

第四届“互动出版杯”数学中国

数学建模网络挑战赛

承 诺 书

我们仔细阅读了第四届“互动出版杯”数学中国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。

我们的参赛队号为：1099

参赛队员（签名）：

队员 1：金日浩

队员 2：周健聪

队员 3：刘硕

参赛队教练员（签名）：

参赛队伍组别：大学组

第四届“互动出版杯”数学中国

数学建模网络挑战赛

编 号 专 用 页

参赛队伍的参赛队号：（请各个参赛队提前填写好）：1099

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）：

2011年第四届“互动出版杯”数学中国 数学建模网络挑战赛

题 目 多指标生物多样性评估模型的实现

关 键 词 多指标生物多样性评价模型 层次分析法 主成分分析法

聚类分析法 归一法

摘 要：

生物多样性是当今国际生态学的三大热点问题之一。尽管相关的大规模科研和考察计划也层出不穷，然而，现在缺少一种能全面考虑不同因素的对生物多样性进行测定的方法。因此，本文建立了多指标生物多样性的评估模型，以利于今后对生物多样性的考察和研究。

生物多样性通常包括遗传多样性、物种多样性、生态系统多样性三个层次，每一个层次包含若干影响生物多样性的因子。通过运用主成份分析法，在每个层次中筛选出若干贡献率较高的影响因子，再运用聚类分析法，对相似程度较高的因子，和对总体影响效果相似的因子整合后，可以提炼出如下八个主要影响因子：（1）群落多样性指数；

（2）物种间多样性相互影响因子；（3）物种密度；（4）生物生产性土地面积；（5）污染指数；（6）水系分布情况指标；（7）总遗传多样性；（8）性状距离。

通过大量的数据统计分析和MATLAB图形处理，可以归纳出以上8个影响因子与整体评价指数的相关性。然后运用层次分析法，以生物多样性评价指标为目标层，以遗传多样性、物种多样性、生态系统多样性为标准层，以上述8个影响因子为方案层，将各个因素两两比较得到比较矩阵，从而求出各影响因子相对目标层的权重，再将归一化处理后的原始数据运用公式 $SS = \sum w_i SD_i$ ($i=1, 2, 3, \dots, 8$)，得出最终的生物多样性评价指数，该指数可以简单、清晰化地描述单位样本范围内的生物多样性，有助于生物科学中对生物多样性的研究考察工作。

本模型在 MATLAB7.0 环境下，对采集到的数据进行统计分析等处理；在 Eclipse 编译环境下，完成对整体流程算法的 C++实现(源代码见附录)。从而对生物多样性进行准确而系统的多指标评估。

参赛队号 1099

所选题目 B 题

参赛密码 _____
(由组委会填写)

英文摘要（选填）

（此摘要非论文必须部分，选填可加分，加分不超过论文总分的 5%）

Abstract

Biodiversity is one of the three most popular topics discussed in International Ecology nowadays. Though several scientific research methods and techniques have been developed, all them are lack of the method in measuring biodiversity on an over-rounded scale. Therefore, an assessment model based on multiple factors to determine biodiversity is established and explained in this article in order to assist later biodiversity researches and measures.

Biodiversity are usually divided into three levels: they are genetic diversity, species diversity, and ecosystem diversity. Each of them is affected by several factors. By using Principal Component Analysis and Cluster Analysis, eight principals factors are determined: (1) Shannon-Wiener Index (2) Mutual Influence between The Diversity of Population Factors (3) Species Density (4) Biological Productive Land Area (5) Pollution Index (6) Indicators of Water System Distribution (7) Total Genetic Diversity Degrees (8) Phenotypic Traits Distance

Depending on the analysis of various data and MATLAB graphic, the relationship of the above eight factors and the overall assessment index can be determined. From Analytic Hierarchy Process, taking biological diversity as the goal hierarchy, taking genetic diversity, species diversity, and ecosystem diversity as standard hierarchy, and taking these eight factors as alternative hierarchy, comparison matrix can be determined by comparing various factors, thus the goal hierarchy can be indicated by calculating weights of each affecting factors; after normalizing the original data using formula $SS = \sum W_i SD_i$, the final biodiversity assessment index can be determined. Therefore, the biodiversity assessment index makes the evaluation to the biodiversity with the scope of sample become simpler, clearer, and more comparative.

This model is processed to statistically analyze collected data under the MATLAB 7.0 language compiler and the implement of the program Algorithm flow with C++ can be done under the Eclipse language compiler. Therefore, the multi-factor assessment of biodiversity can be made accurately and systematically.

一、问题重述及分析

百年以来，随着人类活动对生态环境影响的加深，越来越多地区的生态系统平衡遭到破坏，成千上万的物种濒临灭绝的危险。据统计，全世界仅畜牧品种（包括鸟类及哺乳动物）就有21%濒临灭绝。这不得不引起各国人民甚至各国政府的重视。各国政府曾经决定到2010年大幅降低生物多样性丧失的速度，但这一目标未能实现，也就更加引发了人们对于物种多样性的未来的担忧，也就更加彰显了考察生物多样性的意义。

迄今为止，几乎所有的考察计划都面临着一个基本的困难：如何评价被考察区域的生物多样性。传统的方法是清点物种数量，但现在有许多科学家认为这种方法具有很大的局限性。有人提出也应当考虑物种相似程度和基因多样性等。尽管如此现在还缺少一种能全面考虑不同因素的对生物多样性进行测定的方法。因此我们希望建立合理的数学模型，设计一个更全面评价生物多样性的方法，用来满足研究需要。

为了建立一个更加全面评价生物多样性的方法，就不能仅从物种数量，物种相似程度的但角度来考虑问题。而是应该找出尽量多的主要影响因子，将其有机统一起来，多角度的评价一个地区的生物多样性。

根据生物多样性的定义：生物多样性是指一定时间内，一定地区或空间的所有生物物种与变异及其生态系统组成的复杂性；及其成分：生物多样性通常包括遗传多样性，物种多样性和生态系统多样性三个层次。我们可以具体地从这三方面以及他们所包含的若干因素来综合考虑影响生物多样性的因子。

二、模型假设

1. 一切统计数据都要排除极端因素（如自然灾害）的影响；
2. 模型中的因子指数均为段时间内的短期因子，暂不考虑时间影响；
3. 根据实施情况和小概率事件发生概率极低的原则，将模型中的一些无界变量理想化为有界变量；
4. 模型中未涉及的变量对生物多样性的影响相对较弱，假设可以忽略不计；
5. 本文引用数据、资料均为可靠真实。

三、符号定义及说明

生物多样性指标的评分(score)-----SS

影响因子 i 在总体中所占权重(Weight)----- W_i -----

影响因子 i 数据标准数据(Standard Data)----- SD_i -----

原始输入数据向量(Original Data vector)-----ODV

标准归一向量(Standard Normal Vector)-----SV

归一后的数据向量(Normalized Data Vector)-----DV

表形性状距离(Phenotypic Traits Distance)----- d

总遗传多样性(Total Genetic Diversity Degrees) ----- H_T

群落的多样性指数(Shannon-Wiener Index) ----- H

一定区域内总个体数量(Individual Number) -----IN

一定区域内物种 i 的所有个体所占总个体区域内总个体的比例 (Proportion of i)

种群密度(Population Density) -----D

种群间多样性互相影响因子 (Mutual Influence between THE Diversity of Population Factors)-----EI

空气污染指数(Air Pollution Index) -----API

水污染指数(Silt Density index) -----SDI

土壤污染指数(Soil Pollution Index) ----- PP

水系分部状况指标(Indicators of Water System Distribution) -----WSD

生物生产性土地面积(Biological Productive Land Area) -----BP

四、数学模型建立流程

首先，通过数据分析、实践经验以及咨询相关资深教授，我们得出相关影响因子 (x) 对生物多样性 (y) 的影响关系 ($y=f(x)$)；然后，通过层次分析法，可以得出各项影响因子对整体的影响权重；进一步，通过归一法，整理数据，计算出标准化数据；最后，运用公式 $SS = \sum_{i=1}^n W_i \cdot SD_i$ 求得生物多样性评价指标。流程图如下：

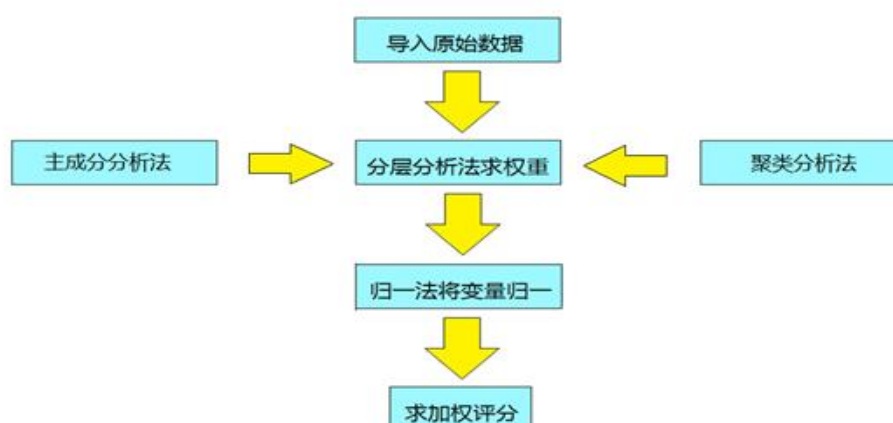


图 4.1

五、模型的建立与求解

5.1 遗传多样性对整体指标的影响分析

5.1.1 性状距离对遗传多样性的影响

定义性状 j 在 i 群体生物中出现总数为 x_{ij} ，群体平均值为 \bar{x}_j ，群体方差为 s^2 ，变异系数 CV （相对于平均值和方差的标准化的差异大小参数）。

$$CV = \sqrt{s^2} / \bar{x} = s / \bar{x} \quad (5.1)$$

CV 属于无量纲参数，可用于不同表型性状间分析比较。

表型性状距离 d 是描述形状差异大小及分布的参数。对某性状或一组性状，两种群体的表型多样性距离 d 有多种，常用的是以下三种定义方式

- 欧氏距离
$$d_{\alpha\beta}^2 = \sum_i (x_{\alpha i} - x_{\beta i})^2 \quad (5.2)$$

- 标准欧氏距离
$$d_{\alpha\beta}^2 = \sum_i (x_{\alpha i} - x_{\beta i})^2 \cdot (s_i^2)^{-1} \quad (5.3)$$

- 积和距离
$$d_{\alpha\beta}^2 = \sum_i Z_{\alpha i} Z_{\beta i} \quad (5.4)$$

式 (5.4) 中 $Z_{\alpha i} = (x_{\alpha i} - \bar{x}_i) \cdot s_i$ ，余类推。不同的表型多样性距离，能反映出表型分化的差异与分布特征。

5.1.2 总遗传多样性指标

在研究总遗传多样度的时候，我们会涉及以下几方面内容：

1. 基因频率与基因谱带频率的方差

实验群体的基因频率估算通常采用巢式设计（Nei, 1987）

设 i 个群体， j 个个体， k 个谱带数值，则有： μ 总均值， S_i 群体效应， $T_{(i)j}$ 群体内个体效应， $\varepsilon_{(ij)k}$ 误差。线性模型是

$$Y_{ijk} = \mu + S_i + T_{(i)j} + \varepsilon_{(ij)k} \quad (5.5)$$

剔除环境及误差后，相应的期望方差 EMS 如下

群体间	$EMS_A = \sigma_w^2 + n\sigma_A^2$	(5.6)
-----	------------------------------------	-------

群体内	$EMS_w = \sigma_w^2$	(5.7)
-----	----------------------	-------

2. 基因多度

基因多度 (genotypic multiplicity) 又称遗传多度, 是遗传变异大小的基础量。等位基因水平数通常以 A 表示, 反映了单基因丰富程度。

设 i 个位点, 每个位点有 a_i 个等位基因, 则基因平均数啊

$$A = \frac{1}{n} \sum_{i=1}^n a_i \quad (5.8)$$

3. 基因杂合度

通常的杂合度 (heterozygosity) 即是观测杂合度 H_0 。其意义是: 群体 j 中 i 个体的第 K 个位点上的杂合度, 又称基因位点杂合度 H_0

$$H_0 = \sum_i \sum_j p_{ij}^k \quad (5.9)$$

如果在一个相对大而稳定的群体中, 基因型在所有位点上符合哈特温伯格平衡定律, 利用 Nei's (1973) 的期望杂合度 H_e 公式来估算杂合度

$$H_e = 1 - \sum_i p_i^2 \quad (5.10)$$

4. 总遗传多样性

Nei 将杂合多样性 H 改称基因多样性, 写作 D_{ST} , 为亚群杂合度 H_T 之差, 继而可直接得出生物种群在一定范围内的遗传多样性指标。

以下为上述总遗传多样性求解的实例操作:

我们已获得蒙古栎 8 个群体和辽东栎 1 个群体的等位酶基因图谱, 详见附录 5。

统计分析结论是:

变因 SV	自由度 df	方差 MS	贡献率 tv	EMS
群体间 A. P	7	45.1632	6.70%	$\sigma_w^2 + n\sigma_A^2$
群体内 W. P	232	14.0754	93.30%	σ_w^2

表 5.1.2.1

由式 (5.8) 求得部分遗传多样性参数

		N	A	Ae	P	H_0	H_e
蒙古栎	种的水平	240	1.828	1.209	52.38	0.092	0.099
			(0.1)			(0.040)	(0.03)
	大兴安岭	30	1.273	1.121	18.18	0.082	0.071
			(0.1)			(0.043)	(0.034)
	嘉荫	30	1.500	1.147	36.36	0.074	0.090
			(0.2)			(0.031)	(0.035)
	双鸭山	30	1.500	1.132	36.36	0.094	0.084

			(0.2)			(0.050)	(0.033)
	蛟河	30	1.364	1.140	27.27	0.098	0.076
			(0.1)			(0.052)	(0.036)
	大青沟	30	1.364	1.141	22.73	0.070	0.080
			(0.1)			(0.033)	(0.037)
	宽甸	30	1.227	1.135	18.18	0.092	0.078
			(0.1)			(0.044)	(0.036)
	雾灵山	30	1.546	1.140	31.82	0.074	0.086
			(0.2)			(0.031)	(0.034)
	赞皇	30	1.591	1.196	40.91	0.121	0.116
			(0.2)			(0.051)	(0.039)
辽东栎	东灵山	33	1.500	1.142	36.36	0.070	0.083
			(0.2)			(0.034)	(0.036)

表 5.1.2.2

再由式 (5.9) (5.10) 可求得点位杂合度，如下

位点		蒙古栎							辽东栎	
杂合度		大兴安岭	嘉荫	双鸭山	蛟河	大青沟	宽甸	雾灵山	赞皇	东灵山
Ata	H ₀	0.3000	0.2000	0.0667	0.0000	0.2667	0.5333	0.3667	0.2002	0.2000
	He	0.4130	0.2418	0.0655	0.0000	0.2966	0.4520	0.4401	0.1876	0.2825
Dia-1	H ₀	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0333
	He	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0333
Dia-2	H ₀	0.8000	0.4333	1.0000	1.0000	0.4333	0.6000	0.1667	0.2000	0.4333
	He	0.4881	0.4921	0.5085	0.5085	0.4001	0.4881	0.1554	0.3977	0.3814
Gpi-1	H ₀	0.4667	0.5333	0.4667	0.4667	0.5333	0.3333	0.5333	0.7000	0.7000
	He	0.4503	0.4503	0.4469	0.5712	0.4901	0.3689	0.5096	0.5588	0.5644
Idh-1	H ₀	0.0000	0.1000	0.0333	0.2000	0.2667	0.0000	0.2667	0.2333	0.1000
	He	0.0000	0.4299	0.2672	0.1876	0.4994	0.0000	0.4316	0.4626	0.3576
Idh-2	H ₀	0.2333	0.2000	0.0333	0.0000	0.0000	0.0000	0.0000	0.0667	0.0000
	He	0.2097	0.1831	0.0333	0.0000	0.0000	0.0000	0.0000	0.0611	0.0655
Mdh-1	H ₀	0.0000	0.1000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0333	0.0333
	He	0.0000	0.0966	0.0655	0.0000	0.0000	0.0000	0.0000	0.0333	0.0333
Mdh-2	H ₀	0.0000	0.0000	0.1667	0.0333	0.0000	0.5667	0.1333	0.0000	0.8000
	He	0.0000	0.0000	0.1554	0.0333	0.0000	0.4130	0.1266	0.0000	0.4994
Mnr-1	H ₀	0.0000	0.0000	0.0000	0.0000	0.0333	0.0000	0.0000	0.0000	0.0333
	He	0.0000	0.0000	0.0000	0.0000	0.0333	0.0000	0.0000	0.0000	0.1554
Pgd-1	H ₀	0.0000	0.0000	0.0000	0.0333	0.0000	0.0000	0.0000	0.0667	0.0000
	He	0.0000	0.0000	0.0000	0.0333	0.0000	0.0000	0.0000	0.0665	0.0000
Pgd-2	H ₀	0.0000	0.0333	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	He	0.0000	0.0333	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Skd	H ₀	0.0000	0.0333	0.3000	0.0433	0.0000	0.0000	0.1667	0.0000	0.3667
	He	0.0000	0.0333	0.2960	0.3452	0.0000	0.0000	0.1576	0.0000	0.6819

表 5.1.2.3

最后，由 $H_T = H_S + G_{ST}$ 得出基因多样性的构成和遗传分化

位点	总遗传多样性 H_T	群体内遗传多样性 H_S	群体间基因多样 DST	基因分化系数 GST	固定指数 FST
Ata	0.2882	0.2621	0.0261	0.0906	0.116
Dia-2	0.4972	0.4348	0.0624	0.1255	0.140
Gpi-1	0.4857	0.4816	0.0041	0.0084	0.025
Idh-1	0.3274	0.2717	0.0557	0.1701	0.184
Idh-2	0.0726	0.0533	0.0193	0.2658	0.057
Mdh-1	0.0247	0.0244	0.0003	0.0121	0.027
Mdh-2	0.1964	0.1576	0.0388	0.1976	0.232
Mnr-1	0.0205	0.0194	0.0011	0.0537	0.074
Pgd-1	0.0085	0.0083	0.0002	0.0235	0.013
Pgd-2	0.00424	0.0042	0.00004	0.0094	0.015
Skd	0.1594	0.1438	0.0156	0.0976	0.089
平均值	0.0993	0.0886	0.0107	0.1073	0.116

表 5.1.2.4

从而求得总遗传多样性为 0.0993

因而我们可以计算出单位样本容量内，全球各区域的总遗传多样性指标，进而求出峰值及谷值，以其作为参照进行归一法处理，从而算出相应的总遗传多样性

综上所述，遗传多样性是生物多样性的本质，在遗传的过程中基因交换、重组和染色体的变异都会决定性的影响一个物种的性状，甚至改变物种本身。故而总遗传多样性指标与生物多样性指标正相关。

5.2 物种多样性对整体指标的影响分析

物种多样性是指地球上动物、植物、微生物等生物种类的丰富程度。为了定量的描述一定区域内的物种多样性，我们不妨引进以下一组变量：

- (1) 群落的多样性指数为 H
- (2) 一定区域内总个体数量为 IN
- (3) 一定区域内物种 i 的所有个体所占总体区域内总个体的比例为 P_i
- (4) 种群密度为 D
- (5) 种群间多样性互相影响因子为 EI

对于群落的多样性指数，根据香农-威纳指数公式（Shannon-Wiener Index）：

$$H = -\sum (P_i) \cdot (\log_2 P_i) \quad (5.11)$$

其中， H =样品的信息含量（彼得/个体）=群落的多样性指数， P_i =样品中属于第 i 种的个体的比例，如样品总个体数为 N ，第 i 种个体数为 n_i ，则 $P_i = n_i / N$ 对于已知的一定区域内物种 i 的所有个体所占总体内总个体的比例 P_i ，我们可以得出其香农-威纳指数 H ；根据香农-威纳指数的意义，我们知道： H 是一个介于 $[0, 1]$ 之间的量。各种之间，个体分配越均匀， H 值就越大。如果每一个体都属于不同的种，多样性指数就最大；如果每一个体都属于同一种，则其多样性指数就最小。因此，香农-威纳指数与物种的多样性呈正相关。

种群密度指单位空间（面积）内的个体数量。根据定义，单位面积的个体数量越多，种群密度就越大，物种多样性就越大。显然，物种多样性与种群密度呈正相关。

此外，影响物种多样性的还有种群间多样性互相影响因子，举例说明，如图 5.2：

数据统计点 (N)	1	2	3	4	5	6	7	8	9
叶子生长高度的多样性 (x)	0.02	0.45	0.55	0.69	0.71	0.76	0.79	1.0	1.19
鸟类物种多样性 (y)	0.34	1.32	1.35	1.84	2.45	2.38	2.47	2.51	2.87

引自 R.H. MacArthur and MacArthur, Ecology 42:594~598(1961)

表 5.2

用 MATLAB 作图可得：

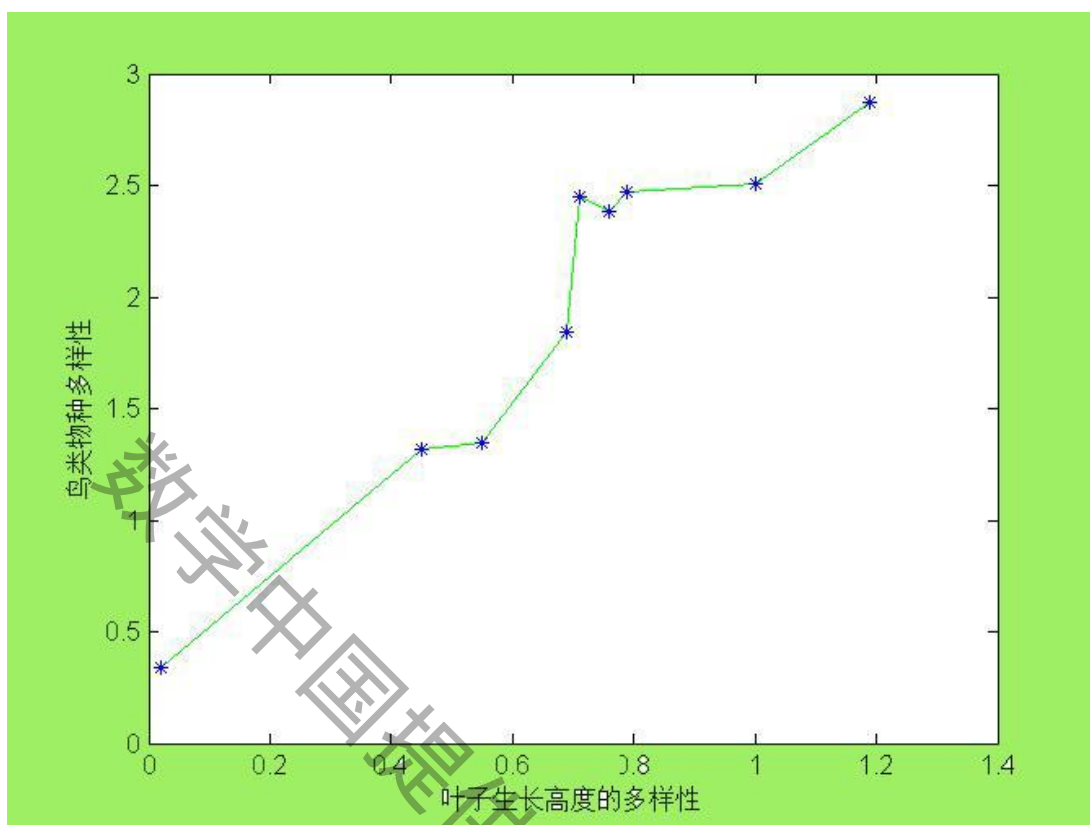


图 5.2

由图中鸟类物种多样性与叶子高度多样性，我们不难得出结论：生物的物种多样性与种群间多样性互相影响因子呈正相关。

5.3 生态系统（环境）多样性对整体指标的影响分析

生态系统是生物多样性存在的基础，即生物生存的环境条件，故生态系统多样性指标对生物多样性的影响是不言而喻的。根据网上搜集的资料分析对比，它的影响指标最主要包括以下几点：

- (1) 水系分布状况指标 (WSD)
- (2) 水污染指标 (SDI)
- (3) 空气污染指标 (API)
- (4) 土壤侵蚀程度 (PP)
- (5) 生物生产性土地面积/生产者提供能量 (BP)

我们可以通过实际数据的整理分析，来得出以上几点对生物多样性的影响。

例一：SDI 值指在一定流域内的水污染指数，下表是某河流中的物种数量与 SDI 值的关系表，共记录了 8 组实际数据

数据组	1	2	3	4	5	6	7	8
SDI 值	1.4	1.8	1.9	2.5	3.4	3.9	4.0	5.0
河流中的	231	210	201	192	191	160	91	42

物种数量								
------	--	--	--	--	--	--	--	--

表 5.3

再由此表我们可以画出河流中物种数量随水质 SDI 值的变化曲线（使用 MATLAB）

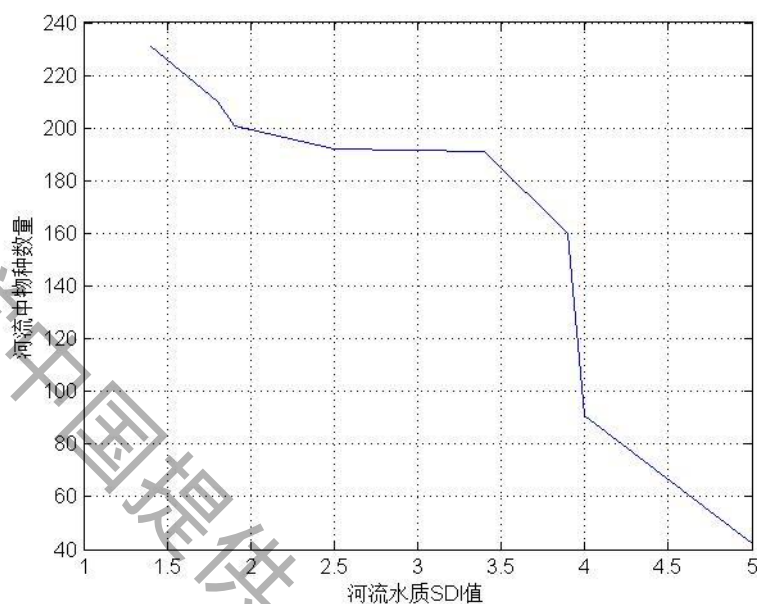
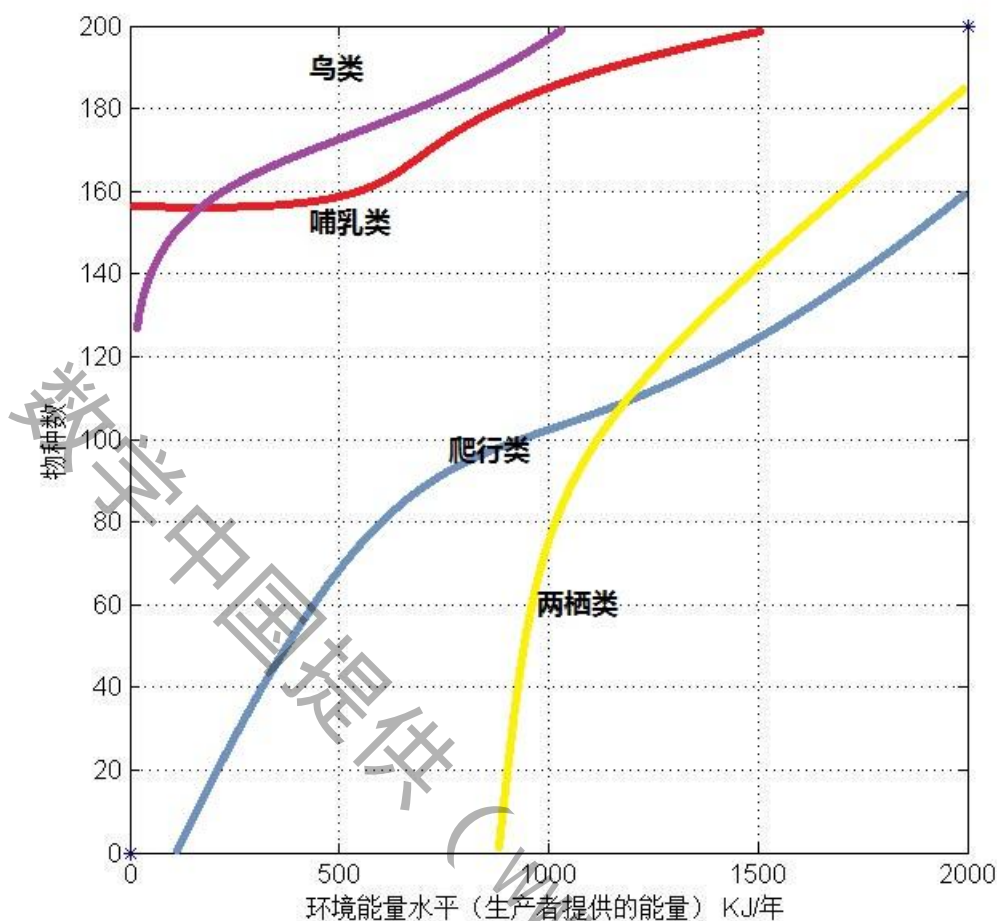


图 5.3.1

显而易见, 河流中的物种数量随着 SDI 值的增加而减少, 由此我们可以得出结论: SDI 值的增加, 会使生物多样性的指标下降。

例二：在生态系统中，生产者起到提供能量的作用，所有消费者和分解者都需要依靠生产者才能存活，所以生产者提供能量能力的高低必然是我们研究生物多样性时的一个指标。BP 值指单位面积内能够提供能量的生产者所占的面积，也称为生物生产性土地面积。

根据 1991 年的这个表格，我们可以看出物种数量与环境提供的能量之间的关系



引自D.J.Currie,Am,Nat.137:29-49(1991)

图 5.3.2

通过上图我们不难看出，当环境能量水平递增的时候，鸟类、爬行类、哺乳类、两栖类动物的物种数量都呈递增趋势，因此，BP 值的增加会引起生物多样性指数的增加。

综合以上两例和类比关系不难看出，以上所举的环境因子对生物多样性的影响力很大，其中 SDI、API、PP 值的增加会导致生物多样性指标的下降，BP、WSD 值的增加会带动生物多样性指标的上升。

5.4 对影响生物多样性因子的主成份分析

影响生物多样性的因子是多种多样的，我们在建立模型时应如何选区相关因子作为模型的参数呢？对此，我们可以对物种多样性影响因子、遗传多样性影响因子以及生态系统(环境)影响因子依次进行主成份分析，从繁多的数据中筛选出所占权重很大的影响因子。

遗传多样性影响因子的筛选就是一个很好的例子，根据查阅书籍和资料可得，影响遗传多样性的因子有很多，其中主要包括：总遗传多样性、性状距离、表形多样性参数、基因谱带频率、多态位点百分率、杂合度、一致度、遗传复杂度、遗传相似性系数、基因分化值等等。在生物学上，我们可以用遗传距离 D 来描述遗传多样性的评价指标，当 D 非常大时，遗传的相似性很低，我们就可以说该地区遗传多样性很高。

由查阅到数据表中剪取其中一部分如下：

	总遗传多样性	等位基因频率方差	基因分化系数	遗传距离
地区 1	0.4972	1.0096	0.0084	0.9180
地区 2	0.0247	14.0754	0.1701	0.2209
地区 3	0.0594	6.1011	0.4032	0.3801

表 5.4

设贡献率向量 $C=[x,y,z]^T$

由以上数据粗略可得

$$\begin{bmatrix} 0.4971 & 1.0096 & 0.0084 \\ 0.0247 & 14.0754 & 0.1701 \\ 0.0594 & 6.1011 & 0.4032 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0.9180 \\ 0.2209 \\ 0.3801 \end{bmatrix}$$

解得： $C=[1.8216, 0.0120, 0.0410]^T$

归一化可知，总遗传多样性的贡献率为 97.2%，等位基因频率方差的贡献率为 0.6%，基因分化系数的贡献率为 2.2%，显而易见的是，第一个主成份，即总遗传多样性的贡献率高达 97.2%，故用总遗传多样性作为影响遗传多样性的因子已经足够了，基因分化系数和基因频率方差的影响能力可以忽略不计。

同理，运用主程序分析法，我们可以提炼出每个大类中的主要影响因子，来组成本数学模型的参变量，经过详细的计算分析，每一个影响因子都是比较重要的，可以假设认为他们具有某一方面的代表性，来参与最后数值的计算，提高了最终求得评价指标的可靠性。

5.5 聚类分析法

在研究各个主要影响因子对生物多样性影响时，即使会遇到数据繁多、整合困难、操作效率不高等问题，但是我们可以灵活地应用聚类分析法，对相似因子进行整合操作，把它们规整为一个统一的因子。具体的操作过程就是，我们首先可以直观分析一些相似因子的影响效果，相互对比，寻找他们之间的相关性；然后统计数据，找出他们之间的实际权重配比；最后，把整合后的新影响因子带回到原来的分析体系中，参与对整体的权值分配计算。

类平均法聚类是聚类分析法的一种方法，通过这种方法，可以改进环境影响因子的结构。首先，从查阅到的统计表中截取出一部分如下：

	空气污染指标	水污染指标	土壤侵蚀程度	亚马逊某流域的物种数量
数据组 1	31	3.4	1.3	19324
数据组 2	64	1.7	1.2	20232
数据组 3	17	3.2	2.5	18723
数据组 4	33	1.9	2.2	21306

表 5.5

直观分析，虽然 4 组数据中的各项污染指标有所差别，但是该组数据中的物种数量却相差不大。因此，通过数据分析可以得出结论，空气污染指标、水污染指标以及土壤侵蚀程度对生物多样性的影响能力基本相同，故我们可以用类平均法聚类，把以上三个因子聚合为综合污染因子，综合影响因子的权值可以用如下公式得出：

$$W_T = (W_1 + W_2 + W_3) \div 3$$

(5.12)

同理，聚类分析法还可以运用到其他的地方，比如物种数量因子是由动物数量、植物数量因子聚合而成等。通过这种聚类整合，可以极大地简化后续运算，使带入计算模型的数据更加精炼可靠。此外，还可以实用最短距离法，最长距离法，重心法来实现聚类分析。

注：类平均法聚类图如下，其他聚类法的图与此大体相同。



图 5.5

5.6 层次分析法

5.6.1 构建层次分析模型

(1) 建立层次结构模型

在研究生物多样性指标的过程中需要建立以下层次结构模型，如图 5.6.1 所示：

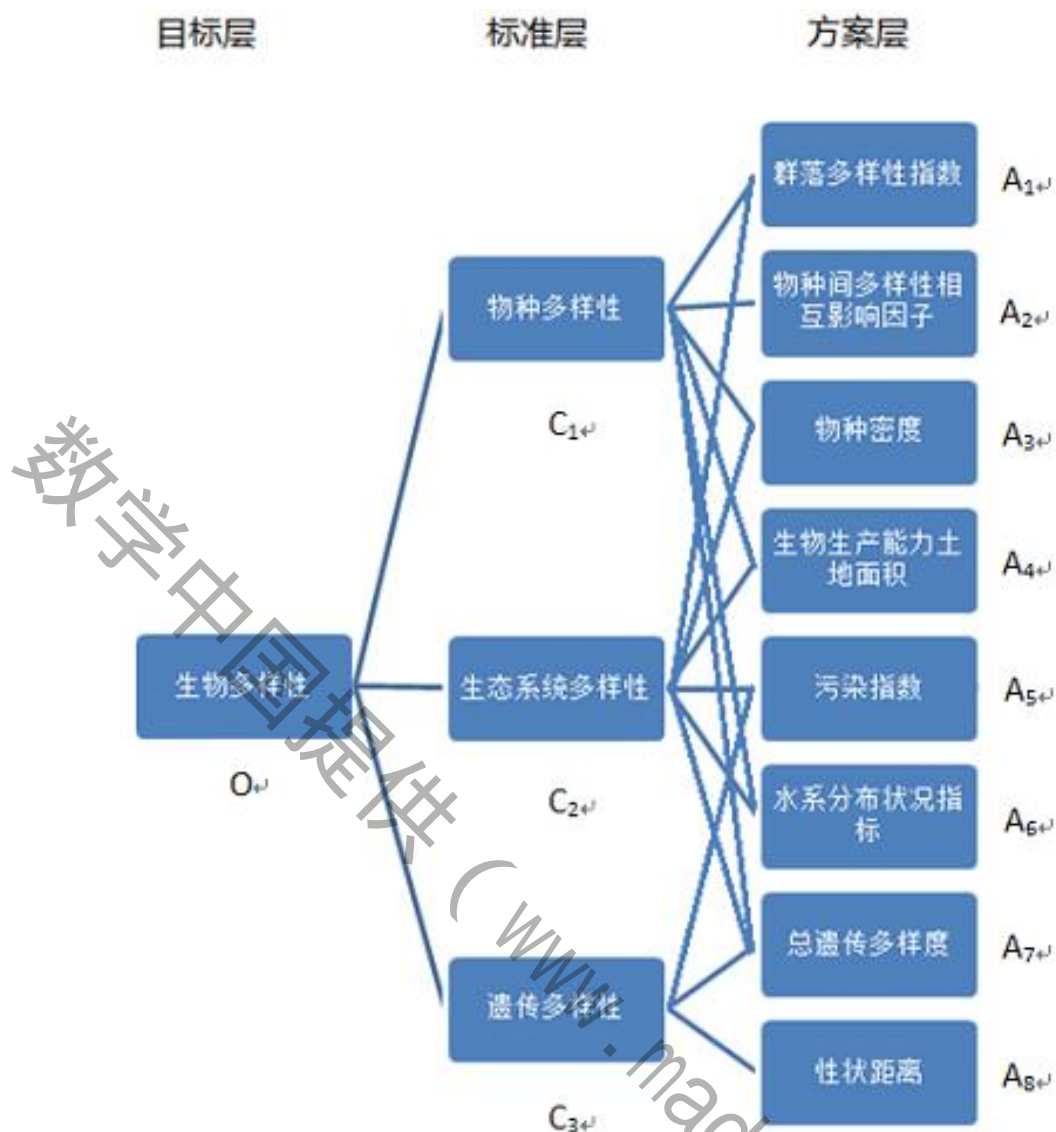


图 5.6.1 生物多样性影响因素层次结构图

(2) 构造成对比较阵

假设要比较某一层 n 个元素 C_1, C_2, \dots, C_n 对上层一个因素 O 的影响，每次抽取两个元素进行相互比较，评定标准见附录 1 构建成对比较矩阵：

$$A = (a_{ij})_{n \times n}, a_{ij} > 0, a_{ji} = \frac{1}{a_{ij}} \quad (5.13)$$

根据上面给出的衡量方法，目标层与标准层判断矩阵建立 O ，标准层与方案层三个判断矩阵 A_1, A_2, A_3 。如下

$$O = \begin{bmatrix} 1 & 1/5 & 1/3 \\ 5 & 1 & 2 \\ 3 & 1/2 & 1 \end{bmatrix}$$

$$A_1 = \begin{bmatrix} 1 & 1 & 4 & 6 & 6 & 5 \\ 1 & 1 & 3 & 5 & 5 & 4 \\ 1/4 & 1/3 & 1 & 3 & 3 & 2 \\ 1/6 & 1/5 & 1/3 & 1 & 1 & 1/2 \\ 1/6 & 1/5 & 1/3 & 1 & 1 & 1/2 \\ 1/5 & 1/4 & 1/2 & 2 & 2 & 1 \end{bmatrix}$$

$$A_2 = \begin{bmatrix} 1 & 1 & 1/2 & 1/2 & 1/3 & 3 \\ 1 & 1 & 1/2 & 1/2 & 1/3 & 3 \\ 2 & 2 & 1 & 1 & 1/2 & 4 \\ 2 & 2 & 1 & 1 & 1/2 & 4 \\ 3 & 3 & 2 & 2 & 1 & 5 \\ 1/3 & 1/3 & 1/4 & 1/4 & 1/5 & 1 \end{bmatrix}$$

$$A_3 = \begin{bmatrix} 1 & 1/5 & 1/3 \\ 5 & 1 & 2 \\ 3 & 1/2 & 1 \end{bmatrix}$$

5.6.2 层次分析模型的求解及一致性检验

根据一致性矩阵的性质，且由 $|A - \lambda I| = 0$ ，求得非零最大特征根 $\lambda_0 = 3.003$

$\lambda_1 = 6.094$ ， $\lambda_2 = 6.060$ ， $\lambda_3 = 3.003$ 。

再根据 $Aw = \lambda w$ 及归一化法则，求出 $\lambda_0, \lambda_1, \lambda_2, \lambda_3$ 所对应的归一化特征向量：

$$\omega_0 = [0.648, 0.230, 0.112]^T$$

$$\omega_1 = [0.369, 0.317, 0.132, 0.052, 0.052, 0.084]^T$$

$$\omega_2 = [0.117, 0.117, 0.196, 0.196, 0.335, 0.048]^T$$

$$\omega_3 = [0.110, 0.582, 0.309]^T$$

一致性检验：根据一致性指标(CI)公式：

$$CI = \frac{\lambda - n}{n - 1}, \quad (5.14)$$

求出 A_0, A_1, A_2, A_3 对应的一致性指标 $CI_0 = 0.002$, $CI_1 = 0.018$, $CI_2 = 0.012$, $CI_3 = 0.002$ ，再根据随机一致性比率公式：

$$CR = \frac{CI}{RI} \quad (5.15)$$

求得与之对应的 $CR_0=0.004$, $CR_1=0.015$, $CR_2=0.010$, $CR_3=0.003$,

其中 RI 表示随机一致性指标，见附录 2。

由于 CR_i 均小于 0.1，认为各矩阵的不一致程度在允许范围之内，可以用其归一化特征向量作为权向量。

再根据总排序权重法则，求得权重分配列向量 W ：

$$W = \begin{bmatrix} 0.369 & 0 & 0.117 \\ 0.371 & 0 & 0.117 \\ 0.132 & 0 & 0 \\ 0.052 & 0 & 0.196 \\ 0.052 & 0 & 0.196 \\ 0 & 0.110 & 0.335 \\ 0.084 & 0.582 & 0.048 \\ 0 & 0.309 & 0 \end{bmatrix} \begin{bmatrix} 0.648 \\ 0.230 \\ 0.122 \end{bmatrix} = \begin{bmatrix} 0.253 \\ 0.255 \\ 0.086 \\ 0.058 \\ 0.058 \\ 0.066 \\ 0.194 \\ 0.071 \end{bmatrix}$$

*注：以上求解过程可用 Matlab 软件实现，语句见附录 3

于是，各个影响因子占总生物多样性指数的比重为：群落多样性指数 25.3%，物种间多样性相互影响因子 8.6%，物种密度 25.5%，生物生产性土地面积 5.8%，污染指数 6.6%，水系分布情况指标 5.8%，总遗传多样性 19.4%，性状距离 7.1%
如图 5.6.2 所示

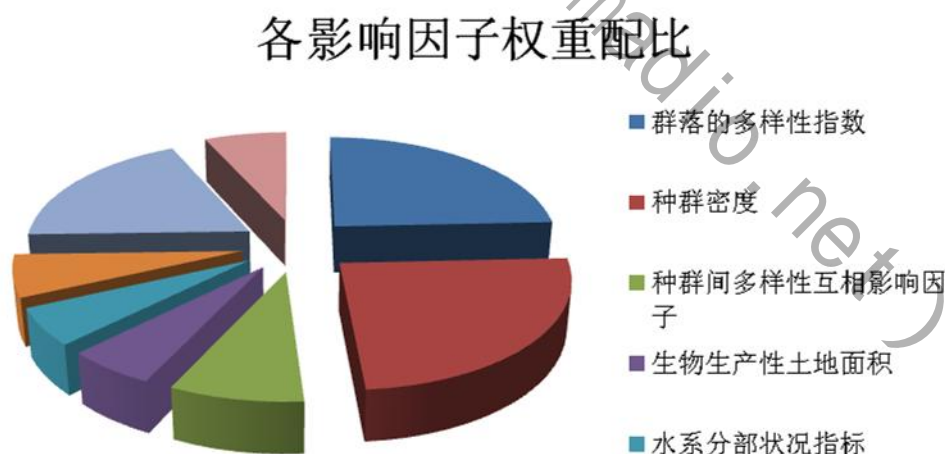


图 5.6.2

5.7 归一化模型的建立

我们在上一部求的一组数据所对应的权重之后，我们需要通过输入数据和所求权重相乘得出最后的生物多样性指标 SS。但是输入数据是毫无规律的，他们的数值范围的波

动性很大，为了使 SS 的值更具有规律性和可比性，我们需要通过设计一种数据归一化算法来实现对输入数据的标准化。

我们的目的是把各项输入数据都化为 0-1 之间的标准小数，因此，我们需要考察每一个影响因子的取值范围，例如：空气污染指数 (API) 的取值范围大约在 0—400 之间，超过 300 的即为严重污染，因此。我们需要把原始的 API 数据乘以 0.002 就可以使 0-400 之间的数转化为 0-1 之间的数，从而实现对输入数据的标准化。

归一化算法如下：

$$W_{ij} = \begin{cases} 0 & , \text{当 } X = X_{\min} \text{ 时} \\ (X - X_{\min}) / (X_{\max} - X_{\min}) & , \text{当 } X \geq X_{\min} \text{ 时} \end{cases} \quad (5.16)$$

由已知数据分析可得所有影响因子的取值范围， X_{\min} 均为 0，故我们可以定义一个标准化向量 SV 与原始输入数据数值向量 OD 做积便可以得到标准化后的数值向量 DV，即：

$$\begin{aligned} SV &= [x, y, z, \dots] \\ DV &= SV \bullet OD \end{aligned} \quad (5.17)$$

算的的标准化后向量 DV 便可以代入下一部中参与运算。

经过数值的严格考证和统计分析，我们可以得出 SV 的一个泛型：

$$SV = [0.0001, 1, 1, 0.002, 1, 0.003, 0.2, 0.25, 1, 1]$$

(注：其中 X 为种群密度乘以项，该项无上限，需要特殊讨论，详细讨论见附录)

5.6 数据汇总，得出生物多样性指标 (说明指标数值的含义)

综合以上求得的所有数据，我们可以通过下边的公式来计算生物多样性指标的评分 SS：

$$SS = \sum_{i=1}^n w_i SD_i \quad (5.18)$$

SS 的值在 0-1 之间，当 SS 的值趋近于 1 时，说明该区域内的生物多样性程度评价很高；当 SS 的值趋近于 0 时，说明该区域内的生物多样性遭到了严重地破坏，生态很不稳定。

通过我们建立的 SS 标准化数值，我们可以清晰地分析、比较、研究、评价一定区域内的生物多样性，为我们以后的研究学习带来了很大方便。

*注：我们可以用 C++ 程序来实现数据分析、得出结论的全过程，源代码见附录 4

六 模型的验证及方法评估

6.1 模型的验证

经过长年大量实际测量评估，我们可以得出本模型主要涉及的十个参变量在方圆十公里内的计数范围，如表 6.1。其中经过层次分析法 API、SDI、PP 的污染性指标影响的权重列向量为： $W_1 = [0.333, 0.333, 0.333]^T$

表形性状距离 d	0-10000
总遗传多样性 H_T	0-1
群落的多样性指数 H	0-1
种群密度 D	0-500
种群间多样性互相影响因子 EI	0-1
空气污染指数 API	0-300
水污染指数 SDI	0-5
土壤污染指数 PP	0-4
水系分部状况指标 WSD	0-1
生物生产性土地面积 BP	0-1

表 6.1 单位样本下的参变量范围表

例如我们获得亚马逊流域（A）与西伯利亚高原（B）的单位样本容量内各项系数指标测评归一后数据，如表 6.2。

	d	H_T	H	D	EI	API	SDI	PP	WSD	BP
A	0.730	0.828	0.847	0.76	0.876	0.238	0.132	0.211	0.867	0.762
B	0.165	0.423	0.443	0.43	0.213	0.013	0.045	0.131	0.324	0.323

表 6.2 A B 区域的具体数值

*注：在数据 API、SDI、PP 的处理中注意到与生物多样性的负相关性，以其峰值为标准，作差求解。

再由权重配比与分配原则，求得 $SS_A=86.36$ 、 $SS_B=43.32$ ，从而得出结论：亚马逊流域的生物多样性远远高于西伯利亚高原。

6.2 主观评价法，聚类分析法以及层次分析法的关系及评估

本文所运用的三种分析方法（主成分分析法、聚类分析法、层次分析法）之间既存在铺垫关系，也存在因果关系。他们之间互相铺垫，互相影响，最终才能得出权重配比关系。

首先，可以说主成分分析法和聚类分析法与层次分析法求权重存在铺垫关系。因为用层次分析法来求比大量的数据的可行性很低，所以，首先我们应该在层次分析法中的方案层中，运用主成分分析法来求得各个相关因子的贡献度，再筛选贡献度比重很大的影响因子，用他们来反映主体。如果经过主成分分析法筛选后的因子仍然很繁杂，我们就可以通过数据对比和日常的经验，运用聚类分析法，将相似程度较大的因子和影响主

能力相同的因子归结为一个影响因子，再以组合因子的形式返回参与方案层的层次分析。最后，经由层次分析法，就可以在避免非常复杂的杂运算和不必要的运算的基础上，实现权值配比。

其次，这三种方法之间也存在着互相影响关系。主成分分析法是通过贡献度来筛选影响因子，而聚类分析法是通过相似性来筛选影响因子。因此，如果把贡献度筛选的标准定得过高，一些相似性的因素就会被集体淘汰，造成影响因子过于单一的误差；相反地，如果把贡献度筛选的标准定得过低，一些相似因子就会被集体选入，也就是无法实现最优化筛选因子的目的。层次分析法中，方案层中因子两两比较而形成的数据，也会受到未聚合的相似因子干扰而浪费了许多计算时间，也可能因为主成分的筛选错误，导致很大的偏差。

通过对三种分析法关系的论述，我们应灵活运用三种分析法针对不同的影响因子组成形式。例如，当相似因子很多的时候，我们就应多运用聚类分析法来简化因子，当影响因子贡献度比较分明的时候，我们就应多运用主成分分析法筛选因子。如何正确灵活运用分析法，将会是本模型完善的重要方向。

七 模型的优缺点评价

优点：

1. 本模型运用了层次分析法，易于将定性分析预定量分析有机统一，使模型的应用范围更加广泛；
2. 本模型运用的模型处理方法不要求复杂的数学基本知识，使读者相对容易理解、接受；从计算角度也非常简便，使过程更加直观。

缺点：

1. 模型有些过于理想化，数据采集比较困难；
2. 只能从已有的方案中择优选择，不能生成更加合理的新方案；
3. 形成比较矩阵的过程较为粗糙，主观因素影响较大，其决策结果往往令人难以彻底信服；
4. 限于所学专业的差异性，考虑相关问题时就难免会有局限性；
5. 资源获取周期过短，不能保证模型中部分较据资料的时效性，究竟是否适合当下课题，还亟待考证。

改进：

1. 还可以考虑其他的影响因子，根据其他因子的影响建立其他配比权重的方法；
2. 静态分析过多，可以考虑增加动态因子；
3. 归一化的过程可以将更加准确的函数关系代替原有的简单的线性处理；
4. 形成比较矩阵的过程可通过权威专家的意见与权威资料改善，以增加该模型的准确性与可信性；
5. 运用层次分析法时，可将标准层中的生态系统多样性与物种多样性合并为一类，统称为“宏观决定因子”，将基因的多样性成为“微观决定因子”。这样，对标准层中的

元素进行归类，可以使结构更加清晰简洁，更能简化之后的计算过程。

八、参考文献

- [1] 周永红, 丁春邦, 《普通生物学》, 高等教育出版社, 2007 年.
- [2] 顾万春, 《统计遗传学》, 科学出版社, 2004 年
- [3] Robert E. Ricklefs, 《The Economy Of Nature》, 高等教育出版社, 2004 年
- [4] 毛文永, 《生态环境影响评价概论 (修订版)》, 中国环境科学出版社, 2003 年
- [5] 程子峰, 徐富春, 《环境数据统计分析基础》, 化学工业出版社, 2006 年
- [6] 邬学军, 周凯, 宋军全, 《数学建模竞赛辅导教程》, 浙江大学出版社, 2009 年
- [7] 李佐锋, 《数学建模》, 东北师范大学出版社, 2005 年
- [8] Stanley B. Lippman, Josee Lajoie, Barbara E. Moo, 《C++ Primer》, 人民邮电出版社, 2006 年
- [9] 艾冬梅, 李艳晴, 张丽静, 刘琳, 《MATLAB 与数学实验》, 机械工业出版社, 2010 年

附录：

附录 1 评价标准规定

尺度	含义
1	二者影响相同
3	影响稍强
5	影响强
7	影响明显得强
9	影响绝对强
2, 4, 6, 8	影响之比在上述两个相邻等级之间
1, 1/2, ..., 1/9	影响之比是上面的相反数

附录 2 随机一致性指标

n	1	2	3	4	5	6	7	8	9	10	11
RI	0	0	0.58	0.9	1.12	1.24	1.32	1.41	1.45	1.49	1.51

附录 3 运用 Matlab 实现正互反矩阵归一化特征向量求解

$b = \text{eig}(A)$

```
[V,D]=eig(A)
CI=(b-n)/(n-1)
CR=CI/RI
If CR<0.1
    Flag=1;
else
    Flag=0
```

附录 4：可以实现模型效果的 C++程序源代码

```
/* Created on: 2011-4-23
/* Author: jinrihao

//Judgement.h

#include <iostream>

#ifndef JUDGEMENT_H_
#define JUDGEMENT_H_

class Judgement //原始导入数据的类
{
public:
    Judgement(istream&);
    //从文件流输入数据
    void Transform_Data();
    //数据归一化
    ~Judgement() {}
private:
    double combinate(Judgement &a)//组合API, SDI, PP的值
    {
        return (1-(a.API+a.SDI+a.PP)/3.0);
    }

    double d;//表面性状距离
    double TGD;//总遗传多样性
    double H;//群落的多样性指数
    double D;//种群密度
    double EI;//种群间多样性互相影响因子
    double API;//空气污染指数
    double SDI;//水污染指数
    double PP;//土壤污染指数
```



```

    double WSD;//水系分部状况指标
    double BP;//生物生产性土地面积
};

#endif

//combine.h

#include<vector>
#include "Judgement.h"
#ifndef COMBINATE_H_
#define COMBINATE_H_
class combination //转化组合后简化的类
{
public:
    combination(judgement
&a):d(a.d), TGD(a.TGD), H(a.H), D(a.D), EI(a.EI), ALL(a.combine()), WSD(a.WSD),
BP(a.BP) {}
    double solve(vector<double> &ivec)//得出最终评比分数
    {
        double temp;

        temp=d*ivec[0]+TGD*ivec[1]+H*ivec[2]+D*ivec[3]+EI*ivec[4]+ALL*ivec[5]+WS
D*ivec[6]+BP*ivec[7];
        return temp;
    }
    ~combination() {}
private:
    double d;//表面性状距离
    double TGD;//总遗传多样性
    double H;//群落的多样性指数
    double D;//种群密度
    double EI;//种群间多样性互相影响因子
    double ALL;//总污染指数
    double WSD;//水系分部状况指标
    double BP;//生物生产性土地面积
};

#endif

//Judgement.cpp

```

```
#include "Judgement.h"
#include <iostream>
#include <vector>
using namespace std;

Judgement::Judgement(istream &oa)
{
    cout<<"正在从文件流中读入数据..."<<endl;
    oa>>d>>TGD>>H>>S>>P>>D>>EI>>API>>SDI>>PP>>WSD>>BP;
}

//从文件流输入数据

void Judgement::Transform_Data(vector<double> &ivec)
//数据归一化
{
    d=d*ivec[0];
    TGD=TGD*ivec[1];
    H=H*ivec[2];
    D=D*ivec[3];
    EI=EI*ivec[4];
    API=API*ivec[5];
    SDI=SDI*ivec[6];
    PP=PP*ivec[7];
    WSD=WSD*ivec[8];
    BP=BP*ivec[9];
}

void Judgement::Output_Score()
//打印返回最终的评定结果
{
    cout<<"表面性状距离:"<<d<<endl;
    cout<<"总遗传多样性:"<<TGD<<endl;
    cout<<"H群落的多样性指数"<<H<<endl;
    cout<<"种群密度:"<<D<<endl;
    cout<<"种群间多样性互相影响因子:"<<EI<<endl;
    cout<<"空气污染指数:"<<dAPI<<endl;
    cout<<"水污染指数:"<< SDI<<endl;
    cout<<"土壤污染指数:"<<PP<<endl;
    cout<<"水系分部状况指标:"<<WSD<<endl;
    cout<<"生物生产性土地面积:"<<BP<<endl;
}
```

```
//main.cpp
```

```
#include <iostream>
#include <vector>
#include <cstdlib>
#include <string>
#include <iomanip>
#include "Judgement.h"
#include "combinate.h"
```

```
using namespace std;
```

```
vector<double> Standard;//存放标准化归一化向量
vector<double> Weight;//存放权重向量
```

```
void Initial_std(vector<double> &ivec)//初始化归一向量
```

```
{
    ivec.push_back(0.0001);
    for(int i=0;i<=1;++i)
        ivec.push_back(1.0);
    ivec.push_back(0.002);
    ivec.push_back(1.0);
    ivec.push_back(0.003);
    ivec.push_back(0.2);
    ivec.push_back(0.25);
    for(int i=0;i<=1;++i)
        ivec.push_back(1.0);
    cout<<"已成功初始化归一向量:"<<endl;
    for(vector<double>::iterator iter=ivec.begin();iter!=ivec.end();++iter)
        cout<<setw(3)<<*iter;
    cout<<endl;
}
```

```
void Initial_Wgt(vector<double> &ivec)//初始化权重向量
```

```
{
    double a[]={0.071,0.194,0.253,0.255,0.086,0.066,0.058,0.058};
    for(int i=0;i<=7;++i)
        ivec.push_back(a[i]);
    cout<<"已成功初始权重向量:"<<endl;
    for(vector<double>::iterator iter=ivec.begin();iter!=ivec.end();++iter)
        cout<<setw(3)<<*iter;
    cout<<endl;
}
}
```

```

int main()    //主函数
{
    Initial_std(Standard);    //初始化
    Initial_Wgt(Weight);
    int count=0;    //用来记录数据的组数
    string file_name;
    cout<<"请输入源数据文件名"<<endl;
    cin>>file_name;
    ifstream Infile(file_name.c_str()); //打开数据文件
    vector<Judgement> it1;
    vector<combination> it2;
    if(!Infile)
    {
        cout<<"Can not open the file!"<<endl;
        abort();
    }
    while(!Infile.eof())
    {
        it1.push_back(Judgement(Infile)); //导入数据到容器
        it1.Output_Score();
        ++count;
    }
    cout<<"共有"<<count<<"组数据"<<endl;
    for(int i=0;i<=count;++i)
        it1[i].Transform_Data(); //归一化数据
    for(int i=0;i<=count;++i)
        it2.push_back(combination(it1)); //组合简化归一化后的数据
    for(int i=0;i<=count;++i)
        cout<<"生物多样性指标的评分(score):"<<it2[i].solve()<<endl;
    // 得出结论
    system("pause");
    return 0;
}

//程序结束

```

附录 5 蒙古栎 8 个群体和辽东栎 1 个群体的等位酶基因图谱

位点 等位基因	蒙古栎								辽东栎
	大兴安岭	嘉荫	双鸭山	蛟河	大青沟	宽甸	雾灵山	赞皇	东灵山

Ata	A	0.133	0.100	0.033	0.000	0.083	0.000	0.032	0.330	0.000
	B	0.750	0.867	0.967	1.000	0.833	0.667	0.683	0.900	0.848
	C	0.117	0.033	0.000	0.000	0.083	0.333	0.000	0.067	0.152
Aco	A	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Amp-1	A	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Amp-2	A	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Dia-1	A	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.985
	B	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.015
Dia-2	A	0.400	0.517	0.500	0.500	0.317	0.400	0.917	0.750	0.228
	B	0.600	0.483	0.500	0.500	0.683	0.600	0.083	0.250	0.712
G6pd-1	A	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
G6pd-2	A	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Gpi-1	A	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Gpi-2	A	0.300	0.300	0.200	0.283	0.150	0.117	0.117	0.250	0.333
	B	0.683	0.683	0.717	0.583	0.683	0.783	0.683	0.600	0.576
	C	0.000	0.000	0.000	0.000	0.000	0.000	0.017	0.000	0.000
	D	0.017	0.017	0.083	0.133	0.167	0.100	0.117	0.150	0.091
	E	0.000	0.000	0.000	0.000	0.000	0.000	0.067	0.000	0.000
Idh-1	A	0.000	0.033	0.033	0.033	0.000	0.000	0.100	0.033	0.091
	B	1.000	0.717	0.850	0.900	0.433	1.000	0.733	0.783	0.742
	C	0.000	0.250	0.017	0.067	0.567	0.000	0.167	0.183	0.167
Idh-2	A	0.117	0.100	0.000	0.000	0.000	0.000	0.033	0.000	0.015
	B	0.883	0.900	0.983	1.000	1.000	1.000	0.967	0.967	0.970
	C	0.000	0.000	0.017	0.000	0.000	0.000	0.000	0.033	0.015
Mdh-1	A	1.000	0.950	0.967	1.000	1.000	1.000	1.000	0.983	0.985
	B	0.000	0.050	0.033	0.000	0.000	0.000	0.000	0.017	0.015
Mdh-2	A	1.000	1.000	0.917	0.983	1.000	0.717	0.933	0.567	1.000
	B	0.000	0.000	0.083	0.017	0.000	0.283	0.067	0.433	0.000
Me	A	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Mnr-1	A	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	B	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.083	0.000
Pgd-1	A	1.000	1.000	1.000	0.983	0.983	1.000	1.000	1.000	0.970
	B	0.000	0.000	0.000	0.017	0.000	0.000	0.000	0.000	0.030
Pgd-2	A	1.000	0.983	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	B	0.000	0.017	0.017	0.000	0.000	0.000	0.000	0.000	0.000
Pgm-1	A	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Pgm-2	A	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Skd	A	0.000	0.017	0.067	0.000	0.000	0.000	0.017	0.050	0.000
	B	1.000	0.983	0.833	0.783	1.000	1.000	0.917	0.817	1.000
	C	0.000	0.000	0.100	0.217	0.000	0.000	0.067	0.133	0.000
等位因子		28	33	33	30	29	27	34	35	33