中国绿色创新效率的测度、因素探究及路径优化

张倩生，黄锦彬，林森峰，郑灿瑶

（广东外语外贸大学数学与统计学院，广州 510006）

摘要：以 2013—2023 年中国 30 个省份（因数据缺失，未包含西藏地区和港澳台地区）为样本，旨在测度各省份的绿色创新效率并识别其重要影响因素。采用超效率基于松弛测度数据包络分析（slack-based measure data envelopment analysis，SBM-DEA）模型测度各区域效率，结合第二代非支配排序遗传算法（non-dominated sorting genetic algorithm II，NSGA-II）模型分析非线性驱动机制，利用 NSGA-II 遗传算法求解全国绿色创新多目标优化模型。实证结果显示，中国绿色创新效率总体缓慢上升但区域差异显著。SHAP 值结果表明产业结构、人口密度与能源强度为关键影响因素，其中产业结构抑制效应最明显。通过多目标模型优化发现，合理配置我国各省市的资源权重可使全国绿色创新效率提升约 11.4%，同时实现 GDP 增长与碳减排效益，为制定地区差异化协同创新政策提供优化路径。

关键词：绿色创新效率；超效率 SBM-DEA；CatBoost；SHAP 可解释性算法；多目标优化

中图分类号：F222；F273；F224 **文献标志码：**A **文章编号：**1671－1807(2026)14－0055－10

在全球可持续发展战略不断深化的背景下，绿色创新是推动经济高质量增长与环境保护协同发展的关键驱动力。中国作为全球最大的发展中国家，已在绿色创新领域取得显著成效，但区域发展不平衡问题仍然突出。因此，在资源有限和环境容量受限的条件下，如何兼顾经济和社会、生态效益的协调，推动绿色创新效率整体提升及区域均衡发展，已成为中国经济高质量发展的重要任务。本文从效率视角系统评估中国绿色创新发展水平，揭示其重要影响因素和内在驱动机制，并探索提升绿色创新效率的有效路径，对于落实“双碳”目标和推动区域协调发展具有重要理论与现实意义。

目前，国内外学者对绿色创新效率的测度方法主要包括随机前沿分析（SFA）的参数法和数据包络分析（DEA）的非参数法[1-2]。数据包络分析因无须预设函数形式且能处理多维投入产出数据，在效率测度领域占据主导地位。栾蕾和张水平[3]采用基于径向和角度的传统 DEA 模型对资源效率进行静态评估，但此类模型因忽略非期望产出约束及非径向松弛问题，往往导致环境约束下的真实效率被扭曲。为克服此缺陷，Tone[4]提出基于松弛测度数据包络分析（slack-based measure data envelopment analysis，SBM-DEA）模型，实现了期望与非期望产出在非径向框架下的统一建模。近年来，学术界进一步运用超效率 SBM-DEA 模型来解决有效单元的排序“截断”问题[5]，并将其应用于减污降碳协同增效等高精尖行业的效率评估中。尽管已有学者利用该模型揭示了中国绿色创新效率的空间演进规律，但在面对数字经济驱动等复杂前因组态及区域异质性的深层路径优化方面，现有研究仍存在量化分析不足及动态配置欠缺等问题。

在厘清绿色创新效率的统计测度后，进一步探究其深层驱动机制至关重要。现有研究主要将影响因素归纳为产业结构、经济发展、金融深化、人口集聚、技术创新及能源强度六大核心维度[6-7]。进入 2025 年后，邢会和霍晓谦[8]从减污降碳协同增效视角，对微观尺度下的驱动机理进行更为精细的时空演化剖析。在实证策略上，现有研究多依赖采用空间计量、Tobit 模型、固定效应模型等线性回归范式[9]。尽管此类方法能揭示变量间的平均相关关系，但在处理复杂的绿色创新系统时面临显著的方法论瓶颈。一是线性假设不符合绿色创新系统的非线性、阶段性或交互效应特征；二是宏观变量间的共变趋势极易引发多重共线性干扰，导致参数估

收稿日期：2026-02-09

基金项目：国家自然科学基金项目（12171103）；广东省自然科学基金项目（2023A1515011354）

作者简介：张倩生（1975—），男，江西吉安人，博士，教授，硕士研究生导师，研究方向为技术经济决策；黄锦彬（1998—），男，广东云浮人，硕士研究生，研究方向为经济统计；林森峰（1997—），男，广东汕头人，硕士研究生，研究方向为经济统计；郑灿瑶（2000—），女，广东潮州人，硕士研究生，研究方向为发展经济学。

计不稳定。为克服上述问题，近年来机器学习方法凭借其强大的高维交互捕捉能力被引入效率预测领域。然而，传统机器学习内部决策过程的“黑箱”难题，使其难以满足政策制定所需的因果逻辑透明度[10]。在路径研究层面，模糊集定性比较分析(fsQCA)凭借处理复杂组态关系的优势得到广泛应用[11]。例如，2024—2025 年的研究多运用该方法探索数字经济驱动旅游业绿色转型的多元路径，以及生物医药企业创新效率的提升组态。然而，fsQCA 本质上属于静态定性分析，在面临区域资源禀赋差异显著且环境规制约束严格的情境下，单纯的组态分析难以实现动态的资源优化配置。由于绿色创新系统的优化伴随着效率最大化、经济效益最大化与能源减排最小化这一“不可能三角”的内生冲突，传统的单目标优化或线性加权方法已难以有效处理此类复杂的权衡关系。更深层次地看，现有路径分析多依赖于 fsQCA 等定性组态方法，虽能解释驱动因素之间的复杂交互，却无法在资源总量有限且环境规制严格的现实约束下，为各决策单元提供精确到数值层面的投入产出调整指引。由于绿色创新系统涉及经济增长、环境保护与创新效率间的内生冲突，亟须引入能够处理多目标权衡的智能搜索算法。因此，本文构建基于第二代非支配排序遗传算法(Non-dominated Sorting Genetic Algorithm II，NSGA-II)的多目标动态优化模型，旨在破解增长、减排，效率的权衡难题，从而将研究视角从现状评估提升至量化决策支持的理论高度。

针对上述研究不足，本文的边际贡献与理论价值主要体现在以下三个维度。首先，在研究框架方面，超越了传统的单一效率测度范式，构建了集效率测度、非线性机理识别、多目标路径优化于一体的全链条中国绿色创新决策支持系统架构。这一闭环框架不仅能精准捕捉环境约束下中国绿色创新效率的时空演变规律，更能深入挖掘其背后的复杂驱动机制，并以此为基础提供可量化的政策优化方案，弥补了现有文献在路径演化研究中缺乏动态优化手段的不足。其次，在方法方面，将具备强非线性拟合能力与可解释性的 CatBoost-SHAP 机器学习模型引入中国绿色创新效率研究。该组合有效解决了传统线性计量模型难以处理多维因素交互效应及变量间共线性问题的瓶颈，通过 SHAP 值量化各驱动因素的边际贡献，从“黑箱”模型中提取出具备解释力的政策依据，实现了模型精度与归因分析深度的协同提升。最后，在实证结论方面，突破了经济增长与碳减排不可兼得的理论局限。通过 NSGA-II 多目标智能算法在多维约束空间内寻求绿色创新路径帕累托最优解，量化论证了在不牺牲经济效益的前提下，通过跨区域资源优化配置可实现绿色创新效率约 11.4%的提升空间，并达成碳减排目标。这一发现为中国在“双碳”目标下实现经济增长与生态改善的“三赢”格局提供了重要的决策工具与量化策略。

1　中国绿色创新效率的统计测度与演变特征

以 2013—2023 年 11 年时间全国 30 个省份(因数据缺失，未含西藏地区和港澳台地区)为研究对象，测评中国绿色创新效率。所取数据来源于《中国统计年鉴》《中国环境统计年鉴》及各省份统计年鉴。绿色创新效率表现为绿色创新投入与经济环境产出之间的比率关系。考虑到绿色创新是对传统创新的继承与发展，其本质特征为资源消耗低、环境污染小的创新活动。基于创新价值链理论及已有研究[12]，将知识创新和创新转化纳入投入范畴，将创新成果、经济效益列为期望产出，以环境污染作为非期望产出的衡量指标。具体投入产出指标如表 1 所示。

表 1　中国绿色创新效率投入产出评估指标

目标层	变量	指标	符号
投入	知识创新	R&D 人员全时当量/人年	x_1
		R&D 经费内部支出/万元	x_2
	创新转化	技术市场成交额/万元	x_3
期望产出	创新成果	绿色科技论文数量/(万人·件$^{-1}$)	y_1
		绿色国内专利授权数量/(万人·件$^{-1}$)	y_2
	经济效益	地区生产总值/万元	y_3
非期望产出	环境污染	二氧化硫排放量/t	b_1

构建包含非期望产出的超效率 SBM-DEA 模型计算中国第 k 省份($k=1,2,\cdots,30$)的绿色创新效率。

$$\min\rho_k=\frac{1-\frac{1}{3}\sum_{i=1}^{3}\frac{s_i^-}{x_{ik}}}{1+\frac{1}{2+1}\left(\sum_{r=1}^{2}\frac{s_r^+}{y_{rk}}+\frac{s_1^{b^-}}{b_{1k}}\right)}$$

$$\text{s.t.}\begin{cases}\sum\limits_{j=1,j\neq k}^{30}x_{ij}\lambda_j+s_i^-\leqslant x_{ik},\ \forall i=1,2,3\\ \sum\limits_{j=1,j\neq k}^{30}y_{rj}\lambda_j-s_r^+\geqslant y_{rk},\ \forall r=1,2\\ \sum\limits_{j=1,j\neq k}^{30}b_{1j}\lambda_j+s_1^{b^-}\leqslant b_{1k}\\ 1+\frac{1}{2+1}\left(\sum\limits_{r=1}^{2}\frac{s_r^+}{y_{rk}}+\frac{s_1^{b^-}}{b_{1k}}\right)>0\\ \lambda_j,s_i^-,s_r^+,s_1^{b^-}\geqslant 0,\forall i=1,2,3\end{cases}\tag{1}$$

式中:$r=1,2;j=1,2,\cdots,30$。

通过上述模型可分别测算出各省份的绿色创新效率值(表2)。从全国效率均值来看,在2013—2023年,中国绿色创新效率呈现较小幅度的上升趋势,由2013年的0.692上升到2023年的0.808,均值为0.762,低于全国均值地区的个数大于高于全国均值的地区个数,这表明省域效率差异较大,发展不均衡,未来还有很大的改进空间。表2中绿色创新效率均值最高的地区为北京(2.443),因为北京作为中国首都,在绿色科技产出与研发人员投入都位于全国前列,而且北京由2013年的2.193上升到2023年的2.781,说明其不断地在完善经济效益与环境协调的方向努力。其次效率较高的地区分别为海南、新疆、上海、浙江。海南依托生态文明建设和自由贸易港政策,在清洁能源和绿色制度创新方面表现突出;新疆通过能源结构优化与可再生能源推广实现了绿色转型;上海和浙江经济基础雄厚、创新体系完善,严格的环保规制促进了技术创新效率提升。相较之下,辽宁和陕西等省份效率偏低,主要受产业结构偏重、技术创新能力不足及污染排放强度较高等因素影响,绿色创新投入产出效率较弱。总体来看,高效率地区多具备政策支持、产业结构优化与创新体系完善等特征,而低效率地区仍需加强绿色技术研发与结构调整。总体来看,高效率地区多具备政策支持、产业结构优化与创新体系完善等特征,而低效率地区仍需加大绿色技术研发力度,推动产业结构调整。此外,低效率地区的效率水平总体呈稳步上升态势,虽在个别年份存在波动,但整体趋势表明其绿色技术创新能力正不断提升。

表2　2013—2023年全国各省份的绿色创新效率评估

省份	2013年	2014年	2015年	2016年	2017年	2018年	2019年	2020年	2021年	2022年	2023年	均值	排名
北京	2.194	2.260	2.260	2.529	2.025	2.637	2.579	2.226	2.539	2.840	2.781	2.443	1
天津	0.384	0.459	0.437	1.061	1.079	1.097	1.106	1.069	1.075	1.038	1.054	0.896	8
河北	0.571	0.480	0.514	0.522	0.551	0.443	0.363	0.396	0.458	0.424	0.440	0.469	23
上海	1.068	1.072	1.092	1.034	1.555	1.360	1.441	1.449	1.283	1.190	1.167	1.246	4
江苏	0.776	0.819	1.012	0.526	0.824	1.028	0.609	1.072	1.064	1.117	1.064	0.901	7
浙江	1.206	1.245	1.336	1.431	1.222	1.164	1.130	1.133	1.072	0.779	0.645	1.124	5
福建	0.573	0.511	0.721	1.130	1.044	1.124	0.500	1.123	1.070	0.531	1.011	0.849	10
山东	0.411	0.381	0.432	0.441	0.459	0.375	0.358	0.392	0.420	0.395	0.434	0.409	27
广东	0.713	1.010	1.009	0.677	1.022	0.846	1.018	0.715	0.802	0.769	1.032	0.874	9
海南	1.525	1.960	1.271	1.288	1.292	1.293	1.680	1.520	1.564	1.554	1.456	1.491	2
东部	0.942	1.020	1.008	1.064	1.107	1.137	1.078	1.110	1.135	1.064	1.108	1.070	—
山西	0.339	0.425	0.435	0.462	0.442	0.394	0.359	0.671	0.571	0.560	0.631	0.481	22
安徽	0.516	0.826	0.620	0.629	0.579	0.625	0.367	0.468	0.419	0.376	0.416	0.531	19
江西	0.338	0.411	0.530	0.722	0.452	0.430	0.322	0.345	0.362	0.320	0.342	0.416	25
河南	0.564	0.435	0.568	0.597	0.800	0.695	0.432	0.482	0.445	0.368	0.427	0.528	20
湖北	0.328	0.384	0.417	0.432	0.461	0.443	0.325	0.373	0.406	0.417	0.414	0.400	28
湖南	0.517	0.509	0.531	0.485	0.481	0.418	0.327	0.371	0.356	0.310	0.317	0.420	24
中部	0.434	0.498	0.517	0.555	0.536	0.501	0.355	0.452	0.427	0.392	0.425	0.463	—
内蒙古	0.154	0.270	0.339	0.410	0.431	0.505	0.494	0.511	0.761	0.639	1.216	0.521	21
甘肃	0.334	0.354	0.342	0.383	0.514	0.572	0.359	0.385	0.391	0.405	0.530	0.415	26
宁夏	0.591	0.515	0.524	0.656	0.567	0.529	0.437	0.513	1.018	0.619	0.726	0.609	16
青海	0.173	0.198	1.004	1.032	0.484	1.030	1.037	1.040	1.096	1.045	1.077	0.838	11
陕西	0.388	0.525	0.431	0.379	0.379	0.411	0.282	0.336	0.419	0.388	0.445	0.399	29
四川	1.036	0.818	0.813	1.014	0.742	0.448	0.338	0.355	0.435	0.404	0.483	0.626	15
新疆	1.268	1.167	1.210	1.360	1.424	1.993	1.507	1.453	1.420	1.263	1.177	1.386	3
云南	0.707	1.049	0.617	0.625	0.585	0.599	0.508	1.015	1.022	0.566	0.707	0.727	13
重庆	0.774	0.762	1.046	1.028	1.002	0.410	0.526	1.012	1.027	0.385	0.443	0.765	12
广西	1.134	1.004	1.203	1.078	1.119	1.071	0.520	0.624	0.496	1.001	1.124	0.943	6
贵州	1.073	1.057	1.059	0.722	0.540	0.497	0.398	0.393	0.442	0.428	0.494	0.646	14
西部	0.694	0.702	0.781	0.790	0.708	0.733	0.582	0.694	0.775	0.649	0.766	0.716	—
辽宁	0.354	0.350	0.426	0.381	0.410	0.361	0.319	0.341	0.460	0.384	0.428	0.383	30
吉林	0.321	0.283	0.373	0.319	0.378	0.437	0.328	0.369	1.027	1.209	1.047	0.554	18
黑龙江	0.437	0.463	0.499	0.489	1.016	0.634	0.391	0.445	0.559	0.470	0.725	0.557	17
东北	0.371	0.365	0.433	0.396	0.601	0.477	0.346	0.385	0.682	0.688	0.733	0.498	—
均值	0.692	0.733	0.769	0.795	0.796	0.796	0.679	0.753	0.816	0.740	0.808	0.762	—

从区域表现来看，东部地区（如北京、海南）绿色创新效率始终保持领先地位，自 2013 年起稳步提升，并在 2021 年达到峰值，2023 年略有回落，依然处于高水平区间；这得益于东部地区绿色技术基础雄厚、创新资源集中及政策引导有力。西部地区（新疆、广西等）紧随其后，整体呈现“上升-调整-回升”的态势。2021 年达到阶段性高点，反映其在资源型经济结构优化与绿色转型推进中的积极成效。中部地区（湖南、湖北等）绿色创新效率则长期处于较低水平，近年来均值维持在 0.4～0.5，波动性较强，2022 年降至最低点，2023 年略有回升；整体反映出该区域在绿色创新投入、成果转化等方面仍存在较大改进空间。东北地区（黑龙江、吉林等）绿色创新效率波动显著，尤其在 2017 年出现快速上升，随后在 2019 年跌至谷底，但在 2021 年以后出现明显反弹，2023 年进一步提高，接近西部地区水平；这表明该区域近年来在绿色技术研发和老工业基地绿色转型方面取得了一定进展。

为深入探究各省份绿色创新效率的时空演化过程，选取 2013 年、2017 年、2021 年与 2023 年 4 个关键年份，对绿色创新效率的空间分布格局进行可视化比较与动态分析发现：2013 年全国整体绿色创新效率水平偏低，地区间差异显著，呈现“东部沿海地区＞西部地区＞中部地区”的梯度分布特征。北京、海南、浙江等地处于领先位置，显示在政策引导、科技资源集聚和绿色基础设施建设等方面的优势；中部地区如江西、湖北、湖南等整体效率偏低，绿色转型起步较晚。2017 年绿色创新效率整体水平有所提升，区域差距有所收敛。一方面，东部地区如上海、福建继续发挥创新集聚优势；另一方面，西部地区如新疆、青海、内蒙古在绿色技术应用、资源节约型产业发展方面实现突破，形成了“点状突破、局部引领”的新格局；中部地区创新水平尽管有所改善，但整体提升幅度有限。进入 2021 年，部分省份绿色创新效率出现波动。一些原本效率较高的地区如浙江和海南仍保持相对优势，但诸如江西、湖南、湖北等省份效率下降，表明其绿色技术创新存在持续性不足、转化能力偏弱等问题。同时，西部地区资源型省份如青海、内蒙古展现出绿色转型的潜力，凸显政策导向与资源生态协同作用。2023 年绿色创新效率总体水平继续提升，呈现“收敛-扩散”双重特征。一方面，北京遥遥领先，进一步巩固其在全国绿色创新格局中的核心地位；另一方面，西部地区如内蒙古、新疆、青海通过绿色矿山建设、清洁能源转型等方式加速追赶，形成较强的区域示范效应。同时，区域差异仍存在，东北地区（如吉林、黑龙江）在绿色制造方面进展明显，而中部地区部分省份（如湖南、江西）绿色创新水平仍偏低，亟须加强研发投入与绿色产业链整合。

基于行标准化地理距离的矩阵，进一步计算 2013—2023 年中国绿色创新效率的全局莫兰指数，计算结果如表 3 所示。2013 年、2017 年和 2021 年 Moran’s I 未通过显著性检验，表明这三年中国绿色创新效率的空间关联性较弱，呈现相对随机分布。其余年份均通过不同程度的显著性检验，且多数年份存在正向显著空间自相关，说明绿色创新效率存在一定的空间集聚特征。尤其是 2015 年和 2018 年，空间相关性最为显著，显示中国绿色创新效率在这两个阶段的空间集聚最为显著。从时间演化看，绿色创新效率的空间自相关强度呈现“先升后降”的波动趋势：2013—2015 年快速增强，2016—2018 年保持较高水平，2019 年以后略有回落。这一趋势反映中国绿色创新在政策推动与区域协同的作用下，初期形成了显著的空间集聚效应；但随着各地创新要素流动性提升、产业结构调整加快，区域差距在一定程度上被缩小，导致空间关联性有所减弱。

2　中国绿色创新效率的影响因素分析

2.1　特征变量选取

将全国各省份绿色创新效率设为目标变量，潜在影响因素作为特征变量，借助机器学习算法和 SHAP 识别各因素对绿色创新效率的影响程度和作用方向。基于文献综述梳理的绿色创新效率影响因素，选取的关键特征变量如表 4 所示。

2.2　算法选择

基于各省份的面板数据及上文测算的绿色创新效率值，按照 80％和 20％的比例划分训练集和测试集，对支持向量机、随机森林等常用的机器学习算法进行训练。同时，为提高模型算法的拟合性能与泛化能力，采用网格搜索调参。最后，为了避免

表 3　中国绿色创新效率的全局自相关莫兰指数

变量	2013 年	2014 年	2015 年	2016 年	2017 年	2018 年	2019 年	2020 年	2021 年	2022 年	2023 年
Moran’s I	0.104	0.225**	0.324***	0.212**	0.024	0.278***	0.217**	0.119*	0.033	0.157*	0.242**

注：*、**、***分别表示 1％、5％、10％的显著性水平。

表 4　影响中国绿色创新效率的特征变量描述性统计

变量名称	变量测度	均值	标准差	最小值	最大值	单位
产业结构	第二产业占 GDP 比例	39.252	7.754	14.912	55.762	
经济发展	经平减处理的人均实际 GDP	6.505	3.264	2.209	20.028	万元
金融发展	金融机构贷款余额占 GDP 比例	1.594	0.432	0.743	2.774	%
人口密度	单位面积的人口	330.890	481.471	5.727	2 748.931	万人/km^2
技术创新	每万人发明专利授权数量	2.957	5.302	0.159	49.348	件/万人
能源强度	单位 GDP 的二氧化碳排放量	16.294	10.942	1.857	55.196	千 t/万元

在单一训练集上的过拟合问题,采用五折交叉验证,并将平均绝对百分误差(MAE)、均方根误(RMSE)和拟合优度(R^2)作为算法性能的评价指标。各算法评估结果如表 5 所示。RMSE 和 MAE 越小,R^2 越大,说明算法拟合效果越好。从表 5 可以看出,CatBoost 机器学习算法取得了最小的 RMSE 和 MAE,以及最大的 R^2,说明拟合效果最好。

表 5　各种机器学习算法预测性能评估结果

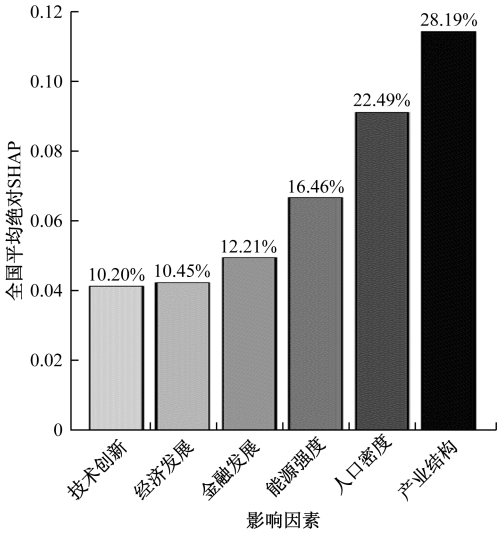
算法	名称缩写	MAE	RMSE	R^2
支持向量机	SVR	0.151	0.232	0.768
极端梯度提升	XGBoost	0.150	0.210	0.810
多层感知器	MLP	0.170	0.222	0.817
类别特征提升	CatBoost	0.138	0.189	0.846
轻量级梯度提升	LightGBM	0.164	0.231	0.769
随机森林	RF	0.145	0.194	0.837

2.3　中国绿色创新效率的影响机制分析

基于训练完成的 CatBoost 模型,进一步引入 SHAP 可解释性算法,打开 CatBoost 的“黑箱子”,揭示各因素对中国省级绿色创新效率的相对重要性及作用机制。通过计算全样本下各变量的 SHAP 均值,量化其对模型预测的贡献。SHAP 算法由 Lundberg 和 Lee[13] 于 2017 年提出,是一种基于博弈论的模型解释方法,能够量化各影响因素对绿色创新效率预测的贡献度并提供可解释性。SHAP 值计算公式为

$$\mathrm{SHAP}(x_i)=\sum_{S\subseteq\{x_1,x_2,\cdots,x_p\}/\{x_i\}}\frac{|S|!(F-|S|-1)!}{|F|!}[f(S\cup\{x_i\})-f(S)] \tag{2}$$

图 1 展示了中国绿色创新效率各影响因素的 SHAP 均值与占比。结果显示,各因素影响程度存在显著差异。其中,产业结构影响最大,SHAP 均值占比达 28.19%。这与中国高耗能产业特征及环境规制倒逼绿色技术投入相关,如钢铁、化工行业的减排改造与工艺创新。产业结构调整不仅影响排放强度,还制约绿色技术创新活力,是提升绿色



柱状图上的数值为该因素 SHAP 均值
占所有因素总体 SHAP 均值的比例

图 1　全国绿色创新效率各影响因素的 SHAP 均值

创新效率的关键路径。除了产业结构,其余影响因素依重要性分别为人口密度、能源强度、金融发展、经济发展与技术创新。

尽管 SHAP 均值能够衡量各变量之间的相对重要性,但无法直观地呈现各变量的具体作用方向。因此,进一步绘制 SHAP 值的单变量依赖图(图 2)。图中横轴为各影响因素的标准化取值,反映样本相对于均值的偏离程度;纵轴为 SHAP 值,用以刻画其对绿色创新效率的边际影响方向与强度。图 2 直观展现 6 个核心变量对中国绿色创新效率的边际影响机制,基于 SHAP 值的变化趋势可揭示其作用方向与强度。

从产业结构[图 2(a)]可知,随着产业结构高级化水平的提升,SHAP 值整体呈下降趋势,说明当前高耗能、高排放产业占比仍然较高,对中国绿色创新效率形成明显抑制。从人口密度[图 2(b)]可知,人口密度的 SHAP 值在低密度区间呈负向波动,在高密度区间逐渐转为正向影响,表现出显著的非线性特征。适度密集的人口有助于知识、资本与技术的集聚,形成创新外部性,从而提升城市和

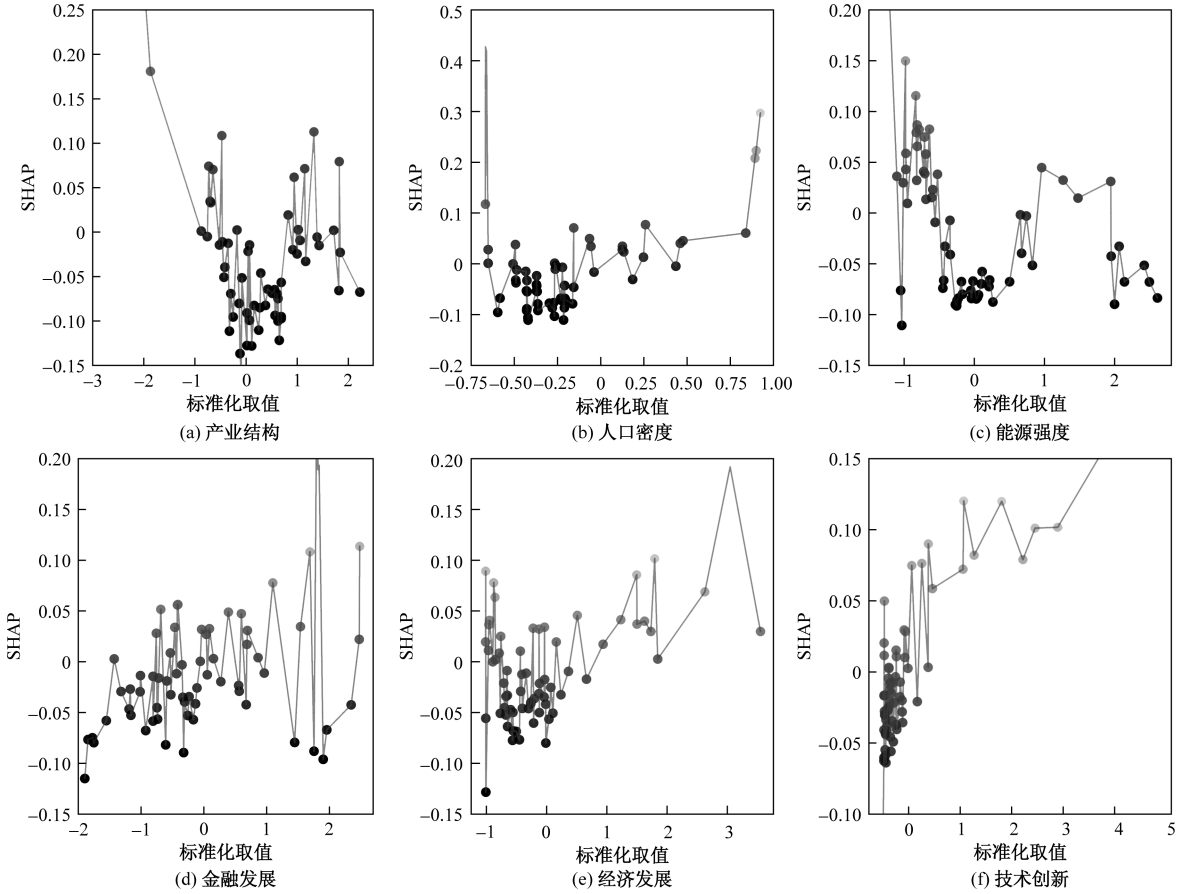


图 2　全国绿色创新效率的 SHAP 值单变量依赖图

区域的绿色创新能力；但当人口密度超过阈值后，过度集聚可能导致交通拥堵、能源紧张和污染加剧等负效应，抑制绿色创新绩效。从图 2(c)可知，能源强度的 SHAP 整体为负，且随能源消耗强度上升而持续下降，表明能源消耗水平与绿色创新效率之间存在显著的反向关系。能源强度高反映经济活动对传统化石能源的高度依赖，以及生产过程中的能源利用效率较低，这种状况不仅限制绿色技术的引入与扩散，也削弱了经济体系的低碳转型能力。从图 2(d)可知，金融发展的 SHAP 波动显著，在金融发展水平较高时显著跃升，说明其对绿色创新效率存在双重影响，既有正向促进，也有负向抑制。早期金融资本可能倾斜于高污染产业，产生负面影响。绿色金融政策实施后，资金转向绿色技术研发与产业发展。国家政策层面持续完善绿色金融标准体系，扩大绿色信贷、绿色债券与碳金融产品的供给，引导资金流向节能环保与绿色技术产业，从而增强金融体系对绿色创新的支持能力。从图 2(e)可知，随着经济发展水平的提升，SHAP 总体呈上升趋势，说明经济增长为绿色创新提供了资本积累、技术支撑和人才保障。该结果与环境库兹涅茨曲线(EKC)假设一致，在经济发展的初期阶段，污染水平随收入增长而上升；当经济达到一定拐点后，公众环保意识增强、产业结构优化与政策调控加强，污染强度下降，绿色创新效率提升。从图 2(f)可知，技术创新的 SHAP 持续上升，对绿色创新效率的驱动作用最为显著，是所有因素中最核心的正向变量。绿色技术创新能够在降低污染与能源消耗的同时，提高资源利用效率和生产率，形成经济增长与生态改善的协同机制。

2.4　中国绿色创新效率空间异质性分析

由于中国区域间的发展阶段、资源禀赋及政策环境存在显著差异，绿色创新效率的影响因素在空间上表现出明显的非均衡性。为了精准识别各区域效率形成的深层逻辑，利用 SHAP 解释框架对东部与中西部地区进行分类解析(图 3 和图 4)。分析图显示，东部地区绿色创新效率受产业结构与能源强度的负向约束最为突出，表现出典型的“产业-能源”双重制约特征，其中产业结构因素影响最大，SHAP 均值占比达 36.09%(图 3)。从发展阶段看，

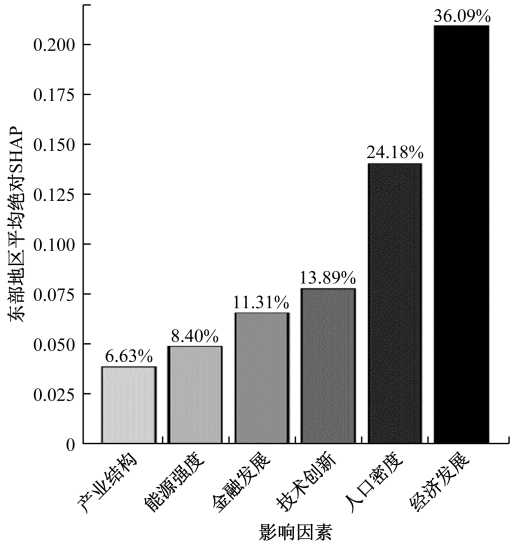


图 3　东部地区绿色创新效率影响因素的 SHAP 均值

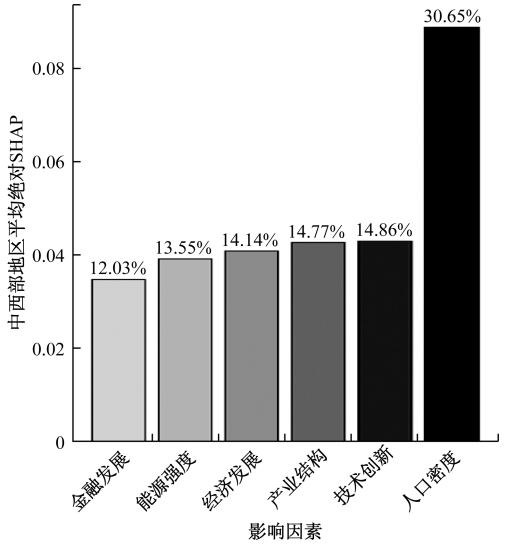


图 4　中西部地区绿色创新效率影响因素的 SHAP 均值

东部地区已率先进入后工业化转型期,虽然高新技术产业集聚,但钢铁、化工等传统重化工业的存量占比依然较高,资源环境依赖度尚未完全脱钩,导致效率提升面临“边际效应递减”的瓶颈。从政策环境看,作为国家“双碳”目标的先行区,东部地区省份面临最严苛的环保规制与能效考核。这种强约束环境虽倒逼了研发投入,但在短期内也因高昂的绿色合规成本对全要素效率产生“挤出效应”。从资源禀赋看,东部地区虽拥有完善的金融体系(SHAP 占 13.39%)与高强度研发投入,但受限于土地与能源成本高企,迫使其必须通过更高层级的产业升级与能源结构优化,如北京、上海的去工业化路径来对冲负向约束。

与东部地区相比,中西部地区的绿色创新效率更受人口密度与技术创新的影响,呈现“人口集聚+创新驱动”的双轮发展模式(图 4)。从发展阶段与资源禀赋看,中西部地区正处于新型城镇化加速期,人口向省会城市(如成都、武汉、西安)的快速集聚产生了显著的“规模经济”与知识溢出效应,为绿色创新提供了必要的人才储备与市场规模。从政策环境看,得益于“西部大开发”与清洁能源转型政策,新疆、青海等省份利用其丰富的可再生能源禀赋,通过能源结构优化实现了绿色创新效率的跨越式提升。从演化机理看,中西部地区仍面临较强的增长压力,经济发展与能源强度对效率的贡献度较为均衡。这表明该区域正处于从“资源依赖”向“创新驱动”转型的窗口期,具有较强的后发优势与提升潜力。

从总体来看,东部地区呈现“产业主导+能效驱动”特征,聚焦第三产业升级与能源结构优化;中西部地区依赖“人口集聚+创新驱动”模式,强化城市群协同发展与绿色技术本地化应用。区域差异源于发展阶段、产业结构及资源禀赋的分化,对后续绿色转型差异化政策设计提出更高要求。因此,在识别关键影响因素后,有必要进一步开展区域层面绿色创新效率的多目标优化研究,这对于实现区域协调发展、提升政策资源配置效率意义重大。

3　中国绿色创新效率的多目标路径优化

基于前文对绿色创新效率的区域异质性分析,如何协调全国经济发展与生态保护的动态平衡,以及如何针对不同省份的地区特点和资源禀赋设计差异化优化系统,成为实现绿色创新可持续发展的关键路径。

3.1　中国绿色创新多目标组合模型构建

在绿色经济背景下,中国各区域面临经济增长、绿色创新效率提升与碳排放最小化的相互制约发展目标。由于各省份在产业结构、资源禀赋和发展阶段上差异明显,传统单一目标优化方法难以兼顾区域发展需求。本文聚焦协调经济效益与绿色创新效率提升,同步优化污染减排、资源节约等环境效益。通过量化不同情境下多目标的协同关系与权衡取舍,为区域精准施策提供理论与决策支持。为确保政策的前瞻性与时效性,使用 2021—2023 年各省份最新数据建模分析,所有数据经过严格处理确保一致性。为增强模型稳健性,目标变量采用三年均值化处理,平滑短期波动,准确反映区

域发展趋势。该策略有助于实现绿色创新效率最大化与碳排放最小化，为制定中国差异化区域协同创新发展政策提供量化依据。

为提高绿色创新效率，实现环境保护和增加经济效益，选取绿色创新效率(GE)、GDP 和碳排放(CE)三个目标变量。通过构建多目标优化模型，对各省份的资源配置权重进行调整，使全国绿色创新效率达到最优均衡状态，模型中优化目标函数如式(3)所示，其中 w_i 为第 i 省份($i=1,2,\cdots,30$)的绿色资源配置权重，GE_i、GDP_i、CE_i 分别为每个省份的三年平均绿色创新效率、GDP 和碳排放量。

$$\max\sum_{i=1}^{30}w_iGE_i;\ \max\sum_{i=1}^{30}w_iGDP_i;\ \min\sum_{i=1}^{30}w_iCE_i \tag{3}$$

若全国加权绿色创新效率小于 0.3 可视为弱有效，若将全国加权效率设置高于 1.2，又会导致局部求解过重。因此将全国加权绿色创新效率水平限制在 0.3～1.2。全国加权 GDP 水平不应低于最低省份的 GDP，同时每年 GDP 增幅为 5%～10%，故将全国加权 GDP 的最低约束水平设定为 3 602.5；为了进一步减少污染，全国加权碳排放量不应高于当前各省份碳排放量的均值 116 747 t。由于每个省份的资源分配最大不可能高于全国资源的 10%，所以每个省份权重不应大于 0.1，也不能低于 0.01。最后，为了控制模型的不确定性且提高分散性，增加熵约束：$-\sum_{i=1}^{30}W_i\ln W_i\geqslant 1.75$，最终构建多目标协同优化模型为

$$\max\sum_{i=1}^{30}w_iGE_i;\ \max\sum_{i=1}^{30}w_iGDP_i;\ \min\sum_{i=1}^{30}w_iCE_i$$

$$\text{s.t.}\begin{cases}0.3\leqslant\sum_{i=1}^{30}w_iGE_i\leqslant 1.2\\ \sum_{i=1}^{30}w_iGDP_i\geqslant 3\ 602.5\\ \sum_{i=1}^{30}w_iCE_i\leqslant 116\ 747\\ -\sum_{i=1}^{30}W_i\ln W_i\geqslant 1.75\\ \sum_{j=1}^{30}W_j=1\\ 0.01\leqslant W_i\leqslant 0.1,i=1,2,\cdots,30\end{cases} \tag{4}$$

使用 MATLAB 算法工具箱的 NSGA-II 算法对上述模型(4)求解，设定的参数如表 6 所示。

表 6　NSGA-II 多目标遗传算法参数设置

参数	种族规模	迭代次数	帕累托前沿比例	交叉概率
数值	200	100	0.35	0.8

3.2　中国多目标绿色创新模型求解与优化前后比较分析

根据各省份的绿色创新效率、GDP 及碳排放量数据，计算每个省份在目标上的占比。由于三个目标的优化方向不同，绿色创新效率和 GDP 是越大越好，可直接计算；碳排放是越小越好，因此需要取倒数后再计算占比。假设决策者对三大目标没有优先序偏好，使用平均法式(5)和式(6)来确定各省份 $i(i=1,2,\cdots,30)$ 的初始资源分配权重 w_i。

$$w_{GE_i}=\frac{GE_i}{\sum_{i=1}^{30}GE_i};\ w_{GDP_i}=\frac{GDP_i}{\sum_{i=1}^{30}GDP_i};$$
$$w_{CE_i}=\frac{1/CE_i}{\sum_{i=1}^{30}(1/CE_i)} \tag{5}$$

$$w_i=\frac{w_{GE_i}+w_{GDP_i}+w_{CE_i}}{3} \tag{6}$$

进而通过表 6 设置的遗传算法参数，并运用 MATLAB 优化工具箱的 NSGA-II 算法，可求解模型(4)获得各省的平均效率 GDP、碳排放及最优权重配置(表 7)，由此可见中国目前的配置存在一定程度的失衡。

在本次多目标创新优化求解中，设定的种群规模为 200，结果产生相应数量的帕累托最优解。在这些帕累托解中，全国平均绿色创新效率在 0.790～0.940、平均 GDP 在 36 100 万～44 500 万元，而平均碳排放则在 31 200～38 700 t。与当前的省份配置相比，这些优化解在绿色创新效率、经济效益获取以及环境效益提升上均实现了显著优化。如表 7 所示，优化后的资源配置权重较初始状态发生了显著调整，这深刻揭示了中国各区域绿色转型路径的差异化政策含义。首先，优化模型结果显示可以对北京(－17.1%)和浙江(－27.5%)等东部地区经济发达省份的绿色资源配置权重进行适度“退坡”。从政策含义上看，这并不意味削弱东部地区的核心地位，而是预示东部地区已进入绿色创新的“边际效益递减”阶段，其政策重心应从单纯的“要素投入驱动”转向以制度创新和技术溢出为主的“存量效能挖掘”，通过发挥辐射带动作用实现跨区域的绿色价值输出。与之相对，中西部地区省份如贵州(＋65.2%)与江西(＋69.6%)的权重显著优化提升，彰显了政策“高杠杆区”的加码信号。这表

表 7　全国绿色创新组合模型优化后的目标值和资源配置权重

省份	平均效率	平均 GDP/万元	平均碳排放/t	比例	优化前	优化后	省份	平均效率	平均 GDP/万元	平均碳排放/t	比例	优化前	优化后
安徽	0.403	44 741.167	52 859.483	W_1	0.023	0.026	江西	0.341	31 080.567	26 260.886	W_{16}	0.023	0.027
北京	2.720	42 115.733	7 984.403	W_2	0.082	0.134	辽宁	0.424	28 868.333	46 815.484	W_{17}	0.020	0.023
福建	0.871	51 895.433	27 055.499	W_3	0.036	0.027	内蒙古	0.872	23 060.633	77 448.878	W_{18}	0.022	0.027
甘肃	0.442	11 070.233	23 238.845	W_4	0.020	0.027	宁夏	0.788	5 002.600	19 386.057	W_{19}	0.026	0.028
广东	0.868	129 968.77	65 520.291	W_5	0.052	0.133	青海	1.073	3 602.500	8 747.731	W_{20}	0.046	0.028
广西	0.874	26 199.200	35 784.758	W_6	0.027	0.027	山东	0.416	87 506.933	116 747.107	W_{21}	0.033	0.026
贵州	0.455	20 127.433	22 853.763	W_7	0.023	0.026	山西	0.587	24 717.500	59 922.680	W_{22}	0.019	0.028
海南	1.525	6 981.633	5 156.824	W_8	0.073	0.030	陕西	0.417	32 248.667	39 933.972	W_{23}	0.021	0.027
河北	0.441	42 109.733	80 586.742	W_9	0.021	0.026	上海	1.213	45 227.000	20 136.055	W_{24}	0.043	0.027
河南	0.413	58 474.633	72 667.193	W_{10}	0.026	0.026	四川	0.441	56 943.700	27 011.702	W_{25}	0.032	0.029
黑龙江	0.585	15 524.533	20 798.985	W_{11}	0.025	0.025	天津	1.056	16 184.867	14 651.092	W_{26}	0.037	0.027
湖北	0.412	52 878.833	52 121.433	W_{12}	0.026	0.025	新疆	1.287	17 826.733	49 678.342	W_{27}	0.028	0.026
湖南	0.327	47 761.667	28 304.071	W_{13}	0.027	0.027	云南	0.765	28 579.600	21 868.950	W_{28}	0.031	0.026
吉林	1.094	13 171.033	16 433.717	W_{14}	0.035	0.027	浙江	0.832	78 218.200	42 754.045	W_{29}	0.040	0.026
江苏	1.081	122 567.970	93 972.784	W_{15}	0.052	0.027	重庆	0.618	28 933.067	17 344.015	W_{30}	0.032	0.011

明中西部地区目前处于新型城镇化与产业结构升级的红利窗口期,具有极高的技术吸收潜力与后发优势,是提升全国绿色创新效率整体均值的核心驱动力。此外,特殊区域的权重变动为战略校准提供了关键指引。海南权重降幅最大(−49.3%),警示其必须补齐绿色制造短板,避免生态红利与工业效能脱钩;而天津权重的上升(+18.9%)凸显其作为京津冀绿色产业链协同枢纽的功能定位。最终,经过多目标模型优化后的全国协同绿色创新效率显著提升至 26.346(表 8),增幅约 11.4%,且在不牺牲经济效益的前提下实现了碳排放总量的显著下降。这一优化结果不仅在实证层面验证了算法驱动全国绿色创新资源精准配置的有效性,更在政策理论层面论证了通过区域差异化协同路径破解中国经济增长-碳减排-绿色效率提升的三角瓶颈问题。

表 8　全国绿色创新多目标模型优化前后结果对比

变量	优化前	优化后
协同绿色创新效率	23.641	26.346
协同经济效益(GDP)	1 193 588.900	1 205 487.994
总环境污染(碳排放)	1 194 045.786	1 042 710.045

4　结论

本文构建了中国绿色创新效率测度-机理分析-路径优化的综合研究框架,系统考察了 2013—2023 年全国 30 个省份(因数据缺失,未含西藏地区和港澳台地区)绿色创新效率的动态演化机理。通过实证分析与模型优化,得出如下重要的结论与理论启示。

(1)中国绿色创新效率在时间维度上呈现稳步提升趋势,在空间维度上,全国绿色创新效率呈现“东部领先、中部滞后、西部趋稳、东北波动”的梯度特征。这一空间格局验证了空间不平衡增长理论在绿色转型领域的投影。东部地区凭借政策红利与创新资源集聚形成了增长极效应;而东北及中部地区受制于重工业为主的产业结构,表现显著的路径依赖特征,陷入了结构性低效率的梯度陷阱。

(2)产业结构、人口密度和能源强度是影响中国绿色创新效率的关键因素。其中东部地区受产业结构和能源强度的显著负向影响,中西部地区主要受人口密度与技术创新的正向影响。从系统演化视角看,SHAP 依赖图揭示了驱动因素对效率影响的非线性响应机制。例如,人口密度与效率之间并非简单的线性正相关,而是存在显著的门槛效应,在不同地区表现从集聚经济向集聚不经济的转变。这一发现丰富了波特假说的内涵,证明了绿色创新是一个多维要素动态耦合的复杂系统,环境规制对技术创新的激发作用受到产业结构与资源禀赋的强烈边际约束。

(3)通过多目标优化模型,可以优化各省份在全国资源的配置,能够在维持经济效益的前提下,显著提升中国绿色技术创新水平并降低碳排放量。这一结果在理论上论证了可持续发展脱钩理论的可行性。通过算法驱动的帕累托优化解改进,可以有效破解经济增长、绿色创新与环境保护协同的瓶颈问题。

针对上述研究结论分析,为整体提升中国绿色创新效率,提出如下针对性的对策建议。

(1)强化区域协同,促进中国绿色创新均衡发展。基于绿色创新驱动经济发展的路径,加大对中西部地区的扶持,通过财政激励引导企业迁移,培

育区域性创新产业集群，提升绿色创新效率以促进经济增长。同时，利用东部沿海区位优势，吸引高新技术产业入驻，借助技术溢出效应辐射带动周边发展，构建“东部引领、中西部协同”的绿色创新发展格局。

(2)推动技术创新，助力中国产业转型。为统筹中国绿色创新与污染减排，因地制宜推广绿色技术。在资源依赖型和高碳排放的中西部地区，加快可再生能源基础设施建设，推动绿色制造与智能制造深度融合，实现碳排放降低。同时，在交通、建筑等重点领域推广绿色技术，驱动全社会向低碳高效转型，提升传统高耗能产业的绿色转型能力。

(3)完善政策体系，优化绿色投入产出资源配置。构建差异化政策引导体系。根据各省份发展阶段、产业结构和资源禀赋的差异制定精准政策，推动中国绿色创新能力评估常态化，动态调整资源配置策略以提升中国绿色创新资源配置效率。同时，鼓励各地结合自身绿色资源配置权重变化制定省级绿色发展优化路径，保障各省份与全国绿色创新协同发展目标的实现。

参考文献

[1] 向小东，陈炎光，长江经济带绿色创新效率评价研究[J]. 长江流域资源与环境，2024，33(3)：472-486.

[2] 郭树华，方鹏飞，毕福芳，基于SFA方法的企业绿色创新效率分析[J]. 工业技术经济，2024，43(7)：112-119.

[3] 栾蕾，张水平. 新质生产力视角下供应链金融对企业绿色创新效率的影响研究[J]. 老字号品牌营销，2025(12)：98-100.

[4] TONE K. A slacks-based measure of efficiency in data envelopment analysis[J]. European Journal of Operational Research，2001，130(3)：498-509.

[5] 刘子豪，王子涵，师路路. 中国省域绿色创新效率测度、差异分析及影响因素[J]. 科技和产业，2024，24(17)：63-68.

[6] 吕承超，邵长花，崔悦，中国绿色创新效率的时空演进规律及影响因素研究[J]. 财经问题研究，2020，12(1)：50-57.

[7] LI Q. Regional technological innovation and green economic efficiency based on DEA model and fuzzy evaluation [J]. Journal of Intelligent & Fuzzy Systems，2019，37(5)：6415-6425.

[8] 邢会，霍晓谦. 中国县域减污降碳协同增效的时空演变及影响因素[J]. 中国环境科学，2025，45(3)：1657-1673.

[9] 丁显有，陈杭盛，田泽，等. 我国高端制造业绿色创新效率评价研究——基于YREB和非YREB比较视角[J]. 生态经济，2022，38(5)：68-74.

[10] RAWAT A，KULSHRESTHA M. Benchmarking water user associations for improved irrigation efficiency：a dea approach with SHAP and machine learning integration [J]. Water Conservation Science and Engineering，2025，10(3)：117.

[11] 侯鹏波，马占新，中国生物医药企业创新效率测度及提升路径研究——基于改进的超效率DEA与fsQCA方法[J]. 软科学，2025，39(9)：111-119.

[12] 兰龙耀. 价值链数字化转型、税收创新激励政策与企业技术创新[J]. 渭南师范学院学报，2025，40(12)：54-64.

[13] LUNDBERG S M，LEE S I. A unified approach to interpreting model predictions[J]. Advances in Neural Information Processing Systems，2017，30：1-10.

Measurement，Factor Exploration and Path Optimization of China's Green Innovation Efficiency

ZHANG Qiansheng，HUANG Jinbin，LIN Senfeng，ZHENG Canyao

(School of Mathematics and Statistics，Guangdong University of Foreign Studies，Guangzhou 510006，China)

Abstract：Using the sample data of 30 Chinese provinces from 2013 to 2023，green innovation efficiency across provinces was measured and its key influencing factors were identified. Methodologically，the super-efficiency SBM-DEA model was employed to assess regional efficiency，the CatBoost-SHAP model was incorporated to analyze nonlinear driving mechanisms，and the NSGA-II genetic algorithm was applied to solve a multi-objective optimization model for national green innovation. The empirical results show that China's green innovation efficiency has been gradually increasing overall，yet significant regional disparities persist. The SHAP values indicate that industrial structure，population density and energy intensity are key influencing factors，with industrial structure exerting the most pronounced inhibiting effect. The multi-objective optimization reveals that rationally allocating resource weights across provinces can improve national green innovation efficiency by approximately 11.4%，while simultaneously achieving GDP growth and carbon emission reduction benefits，thereby offering an optimal pathway for formulating regionally differentiated and synergistic innovation policies.

Keywords：green innovation efficiency；super-efficiency SBM-DEA；CatBoost；SHAP additive explanations；multi-objective optimization