

教育公平的多维测度与资源配置研究

——基于四川省各地级市的实证分析

王文娟¹, 李思潜¹, 闵夕芮², 张诗雨¹, 梁雯¹, 刘芋良², 屈小兵^{1*}

¹乐山师范学院数学与统计学院, 四川 乐山

²乐山师范学院物理与光电工程学院, 四川 乐山

收稿日期: 2025年10月27日; 录用日期: 2025年11月18日; 发布日期: 2025年12月1日

摘要

基于社会经济的快速增长和人们对高质量生活的强烈需求, 城乡以及各地区城市间教育公平问题日益严峻。四川省作为国家区域协调发展战略的关键实施区域, 面临着教育资源配置效率的差异以及地方经济发展状况的不同。为了有效应对和优化教育资源分配不均的问题, 本研究以四川21个市州的面板数据为研究对象, 采用DEA-EM模型、在多区域分析中引入双向固定效应, 并运用模拟退火算法优化DEA效率值的权重分配, 避免陷入局部最优解, 结合K折交叉验证提升模型稳健性, 突破了传统线性分析的局限性, 为教育公平研究提供了更为动态和精准的分析框架。

关键词

教育公平, 资源配置效率, DEA-EM模型, 双向固定效应, 模拟退火算法

A Multidimensional Measurement and Resource Allocation Study on Educational Equity

—An Empirical Analysis Based on Prefecture-Level Cities in Sichuan Province

Wenjuan Wang¹, Siqian Li¹, Xirui Min², Shiyu Zhang¹, Wen Liang¹, Yuliang Liu², Xiaobing Qu^{1*}

¹School of Mathematics and Statistics, Leshan Normal University, Leshan Sichuan

²School of Physics and Optoelectronic Engineering, Leshan Normal University, Leshan Sichuan

Received: October 27, 2025; accepted: November 18, 2025; published: December 1, 2025

*通讯作者。

文章引用: 王文娟, 李思潜, 闵夕芮, 张诗雨, 梁雯, 刘芋良, 屈小兵. 教育公平的多维测度与资源配置研究[J]. 统计学与应用, 2025, 14(12): 51-61. DOI: 10.12677/sa.2025.1412345

Abstract

With the rapid growth of the social economy and people's strong demand for a high-quality life, the issue of educational equity between urban and rural areas and among cities in various regions has become increasingly severe. As a key implementation area of the national regional coordinated development strategy, Sichuan Province is confronted with differences in the efficiency of educational resource allocation and the varying conditions of local economic development. To effectively address and optimize the uneven distribution of educational resources, this study takes the panel data of 21 prefectures and cities in Sichuan as the research object, adopts the DEA-EM model, introduces two-way fixed effects in multi-regional analysis, and uses the simulated annealing algorithm to optimize the weight distribution of DEA efficiency values to avoid getting stuck in local optimal solutions. It also combines K-fold cross-validation to enhance the robustness of the model, breaking through the limitations of traditional linear analysis and providing a more dynamic and precise analytical framework for educational equity research.

Keywords

Educational Equity, Resource Allocation Efficiency, DEA-EM Model, Bidirectional Fixed Effects, Simulated Annealing Algorithm

Copyright © 2025 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

在经济快速增长下，人们对城乡教育公平愈发关注，在关于教育资源分配的研究中，姚文佳、李建华[1]提出实现教育资源的分配正义，首先需实现基础教育资源分配的均等化，通过教育资源的分配保障受教育者选择的自由。盛艳玲[2]等通过构建以各居住小区教育质量方差最小化为目标的函数模型，结合ArcGIS网络分析技术，提出优化方案，提出实现均衡发展为目标的“多校划片”政策建议。吴云勇、姚晓林[3]发现经济环境是教育政策目标得以实现的物质基础；政治环境决定了教育政策改革的性质及发展方向；文化环境是教育政策改革有效实施的非正式制度约束。刘婧[4]分别从建设课程资源服务平台、实现数字教育资源供给模式、创新授课模式等方面对高等教育资源进行合理化分配，以加速高等教育的全面发展，促进教育资源更加合理化应用等。这些研究在基础教育均等化、空间资源优化、政策环境分析等方面为构建差异化的教育公平测度框架与政策优化提供了重要参考。

研究主要选取财政性教育经费、学校数量、师生比、专任教师数、在校生数作为影响教育公平程度的主要特征，同时控制政治特征、家庭特征、人口密度以及其他特征，保证研究的准确性与可行性。

本文引入DEA-EM模型和多层级交互建模，对四川21个市州的面板数据进行教育资源配置优化。选取交叉验证与贝叶斯信息准则优化模型，将DEA效率分析与EM聚类算法相结合，并引入模拟退火算法对DEA效率值进行全局优化，以多区域分析与双向固定效应模型解决地区和时间层面的偏差问题，对教育资源配置进行优化模拟。

2. 基本理论

本文依托面板数据，整合数据包络分析(DEA)、期望最大化(EM)聚类算法，构建DEA-EM融合模型，

并结合多区域分析与双向固定效应、K 折交叉验证及模拟退火算法构建教育资源配置优化理论框架，以精准评估资源利用效率与分配公平性。

2.1. DEA-EM 融合模型

DEA-EM 融合模型是整合 DEA 效率评价与 EM 聚类分析的复合型模型，可同时评估多决策单元 (DMUs) 的资源利用效率，并依据效率值对决策单元聚类分组[5]。其中，DEA 应用 CCR 模型的线性规划形式计算效率值，具体如下：

目标函数：

$$\max \theta. \quad (1)$$

约束条件：

$$\text{s.t.} \begin{cases} \sum_{j=1}^n \lambda_j x_{ij} \leq x_{i_0}, \\ \sum_{j=1}^n \lambda_j y_{rj} \geq \theta y_{r_0}, \\ \lambda_j \geq 0. \end{cases} \quad (2)$$

上式(2)中， i 为输入变量， r 为输出变量， θ 表示决策单元的相对效率值 ($0 \leq \theta \leq 1$)； x_{ij} 和 y_{rj} 分别表示第 j 个决策单元的第 i 项输入和第 r 项输出； λ_j 反映第 j 个决策单元对当前单元的贡献。 i_0 、 r_0 分别对应被评估决策单元 (DMU) 的输入与输出。

通过 EM 算法对 DEA 效率结果进行聚类分析，假设存在 k 个潜在类别，步骤如下：

Step 1: 计算后验概率 (E 步)

$$\gamma(z_{nk}) = \frac{\pi_k \cdot N(\theta_n / \mu_k, \sigma_k^2)}{\sum_{j=1}^k \pi_j \cdot N(\theta_n / \mu_j, \sigma_j^2)}. \quad (3)$$

Step 2: 更新参数 (M 步)

$$\begin{cases} N_k = \sum_{n=1}^N \gamma(z_{nk}), \\ \pi_k = \frac{N_k}{N}, \\ \mu_k = \frac{1}{N_k} \sum_{n=1}^N \gamma(z_{nk}) \theta_n, \\ \sigma_k^2 = \frac{1}{N} \sum_{n=1}^N \gamma(z_{nk}) (\theta_n - \mu_k)^2. \end{cases} \quad (4)$$

上述公式(3)、(4)中， θ_n 为第 n 个区县的 DEA 效率值； π_k 、 μ_k 、 σ_k 分别为第 k 类别的混合系数、均值和标准差； N_k 表示第 k 类混合系统中的样本数量， $\gamma(z_{nk})$ 是样本 n 属于第 k 类的概率或权重。

2.2. 多区域分析与双向固定效应

多区域分析聚焦不同区域 (i) 在不同时间 (t) 的教育资源配置差异，通过横纵向对比捕捉区域异质性；双向固定效应模型纳入区域与时间固定效应，消除地区差异和时间趋势干扰，借助面板数据回归精准估计变量对教育资源配置的影响[6]，核心回归公式为：

$$Y_{it} = \alpha + \sum_k \beta_k X_{kit} + \mu_i + \lambda_t + \varepsilon_{it}.$$

(5)

2.3. *K* 折交叉验证

研究选择 *K* 折交叉验证将数据集均匀地分成 *K* 个子集，每个子集轮流作为测试集，其余 *K* − 1 个子集作为训练集；对于第 *i* 折，在训练集 *D* − *D_i* 上训练模型，在验证集上 *D_i* 进行验证，得到该折的均方误差 *MSE_i*，经过 *K* 次折叠验证后，计算平均均方误差 *MSE_{avg}* 作为模型性能的评估指标以及均方误差的标准差 *R²*，评估模型拟合与预测性能，减少数据划分偶然性误差，保障模型稳定性，核心计算公式为：

$$MSE_{avg} = \frac{1}{k} \sum_{i=1}^k MSE_i; i = 1, 2, \cdots, k.$$

(6)

$$R^2 = \sqrt{\frac{\sum_{i=1}^k (MSE_i - MSE_{avg})^2}{k - 1}}.$$

(7)

如表 1 所示，模型在交叉验证中表现良好，平均 *R²* 达到 0.742，能较好地解释数据的变异；平均 *RMSE* 仅为 2.554，预测误差较小。整体来看，模型既具备较好的拟合能力，表现出稳定的预测性能，能够在实际应用中提供较为可靠的结果。

Table 1. Model fitting degree and prediction effect under different folds of cross-validation
表 1. 不同折交叉验证下的模型拟合程度和预测效果

模型参数	验证结果
<i>R²</i> scores	[0.76, 0.73, 0.78, 0.71, 0.73]
RMSE scores	[2.74, 2.31, 2.49, 2.39, 2.84]
平均 <i>R²</i>	0.742
平均 RMSE	2.554

2.4. 模拟退火算法

如图 1 所示，模拟退火算法(SA)是一种用于全局优化的随机算法，能够有效避免陷入局部最优解，从而在更广泛的范围内搜索最优解。在 DEA 模型中，该算法被用于寻找最优权重或优化效率值，尤其在比较多决策单元(DMUs)的效率时，能够充分发挥其优化能力[7]。

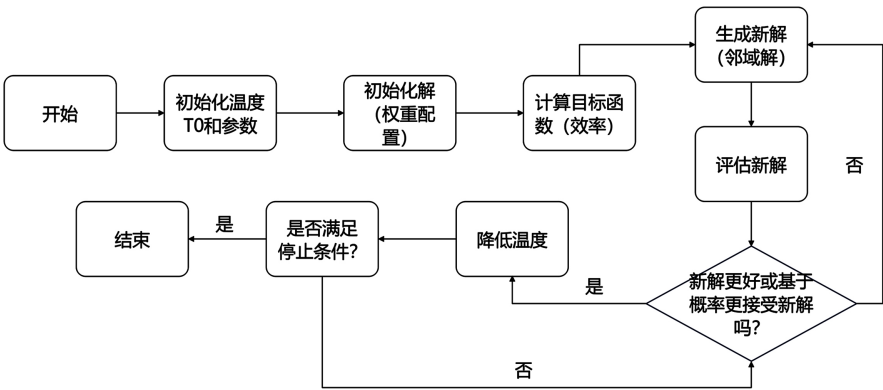


Figure 1. Flowchart of simulated annealing algorithm
图 1. 模拟退火算法流程图

3. 数据来源及模型应用

3.1. 数据来源

本研究采用 2017~2023 年的教育资源数据，主要来源于国家数据、国家统计局、四川省统计局、中国互联网数据平台及四川各地统计局。部分地区数据如下(表 2)。

Table 2. Data of educational resources in various regions

表 2. 各地区教育资源数据

区县	财政性教育经费	对数标准化	师生比	学校数量	专任教师数(人)	资金投入(亿元)
西昌市	19.361	2.963	0.051	147	8178	0.0601
阿坝县	3.806	1.337	0.045	28	1296	1.2103
天全县	2.634	0.968	0.091	19	1279	2.6857
会理市	4.661	2.671	0.065	71	3940	3.0162
名山区	3.459	1.241	0.072	35	1885	3.4745
康定市	3.668	1.299	0.083	30	1685	3.5631
雨城区	4.094	1.409	0.076	39	3111	4.1142
高坪区	8.387	2.127	0.075	67	4617	8.3865
阆中市	10.302	2.332	0.094	82	5156	8.5112
江油市	8.862	2.182	0.064	56	4418	9.0299
三台县	14.434	2.671	0.077	118	7562	12.8603
仪陇县	13.446	2.599	0.089	101	7108	13.2044
金堂县	6.406	1.857	0.072	88	5800	14.3002
锦江区	17.893	2.884	0.075	56	7061	17.8701
武侯区	20.571	2.984	0.079	76	8257	23.6031

3.2. 数据标准化

对选定特征数据进行标准化处理，消除各变量量纲与数值范围差异，保障模型分析中变量贡献的公平性、准确性与科学性。研究采用对数标准化，将大范围数据压缩到较小区间，使数据更集中，便于分析处理。

图 2 展示了我们对各个城市的数据集进行对数标准化处理后各种特征的密度分布图。经过对数标准化后数据密度分布更接近正态，同时弱化了极端值影响，降低了高方差数据中异常值的干扰，故处理后的数据对分布提取有一定的帮助。

3.3. 数据概论

从数据图 3 中看出，各地财政性教育经费波动区域投入差距大，学校数量虽有波动，但多数地区差异不大，这意味着尽管学校数量较均衡，仍可能因财政经费与师资不同，导致部分地区教育资源无法满足学生需求；资金投入与财政性教育经费关联密切，如凉山、甘孜等地区投入较高，而部分地区资金投入少，可能引发教育设施陈旧、师生比失衡等问题，影响教育质量与公平；专任教师数量对教育公平影响显著，即便部分地区财政经费充足，师资不足仍可能造成教育质量差异[8]。

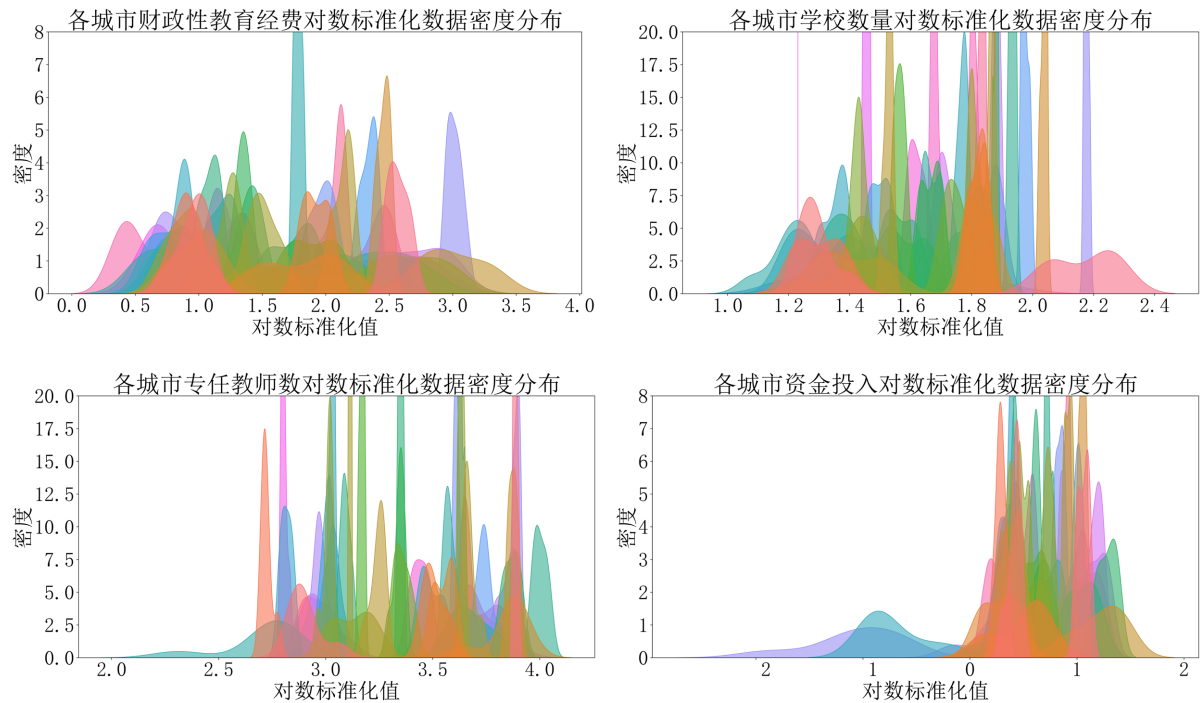


Figure 2. Density distribution of data from each city after logarithmic standardization
图 2. 各市数据对数标准化后密度分布

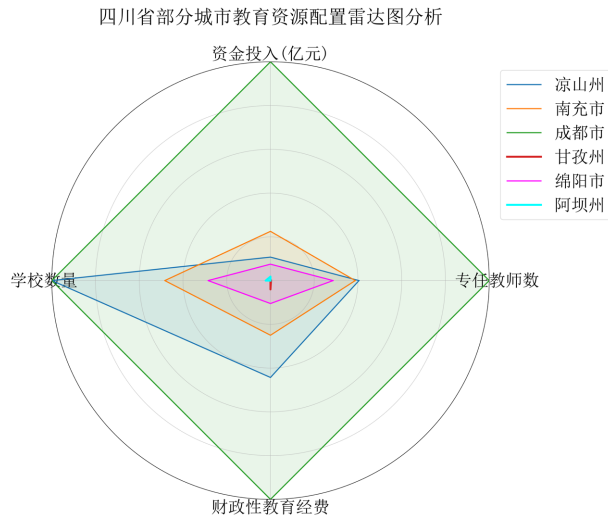


Figure 3. The different impact of the characteristics of educational investment on educational equity in various regions
图 3. 教育投入特征对各个地区的教育公平产生的不同影响

3.4. 模型应用

随机森林分析教育资源特征的相关性,揭示了四川各市教育资源分配的关键特征重要性差异显著[9]。研究分别以不同的特征为因变量,探究财政性教育经费、学校数量、资金投入以及师生比等特征对教育系统关键指标的影响程度。图 4 为特征重要性 3D 图,从紫色到黄色,特征重要性数值由 0.0 (重要性最低)逐步递增至 1.0 (重要性最高),反映出四川教育资源在城市间、特征间存在不均衡分配的状况,为深入探究四川教育资源公平提供了丰富且直观的数据线索与研究方向[10]。

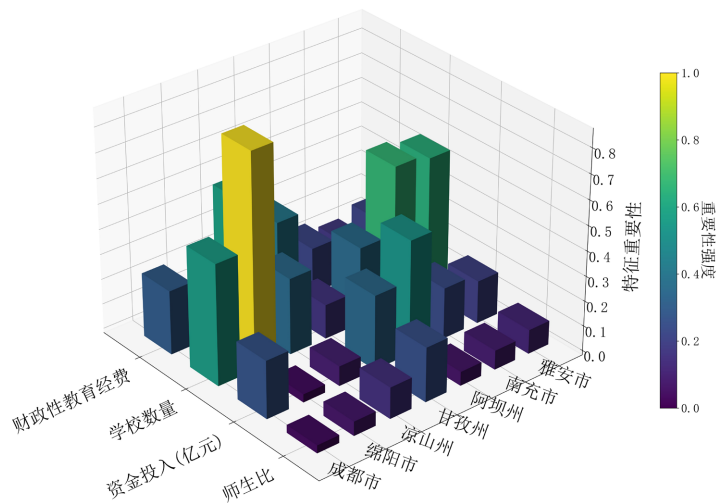


Figure 4. 3D map of the importance of features in each city
图 4. 各城市特征重要性 3D 图

将 DEA-EM 模型应用到实际问题数据中，通过 DEA 效率值分布、地区聚类效率以及在不同聚类类别中的分布，从直方图、核密度估计(KDE)线等图中可以清晰看出四川省教育资源分配情况。

如图 5 所示，DEA 效率值大多数集中在较低的区间(0.0~0.4)，尤其是效率值接近 0.2 时聚集最多。表明这些区域在某些方面的效率较差，大部分区域的教育投入和产出并不充分，效率较低；同时每个聚类类别的效率分布差异显著，这种分布表示不同类别的地区效率差异较大[11]。

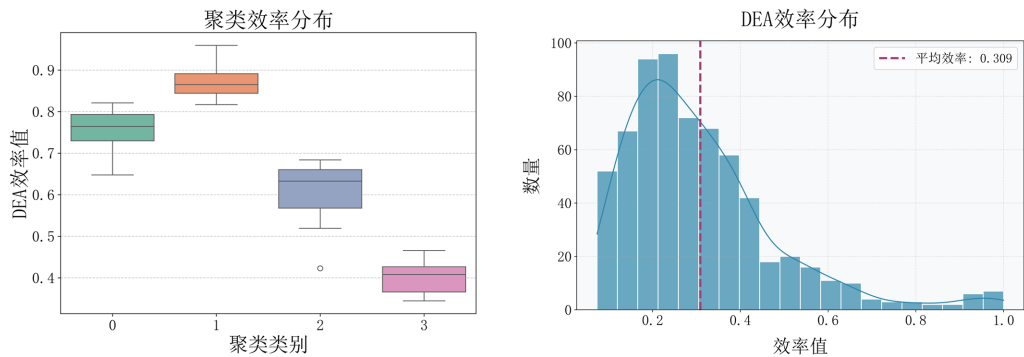


Figure 5. DEA efficiency values and cluster efficiency distribution
图 5. DEA 效率值和聚类效率分布

表 3 中所展示出的聚类结果较为合理，整体效率分布离散，且大量样本集中于低效率区间，四个聚类的效率分布区间界限相对分明。其中，聚类 3 (高效集群)的分布位置最高且跨度较大；而聚类 2 (低效集群)则紧密分布在 0.2 水平线以下；聚类 1 和聚类 0 的分布存在部分重叠，但聚类 1 的整体中位水平明显高于聚类 0。

Table 3. Analysis of clustering features
表 3. 聚类特征分析

Cluster	DEA 效率				资金投入(亿元)	师生比
	mean	std	min	max		
0.0	0.347	0.051	0.265	0.460	2.920	0.060

续表

1.0	0.564	0.070	0.469	0.751	2.045	0.075
2.0	0.180	0.050	0.073	0.265	8.887	0.070
3.0	0.912	0.080	0.761	1.000	1.545	0.087

我们得出系统性结论：资源投入规模并非效率的决定性因素，单纯增加资金投入并不能保证效率提升，甚至可能因管理不善而导致效率损失；高效集群的极高师生比也是显著的特征，暗示了优化人力资本结构、提升人力资源利用效率可能是驱动效率改善的核心因素；对大多数低效和中低效单元，改进重点不应局限于争取更多资源，而应转向优化现有资源配置结构，尤其是提升人力资源的利用效率，并着力消除资源冗余。

通过残差与预测值以及实际值与模型预测值之间的关系分析 DEA 模型的效果和拟合情况[12]。理想的残差图应该表现为随机分布，且残差的平均值应为零，红色虚线是理想的完美预测线，模型的预测效果越接近这条线，则模型越准确(图 6)。

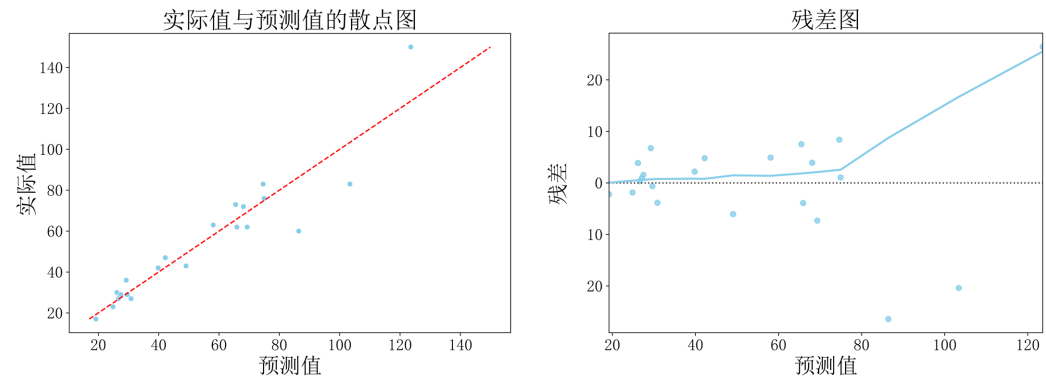


Figure 6. Visualization of regression analysis
图 6. 回归分析可视化图

通过拟合结果发现，而研究大部分点靠近红色预测线，回归模型的预测效果较好，实际值和预测值大体一致，但也存在一些偏差较大的异常点，随着预测值增大，误差或残差也增大，表明模型在较大值的预测上可能出现问题。

为提高模型的准确性，采用模拟退火算法对模型进行优化，如图 7 所示。

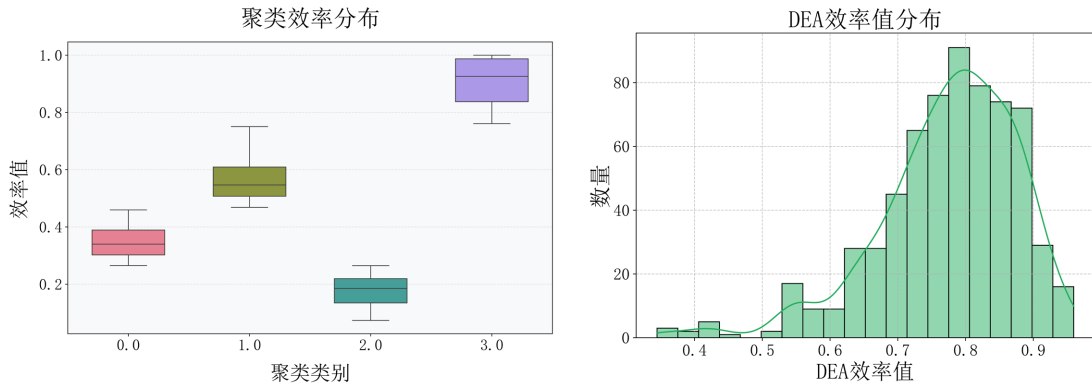


Figure 7. The results of DEA-EM model optimization
图 7. DEA-EM 模型优化后结果图

原始平均效率约为 0.3, 优化后平均效率提升至约 0.75, 说明优化措施对提高效率有积极作用。不同聚类类别的 DEA 效率值分布差异明显, 类别 0 的效率值整体较高, 集中在 0.3~0.4 左右; 类别 1 效率值次之, 集中在 0.6~0.7 左右; 类别 2 效率值整体偏低; 类别 3 效率值集中在 0.8~1.0 左右, 这反映出不同聚类群体在资源利用等方面的效率存在显著差别。从直方图和拟合曲线看, DEA 效率值主要集中在 0.7~0.8 之间, 大部分对象的效率处于中等水平, 同时也存在少量效率较低(如接近 0.4)和较高(如接近 0.9)的情况。

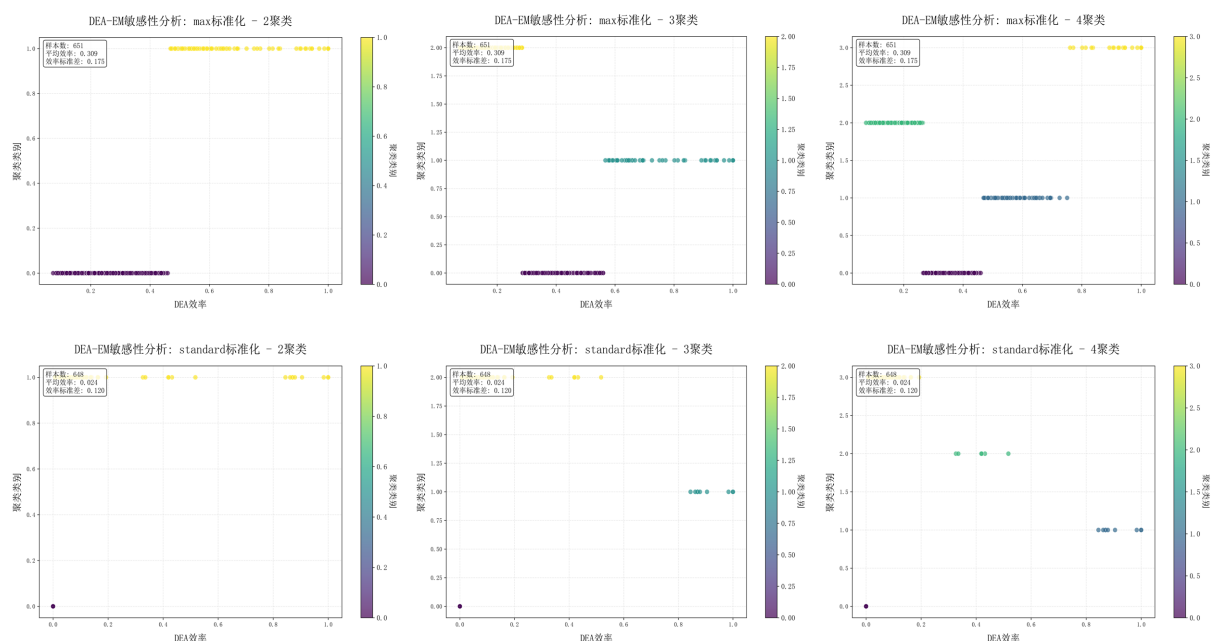
优化前效率值分布较广且有较多较低值, 优化后效率值集中在较高区间, 验证了优化的有效性。通过一系列分析可知, 基于模拟退火优化的聚类及资源分配等操作, 有效提升了教育资源相关的 DEA 效率, 且不同聚类类别在效率表现上存在明显差异, 可为后续针对不同类别制定差异化资源优化策略提供依据。

3.5. 模型检验

通过敏感性分析发现, DEA-EM 分析的敏感性较强。max、standard、minmax 这三种标准化方式大幅改变了 DEA 效率的整体特征(如平均效率、标准差、分布形态等), 进而明显影响了聚类类别与 DEA 效率之间的关联模式。无论是 2 类、3 类还是 4 类的聚类划分, 聚类数的变化会直接调整对 DEA 效率细分的“颗粒度”, 聚类数越多, 效率与类别间的对应关系就越精细, 不过也可能存在过拟合风险, 这使得结果会随着聚类数的改变而有较显著的变化。

这样的结果意味着, 在实际应用 DEA-EM 方法时, 标准化方法以及聚类数的选择会对最终分析结果产生较大影响, 故需要结合具体的研究目的和数据特点, 谨慎地去选取合适的标准化方法与聚类数, 这样能够让分析结果更具有准确性与可靠性。

比较图 8 不同聚类数设置下样本的聚类标签, 可以发现多数样本在不同参数设置下仍被归类于逻辑一致的簇中, 聚类边界随簇数增加而细化, 而非剧烈变化, 进一步印证了模型在 DEA 效率维度上的聚类是稳定且合理的。在不同的聚类数设置下, 模型仍能保持一致的结果, 这表明研究使用的模型具备较强的可靠性。



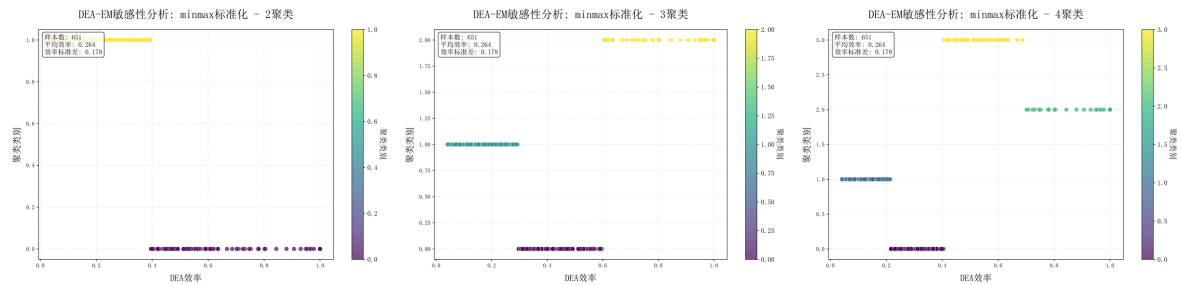


Figure 8. Sensitivity analysis results of the DEA-EM fusion model
图 8. DEA-EM 融合模型敏感性分析结果

4. 结论与建议

4.1. 研究结论

- 1) 经权重分析，学校数量、财政教育经费投入及专任教师数量三项指标权重占比较为突出，在教育资源配置中发挥核心支撑作用。这提示政策制定者在优化资源分配策略时应参考以上三个指标以提升教育资源投入的实效性。
- 2) 区域间教育资源配置效率呈现显著差异，经济欠发达地区(如凉山州)尤为突出。同时城乡过渡地带及流动人口集中区域，存在“基础设施短缺－师资稀释－教育质量滑坡”的恶性循环，优化学校布局、动态调控师生比是破解关键。
- 3) 教育资源分配缺乏精确性，导致资源的浪费和不公平。比如凉山这些地区的财政投入相对较大，但由于资源分配的不精确和滞后性，教育资源未能按需分配，可能导致了资金浪费和地区之间的教育质量差距。
- 4) 四川部分县域教育资源投入与产出效率不匹配，如凉山州等地虽资金投入较高，但 DEA 效率值偏低(见表 3)，可能存在资金使用效率低下、配套机制不完善等问题，这制约了教育资源质量的空间均衡提升。

4.2. 政策建议

- 1) 财政投入从“人均均摊”转向“按需倾斜”，建立资源倾斜机制，优先向配置低效区域定向投放资源，引入“教育机会平等系数”(结合教育基尼系数、平均受教育年限)动态调节，依托数字化技术搭建“教育资源共享平台”；加大对凉山州等资源紧张地区的转移支付，强化资金绩效监管与透明化。
- 2) 构建区域师资流动与数字资源共享机制，将跨区任教纳入职称评定、科研支持、专项补贴体系；组建教育协作体，通过远程教学、在线课程等形式，推进课程共建与直播共享，实现优质资源向薄弱地区常态输送。
- 3) 按区域差异制定分层配置政策：资源薄弱地区重点完善校舍、强化师资培训、调控师生比；加大贫困地区财政投入，统筹资金用于校舍维护与智慧设备引入，构建城乡一体数字教育生态。

4.3. 研究局限

- 1) 所用模型均为量化工具，高度依赖数据质量，假设简化了动态复杂性和参数设置敏感的现实，导致结果与显示存在一定偏差，一定程度上制约了模型应用。
- 2) 教育公平不应该止步于资源层面，更应该延伸至教育机会、教学过程与最终结果的全维度公平。研究主要聚焦于财力、物力与人力等资源的配置效率，对教学质量、课程内容、学生参与度及综合素质发展等非资源性维度的测度尚显不足，而“软性”因素也是影响教育结果公平的关键之一。

3) 学界对教育基尼系数理论适用性仍有异议:一方面,该系数无法完全反映教育质量差异(如教师学历、课程设置);另一方面,教育公平的“多维性”(如机会、过程、结果公平)难以通过单一指标全面展现。

基金项目

中国高等教育学会“2024年高等教育科学研究规划课题”(24SX0203),2025年国家级大学生创新创业训练计划项目(202510649056)。

参考文献

- [1] 姚文佳, 李建华. 教育资源的分配正义——基于罗尔斯的视角[J]. 现代大学教育, 2023, 39(2): 13-18.
- [2] 盛艳玲, 马瑞苗, 乔纪纲. 不同入学政策教育资源分配及优化路径——以广州市海珠区为例[J]. 教育学报, 2022, 18(4): 98-113.
- [3] 吴云勇, 姚晓林. 中国教育发展的政策环境影响与未来改革的总体要求[J]. 现代教育管理, 2021(5): 69-76.
- [4] 刘婧. 教育信息化 2.0 背景下高等教育资源配置探析[J]. 黑龙江高教研究, 2019, 37(10): 89-93.
- [5] Soltanifar, M., Ghiyasi, M., Emrouznejad, A. and Sharafi, H. (2024) A Novel Model for Merger Analysis and Target Setting: A CSW-Inverse DEA Approach. *Expert Systems with Applications*, **249**, Article 123326. <https://doi.org/10.1016/j.eswa.2024.123326>
- [6] 邵思奇. 具有固定效应的异质动态面板数据模型的两阶段最小二乘估计及检验[D]: [硕士学位论文]. 郑州: 郑州大学, 2023.
- [7] 江海, 资智洪, 袁杰. 我国跨区域协同创新下的科技成果转化效率提升路径研究[J]. 中国高校科技, 2024(8): 28-35.
- [8] Enchikova, E., Neves, T., Toledo, C. and Nata, G. (2025) A Long Road to Educational Equity: Tracking Trends through PISA 2000-2018. *International Journal of Educational Research Open*, **8**, Article 100445. <https://doi.org/10.1016/j.ijedro.2025.100445>
- [9] Gamburd, M.R. (2024) Sama in the Forest: Directed by Carlos G. Gómez, 2023, 75 Minutes, Color. Distributed by Documentary Educational Resources. *Visual Anthropology Review*, **40**, 55-57. <https://doi.org/10.1111/var.12317>
- [10] 谢甜. 教育事业统计指标体系构建及其在教育资源配置优化中的应用探讨[J]. 产业与科技论坛, 2025, 24(15): 224-226.
- [11] Bondi, S., Parsons, T.U. and Kanagala, V. (2025) Fostering Equity and Inclusion in Graduate Education: Key Strategies and Perspectives. Taylor & Francis.
- [12] Pang, F., Li, Y., Miao, G. and Shi, Y. (2025) Optimizing Education Resource Allocation Using Grey Model Forecasting of School Age Populations. *Scientific Reports*, **15**, Article No. 35108. <https://doi.org/10.1038/s41598-025-19182-5>