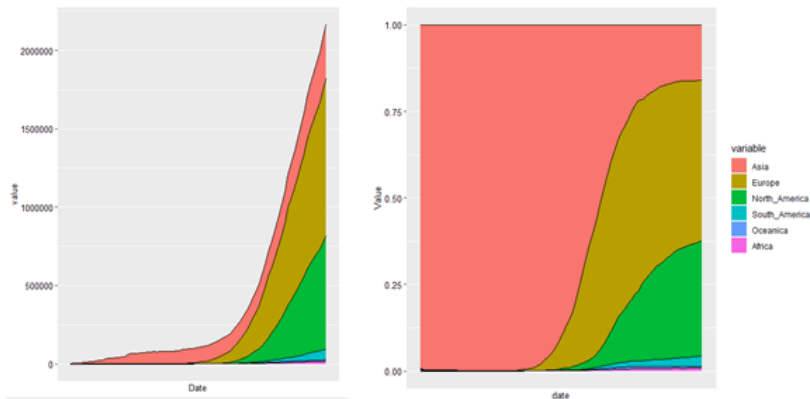


Figure 4: 堆积面积图与百分比堆积面积图.



## 3.2 新增确诊

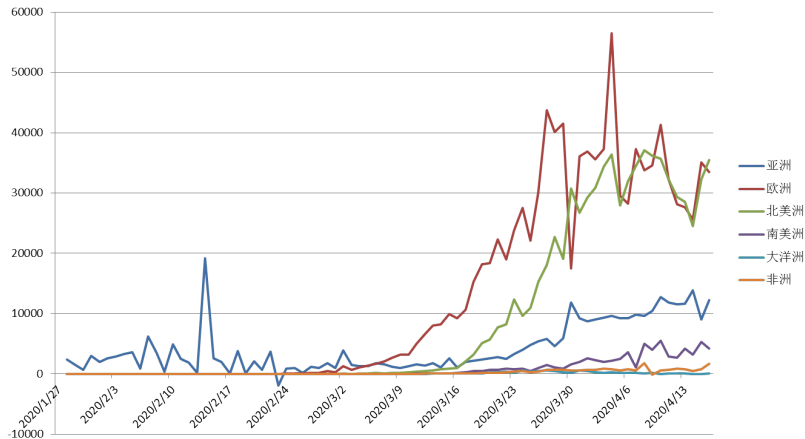
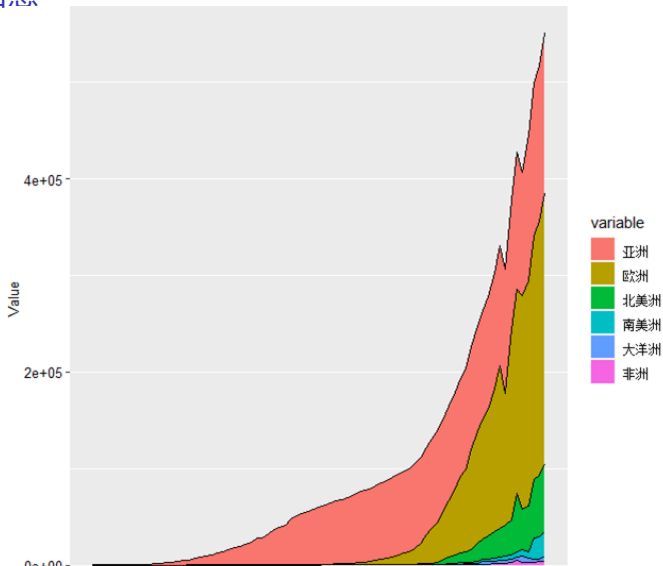
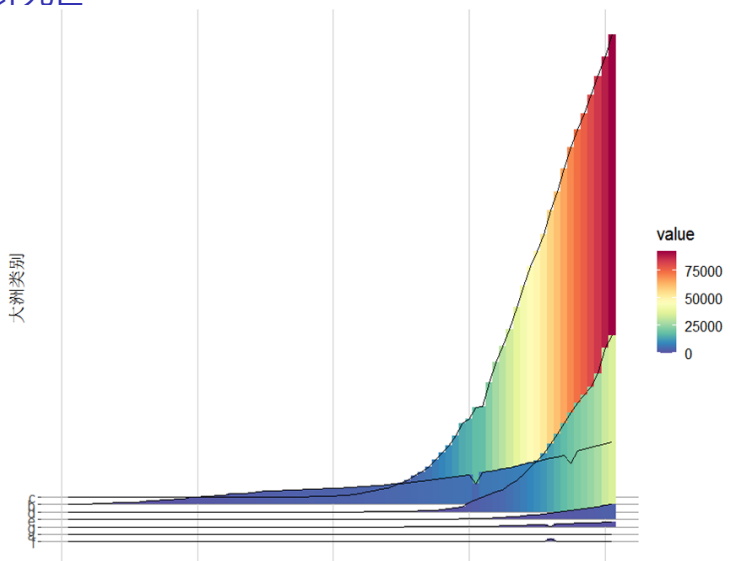


Figure 5: 新增确诊.

### 3.3 累计治愈



## 3.4 累计死亡



## 四. 六大洲发展趋势对比实证分析

## 4.1 提要

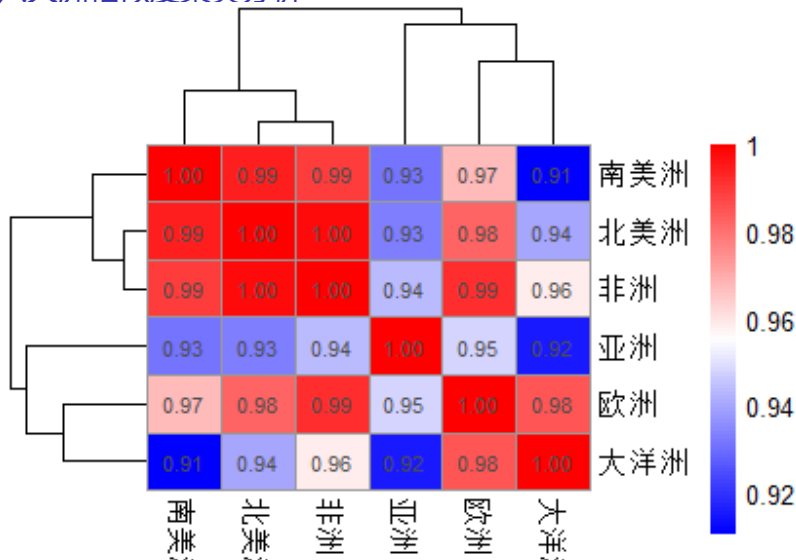
- 六大洲相似度聚类分析
- 多项式曲线拟合分析
- 基于阻滞增长模型的拟合分析

## 4.2 六大洲相似度聚类分析

我们将对疫情的全球化地域性差异做综合评价，对疫情的严重程度做合理的分类，采用系统聚类的方法对累计确诊，死亡，治愈和各大洲之间疫情相似度的发展情况进行聚类。

首先考察的是各大洲之间的相似性，即对相近的变量进行聚类，可以用来评判大洲之间疫情发展趋势的相似程度，这里采用的是系统聚类法中的类平均法。

## 4.2 六大洲相似度聚类分析



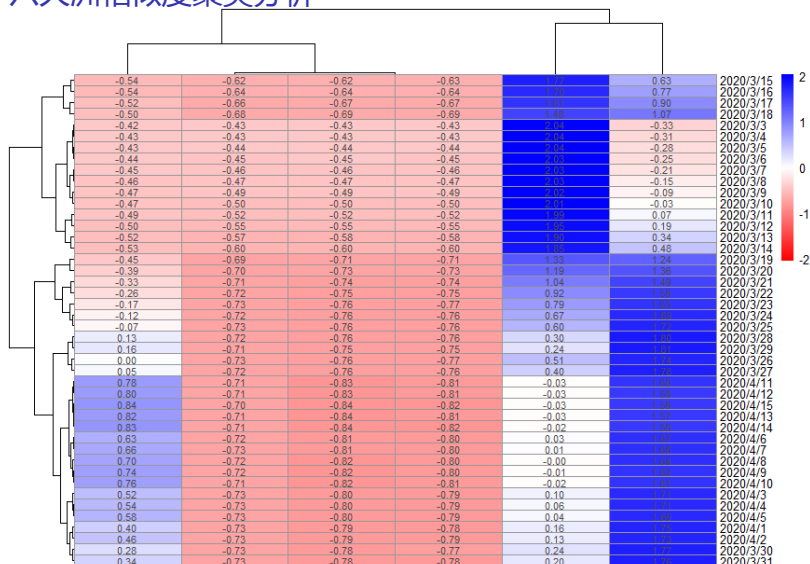
## 4.2 六大洲相似度聚类分析

将系统聚类的过程采用聚类热图进行直观展示，并在此基础上确定合理的分类数。

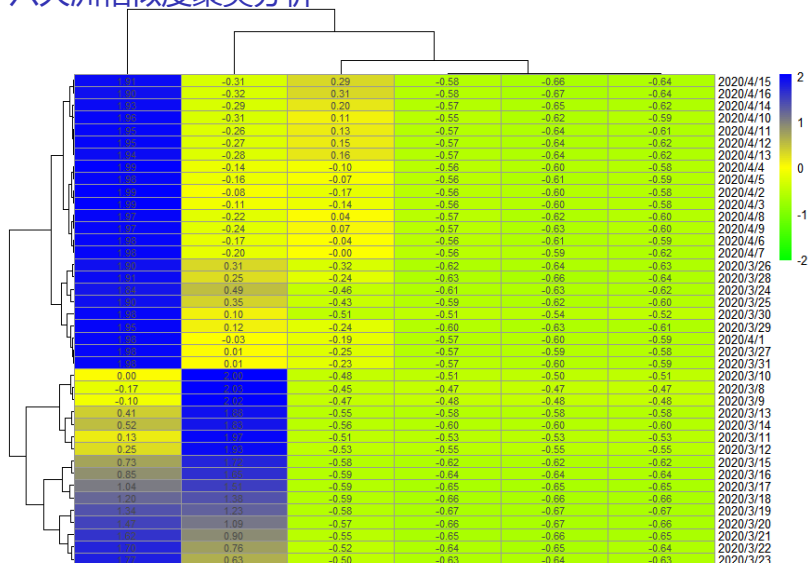
由图看出，确诊病例整体呈上升趋势，按照严重程度可以划分为三类，即欧洲，亚洲为第一类，大洋洲、南美洲和大洋洲为第二类，非洲为第三类。确诊病例按照时间可以划分为两个时间段：3月3日至3月18日，3月19日至4月16日。



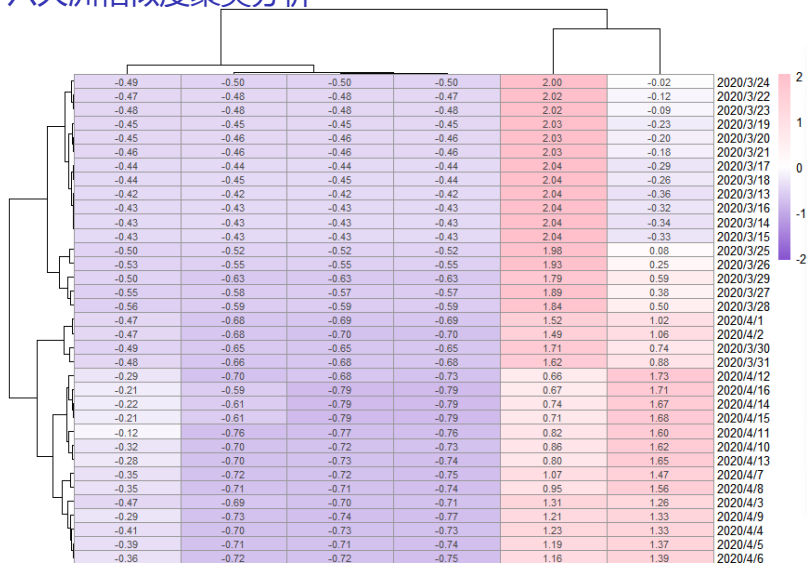
## 4.2 六大洲相似度聚类分析



## 4.2 六大洲相似度聚类分析



## 4.2 六大洲相似度聚类分析



## 4.3 多项式曲线拟合分析

多项式拟合是用一个多项式展开去拟合包含数个分析格点的一小块分析区域中的所有观测点，得到观测数据的客观分析场。根据所给定的这些离散数据点作出接近这些点的线，为拟合线。理想的拟合线是穿过了所有数据点的线，而拟合线的目的是找到这样一个综合靠点距离最近的函数，可得出多项式系数与数据点集的直接关系。本文利用 1 月 27 日至 4 月 16 日全球六个州确诊人数、治愈人数以及死亡人数分别拟合了多项式曲线，得到以下多项式曲线的表达式：

$$y = 2834 + 8154x_1 - 221.5x_1^2 + 2.238x_1^3$$

$$y = 147.3 + 2511x_2 - 242.9x_2^2 + 4.676x_2^3$$

$$y = 4265 + 9608x_3 - 452.8x_3^2 + 5.63x_3^3$$

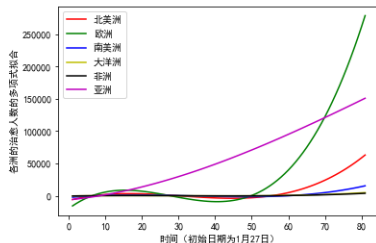
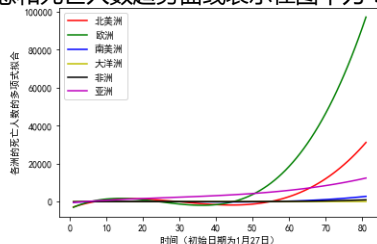
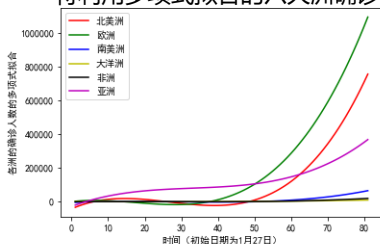
$$y = 4940 + 1031x_4 - 45.51x_4^2 + 0.5365x_4^3$$

$$y = 545.2 - 67.28x_5 + 0.8168x_5^2 + 0.0175x_5^3$$

$$y = -757.1 + 184.2x_6 - 9.356x_6^2 + 0.1251x_6^3$$

## 4.2 多项式曲线拟合分析

将利用多项式拟合的六大洲确诊、治愈和死亡人数趋势曲线表示在图中为：



## 4.3 基于阻滞增长模型的拟合分析

为了更好的刻画各大洲疫情的发展态势，做了以下 Logistic 模型进行预测。六大洲均出现病例的初始日期为 2 月 27 日，故将初始值将设定为疫情爆发日期为 2 月 27 日各大洲的人数，并且每隔 20 天进行一次拟合。基于亚洲、欧洲和北美洲的增长趋势和初始人数较大，故将亚洲、欧洲和北美洲作为一组，而南美洲、大洋洲和非洲的初始人数均较少，将其作为另外一组进行比较。

参数设置：

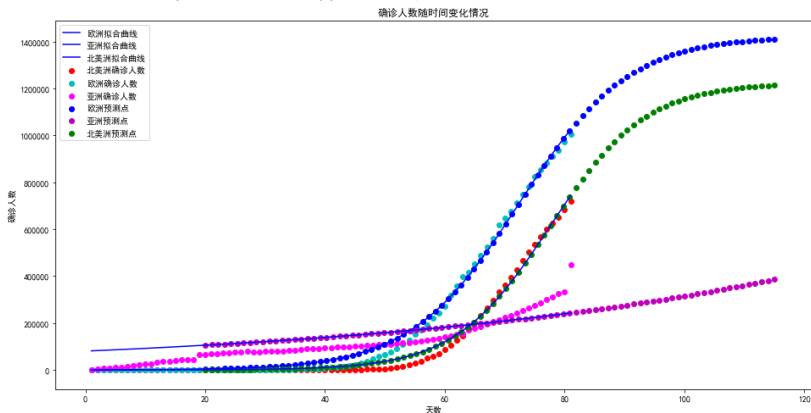
$K$ ：环境容量，即增长到最后， $P(t)$  所能达到的极限。

$p_0$ ：初始容量，即  $t=0$  时刻的数量。

$r$ ：增长速率， $r$  越大则增长越快，越快接近于  $K$  值， $r$  越小，则增长越慢，越慢接近于  $K$  值。

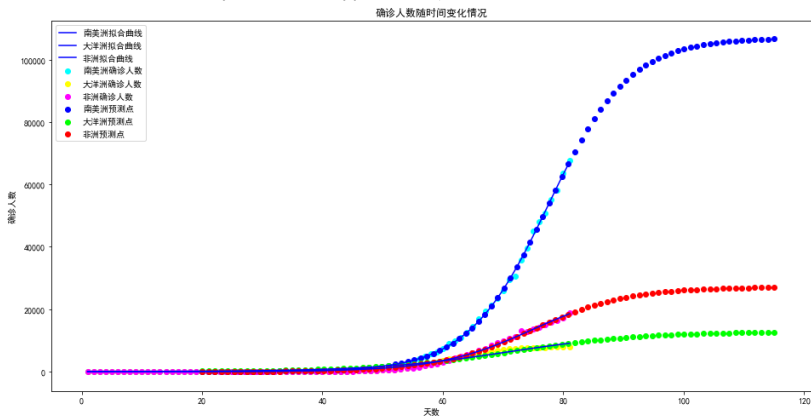
## 4.3 基于阻滞增长模型的拟合分析

亚洲、欧洲和北美洲阻滞增长模型拟合如下：



## 4.3 基于阻滞增长模型的拟合分析

南美洲、大洋洲和非洲阻滞增长模型拟合如下：





## 4.3 基于阻滞增长模型的拟合分析

结论：根据阻滞增长模型拟合的结果来看，除亚洲外，各大洲确诊人数均处于上升阶段。其中，该模型估计欧洲将于 2020 年 2 月 27 日开始 130 后达到感染高峰，约为 140 万人；北美洲将在 120 天左右达到感染高峰，约为 120 万人；南美洲将在 120 天左右达到感染高峰，约为 13 万人；大洋洲将在 105 天达到高峰，约为 2 万人以上，非洲将在 100 天左右达到疫情高峰，即约为 1.5 万人左右；而亚洲由于防疫管控措施的尽早实施，故难以以组织增长模型拟合出较好的结果，但亚洲人数增长速率已明显放缓，可见在疫情蔓延早期进行有效控制是有利于遏制疫情的发展的。

## 4.4 局限性

- 多项式拟合局限性：多项式拟合只能利用现有的疫情数据来拟合曲线的发展态势，并不能准确的预测未来具体的感染人数，也无法结合传染病传播特征进行修正，仅提供一个大致的发展态势分析。

- 阻滞增长模型局限性：

- (1) 传播初期疾病基本没有控制，但是中后期社区及医疗单位对疾病进行了严格控制，传播力度可能降低

- (2) 感染基数较大，部分病例死亡或者痊愈，降低确诊人数，logistic 模型没有考虑到

- (3) 试剂盒的检测周期和用于检测的试剂盒的量都在变化。

因此，logistic 增长模型只是对疾病进行预估，并不能准确判断，也并不是最佳模型。接下来，我们将尝试做传染病动力学的 SEIR 模型进行更进一步的研究。

**谢谢观看！**