

报名号 #1069

第二届“数学中国杯”数学建模网络挑战赛

承 诺 书

我们仔细阅读了第二届“数学中国杯”数学建模网络挑战赛的竞赛规则。

我们完全明白，在竞赛开始后参赛队员不能以任何方式（包括电话、电子邮件、网上咨询等）与队外的任何人（包括指导教师）研究、讨论与赛题有关的问题。

我们知道，抄袭别人的成果是违反竞赛规则的，如果引用别人的成果或其他公开的资料（包括网上查到的资料），必须按照规定的参考文献的表述方式在正文引用处和参考文献中明确列出。

我们郑重承诺，严格遵守竞赛规则，以保证竞赛的公正、公平性。如有违反竞赛规则的行为，我们将受到严肃处理。

我们允许数学中国网站(www.madio.net)公布论文，以供网友之间学习交流，数学中国网站以非商业目的的论文交流不需要提前取得我们的同意。

我们的参赛报名号为：

1069

参赛队员（签名）：

队员 1：张田

队员 2：岳晓蕊

队员 3：梁浩楠

参赛队教练员（签名）：

参赛队伍组别：大学组

报名号 #1069

第二届“数学中国杯”数学建模网络挑战赛

编号专用页

参赛队伍的参赛号码：（请各个参赛队提前填写好）：

1069

竞赛统一编号（由竞赛组委会送至评委团前编号）：

竞赛评阅编号（由竞赛评委团评阅前进行编号）：

报名号 #1069

2009 年 第二届“数学中国杯”

数学建模网络挑战赛

题 目
关 键 词

2030 “完美风暴”

主成份分析 灰色 Gompertz 预测模型
灰色关联度模型 危机预测评价模型

摘 要:

本题属于现在比较热门的长期预测问题。由单因素上升到多种因素相交错从而影响了全局的变化。本文才用主成分分析法和聚类分析法先对全球的主要国家进行筛选，得到了一个可以代表全球的典型国家群体，并以这个国家群体为样本进一步地讨论研究了“完美风暴”发生的可能性。建立了灰色 Gompertz 预测模型，对主要因素进行了预测，结果如下：

| | 灰色 Gompertz 预测模型预测值 | “完美”风暴发生的临界值 |
|-------------|---------------------|--------------|
| 人口（亿） | 82 | 83 |
| 能源需求量（亿吨） | 10 | 10.12485 |
| 淡水需求量（亿立方米） | 13 | 12.69775 |
| 粮食（亿吨） | 1.767 | 1.5537 |

注，本表中的时刻均为 $t=2030$ 年

由表中预测数据可以看出，预测的值和“完美”风暴发生的临界值非常接近，说明了本文预测的准确性。

本文也提出了危机预测评价模型：

$$f(u_1, u_2, u_3, u_4, u_5) = \begin{cases} \text{"完美"风暴发生} & f \geq f_0 \\ \text{"完美"风暴不发生} & f < f_0 \end{cases}$$

即 $f=0.2016*U_1+0.1924*U_2+0.1030*U_3+0.0933*U_4+0.0694*U_5$

其中 $f_0=7.926187625$

根据我们的预测值，带入上述危机预测评价模型可得：

$$f=8.1042878 > f_0$$

所以本文认为，到了 2030 年“完美”风暴会发生。

参赛队号 1069

所选题目 C

参赛密码 _____
(由组委会填写)

报名号 #1069

Abstract

This problem is now a hot issue of long-term forecast. Factor increasing from single-phase stagger to a variety of factors and then affect the overall situation changes. This article uses the principal component analysis and cluster analysis to the major countries of the world should be screened being representative of a typical group of countries around the world and sample the country groups to discuss further the “perfect storm” that maybe occurred. We establish the Gompertz gray prediction model, and the main factors for the prediction results are as follows:

| | The value getting from the Gompertz gray prediction model | The threshold of the perfect storm |
|--|---|---------------------------------------|
| Population (million) | 82 | 83 |
| Energy demand (million tons) | 10 | 10.12485 |
| Fresh water demand (100 million cubic meters) | 13 | 12.69775 |
| Grain (million tons) | 1.767 | 1.5537 |

Remark, the time in this table is $t = 2030$ (Y)

From what can be seen in the table above, the Forecast data of the predicted value and “perfect” storm that may occur are very close, illustrating the accuracy of the forecasts.

In this paper, the crisis also raise the forecast evaluation model:

That is,

$$f = 0.2016 * U_1 + 0.1924 * U_2 + 0.1030 * U_3 + 0.0933 * U_4 + 0.0694 * U_5$$

In which $f_0 = 7.926187625$

According to our forecast, the forecast evaluation of the crisis into the above-mentioned model available:

$$f = 8.1042878 > f_0$$

Therefore this article believes that in 2030 the “perfect” storm will happen.

报名号 # 1069

1. 问题的重述

英国科学家约翰·贝丁顿于3月18日在英国《卫报》发表文章警告称，气候变化和人口增长将导致食品、水和能源短缺，进而引发大规模移民、公共骚乱和国际冲突。如果未来数年内没有充分准备的话，到2030年，世界将面临“完美风暴”，出现大的动荡。

贝丁顿说，人口的持续增长将会在未来20年里引发对食品、水和能源的大量需求，与此同时，各国政府还必须应对气候变化。所有这一切都将同时到来。据贝丁顿教授说，再过20多年，全球人口将增加到大约83亿。仅此一点，就意味着全球资源需求将在未来变得更大。贝丁顿博士警告说，资源短缺压力将急剧增加，气候变化将使问题恶化。目前，全球的粮食储存量太小，只有年消费量的14%，一旦发生干旱或者洪灾，粮食就会严重短缺。他说：“我们的粮食储备是50年来最低的，到2030年，我们对粮食的需求会增加50%，同时，我们对能源的需求也会增加50%，淡水的需求会增加30%。”编者按：人口问题、粮食问题、淡水资料问题、能源问题、环境气候问题及移民问题这些历年来都是数学建模赛题考察的热点，然而这些问题如果集中在一起将引发不可预想的恶果。正如约翰·贝丁顿所说，如果这些问题集中在一起同时爆发，谁也无法保证不会发生世界范围内的动荡或战争。

问题1：

收集上述问题的有效数据论证约翰·贝丁顿的观点，建立危机预测模型，并分析其影响因子之间的关系，根据时间的推移得出最坏的结果或判定“完美风暴”发生的可能性（可建立多个问题子模型，影响因子可根据收集数据的多少作出调整。引用各国数据均以英文单词的第一个大写字母标注国家）。

问题2：

针对“完美风暴”发生或不发生，试写一篇可持续发展报告，向社会阐述你的观点，此报告可作为网上公开展示的新闻稿。

2. 问题的分析

2.1 背景分析

世界人口在不断增长，全球也在不断升温，粮食紧缺、能源短缺等一些列问题不断出现，人类面临着严峻的形势。我们既不能过于悲观，相信所谓的世界灭亡论；但也不可太乐观，坐吃山空。正如约翰·贝丁顿所说，我们要认真地去听取科学的意见。科学的预测以达到未雨绸缪的目的。

2.2 问题1的分析

“完美风暴”原指由一组因素将局势急剧恶化的事件，这里理解为人口增长、气候变化进而食品、能源和水了短缺，2030人类的生存与发展局势全面恶化。建立危机模型，并分析影响因子间的关系，判定“完美风暴”发生的可能性。世界由许多国家组成，全部都考虑显然不易实现，所以首先利用主成份分析法和聚类分析法选出适量、适当的国家，即取出样本来代表世界的水平。利用熵值法对样本中不同的个体赋予不同的权重，进而构成了一个整体。之后，利用灰色Gompertz模型进行预测，并用BP神经网络模型进行检验。

2.3 问题2的分析

在第一问中我们验证了“完美风暴”的论断，但是单纯的依靠这一问中的理论分

报名号 #1069

析与计算结果，显然也是不科学的。也就是说，第一问中的结果为“完美风暴”的发生提供了理论上的证明，但我们并不能因此而说它必然发生。综合考虑如模型本身的缺陷、数据的不准确性、2030年之前可能发生移民等大的变故各方面因素，来更加客观又贴合实际的判断“完美风暴”发生的可能性。其实，数学建模是数学运用于并指导实践的一种形势，仅是第一问中的预测就显得太数学化，我们主要的目标还是走可持续发展道路，减缓甚至减除人类的危机。

3. 问题的假设

- 3.1 假设我们查到的数据真实可靠，对最终结果不会产生太大误差；
- 3.2 假设人口、气候共同影响粮食需求、水资源需求、能源需求，研究期间，不考虑其他因素；
- 3.3 由于数据查询的困难性及概念的模糊性，假设可以用谷物量、人均水资源消耗量能源消耗量分别代表粮食需求、水资源需求、能源需求，温度代表气候变化；
- 3.4 筛选出的8个国家作为样本可以代表整个世界，其他国家则忽略不计；
- 3.5 假设聚类分析中，关联度较大的即为较相似的，在对“完美风暴”发生的贡献上是等价的，一类的可以只取个别的做样本；
- 3.6 假设经主成分分析中引用的8个指标足够权威，以其为指标误差可以接受；
- 3.7 假设研究期间没有发生重大变革，如近于全球性的自然灾害、传染病、大规模的移民等

4. 符号说明

| 符号 | 符号表示的意义 |
|---|---------------------------------|
| m | 主成分个数 |
| $score_i$ | 第 i 个国家的综合得分 |
| $factor_{ik}$ | 第 i 个国家在第 k 个主成分上的得分 |
| $corrcoefX$ | 和主成分分析法相关的相关系数矩阵 |
| $corroefX(i,1:n)$ | 相关系数矩阵的第 i 列 |
| β_k | 各主成分归一化后的权重 |
| w_i | 熵值法中每个 p_{ij} 的权重系数 |
| $P = (p_{ij})_{8 \times 8}$ | 第 i 个国家的第 i 个指标的值 |
| r_{jk} | 指标 j 与指标 k 的相关系数 |
| x_{ij} | 第 i 个样品的第 j 个指标的观测数据 |
| $e_j = -k \sum_{i=1}^n p_{ij} \ln(p_{ij})$ | 第 j 项指标的熵值 |
| $D_{ij} = \min_{X_i \in G_i, X_j \in G_j} d_{ij}$ | 类 G_i 与 G_j 之间的距离为两类最近样品的距离 |
| q_i | 第 i 个影响“完美”风暴发生因素的权重 |
| $U = (u_{1 \times 5})$ | 5 个影响“完美”风暴发生的因素归一化后的值 |
| $f_0 = 7.926187625$ | “完美”风暴发生的临界值 |

报名号 #1069

5. 预备模型的建立

5.1 主成份分析

判定“完美风暴”发生的可能性，虽然是全球性的问题，但考虑所有的国家显然是不可行的。主成分分析是把原来多个变量划为少数几个综合指标的一种统计分析方法，即一种降维处理技术。

5.1.1 观测 29 个对象，记第 i 个观测对象 8 个指标的观测值分别为： $X_{i1}, X_{i2}, \dots, X_{i8}$ ，则所有 29 个对象 8 个指标的观测值可以表示为矩阵形式：

$$X = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1p} \\ X_{21} & X_{22} & \cdots & X_{2p} \\ \cdots & \cdots & \cdots & \cdots \\ X_{n1} & X_{n2} & \cdots & X_{np} \end{bmatrix}$$

其中， n ($=29$) 为观察对象， p ($=8$) 为指标或变量。如图：

表 1 世界主要国家的一些评价指标

| 国家 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|-------|----------|-----|-------|----------|-------------|----------|-------|-------------|
| 中 国 | 960 | 89 | 2010 | 131179.8 | 26680.71223 | 36320000 | 2156 | 44535.51563 |
| 印 度 | 328.726 | 150 | 820 | 110981.1 | 9062.679839 | 8608800 | 1152 | 23913 |
| 印度尼西亚 | 190.457 | 104 | 1420 | 22304.16 | 3644.589015 | 14520.8 | 12867 | 6601.064453 |
| 伊 朗 | 164.82 | 58 | 3000 | 6915.254 | 2228.894925 | 22237.9 | 1300 | 2280.992188 |
| 日 本 | 37.791 | 40 | 38410 | 12756.49 | 43401.33396 | 63117.6 | 3365 | 1174.147827 |
| 韩 国 | 9.926 | 41 | 17690 | 4841.808 | 8880.242033 | 22153.6 | 1344 | 665.309082 |
| 马来西亚 | 32.974 | 29 | 5490 | 2576.66 | 1489.403576 | 8505.2 | 22882 | 223.3999939 |
| 菲 律 宾 | 30 | 39 | 1420 | 8458.976 | 1169.313464 | 3536.9 | 5767 | 2140.901855 |
| 新 加 坡 | 0.0699 | 44 | 29320 | 439.3208 | 1321.584722 | 2108.9 | 211 | 300.546473 |
| 泰 国 | 51.312 | 79 | 2990 | 6472.442 | 2062.470349 | 11544.1 | 3269 | 3314.629395 |
| 土 耳 其 | 78.356 | 55 | 5400 | 7293.532 | 4027.099382 | 10415.2 | 3150 | 3459.828369 |
| 埃 及 | 100.145 | 169 | 1350 | 7539.691 | 1074.840371 | 8252.7 | 500 | 2273.643311 |
| 南 非 | 121.909 | 16 | 5390 | 4739.09 | 2549.915648 | 17093.9 | 955 | 945.3253784 |
| 加 拿 大 | 998.467 | 22 | 36170 | 3255.646 | 12514.63397 | 38952.5 | 88238 | 5089.450195 |
| 墨 西 哥 | 195.82 | 51 | 7870 | 10422.14 | 8391.819264 | 20427.4 | 3967 | 3195.916504 |
| 美 国 | 963.203 | 21 | 44970 | 29898.81 | 132018.193 | 293130.6 | 9446 | 34656.19531 |
| 阿 根 廷 | 278.04 | 58 | 5150 | 3912.046 | 2140.577464 | 8561.5 | 7123 | 3355.633301 |
| 巴 西 | 851.488 | 40 | 4730 | 18869.44 | 10679.6227 | 21191.1 | 29066 | 5901.74707 |
| 委内瑞拉 | 91.205 | 10 | 6070 | 2702.092 | 1818.618429 | 8690.9 | 27185 | 407.8241882 |
| 法 国 | 55.15 | 11 | 36550 | 6103.751 | 22307.21446 | 24889.5 | 2932 | 6181.269043 |
| 德 国 | 35.703 | 22 | 36620 | 8241.144 | 29066.81246 | 42108 | 1297 | 4347.47998 |
| 意 大 利 | 30.134 | 30 | 32020 | 5857.1 | 18447.48878 | 25701.6 | 3114 | 2014.542969 |
| 波 兰 | 31.269 | 43 | 8190 | 3811.241 | 3387.329085 | 12708.7 | 1404 | 2177.593018 |
| 俄罗斯联邦 | 1709.824 | 21 | 5780 | 14236.76 | 9869.395886 | 86974.3 | 30135 | 7686.60498 |
| 西 班 牙 | 50.537 | 30 | 27570 | 4354.63 | 12239.88478 | 16849.5 | 2562 | 1935.315552 |
| 乌 克 兰 | 60.355 | 35 | 1950 | 4657.115 | 1061.108941 | 18286.5 | 1128 | 3369.800049 |
| 英 国 | 24.361 | 21 | 40180 | 6036.08 | 23450.14657 | 31646 | 2408 | 2082.040771 |
| 澳大利亚 | 774.122 | 22 | 35990 | 2052.094 | 7681.776681 | 14925.2 | 24202 | 1633.150024 |
| 新 西 兰 | 27.053 | 14 | 27250 | 412.5376 | 1038.732083 | 2011.7 | 79778 | 80.31999969 |

报名号 #1069

其中:

指标 1: 国土面积 (万/平方公里)

指标 2: 总悬浮颗粒物 (微克/立方米)

指标 3: 人均国民总收入 (美元)

指标 4: 人口/万人

指标 5: 国内生产总值 (美元)

指标 6: 能源储量/万吨

指标 7: 人均淡水量/立方米

指标 8: 粮食 (谷物) /万吨

5.1.2 对原始数据进行无量纲化处理

由于上述指标是有量纲的, 有的指标是绝对数, 有的指标是相对数, 有的指标数量等级大, 有的指标数量等级小, 不进行无量纲化的处理, 无法进行综合评价。所以在综合评价之前必须消除各指标的量纲影响, 即通过数学变换消除原始指标的量纲影响。以统计指标为基础进行综合评价的研究中, 本文采用采用 Z-score 法来消除量纲。Z-score 计算方法如下:

$$Z_{ij} = (X_{ij} - \bar{X}_j) / S_j \cdots \cdots 5.1.2-1$$

式中 $\bar{X}_j = \sum_{i=1}^n X_{ij} / n$; $S_j^2 = \sum_{i=1}^n (X_{ij} - \bar{X}_j)^2 / (n-1)$; $i=1, 2, \cdots, n$; $j=1, 2, \cdots, p$ 。标准化处理后, 指标的方差为 1, 均值为 0。

5.1.3 求指标的相关系数矩阵

$R = (r_{jk})_{p \times p}$; $j, k=1, 2, \cdots, p$ 。 r_{jk} 为指标 j 与指标 k 的相关系数, 且:

$$r_{jk} = \frac{\sum_{i=1}^n (X_{ij} - \bar{X}_j)(X_{ik} - \bar{X}_k)}{\sqrt{\sum_{i=1}^n (X_{ij} - \bar{X}_j)^2 \sum_{i=1}^n (X_{ik} - \bar{X}_k)^2}} \cdots \cdots 5.1.2-2$$

5.1.4 求相关矩阵 R 的特征值与特征向量, 确定主成份。确定主成份由特征方程式 $|\lambda I - R| = 0$, 可求得 P 个特征根 λ_g ($g=1, 2, \cdots, p$)。将特征根按大小顺序排列为

$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$ 。主成份方差的大小描述了各个主成份在描述被评价对象上所起作用的大小。由特征方程式, 特征值 λ_g 对应的一个标准特征向量为 L_g , 记

$L_g = (l_{g1}, l_{g2}, \cdots, l_{gp})^T$ 。令 $C_g = \sqrt{\lambda_g} L_g$ ($g=1, 2, \cdots, p$), 则 C_g 表示各指标在第 g 个主成分上的载荷, $C = [C_1, C_2, \cdots, C_p]$ 为因子载荷矩阵。将原指标变量转换为主成份

$F_g = (X_1, X_2, \cdots, X_p) C_g$ 。 F_1 称为第 1 主成份, F_2 称为第 2 主成份, \cdots, F_p 成为第 P 主成份。

5.1.5 求方差贡献率以确定主成份个数。主成份分析法是选取尽量少的 K 个主成份 ($K < P$) 来综合评价, 同时还要使损失的信息尽可能少。K 值为方差贡献率

$$\lambda_g / \sum_{g=1}^p \lambda_g \geq 80\%。$$

5.1.6 对主成份进行综合评价

报名号 # 1069

记 $V = [V_1, V_2, \dots, V_n]^T$; $U = (u_{ij})_{n \times k}$; $Z = (z_{ij})_{n \times p}$; $A = (L_1, L_2, \dots, L_k)$;
 $B = [\beta_1, \beta_2, \dots, \beta_k]^T$ 。通过 $U=ZA$ 可计算出每个样本在各主成分上的得分, 通过 $V=UB$ 可计算出每个样本综合得分, 进而可对各样本进行综合评价。其中: u_{ij} 表示第 i 个样本的综合得分。 Z 为原始数据标准化后得到的矩阵, L_g ($g=1, 2, \dots, k$) 表示相关系数矩阵 R 的特征根 λ_g 的一个标准特征向量 β_g ($g=1, 2, \dots, k$) 表示 λ_g 的方差贡献率。

5.1.7 建立国家的评价指标体系:

$$\begin{cases} score_i = \sum_{k=1}^m factor_{ik} \\ factor_{ik} = \beta_k \cdot \text{corrcoef}X(i, 1:n)^T \cdot L_{\lambda_k} \end{cases} \quad (i=1, 2, 3, \dots, 29)$$

从而得到这 29 个国家的排序, 如下表所示:

表 2 世界主要国家的评价体系得分及排名

| 国家 | 个主成份得分 | | | | 综合得分 | 排名 |
|-------|---------|---------|---------|---------|---------|----|
| 中 国 | 1.3202 | -0.0581 | 0.0742 | -0.2962 | 1.04 | 1 |
| 印 度 | 0.6086 | 0.0307 | 0.012 | -0.0587 | 0.5926 | 2 |
| 俄罗斯联邦 | 0.1697 | -0.1644 | 0.1534 | 0.0705 | 0.2293 | 3 |
| 埃 及 | 0.0967 | 0.0662 | -0.0025 | 0.0436 | 0.2041 | 4 |
| 印度尼西亚 | 0.1283 | 0.0077 | 0.0271 | 0.025 | 0.1881 | 5 |
| 巴 西 | 0.1281 | -0.0988 | 0.096 | 0.0378 | 0.163 | 6 |
| 泰 国 | 0.0625 | 0.0175 | 0.001 | 0.0189 | 0.0998 | 7 |
| 阿 根 廷 | 0.0596 | -0.0148 | 0.0228 | 0.0226 | 0.0901 | 8 |
| 伊 朗 | 0.0579 | 0.0046 | 0.0053 | 0.019 | 0.0868 | 9 |
| 土 耳 其 | 0.0581 | -0.0029 | -0.0006 | 0.0141 | 0.0687 | 10 |
| 墨 西 哥 | 0.0708 | -0.0238 | 0.0023 | 0.0186 | 0.0679 | 11 |
| 菲 律 宾 | 0.043 | 0.0037 | 0.0093 | 0.0067 | 0.0627 | 12 |
| 乌 克 兰 | 0.044 | 0.002 | 0.0006 | 0.0094 | 0.056 | 13 |
| 波 兰 | 0.0357 | -0.0059 | -0.0075 | 0.009 | 0.0313 | 14 |
| 加 拿 大 | 0.0495 | -0.2433 | 0.2046 | 0.0157 | 0.0282 | 15 |
| 马来西亚 | 0.0067 | -0.0269 | 0.0435 | -0.0002 | 0.0231 | 16 |
| 南 非 | 0.0273 | -0.0151 | 0.0021 | 0.0066 | 0.0209 | 17 |
| 委内瑞拉 | 0.0022 | -0.0461 | 0.0561 | -0.0036 | 0.0086 | 18 |
| 澳大利亚 | 0.0423 | -0.1409 | 0.0648 | 0.021 | -0.0128 | 19 |
| 韩 国 | 0.0277 | -0.0289 | -0.0211 | 0.0064 | -0.0159 | 20 |
| 美 国 | 0.4008 | -0.4233 | -0.1206 | 0.102 | -0.0412 | 21 |
| 新 加 坡 | 0.0049 | -0.0317 | -0.0236 | -0.0014 | -0.0518 | 22 |
| 西 班 牙 | 0.0288 | -0.0627 | -0.027 | 0.0033 | -0.0576 | 23 |
| 新 西 兰 | -0.0439 | 0.1359 | 0.1486 | -0.0292 | -0.0604 | 24 |
| 意 大 利 | 0.0341 | -0.0808 | -0.0375 | 0.0038 | -0.0804 | 25 |
| 法 国 | 0.0523 | -0.112 | -0.0453 | 0.0011 | -0.1039 | 26 |

报名号 # 1069

| | | | | | | |
|-----|--------|---------|---------|--------|---------|----|
| 德 国 | 0.0567 | -0.113 | -0.0573 | 0.0068 | -0.1068 | 27 |
| 日 本 | 0.0666 | -0.1304 | -0.0696 | 0.0172 | -0.1162 | 28 |
| 英 国 | 0.0317 | -0.1067 | -0.0505 | 0.0011 | -0.1245 | 29 |

通过对上述排序结果的分析，我们发现表 1 结果非常合理：

不难看出，容易产生“完美风暴”的国家往往是发展中国家，因为它们的制度还不够健全、社会保障等还不够完善，经济条件也不太好，而人口往往还比较多；相反像英国、日本等国家，地理面积比较小，且高度发达，容易治理，即使出现了一些危机兆头，也会比较快的解决。基于这样的思考与认识，我们再来看表 2 时，会发现其结果是非常令人满意的，这也为后面的预测奠定了良好的基础，尽可能的减小初始误差。

为更直观的描述表 2 中的结果，我们用图 1 表示，可方便下一步聚类分析

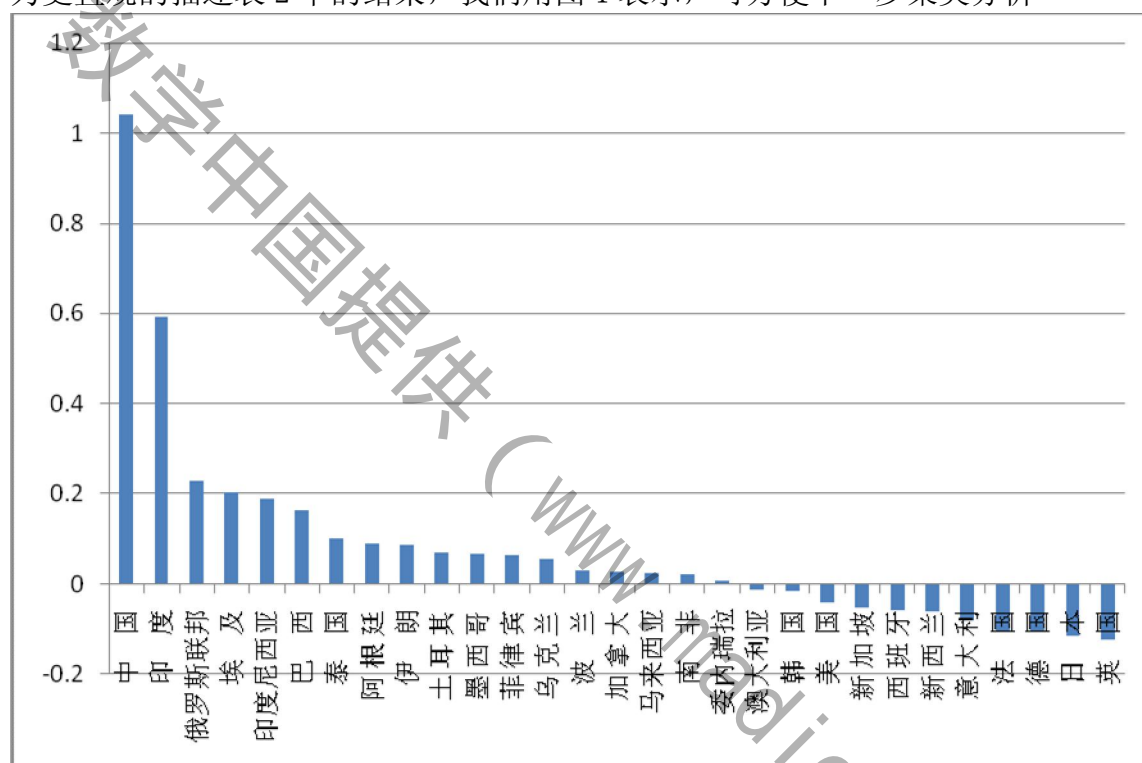


图 1 29 个主要国家的评价体系指标得分排名

5.2 聚类分析

主成份分析法，我们初步筛选出了 29 个国家并对其进行了合乎常理的排名，但考虑到国家的数目仍然太多，并且一些国家在综合评判时比较相似，即对“完美风暴”发生与否的影响差不多，我们吸取分层抽样的思想，再进一步通过聚类分析，找出典型国家代表这 29 个国家，进而代表全球。聚类分析是研究分类问题的一种多元统计方法，类即指相似元素的集合。

5.2.0 相关理论：

设有 n 个样品，每个样品测得 p 项指标（变量），原始资料矩阵为：

$$x_1 \quad x_2 \quad \cdots \quad x_p$$

报名号 # 1069

$$X \equiv \begin{matrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{matrix} \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}$$

其中 x_{ij} ($i=1, \dots, n$; $j=1, \dots, p$) 为第 i 个样品的第 j 个指标的观测数据, 第 i 个样品 X_i 为矩阵 X 的第 i 行所描述, 所以任何两个样品 X_k 与 X_l 的相似性, 可以通过矩阵 X 中的第 k 行与第 l 行的相似程度来刻画; 任何两个变量 x_k 与 x_l 之间的相似性, 可以通过第 k 列与第 l 列的相似程度来刻画。

由于上述指标是有量纲的, 有的指标是绝对数, 有的指标是相对数, 有的指标数量等级大, 有的指标数量等级小, 不进行无量纲化的处理, 无法进行综合评价。所以在综合评价之前必须消除各指标的量纲影响, 即通过数学变换消除原始指标的量纲影响。以统计指标为基础进行综合评价的研究中, 本文采用采用 Z-score 法来消除量纲。Z-score 计算方法如下:

$$Z_{ij} = (X_{ij} - \bar{X}_j) / S_j$$

式中 $\bar{X}_j = \sum_{i=1}^n X_{ij} / n$; $S_j^2 = \sum_{i=1}^n (X_{ij} - \bar{X}_j)^2 / (n-1)$; $i=1, 2, \dots, n$; $j=1, 2, \dots, p$ 。标准化处理后, 指标的方差为 1, 均值为 0。

用于聚类分析的最短距离法:

定义类 G_i 与 G_j 之间的距离为两类最近样品的距离, 即 $D_{ij} = \min_{X_i \in G_i, X_j \in G_j} d_{ij}$, 设类 G_p 与 G_q 合并成一个新类记为 G_r , 则任一类 G_k 与 G_r 的距离是

$$D_{ij} = \min_{X_i \in G_i, X_j \in G_j} d_{ij} = \min \left\{ \min_{X_i \in G_k, X_j \in G_q} d_{ij}, \min_{X_i \in G_k, X_j \in G_p} d_{ij} \right\} = \min \{ D_{kp}, D_{kq} \}。$$

5.2.1 定义样品之间的距离, 计算样品两两距离, 得一距离矩阵记为 $D_{(0)}$, 开始每个样品自成一类, 这时显然 $D_{ij} = d_{ij}$ 。

5.2.2 找出 $D_{(0)}$ 的非对角线最小元素, 设为 D_{pq} , 则将 G_p 与 G_q 合并成一个新类, 记为 G_r , 即 $G_r = \{G_p, G_q\}$ 。

5.2.3 给出计算新类与其他类的距离公式 $D_{kr} = \min \{D_{kp}, D_{kq}\}$, 将 $D_{(0)}$ 中第 p, q 行及 p, q 列用上面的公式并成一个新行新列, 新行新列对应 G_r , 所得到的矩阵记为 $D_{(1)}$ 。

5.2.4 对 $D_{(1)}$ 重复上述对 $D_{(0)}$ 的第 2、3 步得 $D_{(2)}$; 如此下去, 直到所有的元素并成一类为止。

如果某一步 $D_{(k)}$ 中非对角线最小的元素不止一个, 则对应这些最小元素的类可以同时合并。

聚类图, 如下图所示

报名号 # 1069

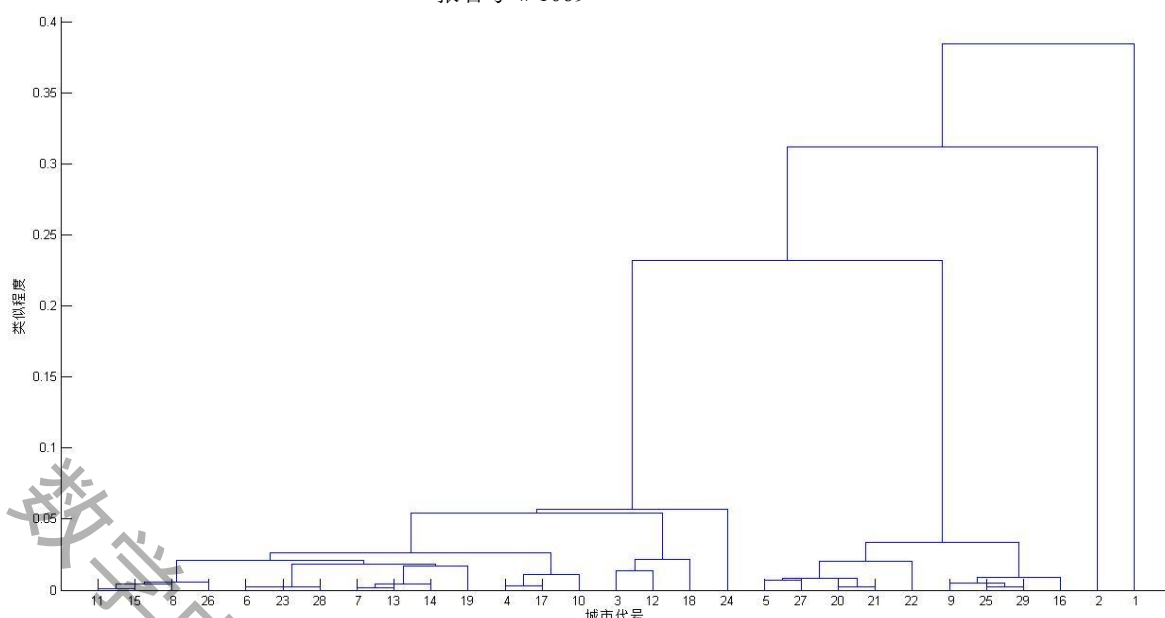


图 2 29 个国家的聚类分析图

最终本文结合分层抽样的原则从 29 个具有代表性的国家里选出了以下 8 个最具有代表性的国家，分别为：

中国 印度 俄罗斯 巴西 土耳其 南非 美国 日本

5.3 熵值法求权系数

1 基本原理

在信息论中，熵是对不确定性的一种度量。信息量越大，不确定性就越小，熵也就越小；信息量越小，不确定性越大，熵也就越大。根据熵的特性，我们可以通过计算熵来判断一个事件的随机性及无序程度，也可以用熵值来判断某个指标的离散的程度，指标的离散程度越大，该指标对综合评价的影响越大。

2 熵值法步骤

(1) 选取 n 个国家， m 个指标，则 x_{ij} 为第 i 个国家的第 j 个指标的数值。（ $i=1, 2, \dots, n; j=1, 2, \dots, m$ ）

(2) 指标的标准化处理：异质指标同质化

由于各个指标的计量的单位并不统一，因此在使用它们计算综合指标之前，我们先要对它们进行标准化处理，即把指标的绝对值转化为相对值，并令 $x'_{ij} = \frac{x_{ij} - \min(x_{1j}, x_{2j}, \dots, x_{nj})}{\max(x_{1j}, x_{2j}, \dots, x_{nj}) - \min(x_{1j}, x_{2j}, \dots, x_{nj})}$ ，从而解决各项不同质指标的同质化问题。而且，由于正向指标和负向指标数值代表的含义不同（正向指标值越高越好，负向指标数值越低越好），因此，对于高低指标我们用不同的算法进行数据标准化。

结合本题，用正向指标，具体如下：

$$\text{正向指标: } x'_{ij} = \left[\frac{x_{ij} - \min(x_{1j}, x_{2j}, \dots, x_{nj})}{\max(x_{1j}, x_{2j}, \dots, x_{nj}) - \min(x_{1j}, x_{2j}, \dots, x_{nj})} \right] \times 100$$

则 x'_{ij} 为第 i 个国家的第 j 个指标的数值。（ $i=1, 2, \dots, n; j=1, 2, \dots, m$ ），为方便起见，人记数据为 $x'_{ij} = x_{ij}$ 。

(3) 计算第 j 项指标下第 i 个国家占该指标的比重：

报名号 #1069

$$p_{ij} = \frac{X_{ij}}{\sum_{i=1}^n X_{ij}}, \quad (i=1, 2, \dots, n, j=1, 2, \dots, m)$$

(4) 计算第 j 项指标的熵值。

$$e_j = -k \sum_{i=1}^n p_{ij} \ln(p_{ij}), \quad \text{其中, } k > 0, \quad k = 1 / \ln(n), \quad e_j \geq 0$$

(5) 计算第 j 项指标的差异系数。对第 j 项指标, 指标值的差异越大, 对方按评价的左右就越大, 熵值就越小, 定义差异系数:

$$g_j = \frac{1 - e_j}{m - E_e},$$

$$E_e = \sum_{j=1}^m e_j, \quad 0 \leq g_j \leq 1, \quad \sum_{j=1}^m g_j = 1$$

(6) 求权值:

$$w_i = \frac{g_i}{\sum_{j=1}^m g_j} \quad (i=1, 2, \dots, n)$$

(7): 计算各国家的综合得分:

$$s_i = \sum_{j=1}^m w_j * p_{ij} \quad (i=1, 2, \dots, n)$$

本文以表 1 中的 8 个指标作为计算各个国家综合得分的指标, 通过 MATLAB 编程, 可得这 8 个国家的权重为:

表 3 8 个主要国家的权重

| 国家 | 中国 | 印度 | 俄罗斯 | 巴西 | 土耳其 | 南非 | 美国 | 日本 |
|-------|--------|--------|--------|--------|--------|--------|--------|--------|
| 熵值法权重 | 0.0757 | 0.0102 | 0.1034 | 0.0937 | 0.1253 | 0.3611 | 0.1246 | 0.1060 |

6. 模型的建立

6.1 灰色 Gompertz 模型——预测模型

灰色系统分析是我国学者邓聚龙教授于 20 世纪 80 年代前期提出的用于控制和预测的新理论、新技术。但是, 从一个动态系统中得到的实际数据, 往往具有一定的波动性。若在建模中仅仅追求与原始数据的拟合, 则在模型中必然承袭这种误差, 而忽视对系统总体的拟合。只从数学角度得出的模型, 容易沿袭这种误差。传统的灰色模型在长期预测中常常不尽如人意, 原因也是忽略了这种实际数据的波动性。同小军等人后来提出了改进的灰色模型——灰色 Gompertz 模型。该模型以严格的数学推理为基础, 考虑了实际数据带有误差, 以绝对误差为标准, 由背景函数导出。对与问题 1 中的预测问题我们建立如下灰色 Gompertz 模型

灰色 Gompertz 模型的计算方法为:

- (1) 令原数据为 $y^{(0)}$, $y^{(0)}(i) > 0$ ($i=1, 2, \dots, n$) 对于 $y^{(0)}$ 采用对数变换进行生成处理, 即令

报名号 # 1069

$$x^{(0)}(i) = \ln[y^{(0)}(i)] \quad (i = 1, 2, \dots, n)$$

- (2) 对于数据序列 $x^{(0)}(i) = \{x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)\}$, 一般利用下述方程组的最小二乘解

$$\begin{bmatrix} x^{(1)}(2) + x^{(1)}(3) & 2 & 1 \\ x^{(1)}(3) + x^{(1)}(4) & 3 & 1 \\ \vdots & \vdots & \vdots \\ x^{(1)}(n-1) + x^{(1)}(n) & n-1 & 1 \end{bmatrix} \begin{bmatrix} p_2 \\ p_3 \\ p_4 \end{bmatrix} = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n-1) \end{bmatrix} - 2 \begin{bmatrix} x^{(1)}(2) \\ x^{(1)}(3) \\ \vdots \\ x^{(1)}(n-1) \end{bmatrix}$$

由 p_2 、 p_3 、 p_4 及 $p_2 = -e^a$ 解得参数 a ;

- (3) 利用 $x^{(0)}(k) = be^{-ak} + c$ 通过线性拟合得到参数 b, c . 从而, 得还原模型为

$$\hat{y}^{(0)}(k) = e^{c+b^*e^{-a^*k}}$$

6.2 灰色关联度分析方法模型

灰色综合评价主要是依据以下模型: $R=Y \times W$

式中, R 为 M 个被评价对象的综合评价结果向量; W 为 N 个评价指标的权重向量; E 为各指标的评判矩阵, (矩阵略)

$\xi_i(k)$ 为第 i 个被评价对象的第 K 个指标与第 K 个最优指标的关联系数。根据 R 的数值, 进行排序。

6.2.1 确定最优指标集

设 $F = [j_1^*, j_2^*, \dots, j_n^*]$, 式中 j_k^* 为第 k 个指标的最优值。此最优序列的每个指标值可以是诸评价对象的最优值, 也可以是评估者公认的最优值。选定最优指标集后, 可构造矩阵 D (矩阵略)

式中 j_k^i 为第 i 个评价对象第 k 个指标的原始数值。

6.2.2 指标的规范化处理

由于评判指标间通常是有不同的量纲和数量级, 故不能直接进行比较, 为了保证结果的可靠性, 因此需要对原始指标进行规范处理。设第 k 个指标的变化区间为 $[j_{k1}, j_{k2}]$, j_{k1} 为第 k 个指标在所有被评价对象中的最小值, j_{k2} 为第 k 个指标在所有被评价对象中的最大值, 则可以用下式将上式中的原始数值变成无量纲值 $C_k^i \in (0,1)$ 。

$$C_k^i = \frac{j_k^i - j_{k1}}{j_{k2} - j_{k1}}, \quad i = 1, 2, \dots, m, \quad k = 1, 2, \dots, n \quad (\text{矩阵略})$$

6.2.3 计算综合评判结果

根据灰色系统理论, 将 $\{C^*\} = [C_1^*, C_2^*, \dots, C_n^*]$ 作为参考数列, 将 $\{C\} = [C_1^i, C_2^i, \dots, C_n^i]$ 作为被比较数列, 则用关联分析法分别求得第 i 个被评价对象的第 k 个指标与第 k 个指标最优指标的关联系数, 即

$$\xi_i(k) = \frac{\min_i \min_k |C_k^* - C_k^i| + \rho \max_i \max_k |C_k^* - C_k^i|}{|C_k^* - C_k^i| + \rho \max_i \max_k |C_k^* - C_k^i|}$$

式中 $\rho \in (0,1)$, 一般取 $\rho = 0.5$ 。

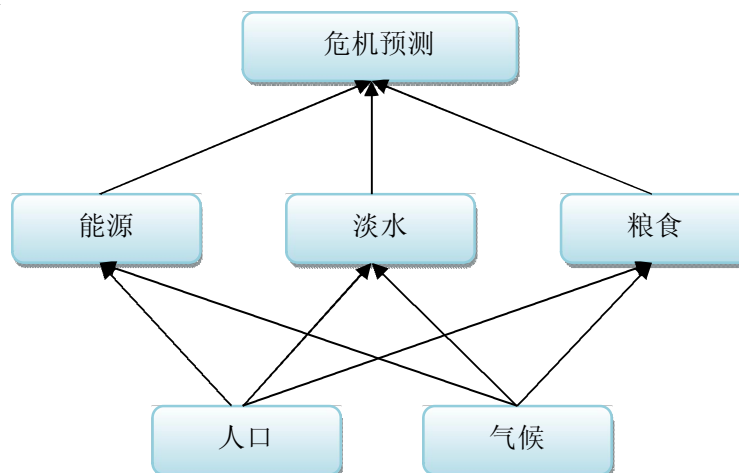
这样综合评价结果为: $R = E \times W$

若关联度 r_i 最大, 说明 $\{C\}$ 与最优指标 $\{C^*\}$ 最接近, 即第 i 个被评价对象优于其他

报名号 #1069

被评价对象，据此可以排出各被评价对象的优劣次序。

6.3 危机预测模型



危机函数 $f = \sum_{i=1}^s q_i * X_i$, 其中 q_i 为每项因素的权重, X_i 为每个因素的实际值

7. 模型的求解

7.1 灰色 Gompertz 模型的求解

通过 MATLAB 编程, 本文分别对人口、气候、能源、粮食和淡水共 6 条时间序列进行灰色 Gompertz 模型预测, 最终得出:

表 3 人口时间序列

| 年份 | 1950 | 1960 | 1970 | 1980 | 1990 | 2000 | 2009 |
|---------|-------------|-------------|--------------|--------------|--------------|--------------|-------------|
| 世界人口(亿) | 25.194 7 | 30.000 0 | 36.965 88 | 44.422 95 | 52.795 19 | 60.855 72 | 68.000 0 |

表 4 人口预测

| 年份 | 2020 | 2030 | 2030 年完美风暴临界值 |
|---------|------|------|---------------|
| 世界人口(亿) | 75 | 82 | 83 |

报名号 # 1069

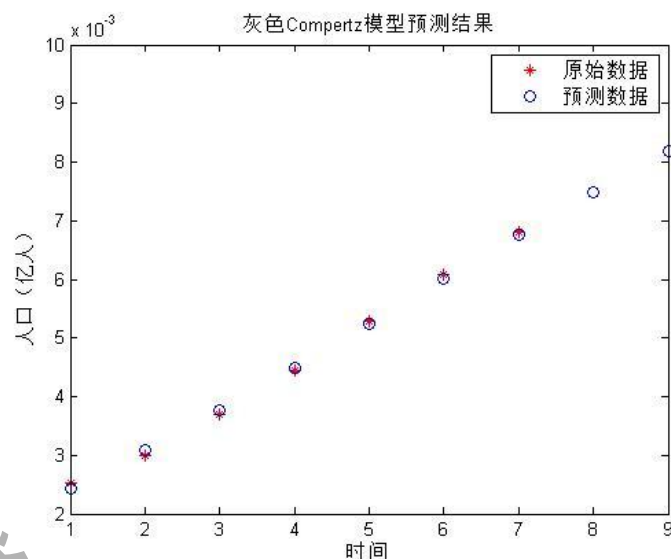


图 3 人口时间序列的预测曲线

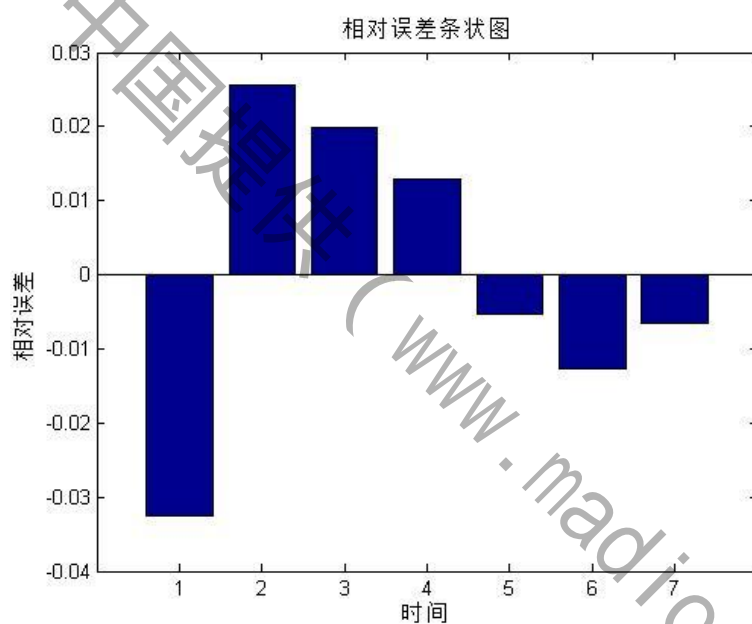


图 4 口时间序列的预测曲线的相对误差分析

表 5 粮食时间序列

单位：万吨

| 年份 | 1976 | 1981 | 1986 | 1991 | 1996 | 2001 | 2006 |
|-----|---------|---------|---------|---------|---------|---------|---------|
| 中国 | 37081.5 | 37746.3 | 39648.2 | 40522.4 | 40732.5 | 42557.6 | 44535.5 |
| 印度 | 23981.4 | 23061.1 | 23000.1 | 23204.8 | 23400.7 | 23632.5 | 23913 |
| 俄罗斯 | 6408.8 | 8362.3 | 6489.9 | 6546.4 | 7001.3 | 7321.5 | 7686.6 |
| 巴西 | 4659.7 | 5632.5 | 6700.5 | 6638.8 | 6324.7 | 6102.1 | 5901.7 |
| 土耳其 | 2665.7 | 2557.1 | 2989.5 | 3079.8 | 3149.6 | 3289.3 | 3459.8 |
| 南非 | 945.3 | 960.3 | 1001.1 | 1074.1 | 1074.7 | 1174 | 1324.5 |
| 美国 | 32531.5 | 34386.6 | 34656.2 | 34687.5 | 34746.5 | 34864.4 | 34981.2 |

报名号 # 1069

| | | | | | | | |
|-----------------------|--------|--------|--------|--------|--------|--------|--------|
| 日本 加权序列 (*10e4) | 1276.9 | 1227 | 989.7 | 1082.3 | 1090.5 | 1102.2 | 1174.1 |
| | 0.9015 | 0.9567 | 0.9694 | 0.9814 | 0.9866 | 1.0088 | 1.0358 |

表 6 粮食预测

| 年份 | 2011 | 2016 | 2021 | 2026 | 2031 | 2030 年完美风暴临界值 |
|---------|-------|-------|-------|-------|-------|---------------|
| 粮食 (亿吨) | 1.097 | 1.175 | 1.291 | 1.472 | 1.767 | 1.5537 |

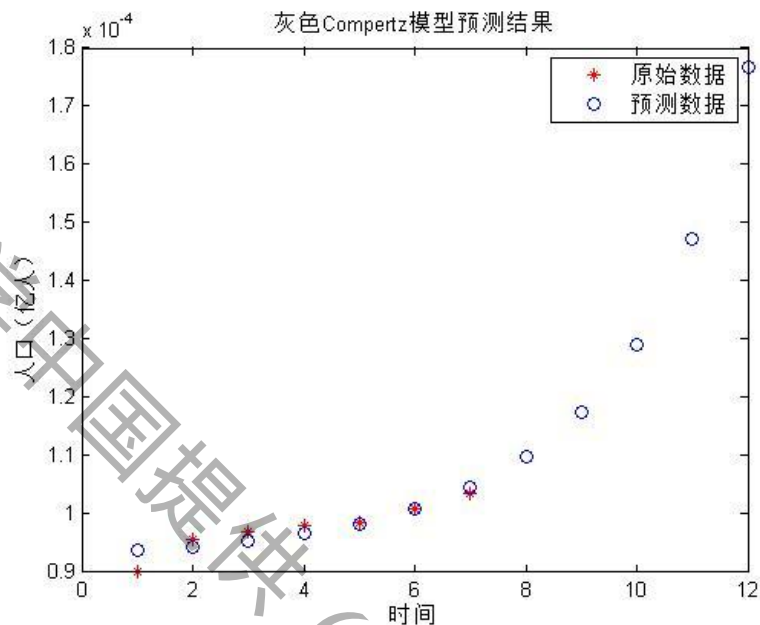


图 5 粮食时间序列的预测曲线

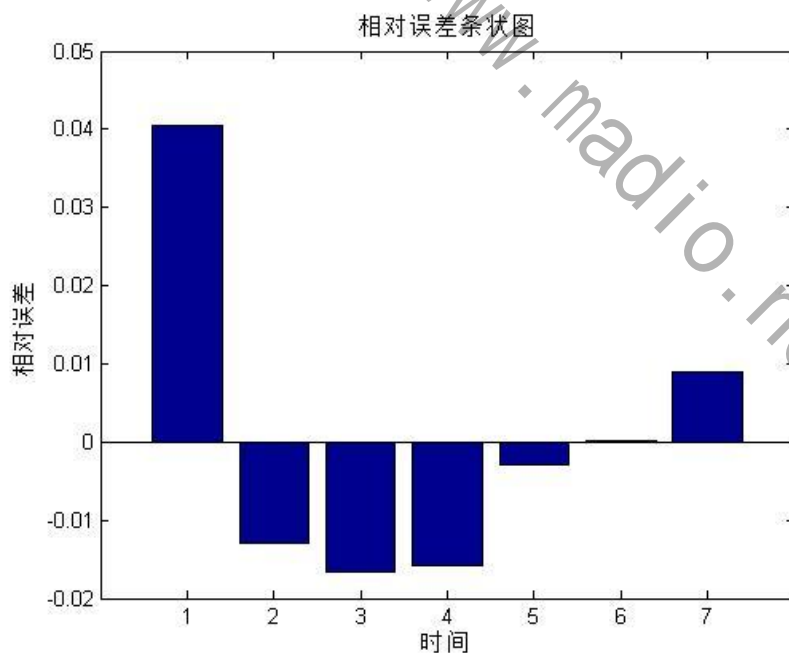


图 6 粮食时间序列的相对误差分析

表 7 能源需求时间序列
单位：万吨

报名号 #1069

| 年份 | 1976 | 1981 | 1986 | 1991 | 1996 | 2001 | 2006 |
|-----------------|--------|--------|--------|--------|--------|--------|--------|
| 中国 | 112258 | 118765 | 122857 | 140938 | 160935 | 171657 | 182317 |
| 印度 | 51198 | 5200 | 53831 | 55339 | 57285 | 58384 | 59030 |
| 俄罗斯 | 61395 | 61567 | 61784 | 63972 | 64153 | 65240 | 66342 |
| 巴西 | 18570 | 19876 | 19066 | 19325 | 20485 | 21111 | 22000 |
| 土耳其 | 7703 | 7638 | 7542 | 7895 | 8191 | 8321 | 8520 |
| 南非 | 11290 | 11301 | 11346 | 11857 | 13113 | 14613 | 16613 |
| 美国 | 230419 | 230000 | 229041 | 228079 | 232589 | 245278 | 249879 |
| 日本 | 52894 | 52124 | 51693 | 51710 | 53320 | 55230 | 56347 |
| 加权序列 (*10e4) | 5.6467 | 5.6493 | 5.7084 | 5.8830 | 6.1714 | 6.5049 | 6.7499 |

表 8 能源需求预测

| 年份 | 2011 | 2016 | 2021 | 2026 | 2031 | 2030 年完美风暴临界值 |
|---------|------|------|------|------|------|---------------|
| 粮食 (亿吨) | 7 | 8 | 9 | 10 | 11 | 10.12485 |

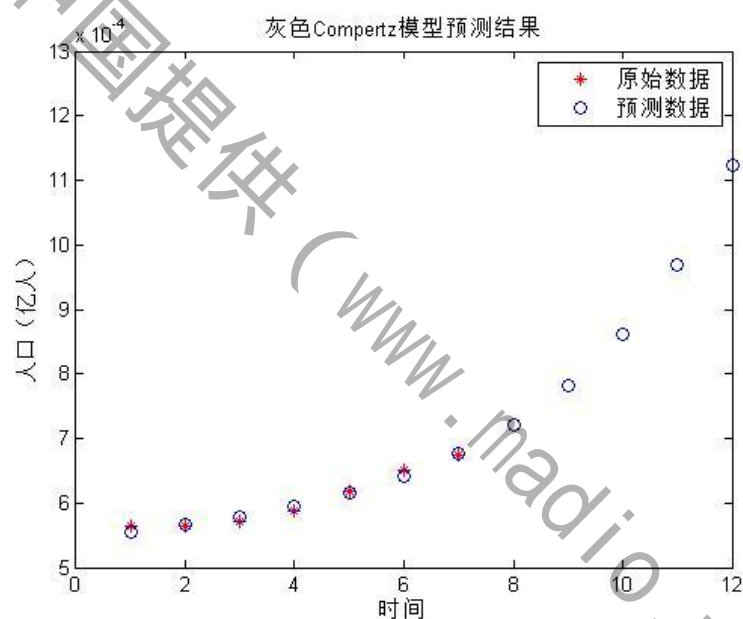


图 7 能源需求量时间序列的预测曲线

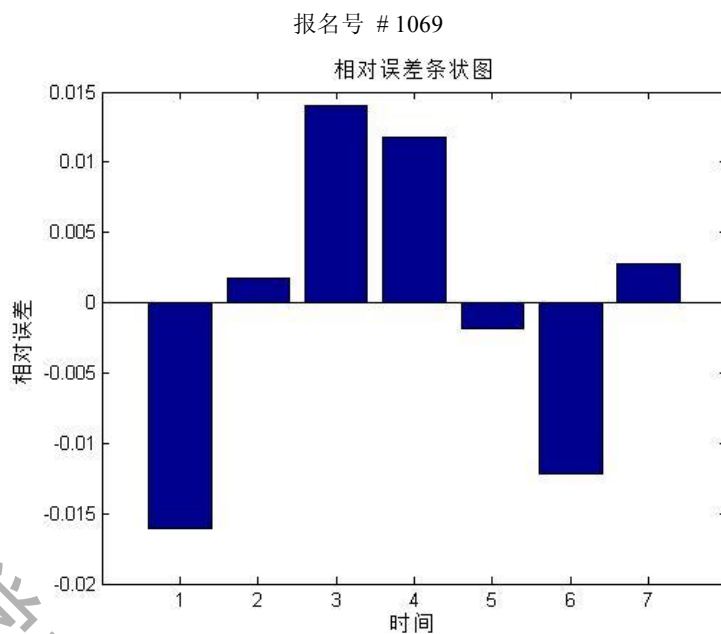


图 8 能源需求量时间序列的相对误差分析

表 9 淡水需求量时间序列
单位：亿立方米

| 年份 | 1976 | 1981 | 1986 | 1991 | 1996 | 2001 | 2006 |
|-----------------|--------|--------|--------|--------|--------|--------|--------|
| 中国 | 1932.1 | 2009 | 2100 | 2131.3 | 2207.2 | 2230 | 2241 |
| 印度 | 1054 | 1152 | 1167 | 1287 | 1489 | 1623 | 1878 |
| 俄罗斯 | 28875 | 29000 | 29981 | 30212 | 30438 | 30845 | 30904 |
| 巴西 | 28768 | 29066 | 29460 | 32457 | 34785 | 37089 | 42944 |
| 土耳其 | 3000 | 3150 | 3165 | 3276 | 3478 | 3612 | 3845 |
| 南非 | 920 | 955 | 984 | 991 | 1001 | 1112 | 1168 |
| 美国 | 8801 | 8978 | 9121 | 9387 | 9446 | 9599 | 8801 |
| 日本 | 3360 | 3365 | 3366 | 3369 | 3375 | 3380 | 3389 |
| 加权序列 (*10e4) | 7.9991 | 8.1008 | 8.2765 | 8.6347 | 8.9209 | 9.2584 | 9.7675 |

表 10 淡水需求量预测

| 年份 | 2011 | 2016 | 2021 | 2026 | 2031 | 2030年完美风暴临界值 |
|--------|------|------|------|------|------|--------------|
| 粮食（亿吨） | 10 | 11.1 | 11.4 | 12 | 13 | 12.69775 |

报名号 # 1069

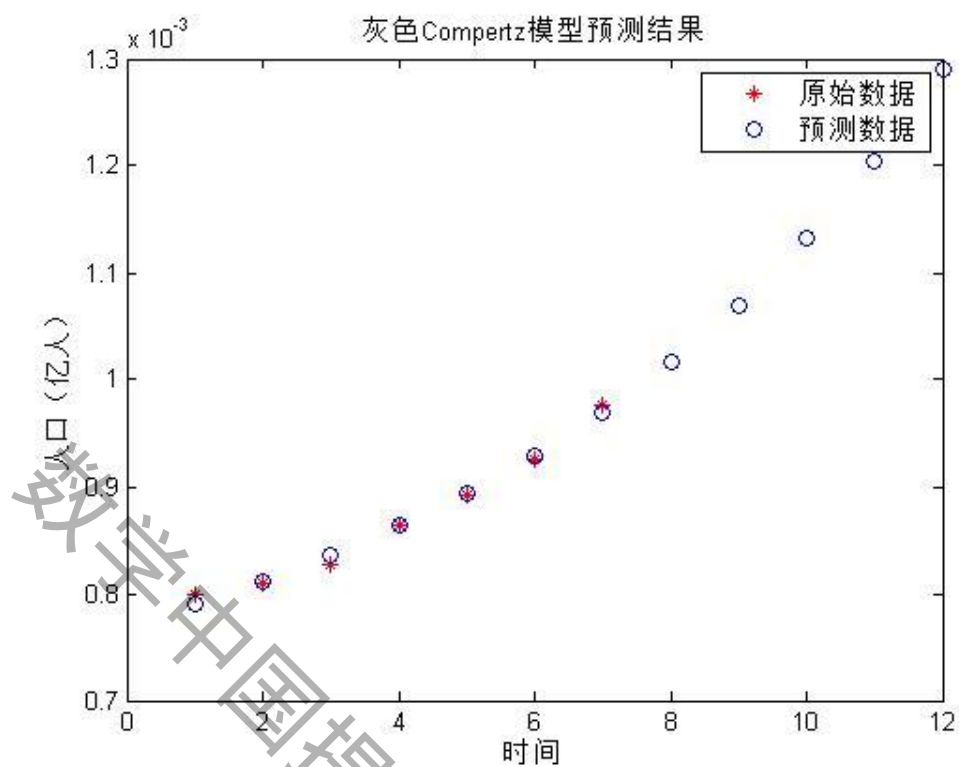


图 9 淡水需求量的时间序列的预测曲线

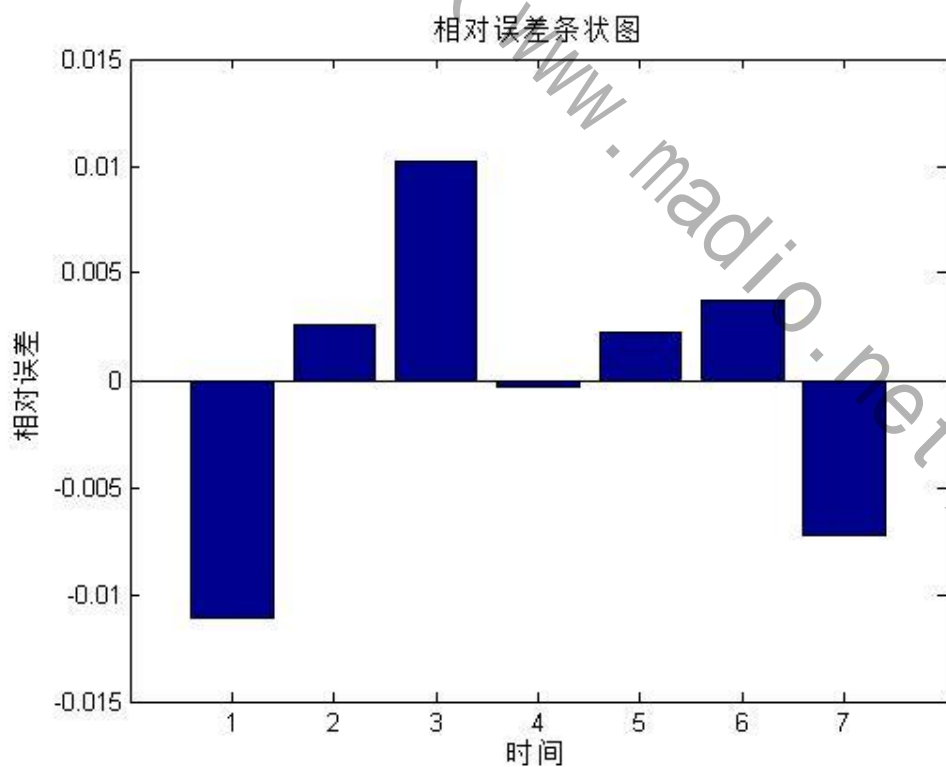


图 10 淡水需求量时间序列的相对误差分析

表 11 气温时间序列

报名号 #1069

单位：摄氏度

| 年份 | 1976 | 1981 | 1986 | 1991 | 1996 | 2001 | 2006 |
|------|--------|--------|--------|--------|--------|--------|--------|
| 中国 | 11.5 | 11.55 | 11.6 | 11.63 | 11.7 | 11.8 | 11.8 |
| 印度 | 26.7 | 26.7 | 26.75 | 26.75 | 26.78 | 26.8 | 26.85 |
| 俄罗斯 | 4.8 | 4.8 | 4.9 | 4.9 | 4.9 | 4.8 | 5 |
| 巴西 | 23.3 | 23.3 | 23.5 | 23.55 | 23.6 | 23.6 | 23.7 |
| 土耳其 | 11.4 | 11.4 | 11.5 | 11.54 | 11.6 | 11.6 | 11.7 |
| 南非 | 17.7 | 17.8 | 18 | 18 | 18.2 | 18.2 | 18.2 |
| 美国 | 12.2 | 12.2 | 12.3 | 12.3 | 12.4 | 12.5 | 12.56 |
| 日本 | 15.2 | 15.2 | 15.3 | 15.3 | 15.4 | 15.5 | 15.58 |
| 加权序列 | 14.773 | 14.813 | 14.954 | 14.966 | 15.079 | 15.100 | 15.159 |
| | 6 | 5 | 7 | 7 | 8 | 3 | 3 |

表 12 气温预测

| 年份 | 2011 | 2016 | 2021 | 2026 | 2031 | 2030 年完美风暴临界值 |
|---------|------|------|------|------|------|---------------|
| 粮食 (亿吨) | 15.3 | 15.5 | 15.7 | 15.8 | 15.9 | — |

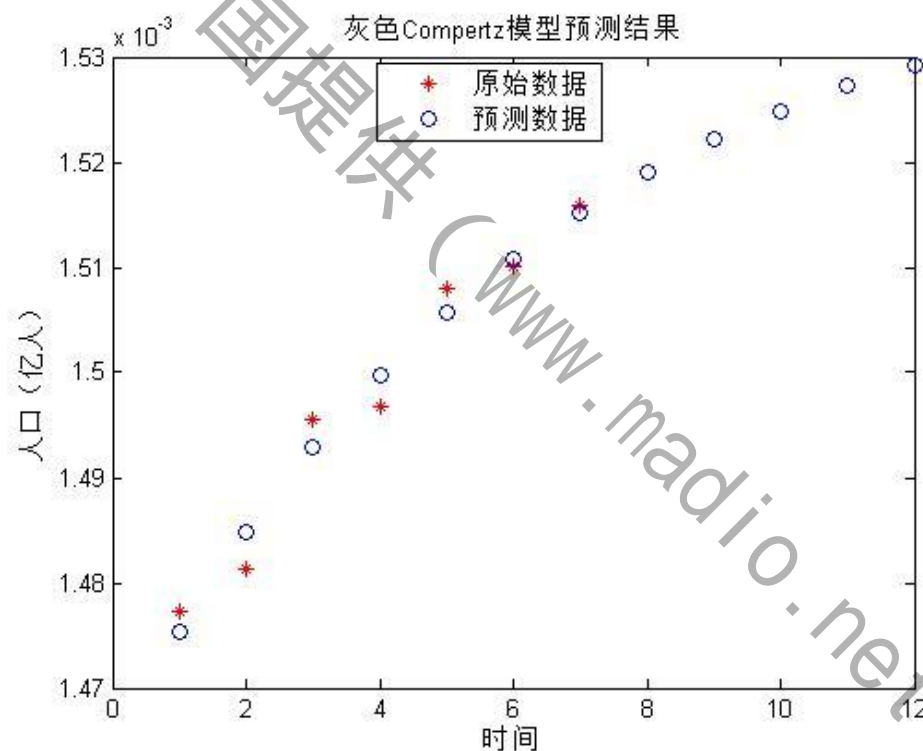


图 11 气温时间序列的预测曲线

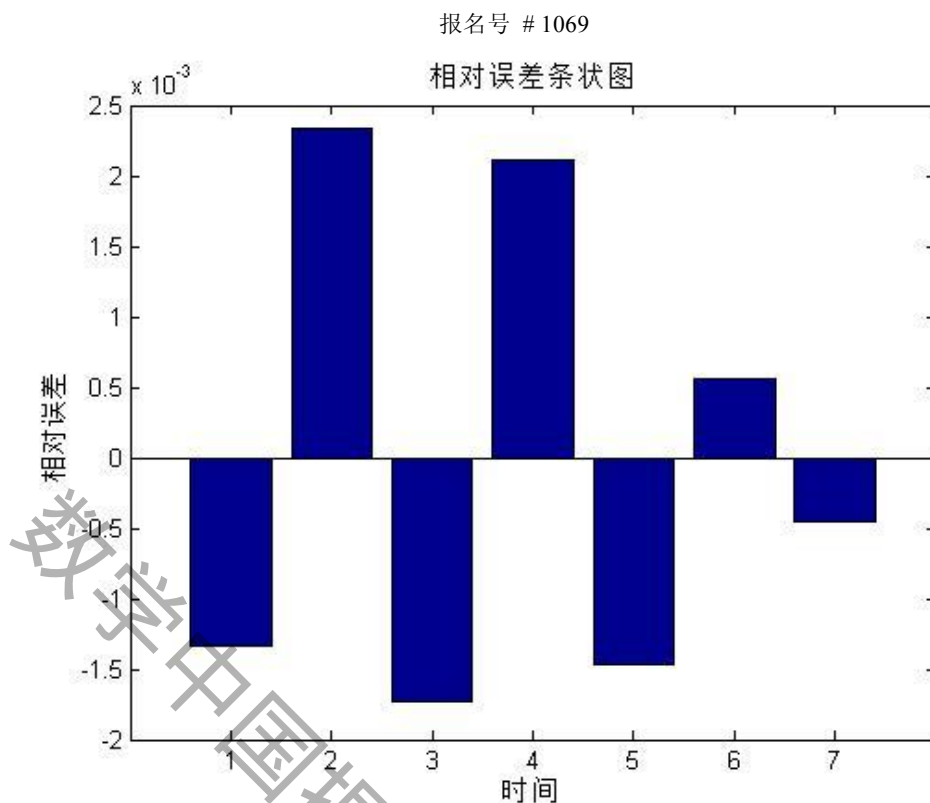


图 12 气温时间序列的相对误差分析

7.2 灰色关联度分析方法模型的求解

表 13 五个因素的加权时间序列

| 年份 | 1976 | 1981 | 1986 | 1991 | 1996 | 2001 | 2006 |
|----|--------|--------|--------|--------|--------|--------|--------|
| 粮食 | 0.9015 | 0.9567 | 0.9694 | 0.9814 | 0.9866 | 1.0088 | 1.0358 |
| 能源 | 5.6467 | 5.6493 | 5.7084 | 5.8830 | 6.1714 | 6.5049 | 6.7499 |
| 淡水 | 7.9991 | 8.1008 | 8.2765 | 8.6347 | 8.9209 | 9.2584 | 9.7675 |
| 气温 | 14.773 | 14.813 | 14.954 | 14.966 | 15.079 | 15.100 | 15.159 |
| 年份 | 6 | 5 | 7 | 7 | 8 | 3 | 3 |
| 年份 | 1975 | 1980 | 1985 | 1990 | 1995 | 2000 | 2009 |
| 人口 | 40.737 | 44.422 | 48.439 | 52.795 | 56.923 | 60.855 | 68.000 |
| | 4 | 95 | 47 | 19 | 53 | 72 | |

表 14 灰色关联度

| | 粮食 | 能源 | 淡水 | 气温 | 人口 |
|----|--------|--------|--------|--------|--------|
| 粮食 | 1 | 0.3438 | 0.3180 | 0.2604 | 0.3591 |
| 能源 | 0.3438 | 1 | 0.3061 | 0.5110 | 0.2942 |
| 淡水 | 0.3180 | 0.3061 | 1 | 0.3915 | 0.3011 |
| 气温 | 0.2604 | 0.5110 | 0.3915 | 1 | 0.3188 |
| 人口 | 0.3591 | 0.2942 | 0.3011 | 0.3188 | 1 |

7.3 危机预测模型的求解

首先本文采用基于判断矩阵 U 的熵值法

其中 $U = \begin{bmatrix} 89 & 131180 & 36320000 & 2156 & 44535.5 \end{bmatrix}$

报名号 #1069

| | | | | |
|-----|--------|---------|-------|---------|
| 150 | 110981 | 8608800 | 1152 | 23913.0 |
| 21 | 14237 | 86974 | 30135 | 7686.6 |
| 40 | 18869 | 21191 | 29066 | 5901.7 |
| 55 | 7294 | 10415 | 3150 | 3459.8 |
| 16 | 4739 | 17094 | 955 | 945.3 |
| 21 | 29899 | 293131 | 9446 | 34656.2 |
| 40 | 12756 | 63118 | 3365 | 1174.1] |

利用前文所述的熵值法，可得五大因素对危机（“完美”风暴）爆发的影响权重为

表 15 五大因素对机（“完美”风暴）爆发的影响权重为

| 因素 | 气候 | 人口 | 能源 | 淡水 | 粮食 |
|------|--------|--------|--------|--------|--------|
| 影响权重 | 0.2016 | 0.1924 | 0.1030 | 0.0933 | 0.0694 |

从而构造线性函数：

$$f=0.2016*U_1+0.1924*U_2+0.1030*U_3+0.0933*U_4+0.0694*U_5$$

考察当 U_i = “完美” 风暴临界值 U_0 时， $f_0=7.926187625$

所以可以建立一个评价系统：

$$f(u_1, u_2, u_3, u_4, u_5) = \begin{cases} \text{"完美"风暴发生} & f \geq f_0 \\ \text{"完美"风暴不发生} & f < f_0 \end{cases}$$

根据我们的预测值，带入上述危机预测评价模型可得：

$$f=8.1042878 > f_0$$

所以本文认为，到了 2030 年 “完美” 风暴会发生。

8. 模型的评价及改进

8.1 模型的评价

不难发现，灰色 Compertz 模型有以下特点：

8.1.1 是精确级差格式；

8.1.2 该级差格式和两项 $x^{(1)}(k)$ ， $x^{(1)}(k+1)$ 或者 $x^{(1)}(k-1)$ ， $x^{(1)}(k)$ 相关，从而更有利于信息的综合利用；

8.1.3 类似地， $x^{(0)}(k+1)$ 的级差格式也可以表达为关于 $x^{(1)}(k-1)$ ， $x^{(1)}(k+1)$ 的关系，只是为了求参数 a ，这时需要解 3 次方程，计算量增大；

8.1.4 通过大量实验发现，该模型具有较强的抗干扰性。

8.2 模型的改进

8.2.1 熵值法的计算量较大，可以探究计算性稍弱但和实际问题联系紧密的方法；

8.2.2 改进的预测模型在现实中偏简单，针对本问题，可以用仿真有关的模型来预测；

8.2.3 考虑到除文中所述的 5 个因素，可以再找多一些方面，使得误差更小。

报名号 #1069

9. 可持续发展报告

又一次警钟！3月18日英国科学家约翰·贝丁顿警告称，气候变化和人口增长将导致食品、水和能源短缺，进而引发大规模移民、公共骚乱和国际冲突。到2030年，世界有可能面临“完美风暴”，出现大的动荡。

贝丁顿引用了美国总统奥巴马说过的一句话：尽管在你听起来不舒服的情况下，也要认真地去听取科学的意见。通过收集大量数据、建立模型，我们验证了他的观点，并且对于“完美风暴”的发生与否给出了判定。

近些年来，我们一直被一些耸人听闻的预言——人口灾难即将到来等所困扰。有些人深信不疑、终日惶恐不安，而有些人却不以为然、非常乐观。我们现在提倡科学发展观，诚然，我们应当科学的、理性的分析人口、资源等问题，客观的认识到我们的处境。

通过数学模型的求解，我们并不能完全相信这个结果。为什么呢？首先仅就数学模型而言，一模型本身就是存在误差，我们都知道好多模型一直都是在不断的改进中；二发挥基础作用的数据的真实性精确性都是值得怀疑的；三起初的假设具有主观性，缺少坚实的理论证明。其次考虑到实际情况，事情一直是在发展变化的，如果从现在开始我们就采取一定措施的话，为时未晚。

具体分析，贝丁顿所提“完美风暴”，在理论上是指由一组因素将局势急剧恶化的事件，这里理解为人口增长、气候变化进而使得食品、能源和水出现了严重短缺，社会混乱，人类的生存与发展局势全面恶化。我们验证了他的观点，但是，完美风暴是否真的会发生呢？现在我们就束手无策吗？

首先，我们必须坚定一个信念，完美风暴并不是一定会发生——人定胜天！我们如何“定”而“胜”——可持续发展。

可持续发展，即既要满足当代人的需求，又不能对后代人满足其需求的能力构成危害的发展。可持续发展的核心是发展，但要求在严格控制人口、提高人口素质和保护环境、资源永续利用的前提下进行经济和社会的发展。发展是可持续发展的前提；人是可持续发展的中心体；可持续长久的发展才是真正的发展。我们并不是束手无策，我们其实有一个非常厉害的法宝就是可持续发展。空喊口号显然是不行的，我们需要付诸实践。

江泽民同志在第四次全国环境保护会议上的讲话中指出：经济发展，必须与人口、资源、环境统筹考虑，不仅要安排好当前的发展，还要为子孙后代着想，为未来的发展创造更好的条件，决不能走浪费资源和先污染后治理的路子，更不能吃祖宗饭、断子孙路；控制人口增长，保护生态环境，是全党全国人民必须长期坚持的基本国策；环境意识和环境质量如何，是衡量一个国家和民族的文明程度的一个重要标志；各级党委和政府要把环境保护工作摆上重要议事日程，每年要听取环保工作的汇报，及时研究和解决出现的问题，这要成为一项制度。

其实就是在强调两个因素——人口和环境，这二者是直接和可持续发展挂钩的，而可持续发展又体现在粮食、能源、水资源供求上，这就正好和我们的数学模型一致了！

诚然，良好的环境是可持续发展的重要基础。保护环境是保护健康、维护生态平衡的迫切需要，同时也具有重要的经济意义。环境的保护是要靠大家的共同努力，是各行各业都必须融入其中的。小学生节约用纸、用铅笔，大学生积极宣传并参加志愿活动，上班族也时刻注意不乱扔垃圾等，其实身边有很多事情都是力所能及的！

我国是世界上人口最多的国家，但是改革开放以来我们已经取得了一定的成绩，既然有了基础与经验，我们又怎会一味的悲观呢？其实，不难发现越是边远的落后的地区，越是人口增长限制性小。所以，我们有必要加强教育，提高大家的素质。

报名号 #1069

总会有人认为自己太“渺小”，不知道或不愿意为全球的可持续发展做些事。其实，要知道，世界是由个体组成的，而且我们可以做些事。比如尽量乘坐公交车。现在，汽车废气的治理已取得相当的成功，但数量的急剧增长，使汽车仍是城市大气污染的主要来源。据报道，近年国内某些大城市也出现过光化学烟雾污染。不仅如此，制造汽车的过程中也要消耗自然资源，也要排放污染物，汽车还产生噪声等危害。我们能做的“小事”还有很多。不使用非降解塑料餐盒、不燃放烟花爆竹、双面使用纸张、节约粮食、拒绝使用一次性用品等等。只要始终不忘我们可以而且必须为我们自己更为子孙后代多做些事，“完美风暴”便会在大家的完美配合中销声匿迹！

最后，我想引用美国总统肯尼迪就职演讲中的两段话：

Ask not what your country can do for you; ask what you can do for your country.

My fellow citizens of the world: Ask not what American will do for you, but what together we can do for the freedom of man.

是的，现在形势很严峻，不是“人类的自由”，而是更基本的“人类的生存”！何以生存，唯有可持续发展！

10. 总结

比较容易看出，此题是一个预测题。用于做预测的模型有很多种，但是进一步分析，可知是一长期预测问题，这样很多短期预测效果不错、但长期不理想的模型，我们最终选择了灰色 Gompertz 预测模型。

此题一个非常大的特点是数据多，处理麻烦。这样就要求我们首先筛选出具有代表性的，看做一个样本。而经过我们的处理，发现与实际情况吻合，效果令人非常满意。进而我们将第一问比较圆满的结出来了。对于第二问，体现了数学建模的本质与意义。单纯的算题，并过分依赖其结果是不恰当的。基于第一问的结果，再加上实际情况的分析，我们就可持续发展给出了自己的观点。我们认为，应当客观理性的认识“完美风暴”问题，既不可过于悲观，也不能太乐观，我们要有信心，也要有决心与恒心来应对它。最后提出了建议，号召大家团结在一起，为全球的美好做出力所能及的事情。我想这就是数模的本质与意义了。来源于实际，经过数学上严谨的推导，最终又回归实际。

参考文献：

- [1]熊和金，徐华中《灰色控制》北京 国防工业出版社
- [2]张智丰《数学实验》北京 科学出版社 2008
- [3]同小军，王中民，周龙一种新型灰色Gompertz 模型
- [4]宋来宗，王志明，数学建模与实验，北京：科学出版社，2005

附：

1、主成份分析的代码

```
%main.  
sj1  
X=sysj;  
v1=bzh(X)  
result=xgxs(v1);
```

报名号 # 1069

```

zcfd(f(v1,result);
%bzh.m,用总和标准化对输入矩阵进行标准化
function y=bzh(X)
colsum=sum(X,1);%对列求和
[m,n]=size(X);%矩阵大小，m 为行数，n 为列数
for i=1:m
    for j=1:n
        y(i,j)=X(i,j)/colsum(j);
    end
end
end
sysj=[960.0 89    2010    131180    26681    36320000    2156    44535.5
328.7    150 820 110981    9063    8608800 1152    23913.0
190.5    104 1420    22304    3645    14521    12867    6601.1
164.8    58    3000    6915    2229    22238    1300    2281.0
37.8    40    38410    12756    43401    63118    3365    1174.1
9.9 41    17690    4842    8880    22154    1344    665.3
33.0    29    5490    2577    1489    8505    22882    223.4
30.0    39    1420    8459    1169    3537    5767    2140.9
0.1 44    29320    439 1322    2109    211 300.0
51.3    79    2990    6472    2062    11544    3269    3314.6
78.4    55    5400    7294    4027    10415    3150    3459.8
100.1    169 1350    7540    1075    8253    500 2273.6
121.9    16    5390    4739    2550    17094    955 945.3
998.5    22    36170    3256    12515    38953    88238    5089.5
195.8    51    7870    10422    8392    20427    3967    3195.9
963.2    21    44970    29899    132018    293131    9446    34656.2
278.0    58    5150    3912    2141    8562    7123    3355.6
851.5    40    4730    18869    10680    21191    29066    5901.7
91.2    10    6070    2702    1819    8691    27185    407.8
55.2    11    36550    6104    22307    24890    2932    6181.3
35.7    22    36620    8241    29067    42108    1297    4347.5
30.1    30    32020    5857    18447    25702    3114    2014.5
31.3    43    8190    3811    3387    12709    1404    2177.6
1709.8    21    5780    14237    9869    86974    30135    7686.6
50.5    30    27570    4355    12240    16850    2562    1935.3
60.4    35    1950    4657    1061    18287    1128    3369.8
24.4    21    40180    6036    23450    31646    2408    2082.0
774.1    22    35990    2052    7682    14925    24202    1633.2
27.1    14    27250    413 1039    2012    79778    80.3]

```

```

%zcfd(f.m,计算各主成分得分、计算综合得分、排序
function z1=zcfd(f(vec1,vec2);
sco=vec1*vec2;
csum=sum(sco,2);
[newcsum,i]=sort(1*csum);
[newi,j]=sort(i);
fprintf('计算得分:\n')
num=length(sco(1,:));
fprintf('各主成分得分    综合得分    排序结果\n')

```

报名号 #1069

```

z1=[sco,csum,j]; %得分矩阵: sco 为各主成分得分; csum 为综合得分; j 为排序结果
disp(z1)
%xgxs.m,计算相关系数、特征值、特征向量、对主成分排序、计算各特征值的贡献率、
挑选主成分、输出主成分个数、计算主成分载荷
function L=xgxs(X);
fprintf('相关系数矩阵:\n')
corrcoefX=corrcoef(X) %计算相关系数矩阵
fprintf('特征向量(eig_vec)及特征值(eig_val):\n')
[eig_vec,eig_val]=eig(corrcoefX) %求特征值(eig_val)及特征向量(eig_vec)
eval=diag(eig_val);
[y,i]=sort(eval); %对特征值进行排序, y 为排序结果, i 为索引
fprintf('特征根排序:\n')
for z=1:length(y)
    newy(z)=y(length(y)+1-z);
end
fprintf('%g\n',newy)
fprintf('\n 贡献率:\n')
newrate=newy/sum(newy)
sumrate=0;
newi=[];
for k=1:length(y)
    sumrate=sumrate+newrate(k);
    if sumrate>0.85 break;
end
end %累计贡献率大于 85%的特征值的序号就是 k
fprintf('主成分数:%g\n',k);
fprintf('主成分载荷:\n')
for p=1:k
    for q=1:length(y)
        L(q,p)=sqrt(eval(i(length(y)+1-p)))*eig_vec(q,i(length(y)+1-p));
    end
end %计算载荷
disp(-L)

```

2、聚类分析法代码

% 主程序, 输入数据, 输出数据并画图

```

X=[    1.0400
      0.5926
      0.1881
      0.0868
     -0.1162
     -0.0159
      0.0231
      0.0627
     -0.0518
      0.0998
      0.0687

```

报名号 # 1069

```

0.2041
0.0209
0.0282
0.0679
-0.0412
0.0901
0.1630
0.0086
-0.1039
-0.1068
-0.0804
0.0313
0.2293
-0.0576
0.0560
-0.1245
-0.0128
-0.0604 ]
fprintf('标准化结果如下:\n')
v1=bzh(X);
v2=absd(v1);
Z=zdljl(v2);
[H,T]=dendrogram(Z)
% absd.m,用绝对值法求距离
function d=absd(vector);
[a,b]=size(vector);
d=zeros(a);
for i=1:a
    for j=1:a
        for k=1:b
            d(i,j)=d(i,j)+abs(vector(i,k)-vector(j,k));
        end
    end
end
end
fprintf('绝对值距离矩阵如下: \n');
disp(d)
% zdlzl.m,最短距离聚类法
function y=zdljl(Matrix);
[a,b]=size(Matrix);
max_Matrix=max(max(Matrix));
for i=1:a
    for j=i:b
        Matrix(i,j)=max_Matrix;
    end
end
disp(Matrix)
Z=zeros(b-1,3);
for k=1:(b-1)
    [c,d]=size(Matrix);

```

报名号 # 1069

```

fprintf('第%g 次聚类:\n',k);
[e,f,g]=min1(Matrix);
fprintf('最小值=%g, 将第%g 区和第%g 区并为一类, 记作 C%g\n',g,e,f,c+1);
Z(k,1)=e;
Z(k,2)=f;
Z(k,3)=g;
for l=1:c
    if l<=min2(e,f)
        Matrix(c+1,l)=min2(Matrix(e,l),Matrix(f,l));
    else
        Matrix(c+1,l)=min2(Matrix(l,e),Matrix(l,f));
    end
end
Matrix(1:c+1,c+1)=max_Matrix;
Matrix(1:c+1,e)=max_Matrix;
Matrix(1:c+1,f)=max_Matrix;
Matrix(e,1:c+1)=max_Matrix;
Matrix(f,1:c+1)=max_Matrix;
Matrix
end
y=Z
p=[960.0 89 2010 131180 26681 36320000 2156 44535.5
328.7 150 820 110981 9063 8608800 1152 23913.0
1709.8 21 5780 14237 9869 86974 30135 7686.6
851.5 40 4730 18869 10680 21191 29066 5901.7
78.4 55 5400 7294 4027 10415 3150 3459.8
121.9 16 5390 4739 2550 17094 955 945.3
963.2 21 44970 29899 132018 293131 9446 34656.2
37.8 40 38410 12756 43401 63118 3365 1174.1];
p=bzh(p);
3、熵值法代码:

```

```

m=[p(1,1) p(2,1) p(3,1) p(4,1) p(5,1) p(6,1) p(7,1) p(8,1)];%每一行的最大值
for i=1:8
    for j=1:8
        if m(i)<p(i,j)
            m(i)=p(i,j);
        end
    end
end
n=[p(1,1) p(2,1) p(3,1) p(4,1) p(5,1) p(6,1) p(7,1) p(8,1)];%每一行的最小值
for i=1:8
    for j=1:8
        if n(i)>p(i,j)
            n(i)=p(i,j);
        end
    end
end
end
for i=1:8

```

报名号 # 1069

```

        for j=1:8
            p(i,j)=(p(i,j)-n(i))/(m(i)-n(i));
        end
    end
    su=[0 0 0 0 0 0 0 0];
    for i=1:8
        for j=1:8
            su(i)=su(i)+p(i,j);
        end
    end
    for i=1:8
        x=su(i);
        for j=1:8
            p(i,j)=p(i,j)/x;
        end
    end
    e=[0 0 0 0 0 0 0 0];
    for j=1:8
        for i=1:8
            if p(i,j)~=0
                e(j)=e(j)-p(i,j)*log(p(i,j))/log(8);
            end
        end
    end
    y=0;
    for i=1:8
        y=y+e(i);
    end
    g=[0 0 0 0 0 0 0 0];
    for i=1:8
        g(i)=(1-e(i))/(8-y);
    end
    z=0;
    for i=1:8
        z=z+g(i);
    end
    w=[0 0 0 0 0 0 0 0];
    for i=1:8
        w(i)=g(i)/z;
    end
    s=[0 0 0 0 0 0 0 0];
    for i=1:8
        for j=1:8
            s(i)=s(i)+w(i)*p(i,j);
        end
    end
    s

```

4、灰色 Gompertz 模型代码：

报名号 # 1069

```

%灰色 Compertz 模型
function compertz
%format long;
close all;
clear;
y0=[14.7736    14.8135    14.9547    14.9667    15.0798    15.1003    15.1593];
%y0=[7.9991     8.1008     8.2765     8.6347     8.9209     9.2584     9.7675];
%y0=[5.6467     5.6493     5.7084     5.8830     6.1714     6.5049     6.7499];
%y0=[0.9015     0.9567     0.9694     0.9814     0.9866     1.0088     1.0358]
%y0=[ 25.19470000 30.00000000 36.96588000 44.42295000 52.79519000 60.85572000
68.000000000];
%y0=[97542    98705    100072    101654    103008    104357    105851    107507    109300
111026    112704    114333    115823    117171    118517    119850    121121    122389    123626
124761    125786    126743    127627    128453    129227    129988    130756    131448    132129];
%y0=[886 1087 1396 1862 2471 3042 4001 5482];
%y0=[231810 232107 235449 242362 246615 257588 264829 273711 279202 287503
292839 305237 317932 328545 339847 348972 353676 364583 372786 380764 388417
396651 403116 407792 412115 415181 416238 419129 423890 428243 443402 448052
450995 454506 458915 465920 470047 474289 479428 479409 484316];
%y0=[235.2 249.9 263.2 293.6 318.9 356.7 363.8 424.2 486.5 586.8 768 1019];
x0(1)=0;
for i=1:length(y0)
    %x0(i)=1/y0(i);
    x0(i)=log(y0(i));
end
n=length(y0);
x1(1)=x0(1);
for i=2:n
    x1(i)=0;
    for j=1:i
        x1(i)=x1(i)+x0(j);%一次累加序列
    end
end
end
A(1,1)=0;
B(1,1)=0;
C(1,1)=0;
for i=1:n-2
    A(i,1)=x1(i+1)+x1(i+2);
    %A(i,1)=x1(i+2);
    A(i,2)=i+1;
    %A(i,3)=(1+(-1)^(i-1))/2;
    A(i,3)=1;
    B(i,1)=x0(i+1);
    C(i,1)=x1(i+1);
end
D=B-2*C;
p=inv(A'*A)*A'*D;
p2=p(1);p3=p(2);p4=p(3);
a=log(-p2);
xx(1)=0;

```

报名号 # 1069

```
for k=1:1:n
    xx(k)=exp(-a*k);
end
fl=polyfit(xx,x0,1);
b=f1(1);c=f1(2);
y1=polyval(p,xx);
%yj=1./y1;
t=1:7;
t1=1:12;
yj=exp(c+b*exp(-a.*t1));
[y0'*1e-4;zeros(7,1)]
yj'*1e-4
ep=0;
for i=1:n
    ep=ep+(y0(i)*1e-4-yj(i)*1e-4)^2;
    ex(i)=(yj(i)-y0(i))/y0(i);
end
ep
ex'*100
plot(t,y0*1e-4,'*r',t1,yj*1e-4,'ob');
legend('原始数据','预测数据')
title('灰色 Compertz 模型预测结果')
xlabel('时间')
ylabel('人口（亿人）')
figure
aa=1:7;
bar(aa,ex)
title('相对误差条状图')
xlabel('时间')
ylabel('相对误差')
%function y=f3(x)
```