

生态保护建设与评价其对环境的影响

摘要

中国坚持绿色协调发展，节约资源，保护环境，实施可持续发展战略，本文利用收集到的数据，进行标准化处理后，建立出森林生态环境评分模型，沙尘暴程度评分模型，以及基于 BRF 神经网络的二氧化碳吸收量预测模型，并给出建设生态保护区的计划和建设。

本文主要工作如下：

问题 1: 我们通过指标的选取，最终得到平均降水量、覆盖面积、林木储蓄、蓄水量、氧气释放量、二氧化碳吸收量、大风天数 7 个指标。对这 7 个指标，每隔 3 年取一次平均值，共得到 13 组数据（1962-2000），通过因子分析降维处理各指标，得到生态环境综合评价模型（1-1），计算各时间段的生态环境情况得分，得到：在修复塞罕坝的第一个三年间（1962~1964），其生态环境情况得分为-4.106，在修复塞罕坝后第 13 个三年间（1998~2000）得分为 9.70，在修复前后生态环境改善幅度为 13.30。

问题 2: 使用问题一的指标选择方法，我们得到评价沙尘暴的 4 个指标：沙尘暴日数，沙尘暴发生频率，相对温度，相对湿度，以及 13 个时间段的数据（1962-2000），在此基础上建立基于主成分分析的综合评价模型，得到沙尘暴程度评价方程（2-1），由此计算得出：在修复塞罕坝的第一个三年间（1962~1964），北京市沙尘暴程度评分为 1.418，在修复塞罕坝后的第 11 个三年间（1992~1994），评分为-0.69，沙尘暴程度改善了 2.10 个指标。

问题 3: 我们通过中国地图，选择出 14 个能够建设生态保护区的地区，收集该地区 7 个评价指标（问题一所得）的数据，进行 Q 型聚类分析，发现南京、长春两个地区的生态环境与其他地区差异最大，我们套用公式（1-1），解得长春的生态环境得分-2.956，南京的生态环境指标为 0.7426。通过观察收集的数据、结合当地可植树面积，我们得出：应在南京建立 1351223.4 亩的森林生态系统，其中主要种植杉树来提高降水量；应该在长春建立 1789419.6 亩的森林生态系统，其中各种类型的树都应该兼顾。通过在这些地方建立生态保护区，能够吸收大量二氧化碳，中国能够很快实现碳中和目标。

问题 4: 我们通过比对亚太地区个国家的环境状况，最终选择在越南的红河河谷上游建立生态保护区，通过将收集到的各项指标带入公式（1-1），得到该地区的生态环境评分为 0.344，通过分析其地理环境，我们决定在此建立覆盖面积为 851223.4 亩，林木储蓄为 11 万立方米的生态保护区。我们以覆盖面积、林木储蓄作为自变量，二氧化碳的吸收量做为因变量，用第一问中收集到的数据作为训练样本，进行 RBF 神经网络预测，得到建立该生态保护区可以吸收温室气体数量为：1.4872 万立方吨。

问题五：我们通过得到的模型制定了建立生态保护区的可行计划，并撰写了非技术性报告（4.5）

关键词: 因子分析、Q 型聚类分析、基于主成分分析的综合评价模型，RBF 神经网络预测

目录

一、	问题重述	4
1.1	问题背景	4
1.2	问题重述	4
1.3	对问题 1 的分析	5
1.4	对问题 2 的分析	5
1.5	对问题 3 的分析	6
1.6	对问题 4 的分析	6
1.7	对问题 5 的分析	6
二、	模型假设	6
三、	符号说明	7
四、	模型的建立和求解	7

4.1 问题 1 模型的建立与求解	7
4.1.1 问题 1 模型指标的选取	7
4.1.2 建立多属性决策模型	8
4.1.2.1 处理数据, 减少突发情况的影响	8
4.1.3 因子分析法求解模型	9
4.1.3.1 标准化所选数据	9
4.1.3.2 求解各系数评定指标的相关系数矩阵	10
4.1.3.3 KMO 和 Bartlett 球形度检验	10
4.1.3.4 主成分法提取初等公因子	11
4.1.3.5 选择 p ($p \leq 7$) 个公因子	11
4.1.3.6 方差最大法旋转因子热力载荷矩阵	12
4.1.3.7 求解生态环境评价指标得分	14
4.1.4 分析修复塞罕坝前后森林生态环境情况得分	14
4.2 问题 2 模型的建立与求解	15
4.2.1 问题 2 模型指标的选取	15
4.2.2 建立基于主成分分析的综合评价模型	16
4.2.2.1 数据准备	16
4.2.2.2 KMO 和 Bartlett 球形度检验	16
4.2.2.3 主成分的选取	17
4.2.2.4 基于主成分综合评价模型的构成	19
4.3 问题 3 模型的建立与求解	20
4.3.1 数据的处理	20
4.3.2 聚类分析模型的建立	21
4.3.3 选择需要改善的地区, 并推荐规模	22
4.4 问题 4 模型的建立与求解	23
4.4.1 选取适宜地区	23
4.4.2 规划生态保护区面积	24
4.4.3 RBF 神经网络求解模型	24
4.4.3.1 RBF 网络结构	24
4.4.3.2 网络学习样本的建立	25
4.4.3.3 网络的训练	25
4.5 问题 5 的求解	25
五、模型的分析与检验	26
5.1 模型分析	26
5.2 模型检验	27
六、模型的评价与推广	27
6.1 模型的优缺点	27
6.2 模型的推广	28
七、参考文献	28
[1] 中国地图矢量图 http://ditu.yjdy.org/read/MDIOMWNjOWLkuK3lm73lnL=/ ..	28
[2] 司守奎 数学建模算法与应用 国防工业出版社 2011 年 8 月版	28

附录： 28

一、 问题重述

1.1 问题背景

在河北省，55 年来，塞罕坝林场的建设者们听从党的召唤，在充满黄沙的荒漠沙地上艰苦奋斗、甘于奉献，创造了荒原变林海的人间奇迹，用实际行动诠释了绿水青山就是金山银山的理念，铸就了牢记使命、艰苦创业、绿色发展的塞罕坝精神。”习近平总书记也高度肯定“塞罕坝人”长期以来的坚守和奉献。塞罕坝精神的背后，是时代的风云变迁和一代代人的艰苦奋斗。

而塞罕坝能够成功的原因，是建立生态保护区时合适的指标选择和合理的因子分析。而采用数学建模算法，探索在建立生态保护区时如何选择恰当的影响因子和评定指标，与生态保护区能否有效果息息相关，建立有效的数学模型可以准确的评定环境指标并针对指标提供有效的解决方法，有很强的现实意义，同时可以积极响应保护环境的时代声音。

1.2 问题重述

Adhere to the concept that lucid waters and lush mountains are invaluable assets, China

insists

on respecting for, being harmony with and protecting nature, giving high priority to conserving resources, protecting the environment and letting nature restore itself, implementing sustainable development strategies, improving the overall coordination mechanism in the field of ecological civilization, building an ecological civilization system, promoting the transformation of economic and social development toward comprehensive green growth and building a beautiful country. With the help of the Chinese government, China's Saihanba Tree Farm has recovered from the desert and has now become an eco-friendly and green farm with the stable sand prevention function.

Since 1962, 369 young people with an average age under 24 have come to this wasteland filled with yellow sand. From then on, they dedicated their lives here, advanced wave upon wave, to planting seeds in the sand and planting green in the crevices of the stones, like nails fastening millions of acres of forest on the wasteland. Planting trees to fix the sand and conserve water sources, they build a green barrier to block wind and sand. Today, the forest coverage in Saihanba area has reached 80%. It supplies Beijing and Tianjin with 137 million cubic meters of clean water each year, sequesters 747,000 tons of carbon, and releases 545,000 tons of oxygen.

With over a half century's struggle, the world's largest artificial forest was built on the earth of Saihanba. Expanding afforestation of 1.12 million mu with more than 400 million trees, the builders created a green sea on the plateau wasteland which is 400 kilometers north of Beijing. On the one hand, there is the historical mission that "civilization will develop followed with zoology thriving". On the other hand, there are new issues encountered on the road of green development. Therefore, Saihanba people now have a higher goal, which is to restore the ecology. Since the 18th National Congress of the Communist Party of China, they have successively launched three major projects, namely, afforestation, natural improvement of artificial forests, and near-naturalization cultivation of natural forests. They have tried to make artificial forests closer to natural ones.

1.3 对问题 1 的分析

问题 1 要求我们选择合适的指标，定量分析出塞罕坝修复后对环境的影响，我们可以查阅资料，得到森林评价指标（平均降水量、覆盖面积、林木储蓄、蓄水量、氧气释放量、二氧化碳吸收量、大风天数），然后对所收集到的数据，每隔 3 年取一次平均值，减少不确定因素的影响，然后用因子分析进行降维处理，建立出塞罕坝对环境影响的模型。然后通过筛选后的指标权重得分来定量计算出对生态环境的情况，并将塞罕坝修复前后生态环境的情况进行对比，得出修复塞罕坝能够改善生态环境情况的具体大小。

1.4 对问题 2 的分析

问题 2 要求我们建立数学模型，选择合适的指标，定量评价北京抗沙尘暴能力。与问题一类似，我们可以查阅资料，得到沙尘暴能力指标（沙尘暴日数，沙尘暴发生频率，相对温度，相对湿度），然后对所收集到的数据，每隔 3 年取一次平均值，减少不确定因

素的影响,通过基于主成分分析的综合评价模型进行降维处理,建立出塞罕坝对北京沙尘暴环境影响的模型,然后通过模型解出的权重得分定量比较塞罕坝修复前后对环境的影响的大小。并进行解释。

1.5 对问题 3 的分析

第三问要求我们确定中国哪些地理位置需要建立生态区域,以及建立的规模。我们可以收集全国的各地区的平均降水量、覆盖面积、林木储蓄、蓄水量、氧气释放量、二氧化碳吸收量、大风天数。通过这 7 个指标对他们进行 Q 型聚类分析,将生态环境指标有明显差异的城市进行分类,然后使用第一问的模型进行打分可以选出生态环境状况不好的地区,再对比收集的数据以及该地区可供植树的面积,规划生态区域的规模。

1.6 对问题 4 的分析

问题 4 要求我们选择合适的目标,定量分析其特征并合理借鉴中国塞罕坝的生态保护模式,讨论该国需要建设哪个地理位置。并确定待建生态区域的数量或规模,并评估其对吸收温室气体和减少碳排放的影响。考虑到越南作为“一带一路”沿线的重要国家,且越南和中国同为正在大力发展工业化的发展中国家,且两国政治和经济结构都十分相似,因此我们组选择了越南这个国家作为第四问分析的国家。越南首都河内作为全国的政治、经济中心,也是越南国内最大的工业中心,其经济约占全国 GDP 的 12.1%,年碳排放约为 952.67 万吨,因此在河内附近规划生态保护区,为越南缓解首都地区的污染和碳排放就尤为重要,通过合理的评估分析,我们团队最终选择了沿红河上游的安沛市和寨忽市附近红河河谷建立生态保护区,并且依照其离河内的距离(寨忽 145 公里,安沛 121 公里),根据第一问的模型,我们可以确定出红河河谷的森林生态系统评分。并结合地型,可以得到需要建立的生态系统规模,用第一问中收集到的数据作为训练样本,进行 RBF 神经网络预测,输入可人为控制的生态环境指标(覆盖面积、林木储蓄)后,可以得到吸收温室气体数量。

1.7 对问题 5 的分析

请向亚太数学竞赛建模组委会(APX4CM)撰写非技术性报告。描述您的模型,并提出建设生态保护区的可行计划和建议。

二、 模型假设

根据生态环境的特点与本题所给的背景,对于本题,我们做出以下假设:

- 1) 所收集的数据真实有效

- 2) 假设在不考虑其他因素的情况下，该生态环境所得的总分数能够反映出当前生态环境的情况；
- 3) 在打分过程中，各因子的贡献率能够代表其在实际生活所占权重。
- 4) 假设对于生态环境的评价指标，能够准确地代表生态环境的好坏；

三、 符号说明

符号	说明
X_1	平均降雨量
X_2	覆盖面积
X_3	林木储蓄
X_4	蓄水量
X_5	氧气释放量
X_6	二氧化碳吸收量
X_7	大风天数
Y_1	沙尘暴日数
Y_2	沙尘暴发生频率
Y_3	相对温度
Y_4	相对湿度

四、 模型的建立和求解

4.1 问题 1 模型的建立与求解

4.1.1 问题 1 模型指标的选取

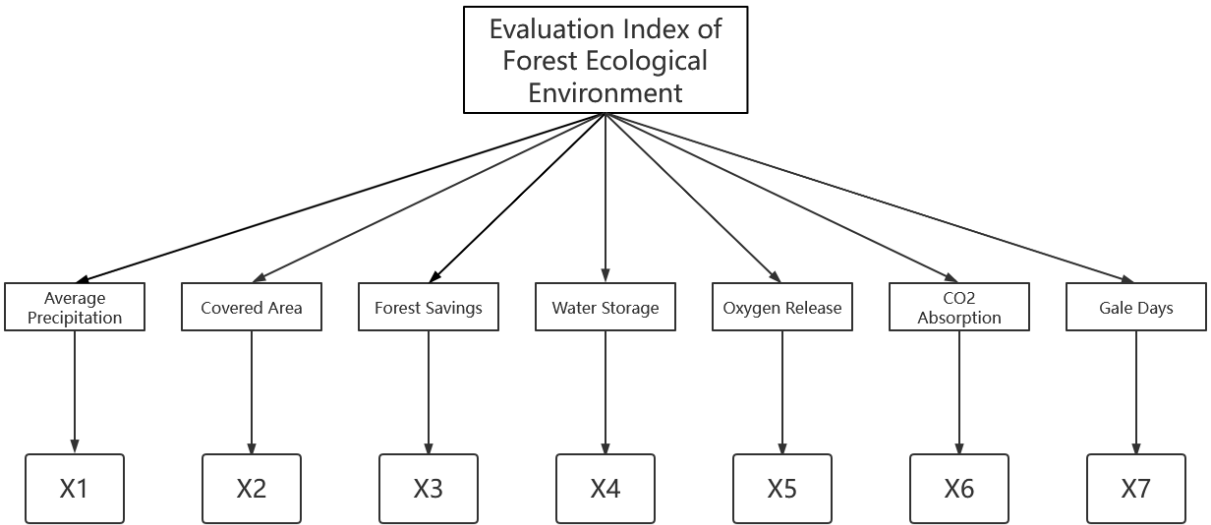
森林生态环境受多个因素的影响，我们不可能收集到所有因素在修复前后的具体数据，因此我们考虑一下几个方面：

(1) 通过咨询专家或翻阅相关文献，得到影响森林生态环境的主要因素
通过查阅资料，我们得到森林生态环境评价指标有：林地利用率，森林覆盖率，林分单位面积蓄积量,林分单位面积生长量,毛竹单位面积立竹量和经济林单位面积产量等木材以外其他林产品产量，生产木质原材料的工业人工林以单位面积蓄积量和每立方米木材容量表示生产量，树种组成，森林蓄积量按粗度级的分配比，木材以外其他林产品质量等级，经济材出材率，造纸材木材干重的得浆率和树皮率等，造林成活率，保存率，用材林林龄组成，单位面积总产值和纯收入，单位面积年产值和年纯收入，对森林开发利用的投资回收期，森林面积平均净增率,森林蓄积平均净增率,毛竹林立株净增率，平均降水量、覆盖面积、林木储蓄、蓄水量、氧气释放量、二氧化碳吸收量、大风天数。

(2) 尽可能减少不稳定因素的影响

问题一要求我们检验塞罕坝修复前后对环境的影响，主要是对比修复塞罕坝的对生态环境情况。因此有些不易统计，或者容易随着时间推移发生大幅度改变的因素，我们选择筛除。

综上 我们选取平均降水量、覆盖面积、林木储蓄、蓄水量、氧气释放量、二氧化碳吸收量、大风天数做为森林生态环境评价指标。



森林生态环境评价指标图 4-1

4.1.2 建立多属性决策模型

4.1.2.1 处理数据，减少突发情况的影响

由于修复的作用效果不会立即体现，所以我们可以不用收集修复前的数据，对比随着修复时间的推移，生态环境情况的变化也能说明修复塞罕坝前后，对当地生态环境的影响。同时我们选择每隔3年，将这段时间内的数据去掉最小值和最大值之后取平均值。经处理后每个所选择指标具有13组数据。

如下图：

	x1	x2	x3	x4	x5	x6	x7
1962~1964	440.5	22.29	33.37	0.09	2.77	1.93	-77
1965~1967	440.5	31.75	36.70	0.10	3.04	2.12	-77
1966~1968	440.5	40.45	43.86	0.12	3.64	2.53	-77
1969~1972	413.6	48.41	54.70	0.15	4.54	3.16	-76
1973~1975	413.6	55.64	69.05	0.18	5.73	3.99	-76
1977~1979	413.6	62.16	86.72	0.22	7.2	5.01	-76
1980~1982	440.8	67.99	107.55	0.25	8.92	6.21	-49
1983~1985	440.8	72.94	139.62	0.29	11.59	8.06	-49
1986~1988	440.8	77.06	210.83	0.33	17.49	12.17	-49
1989~1991	440.8	81.35	259.13	0.37	21.5	14.96	-28.3
1992~1994	513	86.36	320.46	0.42	26.59	18.5	-28.3
1995~1997	513	91.51	398.72	0.48	33.08	23.01	-28.3
1998~2000	513	95.97	502.89	0.56	41.73	29.02	-28.3

13 组数据表 4-1

4.1.3 因子分析法求解模型

4.1.3.1 标准化所选数据

由于不同指标的量纲不同，我们为了统一量纲，进行数据的标准化，步骤如下：

(1)：先计算出各指标变量的算术平均值：

$$\overline{X_j} = \frac{\sum_{i=1}^{13} X_{ij}}{4} \quad (4-1)$$

$$\text{标准差: } S_j = \sqrt{\frac{\sum_{i=1}^{13} (x_{ij} - \overline{X_j})^2}{4-1}} \quad (4-2)$$

(2)：将指标进行标准化处理：

$$Z_{ij} = \frac{X_{ij} - \overline{X_j}}{S_j} \quad (4-3)$$

Z_{ij} 为第 i 个时间段的第 j 个指标经标准化后的值

4.1.3.2 求解各系数评定指标的相关系数矩阵

利用标准化处理后得到的标准化指标:平均降水量 Z_1 、覆盖面积 Z_2 、林木储蓄 Z_3 、蓄水量 Z_4 、氧气释放量 Z_5 、二氧化碳吸收量 Z_6 、大风天数 Z_7 。求解各评定指标的相关系数:

$$Y_{ij} = \frac{\sum_{k=1}^{13} Z_{ki} * Z_{kj}}{4-1} \quad i, j = 1, 2, 3, 4, 5, 6, 7 \quad (4-4)$$

式中 $Y_{ii} = 1, Y_{ij} = Y_{ji}, Y_{ij}$ 为第 i 个指标与第 j 个指标的相关关系, 的到相关系数矩阵

$$R = (Y_{ij})_{7 \times 7}$$

	x1	x2	x3	x4	x5	x6	x7
1962~1964	-0.28649	-1.79882	-0.91696	-1.21731	-0.91686	-0.91689	-0.99172
1965~1967	-0.28649	-1.39225	-0.89526	-1.1511	-0.89567	-0.89544	-0.99172
1966~1968	-0.28649	-1.01834	-0.84862	-1.01867	-0.84856	-0.84916	-0.99172
1969~1972	-1.01248	-0.67624	-0.778	-0.82003	-0.77789	-0.77803	-0.94597
1973~1975	-1.01248	-0.36551	-0.68451	-0.62139	-0.68446	-0.68433	-0.94597
1977~1979	-1.01248	-0.08529	-0.5694	-0.35653	-0.56905	-0.56917	-0.94597
1980~1982	-0.2784	0.165266	-0.4337	-0.15789	-0.434	-0.4337	0.289281
1983~1985	-0.2784	0.378006	-0.22478	0.10696	-0.22437	-0.22484	0.289281
1986~1988	-0.2784	0.555074	0.239132	0.371815	0.238864	0.239167	0.289281
1989~1991	-0.2784	0.739448	0.55379	0.636669	0.553706	0.554148	1.236308
1992~1994	1.670169	0.954767	0.953332	0.967737	0.953344	0.953802	1.236308
1995~1997	1.670169	1.176102	1.463168	1.365019	1.462901	1.462965	1.236308
1998~2000	1.670169	1.367783	2.141798	1.894727	2.142049	2.141473	1.236308

相关系数矩阵表 4-2

4.1.3.3 KMO 和 Bartlett 球形度检验

接下来对各指标变量进行 KMO 和 Bartlett 球形度检验, 判断变量是否适合进行因子分析。其中 KMO 检验是对原始变量之间的相关系数大小进行检验; Bartlett 球形度检验 是检验各个变量之间的相关性程度, 指标之间的相关性越强, 说明越适合做因子分析。

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.832
Bartlett's Test of Sphericity	Approx. Chi-Square	397.964
	df	21
	Sig.	.000

KMO 和 Bartlett 球形度检验图 4-2

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

KMO 检验的结果显示, KMO 的值为 0.835, 同时, Bartlett 球形检验的结果显示, 显著性 P 值为 $0.000 < 0.05$, 水平上呈现显著性, 拒绝原假设, 各变量间具有相关性, 因子分析有效, 程度为适合。

4.1.3.4.主成分法提取初等公因子

用主成分分析法, 计算初等因子载荷矩阵 A,

$$A = [\sqrt{\lambda_1} k_1, \sqrt{\lambda_2} k_2, \dots, \sqrt{\lambda_7} k_7] \quad (4-5)$$

式中 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_7 \geq 0$ 是相关系数矩阵的特征值

因子载荷矩阵 $[k_1, k_2, \dots, k_6, k_7]$ 为其对应的特征向量

得到 7 个初等因子 ($y_1, y_2, y_3, \dots, y_7$)

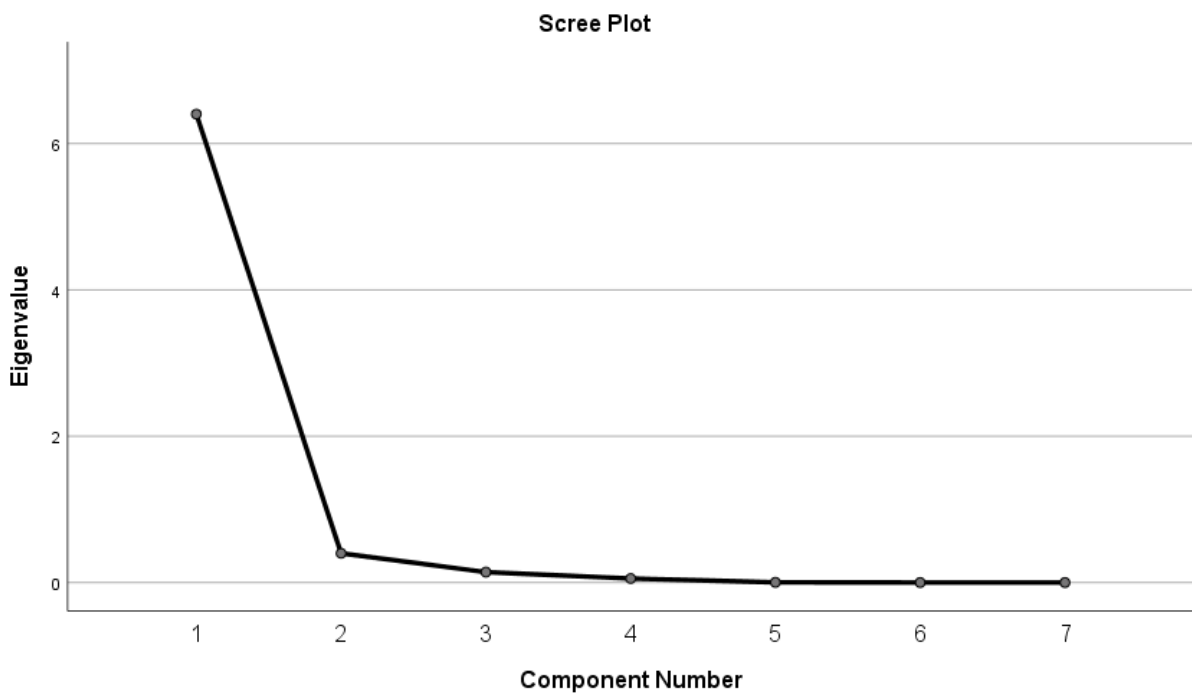
4.1.3.5 选择 p ($p \leq 7$) 个公因子

(1) 计算特征值 λ_j ($j = 1, 2, \dots, 7$) 的信息贡献率和累计贡献率

称为 $k_j = \frac{\lambda_j}{\sum_{k=1}^7 \lambda_k}$ $j = 1, 2, \dots, 7$ 为 y_i 的信息贡献率, 同时有

$$\alpha_p = \frac{\sum_{k=1}^p \lambda_k}{\sum_{k=1}^7 \lambda_k} \quad \text{为主成分 } y_1, \dots, y_p \text{ 的累计贡献率}$$

根据碎石图和总方差解释检验观察特征值的变化, 确定选取的初等公因子数:



碎石图 4-3

Total Variance Explained						
Component	Total	Initial Eigenvalues		Extraction Sums of Squared Loadings		
		% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	6.402	91.460	91.460	6.402	91.460	91.460
2	.399	5.700	97.160			
3	.142	2.033	99.192			
4	.055	.781	99.973			
5	.002	.027	100.000			
6	4.474E-8	6.391E-7	100.000			
7	1.632E-8	2.331E-7	100.000			

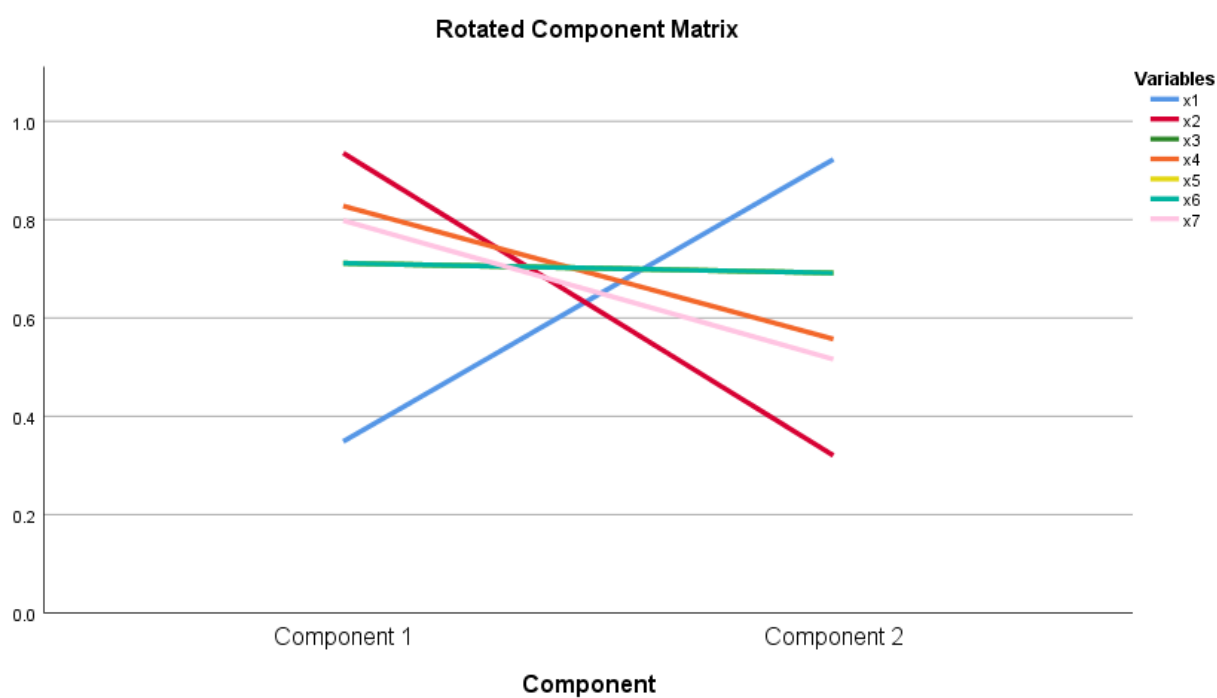
Extraction Method: Principal Component Analysis.

方差解释表 4-3

方差解释表中，在主成分 2 时，总方差解释的特征根低于 1.0，变量解释的贡献率达到 97.16%，即降维后得到的两个公因子概况了 95%以上的原始指标信息，具有较强代表性和有效性，符合公因子提取条件。结合碎石图分析结果确定最终选取的公因子为 X1 和 X2

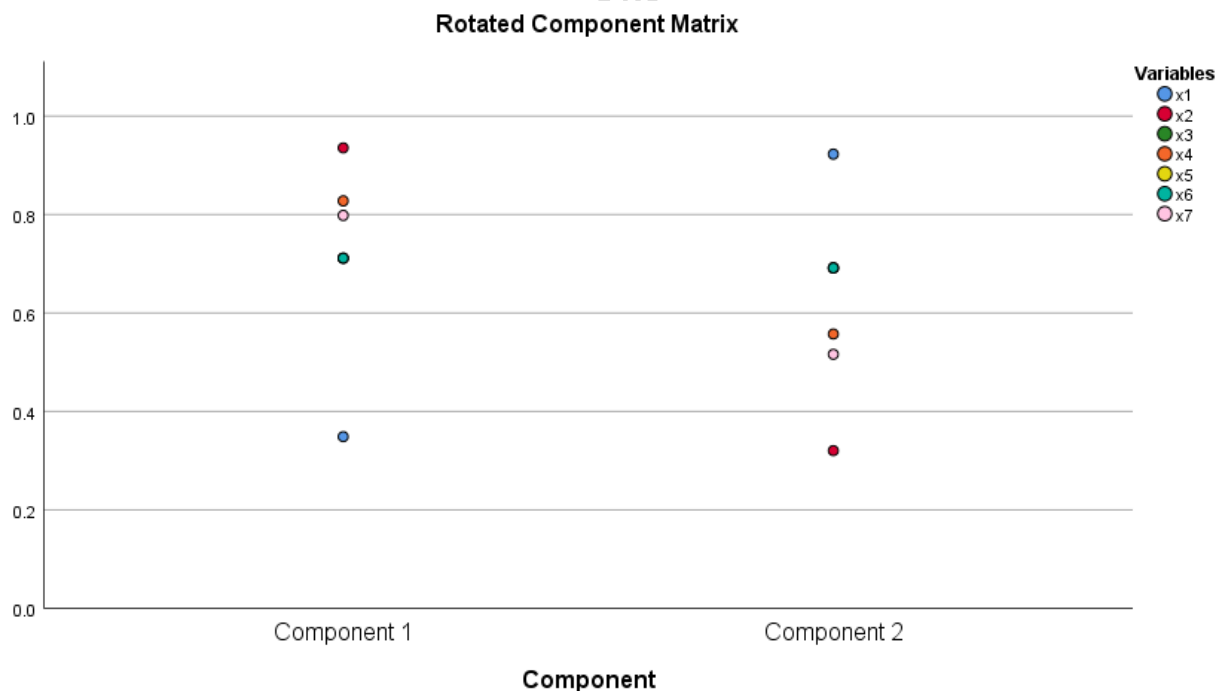
4.1.3.6 方差最大法旋转因子热力载荷矩阵

得到初等因子模型后，由于其中的公因子不一定可以反映出问题的实质特征，需要通过因子旋转使每个公因子上的载荷分配更清晰，从而减少解释公因子实际意义的主观性。利用最大方差法旋转因子，得到旋转后的因子矩阵：



因子矩阵图 4-4

使用因子载荷象限分析



因子载荷象限分析图 4-5

参照旋转后的因子载荷矩阵与因子旋转图，对降维得到的两个公因子进行定义和解释：第一公因子与覆盖面积 X2 蓄水量 X4 大风天数 X7 的相关性高，归纳为评价得分因子 F1 ；

第二公因子与平均降水量 X1 蓄水量 X4 氧气释放量 X5 的相关性高，归纳为环境改善因子 F2 ；

4.1.3.7 求解生态环境评价指标得分

我们将评价得分因子 F1、境改善因子 F2 表示为原始指标变量的线性组合，以旋转后每个公因子的方差累积贡献率作为权重系数得到各因子的得分函数。结果如下图所示：

Component matrix table			
Name	Component		
	Component1	Component2	
X1	0.054	2.314	
X2	0.146	0.807	
X3	0.111	1.737	
X4	0.129	1.4	
X5	0.111	1.736	
X6	0.111	1.736	
X7	0.125	1.297	

各因子的得分函数表 4-4

将每个公因子得分进行加权求和即可得到因子综合评价得分，即各帮扶单位的绩效评价指数为：

$$F = (0.546/0.972) \times F1 + (0.425/0.972) \times F2 \quad (1-1)$$

4.1.4 分析修复塞罕坝前后森林生态环境情况得分

将 13 个时间段的各指标数据带入公式(1-1)中，得到各时间段森林生态环境情况得分如下表：

index	Compre	rank	X1	X2	X3	X4	X5	X6	X7
13	9.708347201498052	1	1.67016886917199	1.36778292626162	2.14179839580597	1.89472728441235	2.14204928207499	2.14147264003934	1.23630822303756
12	7.587353280692391	2	1.67016886917199	1.17610223096602	1.46316826841051	1.3650185812433	1.46290122396652	1.46296476313325	1.23630822303756
11	5.96173039669015	3	1.67016886917199	0.954766898954286	0.953332482963935	0.96773705386652	0.953343894819232	0.953801614240321	1.23630822303756
10	2.622697157162172	4	-0.278396078751975	0.739448449754519	0.553789565490919	0.636889114385889	0.55370648259124	0.554148056129244	1.23630822303756
9	0.9803255312409576	5	-0.278396078751975	0.555073969302024	0.239132417048939	0.371814762801347	0.238864434843843	0.239166862024751	0.28928134339222
8	-0.4222607552958276	6	-0.278396078751975	0.378005703692635	-0.224775151078461	0.100900411216826	-0.224369500744596	-0.22483769273133	0.28928134339222
7	-1.212353381390273	7	-0.278396078751975	0.165265918555141	-0.433699680075031	-0.157893940367695	-0.434002485154957	-0.433696190614237	0.28928134339222
6	-3.3450595921861113	8	-1.01248148849618	-0.0852942728290191	-0.569399646991107	-0.356534704056086	-0.569046954512943	-0.569171973024772	-0.945971108319087
5	-3.9328576406681135	9	-1.01248148849618	-0.365509100929314	-0.684513349122087	-0.621389055640608	-0.684462867278198	-0.684326388073726	-0.945971108319087
3	-4.1667101832404105	10	-0.286492609006507	-1.01834088602801	-0.848617149728649	-1.01867058301739	-0.848557600303425	-0.84915525667321	-0.99172119912321

森林生态环境情况得分表 4-5

由此我们可以得出，在修复塞罕坝后的第一个三年间（1962~1964），森林生态环境情况评分为-4.106，即不符合森林生态环境标准

在修复塞罕坝后的第 8 个三年间（1983~1985）生态环境评分为 0.9，开始符合森林生态环境标准

在修复塞罕坝后的第 13 个三年间（1998~2000）生态环境评分为 9.70，森林生态环境良好，改善幅度为 13.30。

4.2 问题 2 模型的建立与求解

4.2.1 问题 2 模型指标的选取

沙尘暴的防护能力无法通过直接的数据收集获得，因此我们考虑收集北京市每一年的沙尘暴程度指标，而沙尘暴程度的影响因素非常之多，我们不可能收集到所有因素在修复前后的具体数据，因此我们考虑以下几个方面：

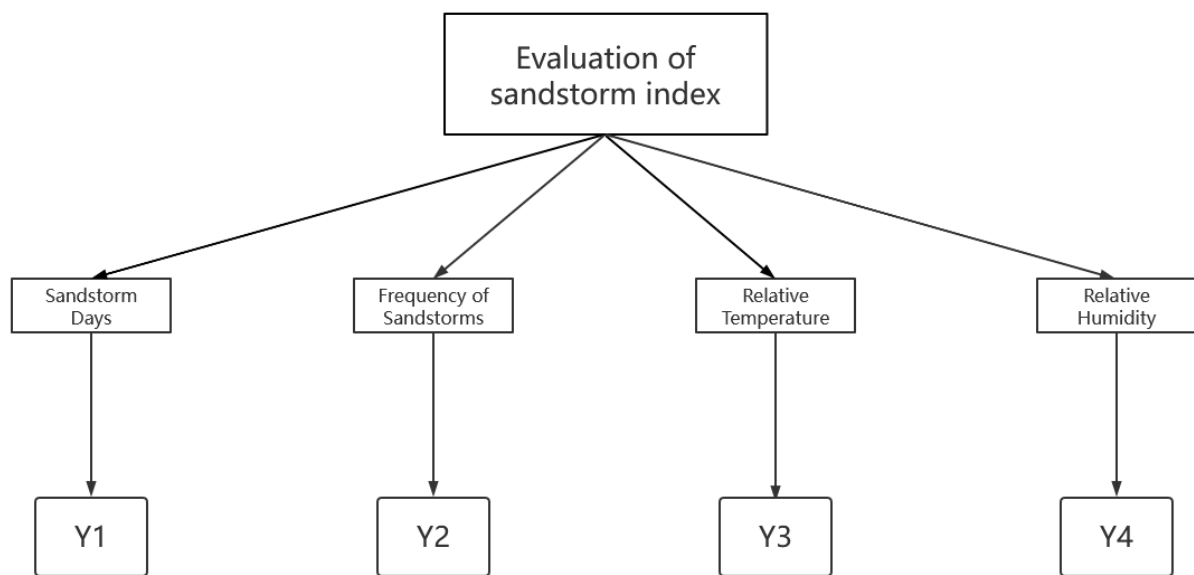
（1）通过咨询专家或翻阅相关文献，得到影响沙尘暴程度的主要因素

通过查阅资料，我们得到沙尘暴评价指标主要有：沙尘暴日数，沙尘暴发生频率，相对温度，相对湿度，空气质量，多云天气占比，恶劣天气占比。

（2）尽可能减少不稳定因素的影响

问题而要求我们检验塞罕坝修复前后对北京抗沙尘暴能力的影响，自变量是塞罕坝是否修复，因此有些不易统计，或者容易随着时间推移发生大幅度改变的因素，我们选择筛除。

综上 我们选取沙尘暴日数 Y1，沙尘暴发生频率 Y2，相对温度 Y3，相对湿度 Y4,4 个指标做为北京抗沙尘暴能力的指标。



沙尘暴能力的指标图 4-6

4.2.2 建立基于主成分分析的综合评价模型

4.2.2.1 数据准备

与问题一所用公式相同，先标准化公式 $(x - \bar{x}) / s$ ，把收集的数据标准化，然后再通过公式 () 求出各评价指标的相关系数，构成相关系数矩阵。

4.2.2.2 KMO 和 Bartlett 球形度检验

结果如下：

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.705
Bartlett's Test of Sphericity	Approx. Chi-Square	72.411
	df	6
	Sig.	.000

KMO 和 Bartlett 球形度检验表 4-6

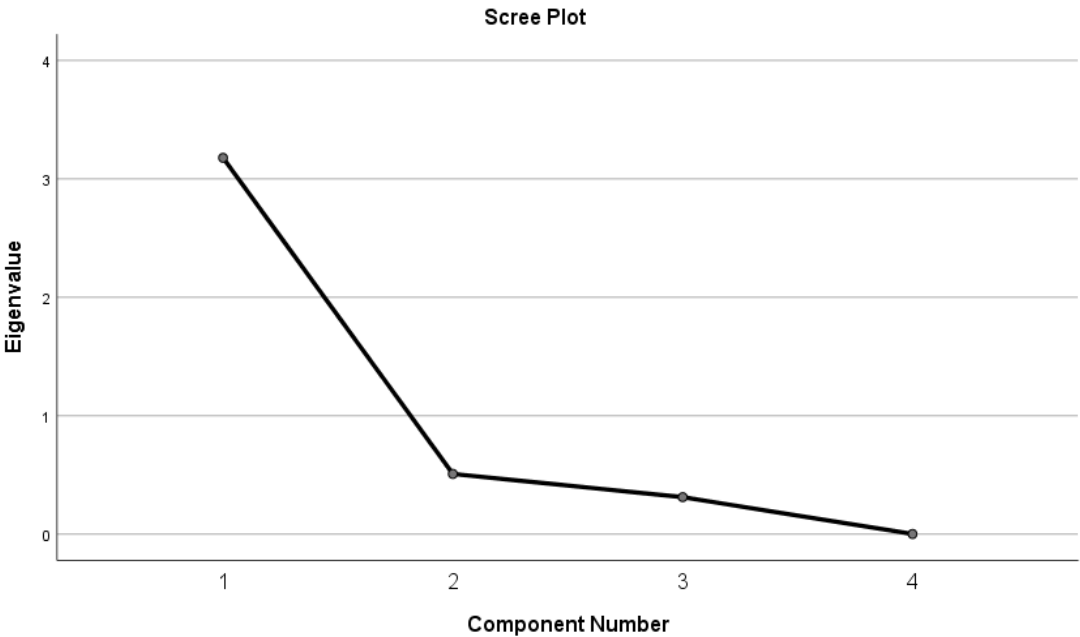
KMO 检验的结果显示，KMO 的值为 0.705，同时，Bartlett 球形检验的结果显示，显著性 P 值为 $0.000 < 0.05$ ，水平上呈现显著性，拒绝原假设，各变量间具有相关性，主成分分析有效，程度为一般。

4.2.2.3 主成分的选取

由碎石图和总方差解释

Total Variance Explained			
Component	Total	Initial Eigenvalues	
		% of Variance	Cumulative %
1	3.178	79.446	79.446
2	.509	12.714	92.160
3	.312	7.808	99.969
4	.001	.031	100.000
Extraction Method: Principal Component Analysis.			

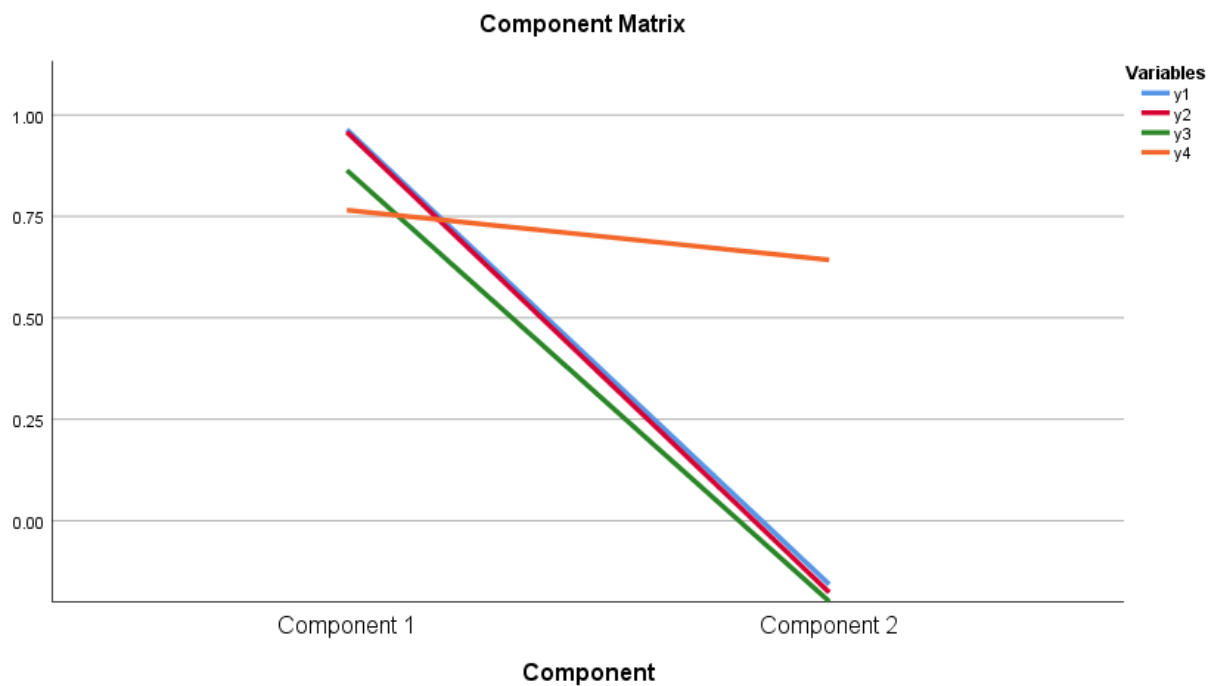
总方差解释表 4-7



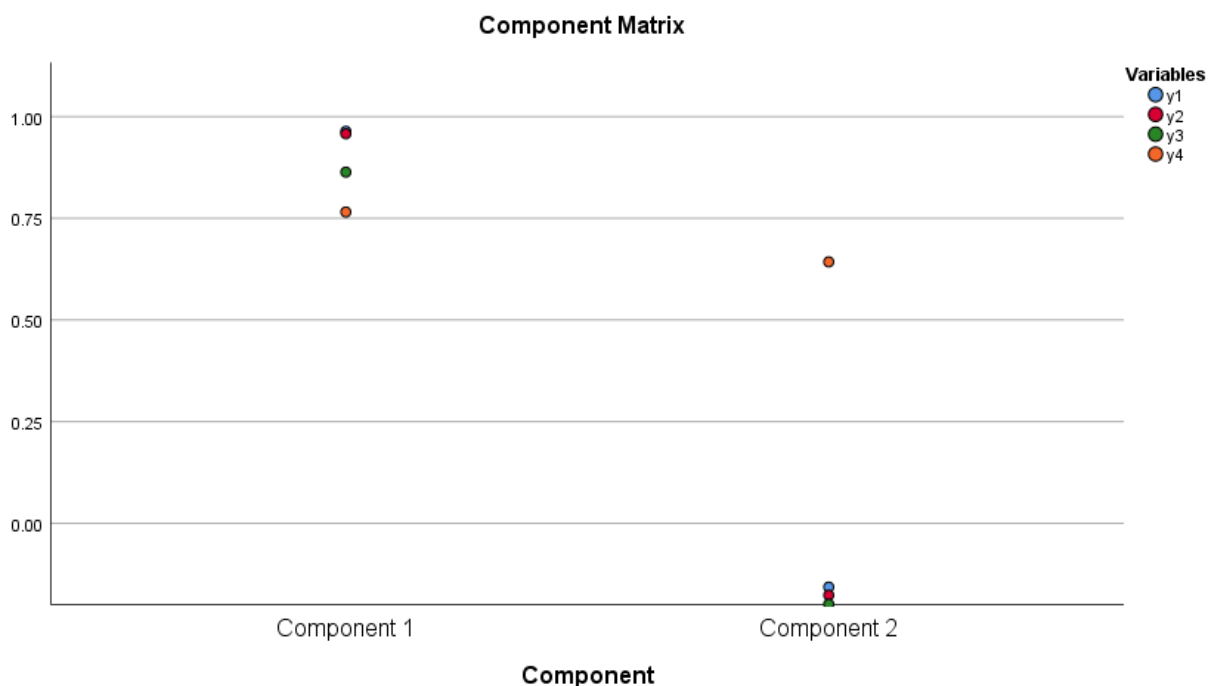
碎石图 4-7

方差解释表中，在主成分 2 时，总方差解释的特征根低于 1.0，变量解释的贡献率达到 92.16 即降维后得到的两个公因子概况了 90%以上的原始指标信息，具有较强代表性和有效性，符合公因子提取条件。结合碎石图分析结果确定最终选取的公因子为 M1 和 M2

其因子载荷矩阵图和因子载荷象限图为



因子载荷矩阵图 4-8



因子载荷象限图 4-9

参照

第一公因子与沙尘暴日数 Y1，沙尘暴发生频率 Y2，相对温度 Y3，相对湿度 Y4 的相关性高，归纳为评价得分因子 F21 ；

第二公因子与相对湿度 Y4 的相关性高，归纳为环境湿度因子 F22 ；

4.2.2.4 基于主成分综合评价模型的构成

我们将评价得分因子 F21、环境湿度因子 F22，表示为原始指标变量的线性组合，以每个公因子的方差累积贡献率作为权重系数得到各因子的得分函数。结果如下图所示：

成分矩阵表为：

Component matrix table		
	Component	
	Component1	Component2
y1	0.303	-0.308
y2	0.301	-0.347
y3	0.272	-0.391
y4	0.241	1.264

成分矩阵表 4-8

由此可以得到

$$F1 = 0.303 \times y1 + 0.301 \times y2 + 0.272 \times y3 + 0.241 \times y4$$

$$F2 = -0.308 \times y1 - 0.347 \times y2 - 0.391 \times y3 + 1.264 \times y4$$

因子权重分析表为：

Principal component weighting results			
Name	Explain variance rate	The rate of total variance explained	The weight
Component1	0.794	0.794	86.205%
Component2	0.127	0.922	13.795%

因子权重分析表 4-9

由上可以得到：

$$F = (0.794/0.922) \times F1 + (0.127/0.922) \times F2 \quad (4-6)$$

把各时间段的 4 个指标带入，可以得到北京市沙尘暴程度的变化，排序结果如下：

Rank	Index	Compre	Component1	Component2
1	2	1.418904832067187	1.7339046084388763	-0.5494665457935304
2	4	0.7620756062071912	0.8921397244300648	-0.05066937585434744
3	5	0.747186468330611	0.8827516733332323	-0.09993371833934764
4	3	0.7368964056940677	0.9239309379875116	-0.4318454400257608
5	6	0.4925533189434863	0.54224085614487	0.18206571418979722
6	7	0.18865368068658425	0.12299856991489876	0.5989194984653539
7	8	0.1588754049334238	0.10422246772123378	0.500390813495353
8	13	0.10723160291693896	-0.02730232550272466	0.9479075459371596
9	12	-0.3985119675081446	-0.7108409052827784	1.5531698675538945
10	11	-0.6902699268949467	-0.8486587446808423	0.29947050750802173

北京市沙尘暴程度变化表 4-10

由此我们可以得出，在修复塞罕坝后的第一个三年间（1962~1964），北京市沙尘暴程度评分为 1.418，属于受沙尘暴影响的地区。

在修复塞罕坝后的第 11 个三年间(1992~1994) 沙尘暴程度评分为-0.69，沙尘暴程度改善了 2.10 个指标

4.3 问题 3 模型的建立与求解

4.3.1 数据的处理

第三问要求我们确定中国哪些地理位置需要建立生态区域，以及建立的规模。我们可以收集全国适宜建设生态环境区的地区近十年的平均降水量、覆盖面积、林木储蓄、蓄水量、氧气释放量、二氧化碳吸收量、大风天数。取平均值之后，将数据标准化。我们选择的城市有：天津、石家庄、太原、呼和浩特、沈阳、大连、长春、哈尔滨、上海、南京、杭州、合肥、福州、南昌

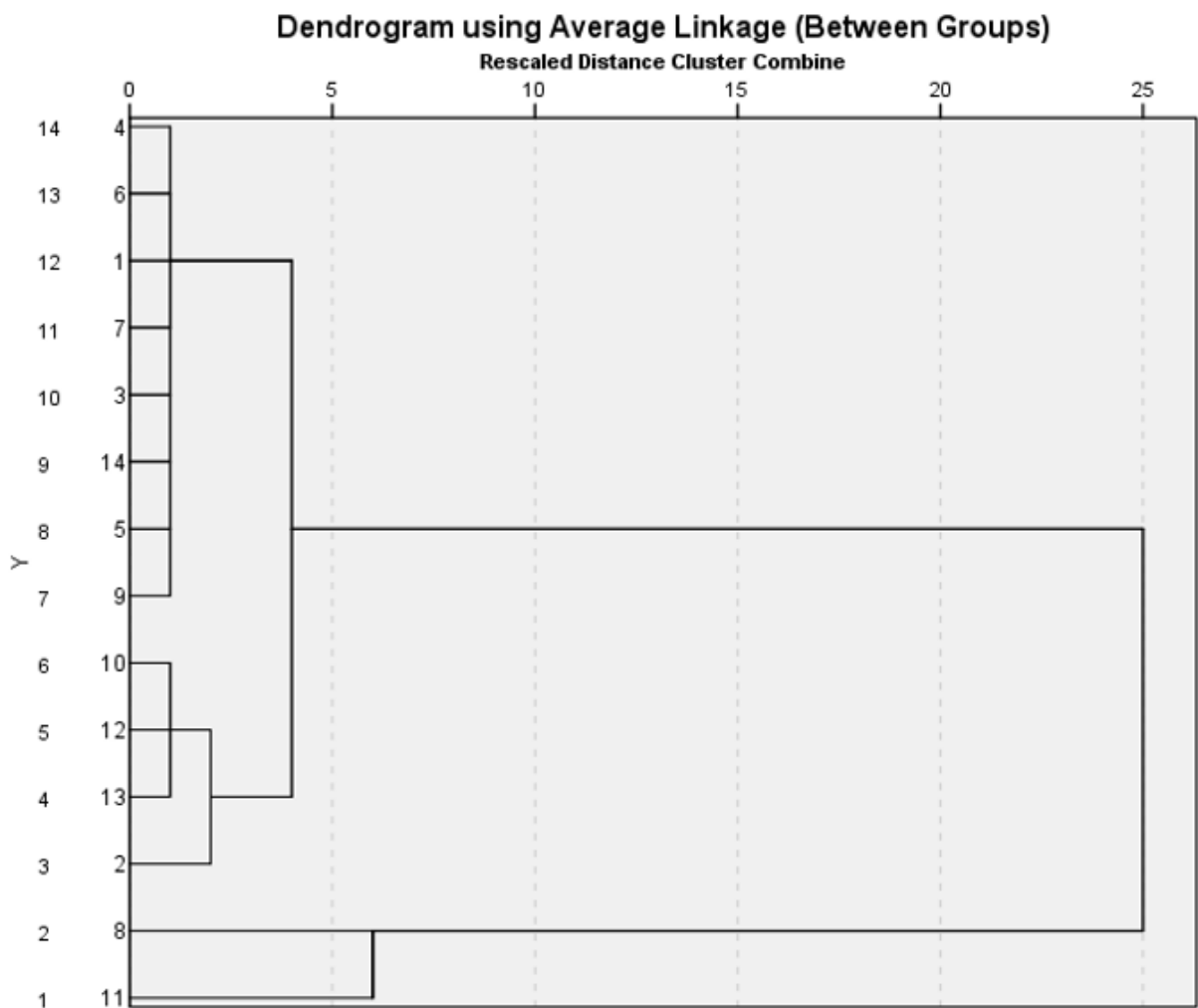
处理有后的结果如图：

Cite Name/Index	x1	x2	x3	x4	x5	x6	x7
Tianjin	-0.40204	-0.61969	-0.74732	-0.4295	1.141358	1.167884	-0.54468
Shijiazhuang	-0.49921	-0.02951	0.738634	-0.56152	-0.54234	-0.54749	1.688272
Taiyuan	-1.17615	-0.50892	-0.76714	-0.6288	-0.71901	-0.72749	-0.93545
Huhehaote	-1.21477	-0.13249	-0.33943	-0.62295	-0.85368	-0.86469	0.125205
Shenyang	-0.88138	-0.4779	-0.69075	-0.34172	-0.47988	-0.47341	-0.09809
Dalian	0.146829	-0.08599	-0.3394	-0.14694	-0.4213	-0.42418	-0.43303
Changchun	-0.98506	-0.57773	-0.75838	-0.40939	-0.46069	-0.4643	-2.2194
Harbin	-0.94603	3.059897	2.563325	0.76989	-0.58916	-0.62426	0.795091
Shanghai	0.822954	-0.5811	-0.70562	-0.29033	2.864131	2.834573	-0.54468
Nanjing	0.632273	-0.49372	0.144118	-0.58598	0.328006	0.347537	0.337335
Hangzhou	1.217529	1.042622	1.517384	1.003428	0.692836	0.710924	-0.67866
Hefei	0.357229	-0.49506	0.100281	-0.38811	-0.18576	-0.18953	0.906738
Fuzhou	1.247818	0.44346	0.070412	-0.38248	-0.20697	-0.17889	1.141199
Nanchang	1.680001	-0.54387	-0.78612	3.014402	-0.56752	-0.56668	0.460148

标准化处理结果表 4-11

4.3.2 聚类分析模型的建立

我们将 7 个指标对 14 个地区进行 Q 型聚类分析，其中样本间相似性采用欧几里得距离度量，类间距离的计算选用类平均法，聚类树形图如下



聚类树形图 4-10

4.3.3 选择需要改善的地区，并推荐规模

通过聚类分析我们选项，城市 8 长春和城市 11 南京与其他地域有明显差异。因此我们选择改善长春和南京两个地区。

通过公式(1-1)我们得到长春的生态环境评分为-2.956,南京的生态环境评分为 0.7426。通过分析收集的数据，我们得到长春的各项指标都很差，需要尽可能多的建立各种森林生态环境系统，而南京的蓄水量比较少，应该多种杉树。

结合地图，分析二者的地型：



二者地形图 4-11

根据可植数面积，我们得到：

应该在南京建立 1351223.4 亩的森林生态系统，其中主要种植杉树来提高降水量。

应该在长春建立 1789419.6 亩的森林生态系统，其中各种类型的树都应该兼顾。

4.4 问题 4 模型的建立与求解

4.4.1 选取适宜地区

问题 4 要求我们选择合适的目标，定量分析其特征并合理借鉴中国塞罕坝的生态保护模式，讨论该国需要建设哪个地理位置。并确定待建生态区域的数量或规模，并评估其对吸收温室气体和减少碳排放的影响。考虑到越南作为“一带一路”沿线的重要国家，且越南和中国同为正在大力发展工业化的发展中国家，且两国政治和经济结构都十分相似，因此我们组选择了越南这个国家作为第四问分析的国家。越南首都河内作为全国的政治、经济中心，也是越南国内最大的工业中心，其经济约占全国 GDP 的 12.1%，年碳排放约为 952.67 万吨，因此在河内附近规划生态保护区，为越南缓解首都地区的污染和碳排放就尤为重要，通过合理的评估分析，我们团队最终选择了沿红河上游的安沛市和寨忽市附近红河河谷建立生态保护区（寨忽距河内 145 公里，安沛距河内 121 公里）。

4.4.2 规划生态保护区面积

通过将收集到的各项指标带入公式（1-1）中，我们得到该地区的生态环境评分为0.344。不算太高，有建立生态保护区的价值。

观察所收集的数据，其森林覆盖率已经很高了，但其余各项指标较低，因此我们可以减少占地，多种植不同种类的树。

通过分析寨忽市地图，我们确定建立面积为 851223.4 亩，林木储蓄为 11 万立方米且均衡种植各种适宜生长的植物的生态保护区。

4.4.3 RBF 神经网络求解模型

4.4.3.1 RBF 网络结构

RBF 神经网络有很强的逼近能力，分类能力，和学习能力，其工作原理是把网络看成对未知函数的逼近，任何函数都可以表示成一组基函数的加权和，即选择各隐层神经元的传输函数，使其构成一组基函数来逼近未知函数。RBF 神经网络由一个输入层，一个隐含层，和一个输出层组成。神经网络的隐层基函数有多种形式，常用函数为高斯函数，设输入层的输入为 $X = [x_1, x_2, x_3, \dots, x_n]$ ，实际输出为 $Y = [y_1, y_2, \dots, y_p]$ 。

输入层实现 $X \rightarrow R(X)$ 的非线性映射，输出层第 k 个神经网络输出为：

$$\hat{y}_k = \sum_{i=1}^m w_{ik} R_i(X), k = 1, 2 \dots p \quad (4-7)$$

式中： n 为输入层节点数； m 为隐含层节点数； p 为输出层节点数； w_{ik} 为隐含层第 i 个神经元与输出层第 k 个神经元的连接权值； R_i 为隐含层第 i 个神经元的作用函数，即

$$R_i(X) = \exp(-||X - C_i||^2 \div 2\sigma_i^2), i = 1, 2 \dots, m \quad (4-8)$$

式中： C_i 为第 i 个基函数中心，与 X 具有相同的维度向量； σ_i 为第 i 个基函数的宽度， m 为感知单元的个数（含隐层节点数）； $||X - C_i||$ 为向量 $X - C_i$ 的范数，它通常

表示 X 与 C 之间的距离； $R_i(X)$ 在 C_i 处有唯一最大值，随着 $||X - C_i||$ 的增大，

$R_i(X)$ 迅速衰减 0。

对于给定的输入，只有一小部分靠近 X 的中心被激活。当确定了 RBF 神经网络的聚类中心 C_i 。权值 w_{ik} 及 σ_i 以后，就可以求出给定某一输入时，神经网络对于的输出。

(1) 确定基函数的中心 C_i ，一般使用 K 均值聚类法，使用一组数据来计算 m 个 C_i ，

使 C_i 尽可能均匀的对数据抽样，在数据密集处 C_i 也密集。

(2) 确定基函数宽度 σ_i ，当基函数中心 C_i 训练完后，可以通过基函数宽度 σ_i 表示与每个中心相联系的子样本集中样本散布的一个测度。常令其等于基函数中心与子样本集中样本模式间的平均距离。

(3) 确定从隐含层到输出层的连接权重 w_{ik} ，RBF 神经网络连接权 w_{ik} 的修正可以采用最小均方误差测度准则进行。

4.4.3.2 网络学习样本的建立

我们利用问题一，我们选取覆盖面积 x_2 、林木储蓄 x_3 作为输入，二氧化碳吸收量 x_6 作为输出，构成两个输入一个输出的网络，将问题一中国塞罕坝测得的 13 个数据作为训练样本集。进行训练。

4.4.3.3 网络的训练

利用 matlab 提供的工具箱实现 RBF 神经网络，解得预测值为 $x_6=1.4872$
即建设该自然保护区，可以额外吸收二氧化碳 1.4872 万立方吨

4.5 问题 5 的求解

4.5.1 关于建设生态保护区的报告

我们团队所建立的模型：

塞罕坝的成功，不仅在于一代代人的坚守，更在于他们能够从中国劳动人民的智慧中学习合理的方法。

我们团队结合塞罕坝及其它成功案例，剔除干扰因素，选择出平均降水量、覆盖面积、林木储蓄、蓄水量、氧气释放量、二氧化碳吸收量、大风天数 7 个指标，并用因子分析进行降维处理，建立出森林生态系统评价模型。然后通过筛选后的指标权重得分来定量计算出对生态环境的情况，并将修复前后生态环境的情况进行对比，得出各种指标能够改善生态环境情况的具体大小。通过该环境评价模型，我们能够通过少量数据，判断出该地区生态环境是否正常。

塞罕坝建设的成功极大地缓解了北京地区的生态缓解，尤其是北京沙尘暴天数得到极大减少。而我们团队的模型将其拓展，将沙尘暴的评判指标细化为所选地区的能见度，沙尘暴风力强度，相对湿度和相对温度。通过这四个指标，建立出基于主成分分析的综

合评价模型，定量计算北京的沙尘暴强度，也可以通过统计这些指标来衡量所选地区的生态环境，量化评判该地区生态环境的改善程度。

塞罕坝的成功是“中国模式”的成功。在其它地区的生态评定中，只要测定平均降水量、覆盖面积、林木储蓄、蓄水量、氧气释放量、二氧化碳吸收量、大风天数 7 个指标，通过对所选地区的 7 个指标进行 Q 型聚类分析，就可以轻松将生态环境指标有明显差异的城市进行分类，再通过生态环境评价模型还可以对其指标进行量化打分，筛选出生态环境状况不好的地区，再对比数据并根据该地区可供植树的面积，规划生态区域的规模。

给 APX4CM 的建议：

为改善环境，应对世界气候变化，我们团队建议 APX4CM 使用我们的指标选取方法（尽可能减少不稳定因素的影响），通过我们建立的环境评价模型，将评分较低的城市统一整理，和我们分析中国国内城市一样，对其进行 Q 类聚类分析，重点分析与其他地区生态环境情况明显差异的城市，再通过该地区的地型，以及所收集到的数据情况，确定好建立的规模，以及主要种植什么类型的树。再通过 RBF 神经网络预测模型，推算出建立此生态保护区能够给当地带来什么影响，均衡得失后决定是否进行建立生态保护区。

五、模型的分析与检验

5.1 模型分析

本文根据数学建模算法，得出了 4 个问题的以下模型：

问题 1: 考虑影响森林生态环境的主要因素且尽可能减少不稳定因素的影响下我们选取平均降水量、覆盖面积、林木储蓄、蓄水量、氧气释放量、二氧化碳吸收量、大风天数做为森林生态环境评价指标。由此我们可以得出，在修复塞罕坝后的第一个三年间（1962~1964），森林生态环境情况评分为-4.106，即不符合森林生态环境标准，在修复塞罕坝后的第 8 个三年间（1983~1985）生态环境评分为 0.9，开始符合森林生态环境标准，在修复塞罕坝后的第 13 个三年间（1998~2000）生态环境评分为 9.70，森林生态环境良好，改善幅度为 13.30。

问题 2: 基于影响沙尘暴程度的主要因素以及尽可能减少不稳定因素的影响下，我们选取尘暴日数，沙尘暴发生频率，相对温度和相对湿度作为北京抗沙尘暴能力的指标。

问题 3: 我们将 7 个指标对中国天津、石家庄、太原、呼和浩特、沈阳、大连、长春、哈尔滨、上海、南京、杭州、合肥、福州、南昌地区进行 Q 型聚类分析，最终我们选择应该重点改善长春市和南京市的环境，通过聚类分析模型，我们得到长春的生态环境评分为-2.956，南京的生态环境评分为 0.7426。并得出应该在南京建立 1351223.4 亩的森林生态系统，其中主要种植杉树来提高降水量。在长春建立 1789419.6 亩的森林生态系统，其中各种类型的树都应该兼顾。

问题 4: 在亚太地区各国中我们重点选择了越南首都河内，通过与前几个问题相同的方法得出该地区的生态环境评分为 0.344，有建立生态保护区的价值。并观察所收集的数据，得出其森林覆盖率已经很高，但其余各项指标较低，最终得出可以减少占地，多种植不同种类的树。并通过分析地图，最终确定建立面积为 851223.4 亩，林木储蓄为 11 万立方米且均衡种植各种适宜生长的植物的生态保护区。最后利用 RBF 神经网络训

练，计算出最终可以额外吸收二氧化碳 1.4872 万立方吨。

5.2 模型检验

各模型的检验均在模型的建立中进行，我们现在对结果进行检验：

根据塞罕坝在生态方面的重要作用，我们选择了选择平均降水量、覆盖面积、林木储蓄、蓄水量、氧气释放量、二氧化碳吸收量、大风天数这 7 个合适的指标，并收集了北京市的相关数据，建立起了塞罕坝对生态环境影响的评价模型，以定量评价塞罕坝恢复后对环境的影响。通过因子分析法后，我们得出，在修复塞罕坝后的第 8 个三年间，森林生态环境情况评分为-4.106，即不符合森林生态环境标准，在修复塞罕坝后的第 8 个三年间，生态环境评分为 0.9，开始符合森林生态环境标准，在修复塞罕坝后的第 13 个三年间，生态环境评分为 9.70，森林生态环境良好，改善幅度为 13.30。这也复合客观事实。

比较分析塞罕坝修复前后的环境条件。

在仔细选择以后我们选取了沙尘暴日数，沙尘暴发生频率，相对温度和相对湿度作为沙尘暴的评定指标，并基于主成分分析法得出在修复塞罕坝后的第一个三年间北京市沙尘暴程度评分为 1.418，属于受沙尘暴影响的地区；在修复塞罕坝后的第 11 个三年间沙尘暴程度评分为-0.69，并得出沙尘暴程度改善了 2.10 个指标。这也与现实契合。

通过聚类分析我们选择的若干城市，长春市和南京市与其他地域有明显差异。因此我们选择改善长春和南京两个地区。并通过公式（1-1）我们得到长春的生态环境评分为-2.956，南京的生态环境评分为 0.7426。再通过分析收集的数据，得出应该在南京建立规模为 1351223.4 亩的森林生态系统，其中主要种植杉树来提高降水量。应该在长春建立 1789419.6 亩的森林生态系统，其中各种类型的树都应该兼顾。这也与现实中两市的主要植树计划相似。

我们组在亚太地区许多国家中选择了越南，并以其首都河内为例子重点研究问题。通过合理的因子评估分析，并用主成分分析选择主要影响因子，结合河内的地理位置，我们团队最终选择了沿红河上游的安沛市和寨忽市以附近红河河谷为重点建立生态保护区，并且通过数据分析，得出应该着重提高林木储量，提高生态多样性而不是只增加林地占地面积的建立生态保护区保护区。这一结论能够有效缓解越南的生态问题，解决方法也符合客观事实。

综上，本文中提出的模型基本正确合理，且有较高的精度和鲁棒性。完成了题目要求。

六、 模型的评价与推广

6.1 模型的优缺点

优点：本文森林生态环境的影响因素选取平均降水量、覆盖面积、林木储蓄、蓄水量、氧气释放量、二氧化碳吸收量、大风天数做为森林生态环境评价指标。沙尘暴程度指标为：沙尘暴日数，沙尘暴发生频率，相对温度，相对湿度。并每隔 3 年取一次数据的平均值，共获得 13 组数据。在此基础上通过因子分析、主成分分析、Q 型聚类分析和 RBF 神经网络预测。将生态分析问题转化为给定部分影响因子下定量计算指标，在

保证算法不过于复杂的条件下提高了模型的精度；对不同地区不同指标的生态保护区问题合理选择指标方法，尽可能增加指标与生态环境的相关性并减少其它无关因子干扰，最后求给定指标下其它指标的最优选择，提高建立生态保护区的生态环境评分。本文使用的因子分析法、主成分分析法和 RBF 神经网络有很好的收敛性，在短时间内可以有效的求解出优化结果，不会出现梯度消失或者梯度爆炸的问题。

缺点：本文仅按照题中所要求或选择的特定条件来建立适应本题的模型，且主要考虑建立生态保护区前自然因子的影响，若人为影响较多，或选择的生态保护区为海洋等无法用树木降水描述的区域，则本文得出的模型还需进一步改进。

6.2 模型的推广

本文根据生态保护区的需求，合理选取了平均降水量、覆盖面积、林木储蓄、蓄水量、氧气释放量、二氧化碳吸收量、大风天数做为森林生态环境评价指标。适用于多种环境，步步精进，登堂入室，并且能够根据需求调整指标或因子模型将保护区的环境增益最大化。本文模型能够考虑诸多现实因素，且鲁棒性较强，可应用于生态保护区、环境评估分析等许多方面，有较强的现实意义。

七、 参考文献

[1] 中国地图矢量图 <http://ditu.yjdy.org/read/MDIOMWNj0WLkuK3lm73lnL=/>
2017 年 10 月 31 日 08:10

[2] 司守奎 数学建模算法与应用 国防工业出版社 2011 年 8 月版

附录：

介绍：支撑材料

1. x1~x7 数据统计.xlsx：第一问七个指标数据表格
2. 第二问数据.xlsx：第二问有关沙尘暴四个指标的数据表格
3. 第三问数据.xlsx：第三问 Y 各个城市地区七个指标的数据表格
4. 第四问数据.xlsx：第四问越南相关指标数据表格
5. rbf.mat: rbf 神经网络预测 matlab 程序
6. quation3.py: 第三问计算 F 值 Python 程序
7. quation4.py: 第四问计算 F 值 Python 程序
8. spss 第一问因子分析相关操作代码
9. spss 第二问主成分分析相关操作代码
10. spss 第三问聚类分析相关操作代码

附录 2

介绍：rbf 神经网络预测 matlab 程序

```

format short
clc,clear;
a= [
440.5 22.29      33.37      0.09  2.77  1.93  -77
440.5 31.75      36.70      0.10  3.04  2.12  -77
440.5 40.45      43.86      0.12  3.64  2.53  -77
413.6 48.41      54.70      0.15  4.54  3.16  -76
413.6 55.64      69.05      0.18  5.73  3.99  -76
413.6 62.16      86.72      0.22  7.2   5.01  -76
440.8 67.99      107.55     0.25  8.92  6.21  -49
440.8 72.94      139.62     0.29  11.59 8.06  -49
440.8 77.06      210.83     0.33  17.49 12.17 -49
440.8 81.35      259.13     0.37  21.5  14.96 -28.3
513   86.36      320.46     0.42  26.59 18.5  -28.3
513   91.51      398.72     0.48  33.08 23.01 -28.3
513   95.97      502.89     0.56  41.73 29.02 -28.3]
a= a' ;P=[
22.29      33.37
31.75      36.70
40.45      43.86
48.41      54.70
55.64      69.05
62.16      86.72
67.99      107.55
72.94      139.62
77.06      210.83
81.35      259.13
86.36      320.46
91.51      398.72
95.97      502.89
]'
[PN,PS1]= mapminmax(P);
T=[1.93  2.12  2.53  3.16  3.99  5.01  6.21  8.06  12.17  14.96  18.5  23.01  29.02]'
[TN,PS2] = mapminmax(T);
net1=newrb(PN,TN)
x=[8.5  11];
xn=mapminmax('apply',x,PS1);
yn1=sim(net1,xn);
y1=mapminmax('reverse',yn1,PS2);

```

介绍：第三问计算 F 值 Python 程序

```

a1=0.546/0.972    a2=0.425/0.972

F1=a1*(-0.402038642-0.619691948-0.747320872-0.42950167+1.141358274+1.167884089-
0.544681051)+a2*(0.054+0.146+0.111+0.129+0.111+0.111+0.125)

F2=a1*(-0.499208629-0.029508331+0.738633548-0.561517002-0.542342172-
0.54748927+1.688272013)+a2*(0.054+0.146+0.111+0.129+0.111+0.111+0.125)

F3=a1*(-1.176145984-0.508920741-0.767137712-0.628802445-0.719014942-0.727485514-
0.935447837)+a2*(0.054+0.146+0.111+0.129+0.111+0.111+0.125)

F4=a1*(-1.214770037-0.132488496-0.339430792-0.622954472-0.853684374-
0.86468828+0.125204868)+a2*(0.054+0.146+0.111+0.129+0.111+0.111+0.125)

F5=a1*(-0.881383472-0.477901134-0.690747822-0.341723371-0.479875386-0.473413096-
0.098090438)+a2*(0.054+0.146+0.111+0.129+0.111+0.111+0.125)

F6=a1*(0.146829484-0.085986516-0.339402885-0.146937748-0.421304747-0.424175621-
0.433033398)+a2*(0.054+0.146+0.111+0.129+0.111+0.111+0.125)

F7=a1*(-0.985058562-0.577725429-0.758381275-0.409391926-0.460693812-0.464304784-
2.219395849)+a2*(0.054+0.146+0.111+0.129+0.111+0.111+0.125)

F8=a1*(-0.94602794+3.059896622+2.563325267+0.769890438-0.589164667-
0.624263564+0.795090787)+a2*(0.054+0.146+0.111+0.129+0.111+0.111+0.125)

F9=a1*(0.822953701-0.581095841-0.705616785-0.29032903+2.864130896+2.834573242-
0.544681051)+a2*(0.054+0.146+0.111+0.129+0.111+0.111+0.125)

F10=a1*(0.632272848-0.493723607+0.14411812-
0.585976029+0.328005824+0.347537232+0.337335409)+a2*(0.054+0.146+0.111+0.129+0.111+0.111+0.125)

F11=a1*(1.217528898+1.042622458+1.517384351+1.003427685+0.692835906+0.710924285-
0.678658235)+a2*(0.054+0.146+0.111+0.129+0.111+0.111+0.125)

F12=a1*(0.357228932-0.495064494+0.100280933-0.38810789-0.18576401-
0.189525675+0.906738441)+a2*(0.054+0.146+0.111+0.129+0.111+0.111+0.125)

F13=a1*(1.247818287+0.443460258+0.070412038-0.38247863-0.206970648-
0.178889746+1.141198512)+a2*(0.054+0.146+0.111+0.129+0.111+0.111+0.125)

F14=a1*(1.680001115-0.543872802-0.786116112+3.014402092-0.567516143-
0.566683298+0.460147828)+a2*(0.054+0.146+0.111+0.129+0.111+0.111+0.125)

print(F1)
print(F2)
print(F3)
print(F4)
print(F5)
print(F6)
print(F7)
print(F8)
print(F9)
print(F10)
print(F11)

```

```
print(F12)
print(F13)  print(F14)
```

附录 4

介绍：第四问计算 F 值 Python 程序

```
a1=0.546/0.972
a2=0.425/0.972
F=a1*(-0.379318808+2.267774876-0.370702588-0.379526855-0.379291271-0.379413629-
0.379521726)+a2*(0.054+0.146+0.111+0.129+0.111+0.111+0.125)
print(F)
```

附录 5

介绍：spss 第一问因子分析相关操作代码

DATASET ACTIVATE 数据集1.

FACTOR

```
/VARIABLES x1 x2 x3 x4 x5 x6 x7
/MISSING LISTWISE
/ANALYSIS x1 x2 x3 x4 x5 x6 x7
/PRINT INITIAL CORRELATION SIG DET KMO EXTRACTION ROTATION FSCORE
/PLOT EIGEN ROTATION
/CRITERIA FACTORS(2) ITERATE(25)
/EXTRACTION PC
/CRITERIA ITERATE(25)
```



```
/ROTATION VARIMAX  
/METHOD=CORRELATION.
```

附录 6

介绍：spss 第二问主成分分析相关操作代码

DATASET ACTIVATE 数据集3.

FACTOR

```
/VARIABLES y1 y2 y3 y4  
/MISSING LISTWISE  
/ANALYSIS y1 y2 y3 y4  
/PRINT INITIAL CORRELATION SIG DET KMO EXTRACTION ROTATION FSCORE  
/PLOT EIGEN ROTATION  
/CRITERIA FACTORS(2) ITERATE(25)  
/EXTRACTION PC  
/CRITERIA ITERATE(25)  
/ROTATION VARIMAX  
/METHOD=CORRELATION.
```

附录 7

11. 介绍：spss 第三问聚类分析相关操作代码

DATASET ACTIVATE 数据集4.

CLUSTER x1平均降水量mm x2覆盖面积亩 x3林木储蓄立方米 x4蓄水量亿立方米 x5氧气释放量万吨
x6二氧化碳吸收量万吨 x7大风天数天

```
/METHOD BAVERAGE  
/MEASURE=SEUCLID  
/PRINT SCHEDULE  
/PLOT DENDROGRAM VICICLE.
```

其他需求请联系Q1423801308