

环境保护与绿色经济间的建模分析

摘要

本文主要研究经济发展与环境之间的关系，建立起时间序列预测模型对中国未来十年内的森林面积进行预测，并基于此给出了退耕还林政策对国家工业水平和国民经济的所造成的影响模型。然后提出基于熵权法的 Topsis 评价模型对全球国家内经济水平、工业水平和森林植被面积进行评价并排名，最终根据评价结果进行分析并向联合国提出相关提议。

针对问题一：我们首先基于附件的全球水平各方面数据进行筛选和清洗，得到只含有中国单个国家的经济与环境数据。之后本文提出了 **BP 神经网络预测**来对中国未来十年的森林面积进行预测。但可能是由于数据不全面以及数据量过少，导致模型的预测误差均大于 10^5 。我们再次根据数据量较少及数据呈现曲线形式的特点提出**灰色预测模型**，求解结果表明，模型平均相对残差为 **0.0048652**，平均级比偏差为 **0.0012958**，具有较好的拟合程度与预测准确率。最终通过求解模型确定中国未来十年的森林面积，详情见附录 11.2.1。

针对问题二：由于附件中并未给出有关于退耕还林政策的有关数据，因此我们利用 **Python 爬虫**获取了自 2002 年开始实施退耕还林政策以来的退耕还林面积。首先我们利用**皮尔逊相关系数**计算出退耕还林政策与国家工业水平、国民经济之间的相关度，得出退耕还林政策与两者间均具有**弱相关性**。考虑到问题一中的灰色预测模型并不适用于大量数据集的预测，于是我们利用 SPSS 软件对未实施政策前的各项数据构建时间序列预测模型，利用预测值与实际值进行对比得出政策对国家工业水平和国民经济的影响分别为 **+36.63%、+4.49%**。

针对问题三：本文首先基于附件所给数据筛选清理出 2020 年世界各国经济水平、工业水平和森林植被面积相关数据。我们首先使用熵权法求解得到经济水平、工业水平和森林植被面积等三个指标的各项权重分别为：**0.839、0.043、0.118**。接着基于熵权法构建 Topsis 评价模型，我们得到了全球各国的**综合得分 W_i** 并进行具体分析，详情见附录 11.2.3。

最后，我们根据上述中模型及所得出的结论提出了一些有关建议和具体措施，最终以一份报告书的形式在本文中呈现。

关键词：BP 神经网络 灰色预测模型 时间序列预测模型 熵权法 Topsis 模型

一、问题重述

伴随经济的飞速发展，同样的，我们的环境形势也日趋严峻。习近平总书记说：“绿水青山就是金山银山”。生态建设是国家发展不可缺失的条件之一，我们的目标是让天更蓝、树更绿、水更清。我们也要有长远的目光，生态建设与经济发展要齐头并进，二者并非是对立的状态，互相影响互相促进，为了让我们的生态与经济的协调发展，对于环境保护与经济发展的协调能力做出了相关的探索。

本文就环境保护问题的解决，提出一下问题：

1. 根据附件中给定的数据，预测中国在 2025-2035 年内的森林面积。
2. 通过分析中国的有关数据并自行查找需要数据，构建数学模型分析退耕还林政策对国家工业水平和国民经济的影响。
3. 根据附件中给出的国家数据构建模型对世界各国经济水平、工业水平和森林植被面积进行评价，并得出结论。
4. 拟向联合国递交一封报告书，书写一份报告书来描述我们的发现，并提出一些具体措施。

二、问题分析

- 针对问题一：

问题一要求我们根据附录中的各项数据预测出中国在 2025 2035 年内的森林面积。考虑到附录中所给的数据为全球各国国家的耗电量（人均千瓦时）、能源使用量（人均千克石油当量）、森林面积（平方公里）、工业增加值（占 GDP 的百分比）、二氧化碳排放量（千吨）、GDP（现价美元）、森林面积（占土地面积的百分比）与农业用地（占土地面积的百分比）的数据。我们首先要在附录所给数据的基础之上整理出只包含中国单个国家的各项数据，并以此建立单变量或多变量预测模型进行求解。

- 针对问题二：

问题二要求我们分析中国退耕还林政策对国家工业水平和国民经济的影响。本题可以首先通过分析退耕还林政策与国家工业水平和国民经济之间的相关度，并在此基础之上建立国家工业水平和国民经济的预测模型。通过比较未实施退耕还林政策与实施状态下的国家工业水平和国民经济水平来具体分析政策的影响。本题的难点在于预测模型的建立，可建立不同的时间预测模型来选择最优拟合与预测值。

- 针对问题三：

问题三要求从附件中全球不同国家数据构建模型对世界各国经济水平、工业水平和

森林植被面积进行评价。首先我们可以以世界各国经济水平、工业水平和森林植被面积作为评价模型的评价指标，并基于此建立 Topsis 评价模型对全球不同国家数据进行评价。其中评价模型的建立是本题的难点，也可以对传统的评价模型进行改进求解。

- 针对问题四：

问题四要求我们拟向联合国递交一份报告书，描述我们的发现，并提出一些具体措施。经过模型建立与每题分析所得结论，我们得到了最终研究结论，同时查询相关文献以及附件或解题过程中使用的数据，撰写、修改并得到问题四的结果。

三、符号说明

符号	意义
$X^{(0)}$	原始数据的非负序列
$X^{(1)}$	原始数据的非负序列的累积生成序列
k	时间点
X	熵权法中由 i 个评价对象与 j 个评价指标构成的评价矩阵
p_{ij}	熵权法中第 j 项指标下第 i 个样本的比重
e_j	熵权法中第 j 项指标的信息熵
d_j	熵权法中的信息效用值
W_j	熵权法中的各个指标的熵权
Z_{ij}	归一化矩阵
X_{ij}	第 i 个评价对象第 j 个评价指标体系的得分情况
Z_j^+	归一化矩阵的最优向量
Z_j^-	归一化矩阵的最劣向量
D_i^+	每级指标与最优向量的欧式距离
D_i^-	每级指标与最劣向量的欧式距离
W_i	与最优值的相对接近程度
p_{ij}	第 j 项指标下第 i 个样本的比重

四、模型假设

1. 假设中国退耕还林政策在未来长时间内还会继续实施。
2. 假设我们从在线数据库收集的数据是准确、可靠和互相一致的。
3. 假设我们所研究的国家是一个整体，并没有考虑到不同国家国内各地区的差异。
4. 假设在模型的提出与验证过程中，我们忽略的国家指标对结果的计算影响不大。

五、问题一的模型建立和求解

5.1 数据预处理

附件中给出了全球不同国家的耗电量（人均千瓦时）、能源使用量（人均千克石油当量）、森林面积（平方公里）、工业增加值（占 GDP 的百分比）、二氧化碳排放量（千吨）、GDP（现价美元）、森林面积（占土地面积的百分比）与农业用地（占土地面积的百分比）的数据。我们使用 Python 中的 Pandas 库对所给数据进行清洗切割与拼接，获得了只含有中国单个国家的上述各项数据。（详情见附录 11.1.1）

5.2 问题一的模型建立

5.2.1 基于 BP 神经网络的森林面积预测模型

BP 神经网络是基于的一种基于 Back-propagation 算法的多层前馈网络，使用梯度下降对人工神经网络进行监督学习，不需要特别提及要学习的功能的特征便能较好的映射输入与输出数据之间的关系。一般来说，BP 神经网络的算法步骤如下：

- 初始化网络及学习参数，如设置网络初始权矩阵、学习因子等；
- 提供训练模式，训练网络直至满足所需的学习要求；
- 前向传播过程：对给定训练模式进行输入，计算网络的输出模式，并与期望模式比较，若有误差，则执行步骤 (4)，否则返回步骤 (2)；
- 反向传播过程：计算同一单元的误差，修正权值和阈值，返回步骤 (2)；根据上述算法步骤，我们将附录 11.1.1 的数据进行拆分，训练集与测试集的比值为 4:1，对整个神经网络进行训练并检验。

具体过程表述如下：

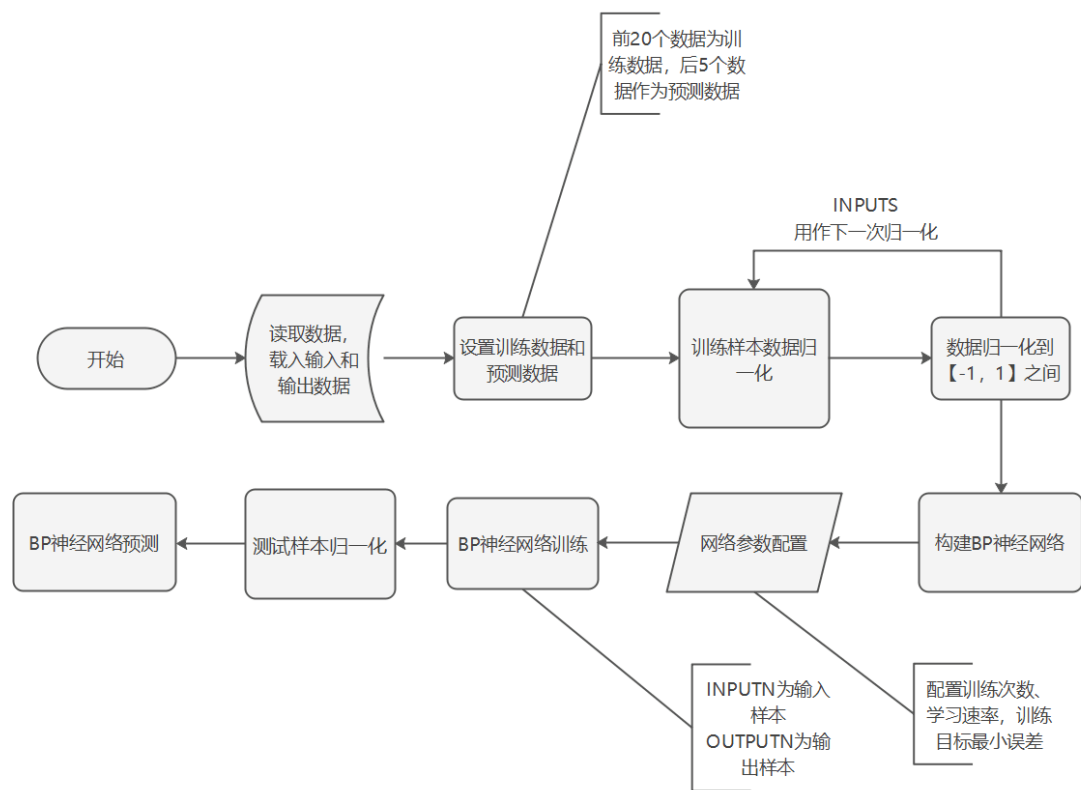


图 1 基于 BP 神经网络的森林面积预测的流程图

(1) BP 神经网络森林面积预测模型结果

预测误差分析：

- 误差平方和 SSE 为：92638478592.0526
- 平均绝对误差 MAE 为：126095.0124
- 均方误差 MSE 为：18527695718.4105
- 均方根误差 RMSE 为：136116.4785
- 平均百分比误差 MAPE 为：6.1346%
- 相关系数 R 为：0.99242

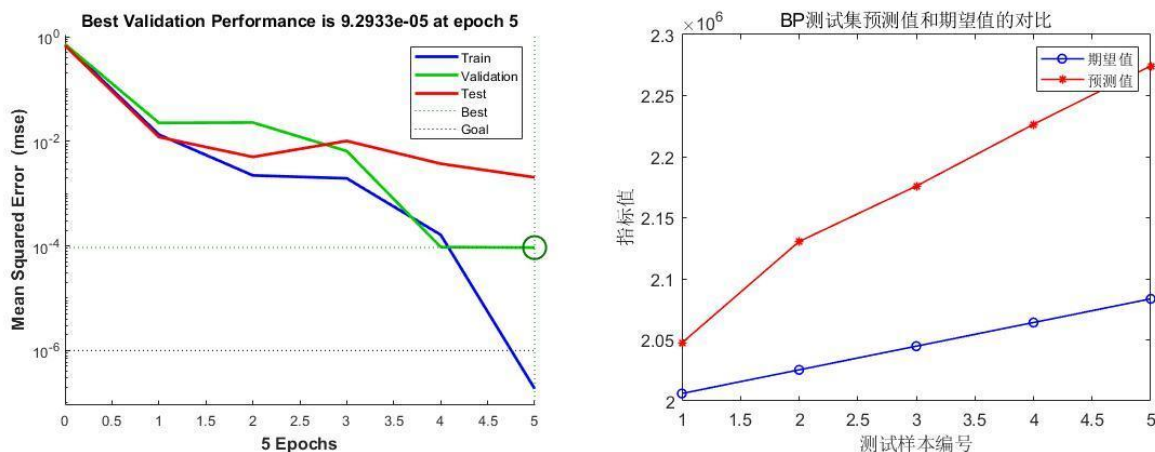


图2 BP神经网络森林面积预测模型结果(1)

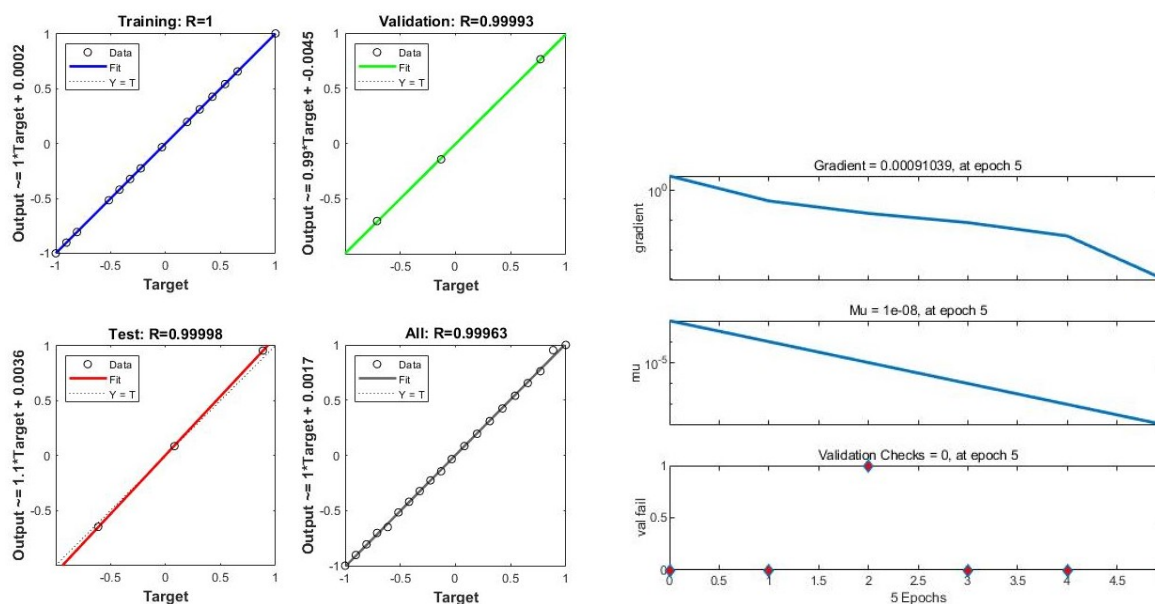


图3 BP神经网络森林面积预测模型结果(2)

(2) **BP神经网络森林面积预测模型分析** 在以上训练生成后的BP神经网络模型中,观察结果图形可得模型中输入结果和期望输出结果二者相对误差均大于1%,预测效果不理想。分析原因,可能是由于数据过少导致神经网络欠拟合所致,故并不采用BP神经网络模型对中国在2025-2035年内的森林面积进行预测。

5.2.2 基于灰色预测的森林面积预测模型

灰色理论是一种计算数据集中各因素之间的关联度并用于有限样本的构建,从而为短期问题提供更好预测优势的方法。灰色预测中系统因素之间具有不确定的关系,一部分系统信息已知,一部分未知。由于附件中所给数据量过少且数据呈现曲线形式,故优

先使用灰色预测法对中国未来森林面积进行预测。灰色模型 GM (1, 1) 是目前使用最为广泛的灰色预测模型之一，此方法至少需要四个观测值。首先对数据应用累加生成算子 (AGO)，然后求解模型的控制微分方程，得到系统的预测值。最后，利用逆累加生成算子 (IAGO) 得到原始数据的预测值。GM(1, 1) 建模过程如下：

$$X^{(0)} = \left\{ x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n) \right\}$$

$X^{(0)}$ 是原始数据的非负序列，其的累积生成序列 (AGO) $X^{(0)}$ 为：

$$X^{(1)} = \left\{ x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n) \right\}$$

$x^{(1)}(1)=x^{(0)}(1)$ ，且

$$x^{(1)}(k) = \sum_{i=1}^k x^{(0)}(i), \quad k = 2, 3, \dots, n$$

连续邻近生成的序列均值 $x^{(1)}$ 为：

$$Z^{(1)} = \left\{ z^{(1)}(1), z^{(1)}(2), \dots, z^{(1)}(n) \right\}$$

$$z^{(1)}(k) = \frac{x^{(1)}(k) + x^{(1)}(k-1)}{2}, \quad k = 2, \dots, n$$

GM (1, 1) 模型的基本形式为：

$$x^{(0)}(k) + ax^{(1)}(k) = b$$

最终的白化方程为：

$$\frac{\partial x^{(1)}}{\partial t} + ax^{(1)} = b$$

求解白化方程，预测值 $x^{(1)}$ 可以计算如下：

$$\hat{x}^{(1)}(k) = \left(x^{(0)}(1) - \frac{b}{a} \right) e^{-a(k-1)} + \frac{b}{a}, \quad k = 2, 3, \dots \quad (1)$$

因此，预测值可以通过以下方式生成，

$$\hat{x}^{(0)}(1) = x^{(0)}(1)$$

$$\hat{x}^{(0)}(k) = \hat{x}^{(1)}(k) - \hat{x}^{(1)}(k-1), \quad k = 2, 3, \dots, n.$$

在等式 (1) 中, k 表示时间点, a, b 为相关参数。方程 (1) 中的参数 a, b 也可以使用最小二乘法估计, 表达式为:

$$[a, b]^T = [B^T B]^{-1} B^T Y$$

其中:

$$B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix}, Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}$$

我们使用附录 11.1.1 中的中国森林面积数据, 构建灰色预测模型, 具体流程如下:

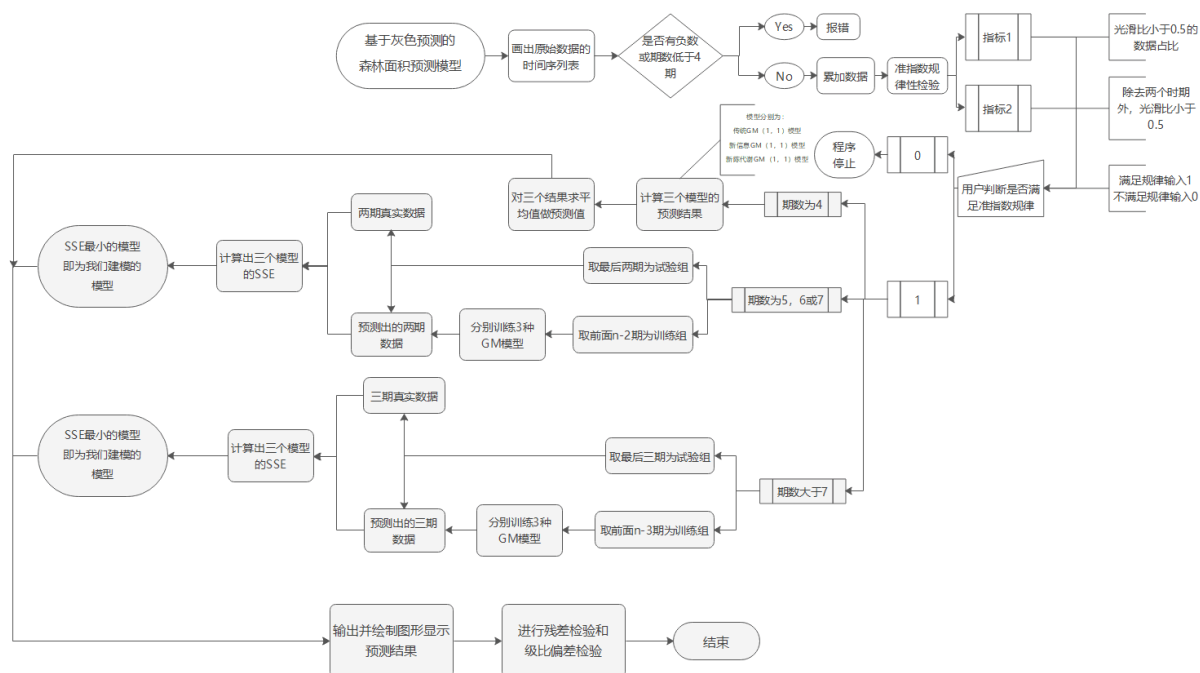


图 4 灰色预测模型的流程图

(1) 灰色预测的森林面积预测模型结果

- 传统 GM(1,1) 对于试验组预测的误差平方和为 3361846443.4237
- 新信息 GM(1,1) 对于试验组预测的误差平方和为 3361219044.7196
- 新陈代谢 GM(1,1) 对于试验组预测的误差平方和为 3243518027.8797

因为新陈代谢 GM(1,1) 模型的误差平方和最小, 所以我们应该选择其进行预测。

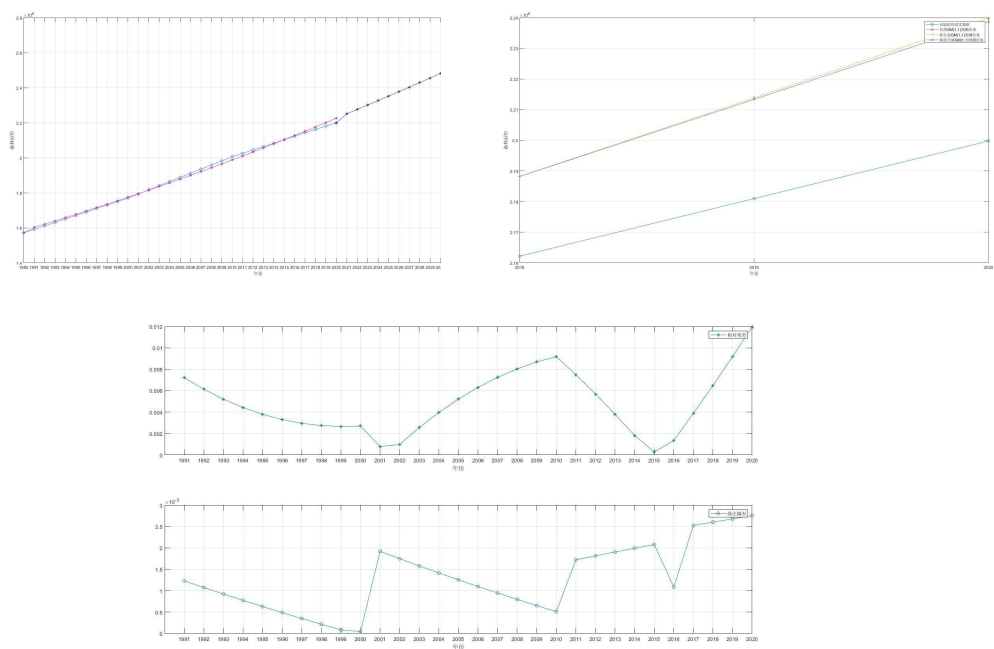


图5 灰色预测的森林面积预测模型效果图

(2)灰色预测的森林面积预测模型分析 经程序运行后得出模型平均相对残差为0.0048652，残差检验的结果表明：该模型对原数据的拟合程度非常不错；平均级比偏差为0.0012958级，比偏差检验的结果表明：该模型对原数据的拟合程度非常不错；从上述的结果图中也可以看出模型具有较好的拟合性，经模型预测后得出 2025-2035 年内的森林面积为：

年份	森林面积（公顷）	年份	森林面积（公顷）
2025	2351971.0019	2031	2508159.6711
2026	2377607.7761	2032	2534859.0553
2027	2403426.7303	2033	2561928.9621
2028	2429412.2762	2034	2589397.9247
2029	2455545.4418	2035	2617295.6867
2030	2481803.5166		

六、问题二的模型建立和求解

6.1 数据的预处理

退耕还林政策从 2002 年开始实施至今，主要两个方面的内容：一是坡耕地退耕还林；二是宜林荒山荒地造林。由于附件中并未给出有关中国退耕还林政策的数据，我们使用 Python 爬虫从国家统计局等官方网站下载了相关数据（数据详情见附录 11.1.2），并且用各个年份的退耕还林的面积大小代表退耕还林政策的影响力度，用每年的工业增加值代表该年单位下的工业水平。考虑到政策在 2012,2013 两年间短暂的暂停实施以及数据收集的不全面性，我们只研究 2003 年至 2017 年中的退耕还林数据。除特别说明外，下文中所提的退耕还林均指在此期间的数据。

6.2 相关度分析

6.2.1 皮尔逊相关系数

皮尔逊相关系数是最常用的两种相关系数之一，它能够反映两个变量之间相关性的强弱。协方差假设变量数据 M 及 N 是样本总体数据，则样本总体均值：

$$E(M) = \frac{\sum_{i=1}^n M_i}{n}$$
$$E(N) = \frac{\sum_{i=1}^n N_i}{n}$$

总体协方差：

$$Cov(M, N) = \frac{\sum_{i=1}^n (M_i - E(M))(N_i - E(N))}{n}$$

两个变量的相关性在一定程度上可以由协方差来反映，当协方差的值大于 0 时，说明两变量呈正相关；当协方差的值小于 0 时，说明两变量呈负相关；当协方差等于 0 时，两变量相互独立。

皮尔逊相关系数的绝对值	对应相关性
0.0-0.2	变量 M 的标准差
0.2-0.4	弱相关性
0.4-0.6	中相关性
0.6-0.8	强相关性
0.8-1.0	极强相关性

6.2.2 结论

我们利用 Python 分析退耕还林政策与国家工业水平、国民经济之间的相关度，并绘制热图，结果如下图所示：

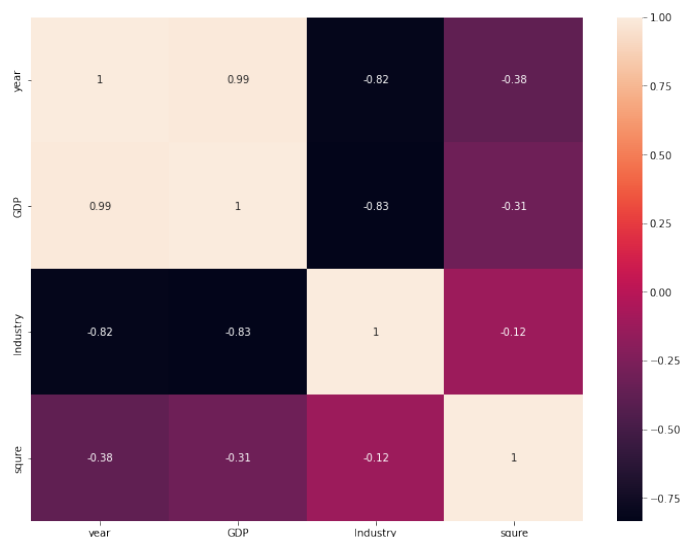


图 6 表示退耕还林政策与国家工业水平、国民经济相关度的热图

根据上述的相关性判断标准以及热图可以看出，退耕还林政策与国家工业水平之间具有微弱相关性，与国民经济（GDP）具有弱相关性。

6.3 退耕还林政策的影响分析

考虑到 2003 年退耕还林政策才开始实施，通过对 1960 年至 2002 年间中国的国家经济与工业水平数据进行拟合，并对 2003 年政策开始实施后未来九年的数据进行预测，在此基础上定量的研究并分析退耕还林政策对国家经济与工业水平的影响。

6.3.1 时间序列预测模型的建立

我们首先利用问题一中所使用的灰色预测模型对数据进行拟合分析，结果如下图所示：

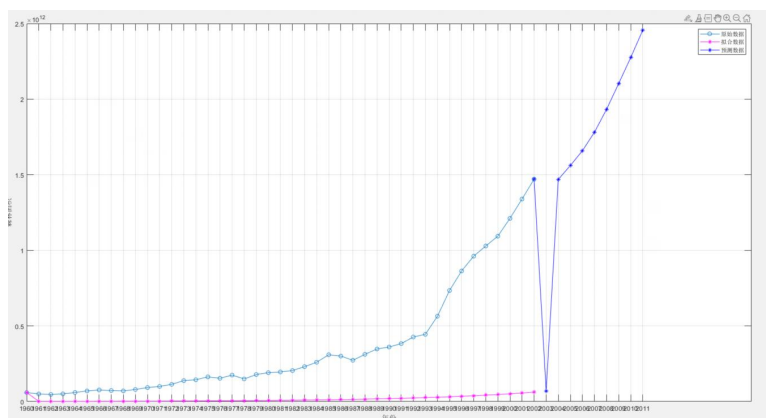


图 7 利用灰色预测模型得到的数据拟合图

从图中可以看出当数据集较大时，灰色预测模型拟合优度较差具有较大的误差，并不适用于预测。再次利用 SPSS 软件时间序列预测功能对附录 11.1.2 中的数据进行时间序列预测，预测结果如下图：

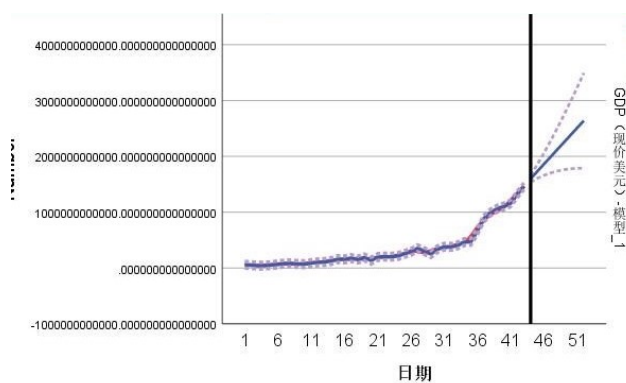


图 8 SPSS 对时间序列的预测结果 (1)

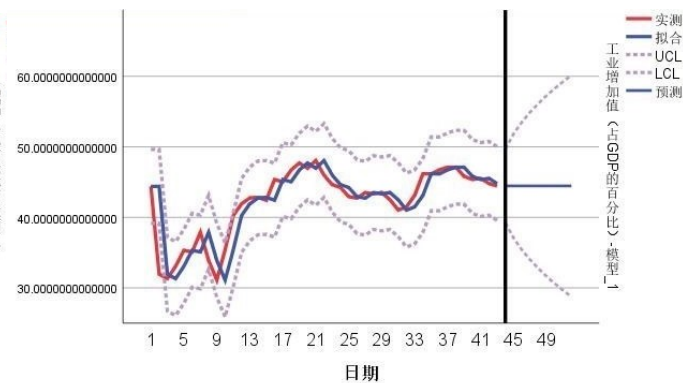


图 9 SPSS 对时间序列的预测结果 (2)

可以看出，模型对数据具有较好的拟合度，具有较好的代表性。

6.3.2 结论

经过时间序列模型预测的数据详情见附录 11.2.2。结果计算后的影响值如下表所示：

预测值	实际值	影响	预测值	实际值	影响
1.6E+12	1.66E+12	-0.03598	44.45058	45.62278	-0.02569
1.73E+12	1.96E+12	-0.11493	44.45058	45.90021	-0.03158
1.86E+12	2.29E+12	-0.18604	44.45058	47.02258	-0.0547
1.99E+12	2.75E+12	-0.27665	44.45058	47.5574	-0.06533
2.12E+12	3.55E+12	-0.40264	44.45058	46.88417	-0.05191
2.25E+12	4.59E+12	-0.51007	44.45058	46.97117	-0.05366
2.38E+12	5.1E+12	-0.5333	44.45058	45.95715	-0.03278
2.51E+12	6.09E+12	-0.58748	44.45058	46.49783	-0.04403
2.64E+12	7.55E+12	-0.65025	44.45058	46.52929	-0.04468
		-0.36637			-0.04493

可以看出退耕还林政策对国内经济具有较大作用的影响，会直接或间接的促进国内经济的发展，具有 36.63% 的促进水平。其原因在于：新一轮退耕还林每亩退耕地补助 1200 元，五年内分三次下达，释放了自由劳动力，同时也为产业的发展提供了原始积累，有利于社会经济的发展。

而对于工业水平的发展，退耕还林政策也有一定的促进的作用，但促进作用不大，仅有 4.49%。其原因在于：我国工业的生产，原料大多来自于自然资源，退耕还林政策的下发，不利于工业原料的开采，但由于技术的升级，相对来说略有发展。

七、问题三的模型建立与求解

7.1 数据的预处理

考虑到年份的不同会对结果产生不同程度的影响，我们利用 Python 中的 pandas 库选取附件中给出的 2020 年的国家、世界各国经济水平、工业水平和森林植被面积数据进行清洗，对空白值进行切除与拼接，获得附录 11.1.3 中的数据。

7.2 基于熵权法的 Topsis 评价模型

7.2.1 基于熵权法的 Topsis 评价模型背景

传统 TOPSIS 法（逼近理想排序法）通过多目标决策的方式找出有限方案中的最优与最劣方案，其中的最优解满足靠近最优方案的同时远离最劣方案。传统 TOPSIS 法中各个指标所占的权重的确定过程中会带有很强的主观性，而熵权法的使用便能很好的解决人为主观性的问题。

7.2.2 熵权法确定指标权重

对于 i 个评价对象与 j 个评价指标构成的评价矩阵 X ：

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1m} \\ x_{21} & x_{22} & \dots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nm} \end{bmatrix}$$

对数据进行标准化：

$$\tilde{y}_{ij} = \frac{x_{ij} - \min \{x_{1j}, x_{2j}, \dots, x_{nj}\}}{\max \{x_{1j}, x_{2j}, \dots, x_{nj}\} - \min \{x_{1j}, x_{2j}, \dots, x_{nj}\}}$$

计算第 j 项指标下第 i 个样本中的比重：

$$p_{ij} = \frac{\tilde{y}_{ij}}{\sum_{i=1}^n \tilde{y}_{ij}}$$

计算第 j 项指标的信息熵：

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

计算信息效用值：

$$d_j = 1 - e_j$$

最终得到各个指标的熵权：

$$W_j = d_j / \sum_{j=1}^m d_j (j = 1, 2, \dots, m)$$

通过 Python 软件进行熵权法确定数据中指标权重，结果如下表所示：

指标	权重
世界各国经济水平	0.8392336225370627
世界各国工业水平	0.04317794141632378
世界各国森林植被面积	0.11758843604661351

7.2.3 熵权法修正的 TOPSIS 评价模型步骤

Step1: 建立归一化矩阵, 将原始矩阵进行标准化, 即

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

其中 i 表示选取的国家个数, j 表示所选的指标体系个数, X_{ij} 表示第 i 个评价对象第 j 个评价指标体系的得分情况。

Step2: 由上述得到的 Z 矩阵, 进而得到最优向量与最劣向量:

$$Z_j^+ = \max_{1 \leq i \leq m} |Z_{ij}|$$

$$Z_j^- = \min_{1 \leq i \leq m} |Z_{ij}|$$

Step3: 计算所选取的评价国家的指标与最优向量的欧式距离和最劣向量的欧式距离:

$$D_i^+ = \sqrt{\sum_{j=1}^n (Z_{ij} - Z_j^+)^2}$$

$$D_i^- = \sqrt{\sum_{j=1}^n (Z_{ij} - Z_j^-)^2}$$

Step4: 最后得到与最优值的相对接近程度:

$$W_i = \frac{D_i^-}{D_i^+ + D_i^-}$$

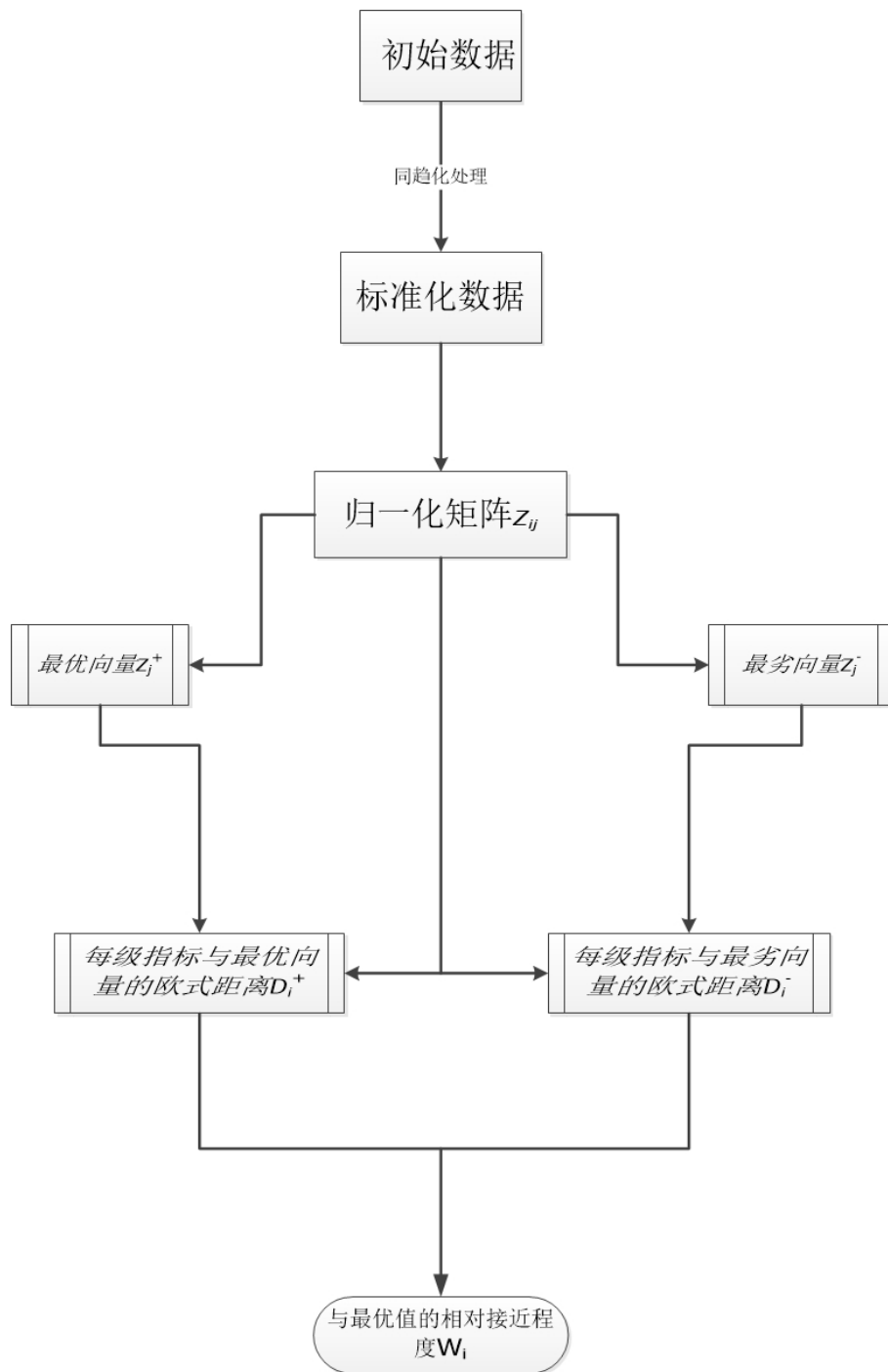


图 11 TOPSIS 算法流程

对于综合得分 W_i 的解释：根据问题三中指标体系，对附录 11.1.3 中的数据进行正向化和标准化处理后，可以发现与最优值的相对接近程度 W_i 表示各国环境保护与经济协调发展的能力，而且当 W_i 的得分越大，表示该国环境保护与经济协调发展的能力越强。

我们通过 MATLAB 对整个过程进行编程求解，得到最终结果见附录 11.2.3.

7.3 评价模型结果分析与发现

通过对世界各国经济水平、工业水平和森林植被面积的综合得分 W_i 进行排序得到各国环境保护与经济协调能力的排名。我们发现有些 GDP 与工业增加值总体水平很高的国家得分与排名却处于下游，比如德国和捷克。分析原因，可能是这些国家过于注重发展国内经济以及工业水平而忽视对环境特别是植被的保护。

与此形成鲜明对比的是，有些国家比如南非和中国，GDP 与工业增加值总体水平很高的同时，综合得分 W_i 也同样位于全球第一梯队。分析原因，可能是这些国家在发展国内经济以及工业水平的同时也注重对环境特别是植被的保护。

八、模型评价

8.1 模型的优点

- 第一问中的森林面积预测模型中我们采用了不同的预测方法对数据进行拟合和预测。我们同时采用 BP 神经网络预测模型与灰色预测模型对中国森林面积数据进行分析，并对模型的拟合度和结果进行比较。该方法在一定程度上弥补了由于数据并不全面的不足，同时也考虑到了单个模型对最终结果求解的局限性。
- 第二问中的基于时间序列预测的影响模型中，我们提出首先通过构建时间序列预测模型来对暂未实施退耕还林政策的时间段进行拟合预测，通过预测值与实际实施退耕还林后的数据进行比较，从而定量地得出政策所带来的影响，具有一定的创新性。
- 第三问中的基于熵权法的 Topsis 评论模型中，我们在传统的 Topsis 方法中加入了由熵权法决定的权重比例。该方法减少了传统 Topsis 方法的主观性的同时使得在数据样本不全面的条件下获得较为准确的得分与结果。

8.2 模型的缺点

- 由于数据的局限性与缺失，一些退耕还林的其它数据并没有包括在分析当中。由于时间有限，我们并没有插入缺失的数据。
- 我们并没有考虑国家内不同区域的影响，忽略这些区域可能会导致我们的预测模型与实际结果之间存在偏差。

8.3 模型的改进

限于问题所给数据集规模较小，因此本文中的预测与评价模型可能会与实际间存在一些误差，未来可以通过更多的数据集（更多年份、更多的指标类）来对模型进行训练，进而才能够进一步提升模型预测的拟合优度和准确率。

九、参考文献

- [1] 李天明, 田晓艳, 占东明. 财务视角下新冠肺炎疫情对贵州工业发展影响的统计测度 [C]//.2020 年 (第七届) 全国大学生统计建模大赛优秀论文集.[出版者不详],2020:335-359.DOI:10.26914/c.cnkihy.2020.045581.
- [2] 王庶, 岳希明. 退耕还林、非农就业与农民增收——基于 21 省面板数据的双重差分分析 [J]. 经济研究,2017,52(04):106-119.
- [3] 刘璨, 武斌, 鹿永华. 中国退耕还林工程及其所产生的影响 [J]. 林业经济,2009(10):41-46.
- [4] 张伟, 薛建宏, 张兴. 退耕还林政策对农户收入的影响及其作用机制 [J]. 农村经济,2019(06):130-136.
- [5] 王馨睿. 关于山西省退耕还林政策对农户劳动收入的影响——基于计量经济学模型分析 [J]. 经济师,2021(06):130-131+133.
- [6] 张帆. 退耕还林政策对河流径流、泥沙、洪水的影响分析 [J]. 陕西水利,2019(02):115-118.DOI:10.16747/j.cnki.cn61-1109/tv.2019.02.044.
- [7] 东梅. 退耕还林对我国粮食安全影响的实证分析 [D]. 南京农业大学,2005.
- [8] 谢涛, 刘晓静, 赵雪敏, 曹清华, 安玲. 降电价政策对工业生产活动影响的实证研究 [J]. 电力与能源,2021,42(04):373-380+385.

十、问题四 报告书

给联合国的一封信

尊敬的联合国秘书长先生:

您好!

伴随着经济全球化范围的扩大,全球经济进一步发展,与之相伴的是日趋严重的环境恶化,诸如臭氧层空洞、土地沙化、资源短缺等环境问题纷纷涌现。我们都知道经济发展不能以牺牲环境为代价,地球是目前唯一适合也是唯一能使人类存活的星球,经济发展与环境保护并非是对立的,而是互相影响互相协调的,二者相辅相成,才能国家的发展注入新鲜血液。

我们团队通过分析中国相关的一些经济与环境方面的数据,包括 GDP、二氧化碳排放量、能源使用量、农业用地和森林面积等方面,构建了灰色测试、基于熵权法修正的 TOPSIS 评价模型等数学模型,得出响应结果并对此进行了讨论分析:

利用灰色预测模型搭建自己的模型对未来 10 年内中国森林面积的预测,我们发现,在政策颁布和民众自主性提高的前提下,森林面积将会日趋增大,且增速基本保持在每年 2-3 万公顷;利用皮尔逊相关系数对数据进行处理、建模后,得出退耕还林对国家工业水平影响较弱,而对国内经济水平有较大影响的结论,这是由于我国工业和经济

组成的本质而决定的，我们要尊重事实，也要主动寻找中和的方法；搭建基于熵权法的Topsis评价模型后，我们得到了不同国家对于环境保护与经济协调能力的强弱排序，发现了一些能很好平衡经济发展与环境保护的国家各国之间加强合作，也发现了一些经济发展的同时忽略了环境建设的国家。但是环境保护是全人类的责任，共同推进生态建设，是大势所趋也是重中之重，对日后环境建设和环境保护具有重要意义。对此我们提出以下建议：

1、坚持植树造林，动员群众参与。

维持树木数量的稳定是环境持续发展的必要因素。因此，我们应该坚持植树造林，大力宣传植树活动，着重宣传“3·12植树节”，动员社区成员、学校学生、机关干部、各级党员积极参与，把保护环境的理念根植于心，把爱护树木的理念埋藏再每个人心底。这样才能有效维持树木数量的稳定持续增长。

2、发展高新技术产业，优化能源结构。

创新的本质是突破，利用现有的条件创造出新的生产要素和生产方式，我们通过发展科技来提升生产力，从而降低对于能源的消耗和浪费。通过创造性转化和创新性发展，可以提高对于资源的利用，从而减少对于环境的开采，同时优化能源结构，开发可再生资源，既有助于经济发展，也有助于维持森林面积的持续增长，有助于环境保护。

习近平总书记多次指出：创新是引领发展的第一动力。抓创新就是抓发展，谋创新就是谋未来。科技创新是我们时代发展的核心，在环境保护这个领域同样显得尤为重要。一方面，通过创新在经济中的转化有利于实现资源的节约与利用；另一方面，在能源的利用上，创新开发对象，也是一种重要的保护环境和促进经济发展的手段。

从消费者的角度来说，构建资源节约型、环境友好型的社会有助于提高我们的资源利用率，减少了金钱与资源的浪费，公民循环意识的形成，有助于循环经济的进一步发展，对于保障国家资源安全，推动实现碳达峰、碳中和，助力工业与经济同步发展，同时促进环境保护与生态文明建设都具有十分重要的意义。

经济发展与环境保护从来都不是对立的，它们二者是相辅相成，互相制约又互相促进的。我们要重视经济与环境的共同发展。

3. 加强“人类命运共同体”思想建设，完善全球环境保护

在新时代治国理政的战略格局和全球一体化的趋势下，合作共赢共同保证生态完整性的重要性愈发凸显。根据我们的编程求解，发现一些综合得分较低的国家拥有着较高的GDP和工业增加值，也有些国家能很好的平衡GDP、工业发展和环境保护。由此，我们建议，协调能力较差的国家应积极向拥有良好协调能力的国家学习，积极开展各类讨论、峰会和座谈会，加强环境保护方面知识的交流沟通，互相借鉴，共同保护地球环境资源。

十一、附录

11.1 构建的数据集

11.1.1 问题一数据集 中国有关数据

时间	能源使用量 (人均千克石油当量)	耗电量 (人均千瓦时)	GDP (现价美元)	二氧化碳 排放量(千吨)	工业增加值 (占 GDP 的百分比)	农业用地 (占土地面积的百分比)	森林面积 (占土地面积的百分比)	森林面积 (平方公里)
1990	766.9953	510.6199	3.61E+11	2173360	41.03292	53.7517	16.67333	1571406
1991	736.8518	548.9538	3.83E+11	2302190	41.4874	54.35426	16.88405	1591266
1992	752.6287	604.6937	4.27E+11	2418180	43.11533	54.66483	17.09477	1611126
1993	788.1287	662.637	4.45E+11	2643530	46.17654	55.07949	17.30549	1630986
1994	816.1629	727.1074	5.64E+11	2763900	46.16306	55.56831	17.51622	1650846
1995	866.8344	770.2802	7.35E+11	3084870	46.75054	55.56812	17.72685	1670706
1996	881.6537	821.0809	8.64E+11	3064880	47.10424	55.57221	17.93759	1690566
1997	871.7563	852.7407	9.62E+11	3128650	47.09905	55.57139	18.14831	1710426
1998	869.3586	870.6172	1.03E+12	3232120	45.79764	55.57078	18.35903	1730286
1999	878.5245	913.9634	1.09E+12	3149200	45.35986	55.57038	18.56975	1750146
2000	898.9873	992.9434	1.21E+12	3344090	45.5361	55.57004	18.7805	1770006
2001	928.8114	1076.549	1.34E+12	3526750	44.79341	55.63284	19.03101	1793615
2002	984.8107	1194.856	1.47E+12	3808330	44.45057	55.69564	19.28152	1817225
2003	1118.432	1379.485	1.66E+12	4413300	45.62278	55.75833	19.53203	1840835
2004	1268.133	1585.839	1.96E+12	5121830	45.90021	55.82049	19.78254	1864445
2005	1393.691	1782.312	2.29E+12	5819120	47.02258	55.88318	20.03305	1888055
2006	1515.174	2039.015	2.75E+12	6431910	47.5574	55.94577	20.28356	1911664
2007	1630.171	2325.927	3.55E+12	6988210	46.88417	56.00835	20.53407	1935274
2008	1672.904	2446.369	4.59E+12	7195010	46.97117	56.07094	20.78458	1958884
2009	1778.434	2612.457	5.10E+12	7715110	45.95715	56.1332	21.03509	1982494

11.1.2 问题二数据集 中国退耕还林面积数据集

年份	GDP（现价美元）	工业增加值（占 GDP 的百分比）	退耕还林面积
1997	9.62E+11	47.09905	
1998	1.03E+12	45.79764	
1999	1.09E+12	45.35986	
2000	1.21E+12	45.5361	
2001	1.34E+12	44.79341	
2002	1.47E+12	44.45057	
2003	1.66E+12	45.62278	3418116
2004	1.96E+12	45.90021	1016557
2005	2.29E+12	47.02258	861187
2006	2.75E+12	47.5574	268853
2007	3.55E+12	46.88417	85294
2008	4.59E+12	46.97117	12030
2009	5.10E+12	45.95715	739
2010	6.09E+12	46.49783	333
2011	7.55E+12	46.52929	59
2012	8.53E+12	45.42298	0
2013	9.57E+12	44.1767	0
2014	1.05E+13	43.08557	403
2015	1.11E+13	40.84134	341836.5
2016	1.12E+13	39.58062	683270

11.1.3 问题三数据集 2020 年全球数据集

国家 (名称)	GDP	工业增加值 (占 GDP 的百分比)	森林植被面积	
2	阿富汗	2.01E+10	14.0317	1.850994
4	安哥拉	5.84E+10	45.94846	53.42695
5	阿尔巴尼亚	1.49E+10	20.06265	28.79197
7	阿拉伯联盟国家	2.44E+12	37.19282	2.812929
8	阿拉伯联合酋长国	3.59E+11	40.86677	4.467756
9	阿根廷	3.89E+11	23.31122	10.44071
10	亚美尼亚	1.26E+10	27.05523	11.53741
12	安提瓜和巴布达	1.37E+09	21.10326	18.45455
13	澳大利亚	1.33E+12	25.4621	17.42131
14	奥地利	4.33E+11	25.45785	47.25097
15	阿塞拜疆	4.26E+10	41.38795	13.69286
16	布隆迪	2.84E+09	10.67899	10.88941
17	比利时	5.22E+11	19.47408	22.7642
18	贝宁	1.57E+10	16.29055	27.80374
19	布基纳法索	1.79E+10	32.59132	22.72076
20	孟加拉国	3.23E+11	29.53989	14.46877
21	保加利亚	6.99E+10	21.92335	35.86035
22	巴林	3.47E+10	40.27633	0.897436
23	巴哈马	9.91E+09	11.64269	50.93506
24	波斯尼亚和黑塞哥维那	1.99E+10	24.55623	42.73262

11.2 计算结果

11.2.1 问题一计算结果

年份	森林面积（公顷）	年份	森林面积（公顷）
2025	2351971.0019	2031	2508159.6711
2026	2377607.7761	2032	2534859.0553
2027	2403426.7303	2033	2561928.9621
2028	2429412.2762	2034	2589397.9247
2029	2455545.4418	2035	2617295.6867
2030	2481803.5166		

11.2.2 问题二计算结果

(1) 对经济（GDP）的影响：			(2) 对工业水平的影响：		
预测值	实际值	影响	预测值	实际值	影响
1.6E+12	1.66E+12	-0.03598	44.45058	45.62278	-0.02569
1.73E+12	1.96E+12	-0.11493	44.45058	45.90021	-0.03158
1.86E+12	2.29E+12	-0.18604	44.45058	47.02258	-0.0547
1.99E+12	2.75E+12	-0.27665	44.45058	47.5574	-0.06533
2.12E+12	3.55E+12	-0.40264	44.45058	46.88417	-0.05191
2.25E+12	4.59E+12	-0.51007	44.45058	46.97117	-0.05366
2.38E+12	5.1E+12	-0.5333	44.45058	45.95715	-0.03278
2.51E+12	6.09E+12	-0.58748	44.45058	46.49783	-0.04403
2.64E+12	7.55E+12	-0.65025	44.45058	46.52929	-0.04468
		-0.36637			-0.04493

11.2.3 问题三计算结果

国家（名称）	GDP	工业增加值（占 GDP 的百分比）	森林植被面积		得分	排名
2	阿富汗	2.01E+10	14.0317	1.850994	0.0006	86
4	安哥拉	5.84E+10	45.94846	53.42695	0.0045	117
5	阿尔巴尼亚	1.49E+10	20.06265	28.79197	0.0023	130
7	阿拉伯联盟国家	2.44E+12	37.19282	2.812929	0.0038	85
8	阿拉伯联合酋长国	3.59E+11	40.86677	4.467756	0.0024	53
9	阿根廷	3.89E+11	23.31122	10.44071	0.0015	202
10	亚美尼亚	1.26E+10	27.05523	11.53741	0.0017	119
12	安提瓜和巴布达	1.37E+09	21.10326	18.45455	0.0017	55
13	澳大利亚	1.33E+12	25.4621	17.42131	0.0026	51
14	奥地利	4.33E+11	25.45785	47.25097	0.0037	186
15	阿塞拜疆	4.26E+10	41.38795	13.69286	0.0025	62
16	布隆迪	2.84E+09	10.67899	10.88941	0.0009	32
17	比利时	5.22E+11	19.47408	22.7642	0.0021	58
18	贝宁	1.57E+10	16.29055	27.80374	0.0022	52
19	布基纳法索	1.79E+10	32.59132	22.72076	0.0024	116
20	孟加拉国	3.23E+11	29.53989	14.46877	0.002	112
21	保加利亚	6.99E+10	21.92335	35.86035	0.0028	191
22	巴林	3.47E+10	40.27633	0.897436	0.0022	178
23	巴哈马	9.91E+09	11.64269	50.93506	0.0036	79
24	波斯尼亚和黑塞哥维那	1.99E+10	24.55623	42.73262	0.0033	107

11.3 程序代码

11.3.1 问题一 BP 神经网络预测 MATLAB 代码

```
clear
close
clc
format short
mydata = xlsread('中国国家数据','Sheet1','A1:H40 ');
input = mydata(:,1:trainN-1);
output = mydata(:,trainN);
N = length(output);
testN = 20;
```



```

trainN = N - testN;
input_train = input(1:trainN,:);
output_train = output(1:trainN);
input_test = input(trainN+1:N,:);
output_test = out(train+1:N);
[inputn,inputps] = mapminmax(input_train,0,1);
[outputn,outputps] = mapminmax(out_train,0,1);
inputn_test = mapminmax('apply',input_test,inputps);
inputnum = size(input,2);
outputnum = size(output,2);
MSE = 1e + 5
transform_func = {'tansig','purelin'};
train_func = 'trainlm';
for hiddennum =
    fix(sqrt(inputnum+outputnum))+1:fix(sqrt(inputnum+outputnum))+10
net = newff(inputn,outputn,hiddennum,transform_func,train_func);
net.trainParam.epochs = 100000;
net.trainParam.lr = 0.01;
net.trainParam.goal = 0.0000000001;
net = train(net,inputn,outputn);
an0 = sim(net,inputn);
mse0 = mse(outputn,an0);
if mse0 < MSE
MSE = mse0;
hiddennum_best = hiddennum;
end
end
net=newff(inputn,outputn,hiddennum_best,transform_func,train_func);
net.trainParam.epochs=1000;
net.trainParam.lr=0.01;
net.trainParam.goal=0.000001;
net=train(net,inputn,outputn);
n=sim(net,inputn_test);
test_simu=mapminmax('reverse',an,outputps);
error=test_simu-output_test;

```

```

figure
plot(output_test,'bo-','linewidth',1.2)
hold on
plot(test_simu,'r*-','linewidth',1.2)
[~,len]=size(output_test);
SSE1=sum(error.^2);
MAE1=sum(abs(error))/len;
MSE1=error*error'/len;
RMSE1=MSE1^(1/2);
MAPE1=mean(abs(error./output_test));
r=corrcoef(output_test,test_simu);
R1=r(1,2);
for i=1:len
    disp([i,output_test(i),test_simu(i),error(i)])
end

```

11.3.2 问题一灰色预测 MATLAB 代码

```

clear;clc
year =[1990:1:2020]';
x0 =
[1571405.9,1591265.86,1611125.82,1630985.78,1650845.74,1670705.7,1690565.66,1710
425.62,1730285.58,1750145.54,1770005.5,1793615.33,1817225.16,1840834.99,1864444.
82,1888054.65,1911664.48,1935274.31,1958884.14,1982493.97,2006103.8,2025471.54,2
044839.28,2064207.02,2083574.76,2102942.5,2124598.67,2143394.7,2162190.4,2180986
.1,2199781.8
];
figure(1);
plot(year,x0,'o-'); grid on;
set(gca,'xtick',year(3:1:end))
xlabel('年份'); ylabel('森林面积');
disp('-----')
disp('准指数规律检验')
x1 = cumsum(x0);
rho = x0(2:end) ./ x1(1:end-1) ;

```

```

figure(2)
plot(year(2:end),rho,'o-',[year(2),year(end)],[0.5,0.5],'-'); grid on;
text(year(end-1)+0.2,0.55,'临界线')
set(gca,'xtick',year(3:1:end))
xlabel('年份'); ylabel('本次测试数据的光滑度');
test_num = 3;
train_x0 = x0(1:end-test_num);
disp(mat2str(train_x0'))
test_x0 = x0(end-test_num+1:end);
disp('试验数据是: ')
disp(mat2str(test_x0'))
result1 = gm11(train_x0, test_num)
result2 = new_gm11(train_x0, test_num)
result3 = metabolism_gm11(train_x0, test_num);
test_year = year(end-test_num+1:end);
figure(3)
plot(test_year,test_x0,'o-',test_year,result1,'*-',test_year,result2,'+-
',test_year,result3,'x-'); grid on;
set(gca,'xtick',year(end-test_num+1): 1 :year(end))
legend('试验组的真实数据','result1的预测结果','result2的预测结果','result3的预测结果')
xlabel('年份'); ylabel('森林面积');
SSE1 = sum((test_x0-result1).^2);
SSE2 = sum((test_x0-result2).^2);
SSE3 = sum((test_x0-result3).^2);
disp(strcat('result1对于试验组预测的误差平方和为',num2str(SSE1)))
disp(strcat('result2对于试验组预测的误差平方和为',num2str(SSE2)))
disp(strcat('result3对于试验组预测的误差平方和为',num2str(SSE3)))
if SSE1<SSE2
    if SSE1<SSE3
        choose = 1;
    else
        choose = 3;
    end
elseif SSE2<SSE3

```

```

        choose = 2;
else
    choose = 3;
end
Model = {'传统GM(1,1)模型','新信息GM(1,1)模型','新陈代谢GM(1,1)模型'};
disp(strcat('因',Model(choose),'的误差平方和最小，选其进行预测'))
predict_num = input('请输入你要往后面预测的期数: ');
[result, x0_hat, relative_residuals, eta] = gm11(x0, predict_num);
if choose == 2
    result = new_gm11(x0, predict_num);
end
if choose == 3
    result = metabolism_gm11(x0, predict_num);
end
disp('对原始数据的拟合结果: ')
for i = 1:n
    disp(strcat(num2str(year(i)), ' : ', num2str(x0_hat(i))))
end
    disp(strcat('往后预测', num2str(predict_num), '期的得到的结果: '))
for i = 1:predict_num
    disp(strcat(num2str(year(end)+i), ' : ', num2str(result(i))))
end

```

11.3.3 问题三熵权法代码

```

df = pd.read_excel('/content/统计表2020年.xlsx')
df.himport pandas as pd
ead
df.shape
df.info
df = df.dropna()
df.info
df.shape
df.head
df.to_excel('/content/001.xlsx')

```

```

import warnings
warnings.filterwarnings("ignore")
import pandas as pd
import numpy as np
def get_entropy_weight(data):
    data = (data - data.min())/(data.max() - data.min())
    data.to_excel("标准化后数据.xlsx", sheet_name="sheet1", index=
False,encoding="utf-8")
    print(data)
    m,n=data.shape
    data=data.as_matrix(columns=None)
    k=1/np.log(m)
    yij=data.sum(axis=0)
9.3.4 问题三Topsis代码
    pij=data/yij
    test=pij*np.log(pij)
    test=np.nan_to_num(test)
    ej=-k*(test.sum(axis=0))
    wi=(1-ej)/np.sum(1-ej)
    wi_list=list(wi)
    return wi_list
def get_score(wi_list,data):
    cof_var = np.mat(wi_list)
    context_train_data = np.mat(data)
    last_hot_matrix = context_train_data * cof_var.T
    last_hot_matrix = pd.DataFrame(last_hot_matrix)
    last_hot_score = list(last_hot_matrix.apply(sum))
    return last_hot_score
if __name__ == '__main__':
    data = pd.read_excel("指标.xlsx", encoding='utf8')
    data= data.iloc[:, 1:]
    mm = data
    wi_list=get_entropy_weight(data)
    score_list=get_score(mm,wi_list)
    print(wi_list)

```

```
print(score_list)
```

11.3.4 问题三 Topsis 代码

```
clear;clc
load 001.mat
[n,m] = size(X);
weigh = input(['输入' num2str(m) '个权数。' '输入' num2str(m) '个权重：'
    ']);
Z = X ./ repmat(sum(X.*X) .^ 0.5, n, 1);
disp('标准化矩阵 Z = ')
disp(Z)
A_P = sum([(Z - repmat(max(Z),n,1)) .^ 2] .* repmat(weigh,n,1) ,2) .^
    0.5;
A_N = sum([(Z - repmat(min(Z),n,1)) .^ 2] .* repmat(weigh,n,1) ,2) .^
    0.5;
S = A_N ./ (A_P+A_N);
disp('最终得分： ')
stand_new = S / sum(S)
[sorted_S,index] = sort(stand_new , 'descend')
```