

# 中国式现代化水平测度及其时空特征

吴晴静

福建师范大学数学与统计学院, 福建 福州

收稿日期: 2023年10月11日; 录用日期: 2023年12月7日; 发布日期: 2023年12月13日

## 摘要

中国式现代化是中国贡献给世界的人类文明新形态, 具有人口规模巨大现代化、全体人民共同富裕现代化、物质文明和精神文明协调发展现代化、人与自然和谐共生现代化、走和平发展道路现代化这几个特征。本文结合这五个重要特征, 建立了五个维度评价指标体系, 采用熵权法, 并运用Dagum基尼系数、Kernel密度估计分析了2012~2020年中国各地区的现代化水平。其中, 由于变量选择过多, 所以结合了随机森林方法进行特征选择后再分析, 发现结果基本一致。我国各省份现代化水平在不断上升, 我国现代化水平总体差异主要来源于地区间的差异, 在三大地带中, 东部地带现代化程度最高, 而中西部现代化程度近似。

## 关键词

中国式现代化, 熵值法, 空间差异, 特征选择

# Measurement of the Level of Chinese-Style Modernization and Its Temporal and Spatial Characteristics

Qingjing Wu

School of Mathematics and Statistics, Fujian Normal University, Fuzhou Fujian

Received: Oct. 11<sup>th</sup>, 2023; accepted: Dec. 7<sup>th</sup>, 2023; published: Dec. 13<sup>th</sup>, 2023

## Abstract

Chinese-style modernization is a new form of human civilization that China has contributed to the world. It has modernization with a huge population, common prosperity for all people, coordinated development of material civilization and spiritual civilization, modernization of harmonious coexistence between man and nature, and modernization of peaceful development. Combining these five important features, this paper establishes a five-dimensional evaluation index system, adopts the entropy weight method, and uses the Dagum Gini coefficient and Kernel density estimation to

analyze the modernization level of various regions in China from 2012 to 2020. Among them, due to the selection of too many variables, the random forest method is combined for feature selection and then analyzed, and the results are found to be basically the same. The modernization level of China's provinces is constantly rising. The overall difference in China's modernization level mainly comes from regional differences. Among the three major zones, the eastern zone has the highest degree of modernization, while the central and western zones have similar degrees of modernization.

## Keywords

Chinese-Style Modernization, Entropy Method, Spatial Difference, Feature Selection

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

现代化是一个复杂的概念，它不仅仅指经济的发展和技术的进步，还包括政治、文化、社会等方面的变革。现代化的进程是一个长期的、漫长的过程，需要不断的投入和努力才能实现。在全球化的背景下，现代化已成为各国共同关注和追求的目标，以满足人民对更好生活和更高质量的服务的需求[1]。

中国式现代化是指在中国特色社会主义理论指导下，中国在现代化进程中所形成的一种独特模式，以经济建设为中心，全面推进现代化建设的过程[2]。中国式现代化的内涵包括经济现代化、政治现代化、文化现代化和社会现代化等方面。在经济领域，中国式现代化注重提高经济发展的质量和效益，在创新驱动、产业升级、绿色发展等方面加大力度，实现经济高质量发展。在政治领域，中国式现代化推动法治化进程，加强国家治理能力和公共服务水平，实现政治体系现代化。在文化领域，中国式现代化注重传承中华文化，加强文化软实力建设，推动文化创新与产业融合，提高文化产业的核心竞争力。在社会领域，中国式现代化注重提高人民群众的生活质量和幸福感，加强教育、医疗、社保等公共服务，实现城乡一体化发展，推进共同富裕[3]。

如何评估中国式现代化的进程，已成为学术研究和政策制定的重要问题。评估现代化进程需要考虑多个方面，如经济发展、社会进步、政治体制等。在经济方面，可以通过 GDP、人均收入等指标来评估经济现代化的进程[4]。在社会方面，可以通过教育、医疗、社保等公共服务的覆盖率和质量来评估社会现代化的进程。在政治方面，可以通过政治体制的改革和民主化进程来评估政治现代化的进程。在文化方面，可以通过文化产业的发展和软实力的提升来评估文化现代化的进程[5]。

总之，中国式现代化是一种独特的、符合中国国情的现代化模式，它在全面推进现代化建设的过程中注重人口规模、全体人民共同富裕、物质文明和精神文明相协调、人与自然和谐共生、走和平发展道路等方面[6]。评估中国式现代化的进程需要考虑多个方面的指标和因素，需要不断的投入和努力才能实现。

本文先结合现代化的五个重要特征，建立了五个维度评价指标体系，然后对收集的 2012~2020 年我国 31 个省级行政区(不含港澳台地区)的面板数据采用熵权法分析了各地区现代化发展水平，并用 Dagum 基尼系数、Kernel 密度估计分析了各地区间和地区内的空间差异。

本文的研究意义在于，通过对中国式现代化水平的测度和时空特征的分析，可以深入了解中国式现代化的实践状况和发展趋势，为中国式现代化建设提供理论和实践支持，同时也可以为国内外学者提供参考和借鉴。

## 2. 数据介绍

本文数据选自中国统计年鉴，截取了 2012 年至 2020 年的数据，包含 31 个省市的多项指标数据，本文将对其进行归纳分类，如表 1 所示。其中属性一栏用来表示指标的正向和负向。

**Table 1.** Evaluation system of modernization level index

**表 1.** 现代化水平指标评价体系

综合指标	二级指标	三级指标	四级指标	属性
人口规模现代化	人口分布	城镇化水平		+
		人口密度		-
	人口增速	人口自然增长率		+
		人力资源综合能力	规模以上工业企业 R&D 人员全时当量	
	城乡居民生活质量	城镇居民家庭恩格尔系数		-
		农村居民家庭恩格尔系数		-
		人均地区生产总值增速		+
	公共基础设施与服务	城市燃气普及率		+
		道路长度		+
		公共汽电车运营数		+
		农村宽带接入用户		+
	全体人民共同富裕的现代化	城乡居民社会养老保险实际领取待遇率		+
参加失业保险率			+	
城镇基本医疗保险年末参保率			+	
中国式现代化	基本社会保障	每万人人均拥有卫生人员数		+
		社会治理能力	公共服务支出占地区生产总值比重 公共安全财政支出占财政总支出	+
	城乡收入差距	泰尔指数		-
	收入与消费	全体居民人均消费支出		+
		年度可支配收入		+
	教育水平现代化	教育经费占地区生产总值比重		+
		高等教育人数比例		+
	物质文明和精神文明相协调的现代化	人均拥有公共图书馆藏量		+
		艺术表演团体演出场次		+
		博物馆参观人次		+
创新能力现代化	电视节目综合人口覆盖率		+	
	R&D 经费支出占地区生产总值比重		+	
	技术市场成交额占地区生产总值比重		+	
	国内实用新型专利申请授权量		+	
生态文明现代化	绿化水平	每万人国内实用新型专利申请授权量		+
		森林覆盖率		+
	建成区绿化覆盖率		+	
	能源消耗	电力消费量 单位地区生产总值电耗		+

Continued

中国 式 现 代 化	生态文明 现代化	资源储备	人均水资源量	+
		污染物排放量	二氧化硫排放量	-
			氮氧化物排放量	-
			氨氮排放量	-
			化学需氧量排放量	-
	环境治理	生活垃圾无害化处理率	+	
	走和平 发展道路 的现代化	对外开放现代化	国际旅游外汇收入	+
			国际旅游收入占地区生产总值比重	+
			接待国际游客	+
			外商投资占比	+
货物进出口总额占地区生产总值比重			+	

同时，为较好地刻画中国式现代化及其时空特征，本文将 31 个省市按东西走向划分为东部、中部及西部，如表 2 所示。

Table 2. Regional division table of 31 provinces and cities

表 2. 31 省市区域划分表

地带	各地带所含区域
东部	北京、天津、河北、辽宁、上海、江苏、浙江、福建、山东、广东、海南
中部	山西、吉林、黑龙江、安徽、江西、河南、湖北、湖南
西部	内蒙古、广西、重庆、四川、贵州、云南、陕西、甘肃、青海、宁夏、新疆、西藏

### 3. 基于全指标的测度分析

#### (一) 指标权重和各地得分排名

本文将全部指标通过熵权法进行计算权重，在表 1 中已将所有指标进行分类，因此每个类别的所有指标权重和为 1，结果如表 3 所示。

其中，熵权法(Entropy Weight Method)是一种常用的多属性决策分析方法[7]，用于确定各个属性在决策中的权重。熵权法基于信息熵的概念，将多个指标的熵值作为权重因子，通过计算每个指标的信息熵和信息熵权重，来确定各个指标的权重。熵权法适用于多属性决策分析中，可以帮助决策者确定各个指标的权重，从而更加科学、客观地评价和选择决策方案[8]。与其他权重确定方法相比，熵权法不需要预先确定指标之间的关系，不受主观因素干扰，具有较高的可靠性和实用性。根据熵的特性，我们可以通过计算熵值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程度，指标的离散程度越大，该指标对总体的影响越大。假设数据有  $n$  行记录， $m$  个变量，数据可以用一个  $n \cdot m$  的矩阵表示( $n$  行  $m$  列)。下面式子中， $i = 1, 2, \dots, n$ ； $j = 1, 2, \dots, m$ 。

#### 1) 数据标准化

将各指标由绝对值变为相对值，同时消除量纲对结果的影响。

对于正向指标：

$$x'_{ij} = \frac{x_{ij} - \min\{x_{1j}, \dots, x_{nj}\}}{\max\{x_{1j}, \dots, x_{nj}\} - \min\{x_{1j}, \dots, x_{nj}\}} \quad (1)$$

对于负向指标:

$$x'_{ij} = \frac{\max\{x_{1j}, \dots, x_{nj}\} - x_{ij}}{\max\{x_{1j}, \dots, x_{nj}\} - \min\{x_{1j}, \dots, x_{nj}\}} \quad (2)$$

2) 确定各指标的信息熵

计算第  $j$  个指标下第  $i$  个样本占该指标的比重:

$$\rho_{ij} = \frac{x'_{ij}}{\sum_{i=1}^n x'_{ij}} \quad (3)$$

计算第  $j$  个指标的熵值:

$$e_j = -\frac{1}{\ln(n)} \cdot \sum_{i=1}^n \rho_{ij} \cdot \ln(\rho_{ij}) \quad (4)$$

3) 确定各指标的权重

$$\omega_j = \frac{1 - e_j}{m - \sum_{j=1}^m e_j}, \text{ 其中 } 0 < \omega_j \leq 1 \quad (5)$$

**Table 3.** Result table of entropy weight method  
**表 3.** 熵权法结果表

熵权法			
项	信息熵值 $e$	信息效用值 $d$	权重(%)
人口自然增长率	0.977	0.023	13.876
城镇化水平	0.988	0.012	7.332
规模以上工业企业 R&D 人员全时当量	0.878	0.122	74.973
人口密度	0.994	0.006	3.819
公共安全财政支出占财政总支出	0.901	0.099	16.735
公共服务支出占地区生产总值比重	0.891	0.109	18.408
参加失业保险率	0.938	0.062	10.47
每万人人均拥有卫生人员数	0.981	0.019	3.253
城镇基本医疗保险年末参保率	0.961	0.039	6.641
城乡居民社会养老保险实际领取待遇率	0.983	0.017	2.822
人均地区生产总值增速	0.991	0.009	1.601
城市燃气普及率	0.997	0.003	0.458
道路长度	0.939	0.061	10.32
公共汽电车运营数	0.949	0.051	8.595
农村宽带接入用户	0.907	0.093	15.73
城镇居民家庭恩格尔系数	0.993	0.007	1.23

## Continued

农村居民家庭恩格尔系数	0.991	0.009	1.517
泰尔指数	0.987	0.013	2.22
每万人国内实用新型专利申请授权量	0.913	0.087	11.617
国内实用新型专利申请授权量	0.877	0.123	16.397
R&D 经费支出占地区生产总值比重	0.949	0.051	6.749
技术市场成交额占地区生产总值比重	0.853	0.147	19.514
全体居民人均消费支出	0.971	0.029	3.853
年度可支配收入	0.962	0.038	5.105
人均拥有公共图书馆藏量	0.937	0.063	8.419
育经费占地区生产总值比重	0.934	0.066	8.77
高等教育人数比例	0.972	0.028	3.661
艺术表演团体演出场次	0.942	0.058	7.659
电视节目综合人口覆盖率	0.996	0.004	0.597
博物馆参观人次	0.942	0.058	7.659
建成区绿化覆盖率	0.997	0.003	0.65
森林覆盖率	0.961	0.039	8.261
电力消费量	0.955	0.045	9.617
人均水资源量	0.667	0.333	70.775
生活垃圾无害化处理率	0.995	0.005	0.963
单位地区生产总值电耗	0.993	0.007	1.526
氨氮排放量	0.993	0.007	1.592
二氧化硫排放量	0.992	0.008	1.7
氮氧化物排放量	0.989	0.011	2.293
化学需氧量排放量	0.988	0.012	2.623
国际旅游外汇收入	0.883	0.117	29.325
国际旅游收入占地区生产总值比重	0.942	0.058	14.633
接待国际游客	0.867	0.133	33.214
货物进出口总额占地区生产总值比重	0.967	0.033	8.362
外商投资占比	0.942	0.058	14.465

由表 3 可知,在人口规模现代维度中,规模以上工业企业 R&D 人员全时当量指标权重最大;在全体人民共同富裕的现代化指标中,公共服务支出占地区生产总值比重指标权重最大;在物质文明和精神文明相协调的现代化维度中,技术市场成交额占地区生产总值比重指标权重最大;在生态文明现代化维度中,人均水资源量指标权重最大;在走和平发展道路的现代化维度中,国际旅游外汇收入指标权重最大。以上几大指标综合考虑了工业水平、社会保障程度、科技发展水平、人均生存资源及国际交流程度,较好地刻画了现代化发展的水平。同时,为考虑各维度的重要性,同样采取熵权法对各二级指标进行权重计算,结果如表 4 所示。

**Table 4.** Weight table of each dimension  
**表 4.** 各维度权重表

项	信息熵值 $e$	信息效用值 $d$	权重(%)
人口规模现代化	0.94	0.06	20.041
物质文明和精神文明相协调现代化	0.938	0.062	20.826
全体人民共同富裕的现代化	0.964	0.036	12.13
生态文明现代化	0.919	0.081	27.115
走和平发展的现代化道路	0.941	0.059	19.887

由表 4 可知, 各个二级指标比较均衡地刻画了现代化水平, 从侧面反映了本文对各指标的分类较为合适。接着我们通过对各个二级指标进行综合得分计算, 可得各省市的综合得分情况, 如表 5 所示。

**Table 5.** Comprehensive score table of provinces and cities  
**表 5.** 各省市综合得分表

地区	人口规模现代化/排名		生态文明现代化/排名		物质文明和精神文明相协调现代化/排名		全体人民共同富裕的现代化/排名		走和平发展的现代化道路/排名		综合得分	综合排名
广东省	0.717	1	0.22	3	0.287	5	0.393	2	0.74	1	0.571	1
西藏自治区	0.174	23	0.79	1	0.122	21	0.361	4	0.096	23	0.4	2
江苏省	0.595	2	0.167	15	0.338	3	0.396	1	0.182	10	0.396	3
浙江省	0.517	3	0.227	2	0.313	4	0.319	5	0.205	7	0.372	4
北京市	0.202	13	0.173	10	0.448	1	0.295	6	0.269	4	0.346	5
上海市	0.211	12	0.135	26	0.359	2	0.198	19	0.344	2	0.285	6
山东省	0.402	4	0.153	21	0.208	7	0.373	3	0.161	13	0.275	7
天津市	0.192	14	0.124	29	0.284	6	0.198	20	0.293	3	0.229	8
福建省	0.292	5	0.22	4	0.156	12	0.211	15	0.231	5	0.22	9
安徽省	0.264	7	0.149	22	0.165	11	0.209	16	0.183	9	0.177	10
四川省	0.189	15	0.171	14	0.182	9	0.266	7	0.113	22	0.174	11
湖北省	0.243	8	0.165	17	0.169	10	0.228	11	0.131	17	0.172	12
陕西省	0.178	20	0.167	16	0.193	8	0.186	25	0.172	11	0.166	13
湖南省	0.239	9	0.178	9	0.146	13	0.223	12	0.132	16	0.165	14
河南省	0.269	6	0.141	25	0.146	14	0.248	9	0.096	24	0.157	15
辽宁省	0.15	28	0.156	20	0.14	16	0.233	10	0.198	8	0.154	16
江西省	0.213	11	0.2	6	0.142	15	0.188	22	0.122	20	0.152	17
广西壮族自治区	0.185	17	0.204	5	0.095	28	0.197	21	0.171	12	0.142	18
重庆市	0.185	19	0.172	12	0.132	18	0.216	14	0.128	18	0.14	19
云南省	0.165	25	0.197	8	0.094	29	0.172	30	0.213	6	0.137	20
河北省	0.223	10	0.145	24	0.11	25	0.26	8	0.085	25	0.132	21
海南省	0.177	21	0.2	7	0.083	31	0.171	31	0.151	14	0.12	22
黑龙江省	0.118	30	0.158	18	0.124	19	0.207	17	0.115	21	0.109	23
吉林省	0.109	31	0.158	19	0.114	24	0.203	18	0.122	19	0.103	24

Continued

青海省	0.176	22	0.172	13	0.123	20	0.188	23	0.043	28	0.103	25
内蒙古自治区	0.152	27	0.149	23	0.093	30	0.182	27	0.139	15	0.099	26
新疆维吾尔自治区	0.185	18	0.131	27	0.096	27	0.22	13	0.061	27	0.093	27
贵州省	0.162	26	0.173	11	0.117	23	0.174	29	0.041	29	0.093	28
山西省	0.168	24	0.131	28	0.103	26	0.188	24	0.07	26	0.085	29
宁夏回族自治区	0.188	16	0.114	31	0.119	22	0.177	28	0.04	30	0.078	30
甘肃省	0.146	29	0.119	30	0.14	17	0.185	26	0.024	31	0.078	31

由表 5 可知, 广东省、西藏自治区、江苏省、浙江省、北京市、上海市、山东省、天津市及福建省等十个省市综合排名靠前, 其中西藏自治区因为人均资源丰富, 污染排放少, 生态文明现代化那一项的得分远远高于其他地区, 并且人口密度小, 有国家政策帮扶, 故在本文的现代化评价体系内其综合得分较高。

(二) 现代化空间发展的差异分析

根据表 2 的划分, 对三大地带进行综合得分计算, 结果如表 6 所示。

Table 6. Comprehensive score table of the three major zones

表 6. 三大地带综合得分表

地带	人口规模现代化	生态文明现代化	物质文明和精神文明相协调现代化	全体人民共同富裕的现代化	走和平发展的现代化道路	综合得分
东部	0.333	0.175	0.248	0.277	0.259	0.281
中部	0.203	0.16	0.139	0.212	0.121	0.14
西部	0.174	0.214	0.125	0.21	0.102	0.141

由表 6 可知, 东部地带现代化程度最高, 中西部地带综合现代化程度略低一些, 且中西部地带现代化差异较小。

Kernel Density Estimation (KDE) 是一种非参数密度估计方法[9], 用于估计随机变量的概率密度函数。它通过对数据进行核函数(Kernel Function)平滑处理来估计概率密度函数, 从而避免了使用参数化概率分布模型的假设。KDE 的基本思想是将每个观测值视为一个独立的概率密度函数, 并通过对所有观测值的概率密度函数进行平均, 来估计总体概率密度函数。具体地说, 对于一个实数随机变量  $X$ , 其概率密度函数可以表示为:

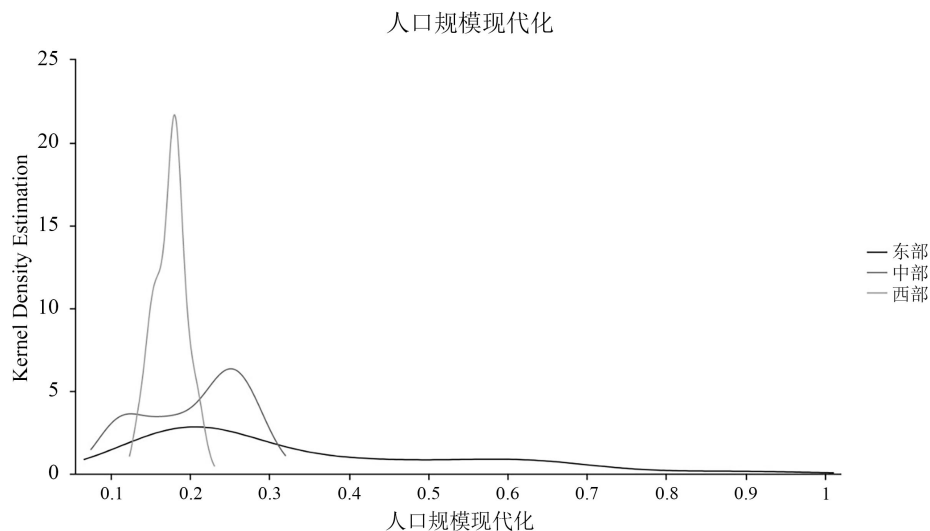
$$f(x) = \frac{1}{n} \sum_{i=1}^n K_h(x - x_i) \tag{6}$$

其中,  $n$  表示观测值的数量,  $x_i$  表示第  $i$  个观测值,  $K_h$  表示核函数,  $h$  表示平滑参数(也称带宽或窗宽), 用于调节核函数的宽度。核函数  $K_h$  必须满足一些条件, 比如非负性、积分为 1 等, 通常使用高斯核函数或 Epanechnikov 核函数等常见的核函数。KDE 的核心思想是通过将每个观测值的概率密度函数加权平均, 来估计总体概率密度函数。权重的大小由核函数  $K_h$  决定, 它决定了每个观测值对最终概率密度函数的贡献大小。当  $h$  较小时, 核函数的宽度较窄, 估计出的概率密度函数较精细, 但容易受到噪声的影响; 当  $h$  较大时, 核函数的宽度较宽, 估计出的概率密度函数较平滑, 但可能会损失一些细节信息。

KDE 的实现步骤通常包括以下几个步骤:

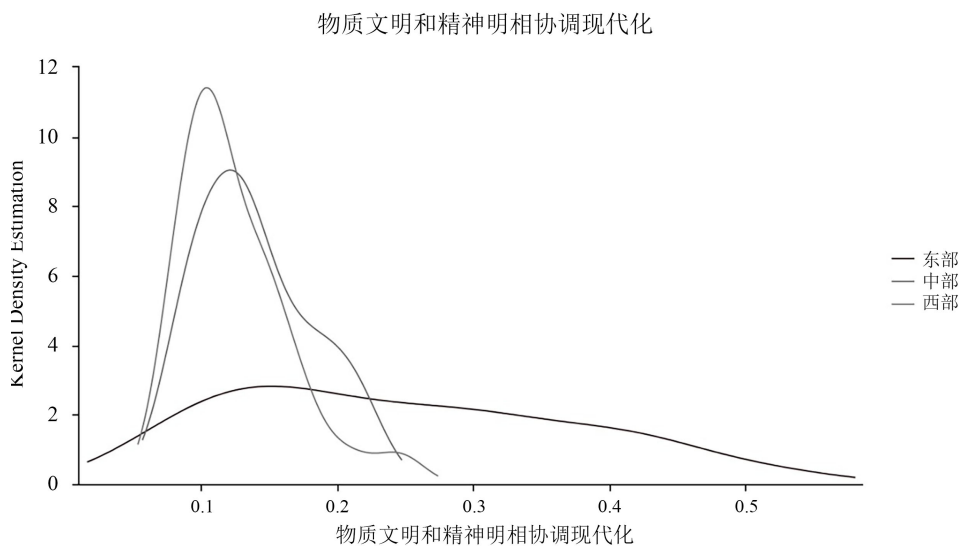


- 1) 选择核函数，通常使用高斯核函数或 Epanechnikov 核函数等常用的核函数；
  - 2) 选择平滑参数  $h$ ，通常使用交叉验证等方法来确定  $h$  的最优值；
  - 3) 对于每个观测值，计算其对应的核函数值；
  - 4) 对所有观测值的核函数值进行加权平均，得到最终的概率密度函数估计值。
- 通过对三大地带的二级指标进行核密度估计，如下所示。



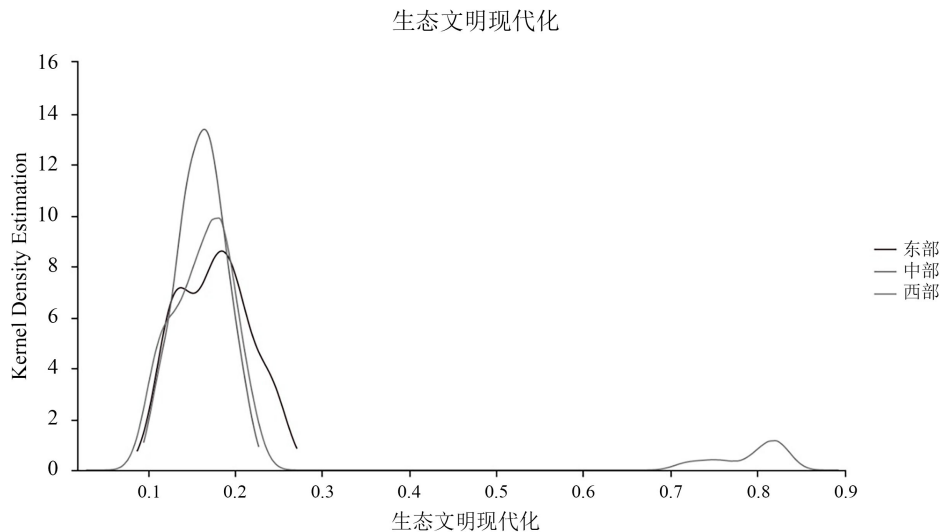
**Figure 1.** Kernel density curve of population scale modernization in the three major zones  
**图 1.** 三大地带人口规模现代化的核密度曲线

由图 1 可知，西部地带各省市的人口规模现代化差异较小，水平低，多集中在 0.1~0.2 之间，而东部地带人口规模现代化分布差异较大，如海南与北京之间的城镇化水平和高级科研人员等指标的较大差异导致此二级指标的得分差异较大。



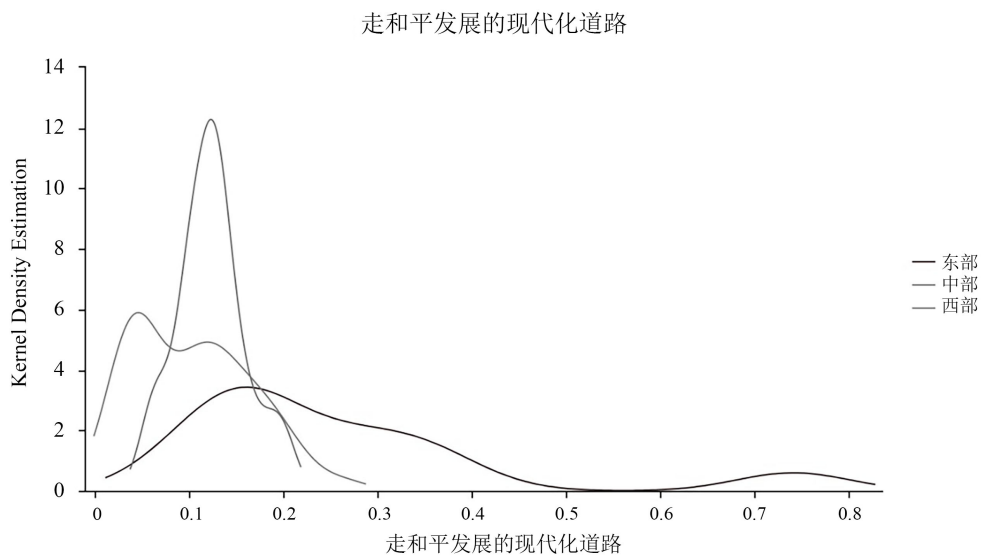
**Figure 2.** Kernel density curve of the coordinated modernization of material civilization and spiritual civilization in the three major zones  
**图 2.** 三大地带物质文明和精神文明相协调现代化的核密度曲线

由图2可知,中部地带和西部地带的物质文明和精神文明相协调现代化程度主要集中在0.1~0.3之间,地带内部差异较小,而东部地带内部各省市差异较大。



**Figure 3.** Kernel density curve of ecological civilization modernization in the three major zones  
**图 3.** 三大地带生态文明现代化的核密度曲线

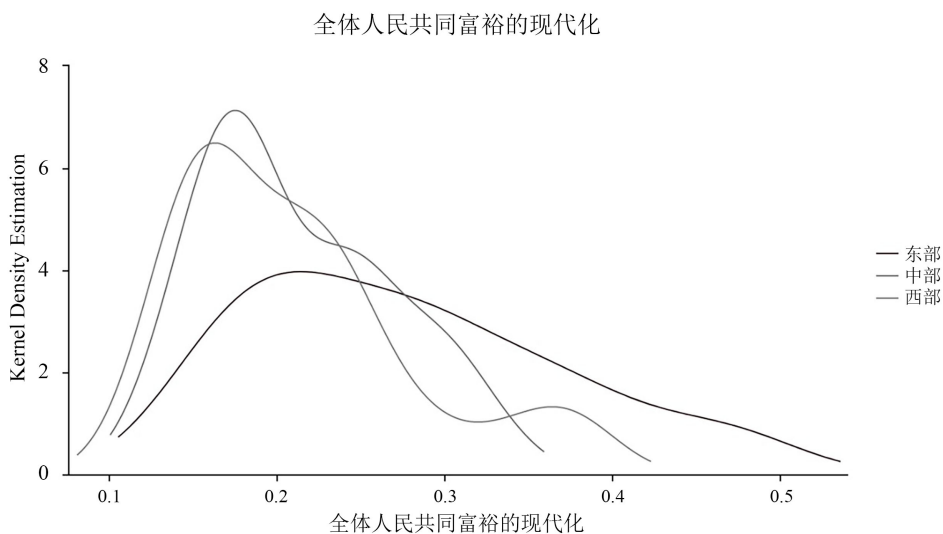
由图3可知,各地带内部的生态文明现代化差异较小,且主要集中在0.1~0.3之间,具体地,近些年各省市均在“绿水青山就是金山银山”的倡导下进行绿色发展,故得分比较接近,其中西藏自治区的得分最高,且远远大于其他地区。



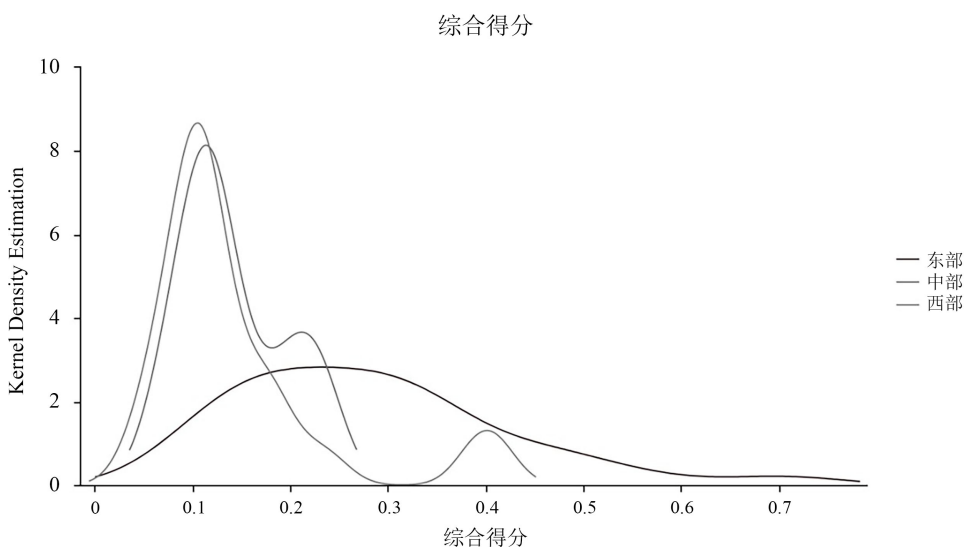
**Figure 4.** Kernel density curve of the modernization path of the three major zones taking away peaceful development  
**图 4.** 三大地带走和平发展的现代化道路的核密度曲线

由图4可知,在走和平发展的现代化道路指标中,东部地带差异较大。

由图5可知,在全体人民共同富裕的现代化指标中,各地带的内部差异均较显著,如东部地带的北京和海南、中部地带的湖南和山西、西部地带的四川和新疆等。



**Figure 5.** Kernel density curve of the common prosperity and modernization of all the people in the three major zones  
**图 5.** 三大地带全体人民共同富裕现代化的核密度曲线



**Figure 6.** Kernel density curve of the comprehensive scores of the three major zones  
**图 6.** 三大地带综合得分的核密度曲线

由图 6 可知，中西部地带内部各省市的综合得分差异较小，而东部地带内部各省市的差异较大。

基尼分解(Gini Decomposition)是一种将总体不平等度量分解为各因素贡献的方法[10]。洛伦兹曲线描述了收入分配的累积分布情况，即横轴表示人口比例，纵轴表示收入比例[11]。而洛伦兹系数则反映了洛伦兹曲线与完全平等状态之间的差距，它等于曲线下方面积与对角线下方面积之差。定义一个变量  $y$  表示某种社会福利指标，如收入、财富或教育水平等。同时，我们假设有  $n$  个人，并且第  $i$  个人的收入为  $y_i$ 。

那么，整个群体的洛伦兹系数可以表示为：
$$G = \frac{\sum_{i=1}^n \sum_{j=1}^n |y_i - y_j|}{2n \sum_{i=1}^n y_i}$$
。其中，分子部分是所有两两个人之间

收入差异的绝对值之和，分母部分是所有人收入之和的两倍。将总体的洛伦兹系数拆分为各因素贡献的和。假设有  $k$  个因素影响了收入分配，如教育、性别、种族等等。那么，我们可以将总体的洛伦兹系数

表示为： $G = G_0 + \sum_{i=1}^k \lambda_i G_i$ 。其中， $G_0$ 是完全平等状态下的洛伦兹系数， $G_i$ 是第  $i$  个因素对总体不平等的贡献， $\lambda_i$ 是第  $i$  个因素对总体不平等的弹性系数。

使用基尼系数(Gini Coefficient)来度量各因素对总体不平等的贡献度。基尼系数等于洛伦兹系数减去完全平等状态下的洛伦兹系数，即：

$$g_i = G - G_0 - \lambda_i G_i$$

这个公式表示了第  $i$  个因素对总体不平等的贡献。如果  $g_j > 0$ ，则说明该因素导致了不平等的增加；如果  $g_j < 0$ ，则说明该因素导致了不平等的降低。

本文对三大地带进行计算基尼系数及贡献率，如表 7 所示。

**Table 7.** Decomposition and contribution rate of Dagum Gini coefficient in Chinese-style modernization level region  
**表 7.** 中国式现代化水平区域 Dagum 基尼系数分解和贡献率

年份	整体基尼系数	地区内基尼系数			地区间基尼系数			贡献率(%)		
		东部	中部	西部	东 - 中	东 - 西	中 - 西	地区内	地区间	超变密度
2012	0.248	0.107	0.035	0.334	0.453	0.465	0.246	24.937	59.869	15.194
2013	0.240	0.099	0.048	0.321	0.421	0.448	0.236	25.095	59.232	15.673
2014	0.246	0.128	0.045	0.303	0.413	0.432	0.235	25.642	58.386	15.972
2015	0.256	0.135	0.065	0.275	0.402	0.416	0.221	26.156	57.298	16.546
2016	0.239	0.129	0.05	0.263	0.368	0.397	0.216	26.051	56.130	17.819
2017	0.243	0.119	0.058	0.235	0.337	0.368	0.196	26.871	53.578	19.552
2018	0.265	0.126	0.055	0.235	0.321	0.359	0.201	28.415	52.146	19.439
2019	0.268	0.152	0.053	0.244	0.326	0.368	0.218	28.486	52.490	19.024
2020	0.267	0.154	0.052	0.241	0.319	0.378	0.226	28.942	55.107	15.951

从表 7 中可以明显看出，2012~2020 年地区间差距贡献率远高于地区内部差距的贡献率。从表中的数据变化趋势来看，地区内部差距的贡献率呈现缓慢递增的趋势，由 2012 年的 24.94% 上升到 2020 年的 28.94%；而地区间差距贡献率由 2012 年的 59.87% 逐渐下降到 2018 年的 52.15%，然后在 2019 年之后大幅上升到 2020 年的 55.11%，地区间差距贡献率整体呈现先下降后递增的趋势。超变密度贡献率在 2012 到 2019 年间整体呈现递增趋势，但在 2020 年又大幅下降到 15.95%。可以得到我国现代化水平总体差异主要来源于地区间的差异。

#### 4. 基于特征选择的测度分析

##### (一) 特征选择

当需要进行大规模数据分析时，可能会面临大量指标和特征的问题，这会导致计算成本高昂、模型过拟合以及模型性能下降等问题[12]。因此，进行特征选择和筛选是一个重要的预处理步骤，可以从原始数据中选择最相关的特征，提高模型的性能和可解释性。对于分析中国式现代化的指标，本文将按照以下步骤进行特征筛选：

1) 数据探索和预处理：首先，需要对数据进行探索和预处理，包括缺失值填补、异常值处理、数据标准化和归一化等。通过减少噪声和冗余信息，提高模型的鲁棒性和可靠性。

2) 特征重要性评估：然后，使用随机森林等算法计算每个指标的重要性得分，以确定哪些指标对于

分析中国式现代化的影响最大。

3) 特征组合和交互：对于一些相关性较高的指标，可以考虑进行特征组合和交互，例如计算这些指标的平均值或比例值，或者构建新的指标作为特征。

4) 模型评估和调优：在进行特征筛选后，需要对模型进行评估和调优，以确定最佳的模型参数和特征组合。可以使用交叉验证和网格搜索等方法来选择最佳的模型参数，或者使用集成学习和深度学习等方法来进一步提高模型性能。

本文将采用随机森林进行特征选择。在分析模型之前，应先引入决策树，随机决策树是根据策略抉择而建立起来的树模型，其构建的目的就是寻找具有决定性作用的特征[13]，再根据特征重要性程度构建一棵树，以程度最大特征为树的结点，逐步递归找出各个分支下的子数据集中排名第二大的结点，以此类推，将所有数据循环划分到所属类中，从而筛选出特征重要性较大的前几项。

随机森林模型就是在决策树的基础上，在训练过程中添加随机属性，在基决策树上每个节点属性中随机选中包含  $n$  个属性的子集合，再从子集合中选最好的属性[14]。对于切分节点的好坏，一般采用子节点不纯度的加权和  $G(x_i, v_{ij})$  来表示：

$$G(x_i, v_{ij}) = \frac{n_{left}}{N_s} H(X_{left}) + \frac{n_{right}}{N_s} H(X_{right}) \tag{7}$$

其中， $x_i$  表示某一个切分点变量， $v_{ij}$  为切分值， $n_{left}$ 、 $n_{right}$  以及  $N_s$  分别表示切分点左右子节点训练样本合数以及当前节点所有训练样本个数， $X_{left}$  和  $X_{right}$  表示训练样本集合， $H(X)$  表示不纯度函数。问题一采用随机森林的回归预测模型，因此  $H(X)$  一般采用均方误差来衡量，即： $H(X_m) = \frac{1}{N_m} \sum_{i \in N_m} (y - \bar{y}_m)^2$ ，

其中  $y_m$  为当前节点样本目标函数的平均值，流程如图 7 所示。

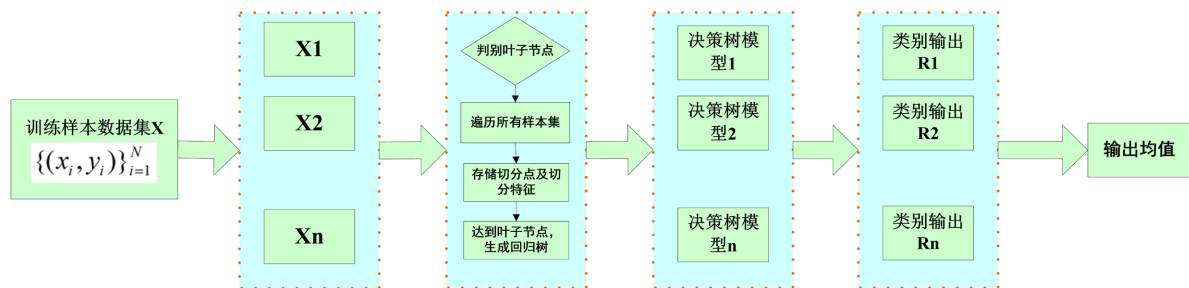


Figure 7. Random forest regression algorithm model

图 7. 随机森林回归算法模型

(二) 指标权重

本文综合以上算法进行了各特征权重计算，显示了各指标对模型贡献的重要程度，其加和值为 1，如表 8 所示。

Table 8. Weight table of each indicator after feature selection

表 8. 特征选择后的各指标权重表

指标	重要性程度
人口自然增长率(%)	4.10%
森林覆盖率(%)	3.95%
单位地区生产总值电耗	3.48%

## Continued

国际旅游收入占地区生产总值比重(百万美元/人民币)	3.17%
R&D 经费支出占地区生产总值比重	3.10%
公共安全财政支出占财政总支出	2.97%
国际旅游外汇收入(百万美元)	2.91%
人均拥有公共图书馆藏量	2.84%
接待国际游客(百万人次)	2.84%
电力消费量(亿千瓦时)	2.82%
人均水资源量(立方米/人)	2.76%
城镇化水平	2.72%
农村宽带接入用户(万户)	2.71%
人口密度	2.71%
公共服务支出占地区生产总值比重	2.68%
道路长度(万公里)	2.67%
参加失业保险率	2.61%
城乡居民社会养老保险实际领取待遇率	2.57%
规模以上工业企业 R&D 人员全时当量(人年)	2.54%
公共汽电车运营数(辆)	2.51%
氮氧化物排放量(万吨)	2.51%
全体居民人均消费支出	2.39%
教育经费占地区生产总值比重	2.29%
技术市场成交额占地区生产总值比重	2.15%
国内实用新型专利申请授权量(项)	2.09%
农村居民家庭恩格尔系数(%)	2.08%
城市燃气普及率(%)	2.05%
建成区绿化覆盖率	2.00%
城镇基本医疗保险年末参保率	1.98%
外商投资占比	1.93%
泰尔指数	1.93%
年度可支配收入	1.80%
城镇居民家庭恩格尔系数(%)	1.78%
每万人人均拥有卫生人员数	1.64%
高等教育人数比例	1.49%
博物馆参观人次(万人次)	1.46%
电视节目综合人口覆盖率(%)	1.45%
二氧化硫排放量(万吨)	1.43%
氨氮排放量(万吨)	1.42%
艺术表演团体演出场次(万场次)	1.09%
货物进出口总额占地区生产总值比重(万/亿)	1.04%

## Continued

每万人国内实用新型专利申请授权量(项)	0.95%
生活垃圾无害化处理率	0.89%
化学需氧量排放量	0.82%
人均地区生产总值增速	0.67%

## (三) 测度分析结果

本文将特征选择后的指标首先通过熵权法进行计算权重, 结果如表 9 所示。同时, 计算经特征选择后的各个维度即二级指标的权重, 结果如表 10 所示。

Table 9. Weight table of four indexes after feature selection

表 9. 基于特征选择后的四级指标权重表

项	信息熵值 $e$	信息效用值 $d$	权重(%)
人口自然增长率	0.977	0.023	13.876
城镇化水平	0.988	0.012	7.332
规模以上工业企业 R&D 人员全时当量	0.878	0.122	74.973
人口密度	0.994	0.006	3.819
城市燃气普及率	0.997	0.003	0.538
道路长度	0.939	0.061	12.133
公共汽电车运营数	0.949	0.051	10.105
农村宽带接入用户	0.907	0.093	18.494
城乡居民社会养老保险实际领取待遇率	0.983	0.017	3.318
参加失业保险率	0.938	0.062	12.31
公共服务支出占地区生产总值比重	0.891	0.109	21.642
公共安全财政支出占财政总支出	0.901	0.099	19.676
农村居民家庭恩格尔系数	0.991	0.009	1.784
全体居民人均消费支出	0.971	0.029	6.048
R&D 经费支出占地区生产总值比重	0.949	0.051	10.594
人均拥有公共图书馆藏量	0.937	0.063	13.217
国内实用新型专利申请授权量	0.877	0.123	25.74
技术市场成交额占地区生产总值比重	0.853	0.147	30.634
育经费占地区生产总值比重	0.934	0.066	13.767
森林覆盖率	0.961	0.039	8.872
建成区绿化覆盖率	0.997	0.003	0.698
电力消费量	0.955	0.045	10.328
人均水资源量	0.667	0.333	76.002
单位地区生产总值电耗	0.993	0.007	1.639
氮氧化物排放量	0.989	0.011	2.462
国际旅游外汇收入	0.883	0.117	37.999
国际旅游收入占地区生产总值比重	0.942	0.058	18.962
接待国际游客	0.867	0.133	43.039

由表 9 可知,规模以上工业企业 R&D 人员全时当量、公共服务支出占地区生产总值比重、技术市场成交额占地区生产总值比重、人均水资源量、接待国际游客等各个维度中权重第一的指标与未经特征选择的大概一致,说明本文通过特征选择有效地剔除了冗余特征。

**Table 10.** Weight table of secondary indicators after feature selection

**表 10.** 基于特征选择后的二级指标权重表

项	信息熵值 $e$	信息效用值 $d$	权重(%)
人口规模现代化	0.94	0.06	17.846
全体人民共同富裕的现代化	0.94	0.06	17.91
物质文明和精神相协调现代化	0.941	0.059	17.528
生态文明现代化	0.926	0.074	22.103
走和平发展的现代化道路	0.917	0.083	24.613

由表 10 可知,与全指标相比,各个二级指标的权重占比差异更小,说明通过特征选择后更有效地选取了具有代表性的特征。结合各个二级指标对各省市进行计算综合得分,如表 11 所示。

**Table 11.** Comprehensive score table of provinces and cities after feature selection

**表 11.** 基于特征选择后的各省市综合得分表

地区	人口规模现代化		全体人民共同富裕的现代化		生态文明现代化		物质文明和精神相协调现代化		走和平发展的现代化道路		综合评价	
广东省	0.717	1	0.365	2	0.196	2	0.243	4	0.892	1	0.607	1
西藏自治区	0.174	23	0.364	3	0.777	1	0.171	7	0.104	16	0.428	2
江苏省	0.595	2	0.375	1	0.131	11	0.216	6	0.155	10	0.364	3
浙江省	0.517	3	0.271	5	0.184	3	0.221	5	0.193	6	0.326	4
北京市	0.202	13	0.216	8	0.115	19	0.477	1	0.256	3	0.316	5
山东省	0.402	4	0.341	4	0.123	14	0.124	14	0.13	13	0.267	6
上海市	0.211	12	0.128	28	0.077	28	0.345	2	0.334	2	0.241	7
福建省	0.292	5	0.168	15	0.176	4	0.097	20	0.246	5	0.199	8
天津市	0.192	14	0.129	27	0.064	30	0.263	3	0.185	7	0.166	9
安徽省	0.264	7	0.175	14	0.104	24	0.136	11	0.139	11	0.16	10
湖北省	0.243	8	0.183	11	0.124	13	0.137	10	0.107	15	0.158	11
四川省	0.189	15	0.229	6	0.134	10	0.1	19	0.083	19	0.152	12
广西壮族自治区	0.185	17	0.157	16	0.161	5	0.078	24	0.178	8	0.143	13
云南省	0.165	25	0.135	26	0.152	7	0.072	28	0.248	4	0.141	14
辽宁省	0.15	28	0.185	10	0.116	18	0.121	15	0.134	12	0.138	15
陕西省	0.178	20	0.141	21	0.118	17	0.141	9	0.158	9	0.137	16
湖南省	0.239	9	0.178	13	0.141	9	0.076	27	0.083	20	0.132	17
河南省	0.269	6	0.209	9	0.104	23	0.071	29	0.04	25	0.13	18



Continued

河北省	0.223	10	0.225	7	0.107	22	0.067	31	0.033	27	0.125	19
江西省	0.213	11	0.138	24	0.158	6	0.102	18	0.051	24	0.113	20
新疆维吾尔自治区	0.185	18	0.183	12	0.084	26	0.095	22	0.053	22	0.105	21
重庆市	0.185	19	0.137	25	0.119	16	0.078	25	0.117	14	0.104	22
海南省	0.177	21	0.125	30	0.143	8	0.086	23	0.098	17	0.102	23
青海省	0.176	22	0.142	19	0.115	20	0.151	8	0.009	29	0.1	24
贵州省	0.162	26	0.141	22	0.125	12	0.115	16	0.023	28	0.091	25
黑龙江省	0.118	30	0.155	17	0.12	15	0.095	21	0.052	23	0.088	26
吉林省	0.109	31	0.142	20	0.11	21	0.111	17	0.065	21	0.085	27
内蒙古自治区	0.152	27	0.128	29	0.104	25	0.068	30	0.089	18	0.078	28
甘肃省	0.146	29	0.152	18	0.066	29	0.129	12	0.003	31	0.075	29
山西省	0.168	24	0.14	23	0.082	27	0.077	26	0.037	26	0.07	30
宁夏回族自治区	0.188	16	0.105	31	0.055	31	0.129	13	0.007	30	0.06	31

Table 12. Comprehensive score table of the three major zones after feature selection

表 12. 特征选择后的三大地带综合得分表

地带	人口规模现代化	全体人民共同富裕的现代化	生态文明现代化	物质文明和精神相协调现代化	走和平发展的现代化道路	综合评价
东部	0.333	0.229	0.13	0.205	0.241	0.258
中部	0.203	0.165	0.118	0.1	0.072	0.117
西部	0.174	0.168	0.168	0.11	0.088	0.134

由表 11 和表 12 可知，与全指标测度分析结论基本一致，其中广东省的现代化程度最高，东部地带的现代化程度综合最高。

### 5. 结论

综合上述，本文中广东省现代化程度最高，西藏自治区也比较靠前，主要因为人均资源较为丰富、国家政策帮扶等因素，而宁夏回族自治区现代化程度较低。而我国现代化水平总体差异主要来源于地区间的差异。在三大地带中，东部地带现代化程度最高，而中西部现代化程度近似。本文还考虑了指标过多是否会影响分析，因此结合随机森林等方法进行特征选择，基于选择后的特征再次进行分析，结果显示，与全指标测度分析结果基本一致，一方面说明了这些指标中确实存在冗余，另一方面也再次论证了本文测度分析结果的稳定性。

### 参考文献

- [1] 回宇. 国家治理现代化视域下思想政治教育现代化研究[D]: [博士学位论文]. 长春: 吉林大学, 2022: 200.
- [2] 刘军. 当代人类文明视域下中国特色社会主义道路研究[D]: [博士学位论文]. 济南: 山东师范大学, 2022: 255.
- [3] 马金保. 中国特色社会主义民生理论研究[D]: [博士学位论文]. 西安: 西北大学, 2019: 148.

- 
- [4] 林炳华. 城镇化、结构转型与政府公共投资[M]. 北京: 中国人民大学出版社, 2017: 462.
- [5] 张媛新. 乡村振兴战略下延寿县乡村治理能力提升研究[D]: [硕士学位论文]. 哈尔滨: 东北农业大学, 2021: 60.
- [6] 卢艳香. 中国式现代化的五重维度探析[J]. 安徽师范大学学报(人文社会科学版), 2023, 51(3): 21-27.
- [7] 张甜, 阎红灿. 基于熵权法的多属性决策算法优化及应用[J]. 华北理工大学学报(自然科学版), 2022, 44(1): 82-88.
- [8] 李艳虹. 基于观点演化的群决策研究及应用[D]: [博士学位论文]. 成都: 西南财经大学, 2021: 185.
- [9] 张卫贞. 基于核密度估计的系统实时剩余寿命预测研究[D]: [博士学位论文]. 太原: 太原科技大学, 2021: 129.
- [10] 余丹. 中国经济不平等研究[D]: [博士学位论文]. 武汉: 中南财经政法大学, 2021: 191.
- [11] 王曦璟. 中国多维不平等的测度及关联性研究[D]: [博士学位论文]. 太原: 山西财经大学, 2017: 159.
- [12] 殷星. 基于 PTN 的网络优化算法研究[D]: [硕士学位论文]. 武汉: 武汉邮电科学研究院, 2021: 75.
- [13] 曹稳. 基于 XGBoost 算法的多因子选股实证研究[D]: [硕士学位论文]. 济南: 山东财经大学, 2021: 68.
- [14] 吕晓玲, 宋捷. 大数据挖掘与统计机器学习[M]. 北京: 中国人民大学出版社, 2016: 239.