

“数据要素×科技创新”驱动新质生产力的空间溢出效应

——基于 GNN-SDM 模型的实证检验

陈义安^{1,2}, 闫 悅¹

(1. 重庆工商大学 数学与统计学院, 重庆 400067;
2. 经济社会应用统计重庆市重点实验室, 重庆 400067)

摘要: 科技创新是引领发展的第一动力, 数据是数字经济时代的核心生产要素, “数据要素×科技创新”将驱动新质生产力持续跃升。采用我国 30 个省份 2011—2023 年的数据, 以数据要素发展水平和科技创新发展水平的耦合协调度衡量“数据要素×科技创新”水平, 利用图神经网络 (GNN) 提取融合特征, 进而构建 GNN-SDM 模型检验, 分析发现: “数据要素×科技创新”不仅显著提升了本地新质生产力发展水平, 还带动了相邻地区的新质生产力发展, 且间接效应大于直接效应, 表明“数据要素×科技创新”的区域联动对新质生产力发展至关重要; GNN 特征揭示出两类空间结构——“抑制本地、带动周边”的反向激励型空间扩散结构 (GNN1) 和“增强本地、抑制周边”的“中心-外围”型空间极化结构 (GNN2), GNN1 值较高地区应优化内部结构与资源配置, GNN2 值较高地区则应强化区域联动; “数据要素×科技创新”对东部地区、“数据要素×科技创新”水平较高地区、新质生产力发展水平较高地区具有更强的新质生产力驱动作用, 发展新质生产力应因地制宜。

关键词: 数据要素; 科技创新; 新质生产力; 耦合协调效应; GNN-SDM 模型

中图分类号: F124; F49 **文献标志码:** A **文章编号:** 1674-8131(2025)06-0041-14

引用格式: 陈义安, 闫悦. “数据要素×科技创新”驱动新质生产力的空间溢出效应——基于 GNN-SDM 模型的实证检验 [J]. 西部论坛, 2025, 35(6): 41-54.

CHEN Yi-an, YAN Yue. The spatial spillover effects of “data elements×technological innovation” on new quality productive forces: An empirical test based on a GNN-SDM model [J]. West Forum, 2025, 35(6): 41-54.

* 收稿日期: 2025-07-04; 修回日期: 2025-11-16

基金项目: 重庆市重大决策咨询课题(2371099); 重庆工商大学研究生科研创新项目(yjscxx2025-269-225)

作者简介: 陈义安(1969), 男, 四川达州人; 教授, 主要从事经济统计、非线性分析理论及其在社会经济中的应用研究; E-mail: 709590113@qq.com。闫悦(2001), 通信作者, 女, 重庆万州人; 硕士研究生, 主要从事社会经济统计学研究; E-mail: yanyue2001@163.com。

一、引言

新质生产力由技术革命性突破、生产要素创新性配置、产业深度转型升级而催生,以劳动者、劳动资料、劳动对象及其优化组合的跃升为基本内涵,以全要素生产率大幅提升为核心标志(王天雨等,2025)^[1]。科技创新作为新质生产力的核心引擎,通过驱动技术突破与生产方式革新,有效促进质量、效率及动力的系统提升,在推动经济增长中发挥着重要作用(周文等,2023)^[2]。数据要素作为新型生产要素,具备可复用性与非消耗性(冯永琦等,2024)^[3],在人工智能与大数据等技术推动下,其处理与应用成本下降,能够有效提升要素配置效率。通过多领域融合与算法驱动,数据要素不断生成认知价值、支撑技术创新(徐翔等,2024)^[4],已成为驱动新质生产力跃升的关键力量。2024年1月,国家数据局等17个部门联合发布《“数据要素×”三年行动计划(2024—2026年)》,明确提出开展“数据要素×科技创新”重点行动。在此背景下,有必要深入探讨“数据要素×科技创新”如何驱动新质生产力发展。

近年来,新质生产力和数据要素成为学术界研究的热点。在理论层面,相关文献大多依托政治经济学的分析框架,对数据要素和科技创新与新质生产力的关系展开研究,均认为科技创新是推动新质生产力发展的核心驱动力(丁任重等,2024;梁昊光等,2024)^[5-6],而数据要素能够有效赋能新质生产力发展(徐翔等,2024;张斌等,2024)^{[4][7]}。在实证方面,科技创新在推动新质生产力发展中的显著作用得到众多经验分析的证实(刘晨等,2024;王文霞等,2025;任宇新等,2025)^[8-10],数据要素对新质生产力的赋能作用也在多方面得到验证(姜汇泳等,2025)^[11],包括数据要素市场化(陆扬等,2024)^[12]、数据要素集聚(姚博伦等,2025)^[13]、数据要素共享(张小宇等,2025)^[14]、数据要素价值化(陈晔婷等,2025)^[15]等。然而,已有实证文献大多分别考察数据要素和科技创新对新质生产力的影响,缺乏将两者有机联系起来的研究。

有鉴于此,本文在已有研究的基础上探究“数据要素×科技创新”对地区新质生产力发展的影响及其空间溢出效应,并采用2011—2023年我国30个省份的数据进行实证检验。本文的边际贡献主要在于:第一,从数据要素与科技创新耦合协调发展的角度探讨地区新质生产力发展的驱动机制,丰富了新质生产力理论研究;第二,在空间杜宾模型(SDM)中引入图神经网络(GNN)提取的融合特征,构建GNN-SDM模型考察“数据要素×科技创新”对地区新质生产力发展的直接影响和空间溢出效应,更深入地揭示了区域间的协同发展机制,并进一步识别出区域间存在的非对称空间扩散机制,不仅为“数据要素×科技创新”对新质生产力的驱动作用提供了经验证据,还为相关空间效应研究提供了方法借鉴。

二、理论分析与研究假设

1.“数据要素×科技创新”对地区新质生产力的影响

新质生产力作为一种极具活力与潜力的现代新型生产力形态,深度融合了高新技术,广泛应用于各个领域。科技创新始终是引领发展的第一动力,而在数字经济时代,数据成为经济增长的关键要素,数据要素与科技创新的协同演进推动着新质生产力的形成和发展。数据要素发展和科技创新发展均能促进地区新质生产力发展,“数据要素×科技创新”即数据要素与科技创新的耦合协调发展,更是驱动新质生产力发展的核心动能。在劳动者层面,“数据要素×科技创新”已成为驱动劳动力质量跃升、塑造“新型劳动者”的核心力量(姜汇泳等,2025)^[11]。作为新型生产要素的数据要素与科技创新深度融合,通过

智能教育平台、个性化学习系统等的创新应用,将数据要素转化为具体、生动的学习场景,使劳动力能够在模拟环境中进行高效学习,极大地增强了学习的针对性和实效性,从而催生具备更高数字化素养的“新型劳动者”。这种协同作用不仅促进了劳动力知识结构的快速迭代和技能水平的持续提升,还激发了劳动力的创新思维和问题解决能力(张斌 等,2024)^[7]。在劳动对象层面,数据要素的引入打破了传统劳动对象的边界,将数据资源纳入生产体系之中,拓展了劳动对象的内涵与价值来源(周文 等,2023)^[2];与此同时,科技创新依托大数据、人工智能、云计算等先进技术,对海量数据资源进行深度挖掘和智能处理,推动数据价值的高效释放与转化(徐翔 等,2024)^[4]。“数据要素×科技创新”不仅极大地丰富了劳动对象的种类和形态,提高了劳动对象的利用效率和附加值,还推动了劳动对象利用方式的数字化和绿色化变革,为新质生产力的形成与壮大奠定了坚实的物质基础和技术支撑(张斌 等,2024)^[7]。在劳动资料层面,“数据要素×科技创新”对劳动资料的演进产生了系统性影响,推动其从传统形态向智能化、高级化方向跃升。“数据要素×科技创新”催生出新型劳动资料,使高级、精密和智能的设备及技术工具逐步取代传统机械,成为新质生产力发展的物质基础(冯永琦 等,2024)^[3];同时,“数据要素×科技创新”通过赋能劳动资料运用方式的深刻变革,使劳动资料的系统性效能倍增,进而显著提升资源配置与利用效率(徐建伟 等,2025)^[16]。更进一步,“数据要素×科技创新”通过重塑劳动者、劳动资料与劳动对象三者间的互动关系,实现生产要素组合模式的系统性优化,这种要素间的深度融合与动态适配,能够打破传统生产要素的组合壁垒,构建更加灵活、高效且可持续的生产函数(张斌 等,2024)^[7],进而推动全要素生产率全面提升,实现新质生产力的持续跃升。

据此,本文提出假设 H1:“数据要素×科技创新”能够显著驱动地区新质生产力发展。

2. “数据要素×科技创新”驱动新质生产力发展的空间溢出效应

一方面,新质生产力的发展具有显著的空间溢出特征。新质生产力以数字化、智能化和绿色化为主要特征,通过高效的信息传递与知识共享,突破传统生产活动的时空约束,促进区域间技术、信息与资源的流动与优化配置。新质生产力的发展在推动本地战略性新兴产业和未来产业加速形成的同时,依托产业链延伸、技术扩散与要素联动带动周边地区的产业升级,由此形成生产力的空间溢出,并促进区域经济的协调发展(丁任重 等,2024)^[5]。另一方面,“数据要素×科技创新”构成一个极具空间穿透力的开放系统,具备显著的空间外溢性。数据要素具有非竞争性、强协同性与可复用性等特征,其与科技创新深度耦合能够形成跨越地理界限的传导机制(姜江泳 等,2025)^[11],如数据驱动的创新成果通过研发网络等平台实现跨区域共享,基于数据互联的产业生态促使技术标准与生产范式沿供应链向周边扩散(张斌 等,2024)^[7],领先地区的实践经验通过公共知识网络产生区域性示范效应。在这两方面空间溢出的共同作用下,“数据要素×科技创新”对新质生产力发展的影响同样具有显著的空间溢出效应。“数据要素×科技创新”可以带动相邻地区的“数据要素×科技创新”,从而促进相邻地区的新质生产力发展;同时,“数据要素×科技创新”驱动本地新质生产力发展,进而通过新质生产力的溢出效应带动相邻地区的新质生产力发展。由此,通过多维空间溢出机制,“数据要素×科技创新”的作用突破地区边界的约束,形成驱动新质生产力的网络化区域联动格局。

据此,本文提出假设 H2:“数据要素×科技创新”对新质生产力的驱动存在显著的空间溢出效应,能够促进相邻地区的新质生产力发展。

三、实证研究设计

1. GNN-SDM 模型构建

由于数据要素、科技创新、新质生产力均存在显著的空间相关性,采用传统的最小二乘法进行模型估计忽略了区域间的空间依赖性,会导致参数估计偏误,本文采用空间计量模型进行实证检验。空间杜宾模型(SDM)能够同时刻画自变量和因变量的空间传导机制,适用于考察区域发展中的交互与溢出效应。然而,SDM 在应用中仍存在一定局限:一方面,其依赖于研究者预先设定的自变量和空间权重矩阵,难以全面刻画区域之间潜在的复杂网络关系与非线性耦合特征;另一方面,其通常假定空间效应以线性方式传导,可能忽略不同区域间的异质性互动机制。为克服这些不足,本文引入图神经网络(GNN)作为特征提取工具,通过自适应学习区域间非线性交互关系和提取高维度的区域嵌入特征,提升空间计量分析的拟合精度和解释能力。基于 GNN 在空间回归分析中的良好表现,Zhu 等(2022)^[17]提出了 SRGCNN 框架,该框架将图神经网络与空间回归模型有机融合,在空间特征传播、半监督数据预测与异质性建模方面较传统模型更具优势。因此,本文将 GNN 学习得到的融合特征嵌入空间计量模型,具体来讲,建模流程主要包括图结构构建、图特征提取、特征降维与 GNN-SDM 模型构建四个步骤:

第一,图结构构建。基于各地区在不同年份的数据要素发展水平(DAT)与科技发展水平(TEC)构建节点特征矩阵 $X_t \in R^{N \times d}$ 。为了构建节点间的图结构,依据地区间的异质性关系生成异质性邻接矩阵 $A_t: A_{ij}^t = \exp \left(-\frac{\|x_i^t - x_j^t\|^2}{\sigma^2} \right)$ 。其中, x_i^t 表示 i 地区在 t 年的标准化向量, σ 为平滑参数。同时,结合空间权重矩阵 W ,年度图可以表示为 $G_t = (\nu, \varepsilon_t)$,节点集合为区域单位 $\nu = \{1, 2, \dots, N\}$,边集合 ε_t 由 A_{ij}^t 与 W 共同定义。以此作为图神经网络的结构输入。

第二,图特征提取。为捕捉区域间的耦合关系及其潜在空间异质性,采用图注意力网络(GAT)对每一年度图进行特征提取。GAT 通过自注意力机制动态分配相邻节点的边权,从而学习节点的表征特征(Velickovic et al. ,2018)^[18]。第 i 层的节点表示更新为: $h_i^{(l+1)} = \sigma \left(\sum_{j \in N_i} \alpha_{ij}^{(l)} W^{(l)} h_j^{(l)} \right)$ 。其中, $\sigma(\cdot)$ 为激活函数, N_i 为节点 i 的邻居集合, $\alpha_{ij}^{(l)}$ 是注意力权重。基于此对 GAT 网络进行训练,为避免过拟合,将隐藏维度设为 4,注意力头数设为 4,学习率为 0.01,每年迭代 500 轮。训练完成后,每个节点在每年获得的隐藏层输出向量即为学习到的融合特征: $Z_t \in R^{N \times d'}$,其综合反映了复杂空间结构中的交互影响。

第三,特征降维。由于 GNN 输出特征维度较高,为降低高维嵌入特征中的冗余信息,提升模型的稳定性并降低多重共线性风险,对每年的融合特征 Z_t 进行主成分分析(PCA),依据方差贡献率提取主要成分,构建融合特征指标,作为最终嵌入空间计量模型的代表性协同特征变量,即: $GNN_{n_{it}} = \sum_{m=1}^{d'} \varphi_m^n z_{it}^m$ 。其中, φ_m^n 表示第 n 个主成分的载荷系数, $GNN_{n_{it}}$ 为提取的第 n 个 GNN 特征变量,用于建模分析。

第四,GNN-SDM 模型构建。为兼顾理论逻辑与数据结构,提升模型结果的解释力与稳健性,将“数据要素×科技创新”与 GNN 提取的融合特征一并纳入 SDM 中,构建 GNN-SDM 模型如下所示:

$$NQP_{it} = \alpha_0 + \rho_0 W \times NQP_{it} + \alpha_1 SYN_{it} + \rho_1 W \times SYN_{it} + \alpha_2 GNN1_{it} + \rho_2 W \times GNN1_{it} + \alpha_3 GNN2_{it} + \rho_3 W \times GNN2_{it} + \alpha_4 C_{it} + \rho_4 W \times C_{it} + \mu_i + \lambda_t + \varepsilon_{it}$$

其中,下标 i 和 t 分别代表地区(省份)和年份,被解释变量 NQP 为地区新质生产力发展水平,解释变量 SYN 为“数据要素×科技创新”(地区数据要素发展水平与科技创新水平的耦合协调度), $GNN1$ 与

$GNN2$ 为生成的主成分融合特征, C 表示控制变量, $\alpha_0 \sim \alpha_4$ 是模型待估参数, 空间自回归系数 ρ_0 反映新质生产力在空间上的依赖性, $\rho_1, \rho_2, \rho_3, \rho_4$ 分别反映 $SYN, GNN1, GNN2$ 、控制变量的空间交互作用, W 为空间权重矩阵, μ_i 和 λ_t 分别为地区(省份)和时间(年份)固定效应, ε_{it} 为随机扰动项。

2. 变量选取与测度

(1) 被解释变量地区新质生产力发展水平(NQP)的测度。劳动者、劳动资料与劳动对象是构成生产力的三大基本要素, 在新质生产力的发展进程中, 这三大要素的结构和内容需同步升级与优化(蒋永穆等, 2025)^[19]。借鉴王珏和王荣基(2024)^[20]、张哲等(2024)^[21]的研究, 本文构建如表 1 所示的指标体系, 通过熵权 Topsis 法计算得到样本地区在样本期间的新质生产力发展水平。

表 1 地区新质生产力发展水平评价指标体系

维度	一级指标	二级指标	衡量方式	属性
劳动者	劳动者质量	受教育程度	人均受教育平均年限	正
		人力资本结构	大学生数量占总人口比重	正
	劳动生产率	人均产值	人均国内生产总值	正
		人均收入	在岗职工平均工资	正
	劳动者意识	就业结构	第三产业就业人员占总就业比重	正
		创业理念	创业活跃度(每百人新创企业数)	正
劳动对象	产业发展	新兴产业	电子商务企业占总企业的比重	正
			软件业务和信息技术服务业收入/国内生产总值	正
		未来产业	机器人密度(机器人数量/总人口)	正
	生态环境	绿色环保	森林覆盖率	正
			环境保护力度(环境保护支出/政府公共财政支出)	正
		污染防治	一般工业固体废物产生量/国内生产总值	负
劳动资料	物质生产资料	数字基础设施	光缆密度	正
			人均互联网宽带接入端口数	正
		能源消耗	能源消耗/国内生产总值	负
			电力消费总量/能源消费总量	正
	无形劳动资料	科技创新	人均专利授权数量	正
			R&D 经费支出/国内生产总值	正
		数字化转型	数字经济指数	正
			企业数字化水平	正

(2) 解释变量“数据要素×科技创新”(SYN)的测度。参考潘宏亮等(2025)^[22]、陶长琪和丁煜(2022)^[23]的研究, 从数据要素基础支撑、数据要素应用水平与数据要素转化水平 3 个维度构建数据要素发展水平评价指标体系如表 2 所示; 参考王洁洁等(2024)^[24]的研究, 从创新投入和创新产出 2 个维度构建科技创新发展水平评价指标体系如表 3 所示。首先, 采用熵权 Topsis 法分别计算得到样本地区在样本期间的数据要素发展水平(DAT)和科技创新发展水平(TEC), 再计算两者的耦合协调度(DAT 和 TEC 的权重相等, 均为 0.5), 即为变量 SYN 。

表 2 地区数据要素发展水平评价指标体系

维度	衡量指标	属性
数据要素基础支撑	人均互联网宽带接入端口数	正
	每万人互联网域名数	正
	每万人互联网网页数	正
	每万人 IPV4 地址数	正
数据要素应用水平	每百家企事业单位网站数	正
	工业机器人安装密度	正
	电子商务销售额占生产总值的比重	正
	有电子商务活动企业占比	正
数据要素转化水平	数字普惠金融—数字支付	正
	数字普惠金融—数字保险	正
	数字普惠金融—数字化程度	正
	高技术产业新产品销售收入占生产总值的比重	正
数据要素转化水平	软件产品收入占生产总值的比重	正
	信息技术服务收入占生产总值的比重	正
	技术市场成交额占第三产业增加值比重	正

表 3 地区科技创新发展水平评价指标体系

维度	衡量指标	属性
创新投入	规模以上工业企业新产品经费占总产值的比重	正
	R&D 经费支出占总产值的比重	正
	规模以上工业企业 R&D 人员全时当量占城镇就业人口的比重	正
创新产出	每万人发明专利申请授权量	正
	规模以上工业企业新产品销售收入占第三产业增加值比重	正
	技术市场成交额占第三产业增加值比重	正

(3)控制变量选取。为缓解遗漏变量问题,参考相关文献的研究(张苏等,2024)^[25],选取以下控制变量:一是工业化水平(*IND*),用工业增加值占地区生产总值的比重衡量;二是政府干预程度(*GOV*),用财政支出与地区生产总值之比衡量;三是城镇化水平(*URB*),用城市人口占总人口的比重衡量;四是信息化水平(*INF*),用电信业务总量与地区生产总值之比衡量;五是金融发展水平(*FIN*),用年末金融业增加值占地区生产总值的比重衡量。

3. 空间权重矩阵设定

借鉴邵帅等(2016)^[26]的研究,本文的空间计量模型采用地理经济距离嵌套矩阵 W ,具体计算方法为: $W = \alpha W_1 + (1-\alpha) W_2$ 。其中,地理距离矩阵 $W_1 = \frac{1}{d_{ij}}$,经济距离矩阵 $W_2 = \frac{1}{|g_i - g_j|}$, d_{ij} 为地区 i 与地区 j 的省会城市之间最短公路里程, g_i 和 g_j 分别为地区 i 与地区 j 的人均 GDP 年均值,为简化分析, α 取值为 0.5。同时,为避免量纲不一致,对权重矩阵 W 进行行标准化处理。通过这种方式构建的空间权重矩阵既反映了地理邻近性,又兼顾了经济相似性,能较好地刻画区域之间的空间关联特征。

4. 样本选择与数据处理

本文以我国 30 个省份为研究样本(限于数据可获得性,不包括西藏自治区和港澳台地区),样本期间为 2011—2023 年。指标测算数据主要来自《中国统计年鉴》《中国工业统计年鉴》《中国环境统计年鉴》《中国能源统计年鉴》《中国科技统计年鉴》《中国人口与就业年鉴》,以及各省份的统计年鉴、国民经济和社会发展统计公报等;数字普惠金融指数由北京大学数字金融研究中心和蚂蚁科技集团研究院组成的联合课题组编制,数字经济指数与企业数字化水平数据来自马克数据网;对部分变量的缺失数据,采用线性插值法和平均值法进行补齐。主要变量的描述性统计结果如表 4 所示。需要说明的是,由于 $GNN1$ 和 $GNN2$ 经过标准化与主成分提取处理,生成的特征变量呈现均值接近于零的分布特征,这是该类嵌入方法捕捉复杂空间结构的自然结果,并不影响其解释性。

表 4 主要变量描述性统计结果

变量类型	变量名称	变量符号	样本量	均值	标准差	最大值	最小值
被解释变量	新质生产力发展水平	NQP	390	0.142	0.086	0.670	0.044
	数据要素发展水平	DAT	390	0.106	0.102	0.739	0.015
	科技创新发展水平	TEC	390	0.131	0.118	0.762	0.004
解释变量	数据要素×科技创新	SYN	390	0.236	0.112	0.707	0.001
	$GNN1$		390	5.360×10^{-10}	0.375	1.174	-0.828
	$GNN2$		390	-7.750×10^{-9}	0.267	1.226	-0.468
控制变量	工业化水平	IND	390	0.331	0.080	0.574	0.100
	政府干预程度	GOV	390	0.257	0.111	0.758	0.105
	城镇化水平	URB	390	0.606	0.200	0.896	0.350
控制变量	信息化水平	INF	390	0.048	0.050	0.285	0.010
	金融发展水平	FIN	390	0.073	0.032	0.198	0.026

图 1 和图 2 展示了新质生产力发展水平和“数据要素×科技创新”水平的变化趋势,可以看出:我国各地区的新质生产力发展水平持续提升,尤其在 2018 年后增速明显加快,其中,东部地区始终保持领先;“数据要素×科技创新”水平也不断提高,东部地区的领先优势同样明显。总体来看,新质生产力发展水平和“数据要素×科技创新”水平提升轨迹基本一致,初步表明二者之间存在显著的正向关联,即“数据要素×科技创新”正在成为驱动新质生产力跃升的重要动力。

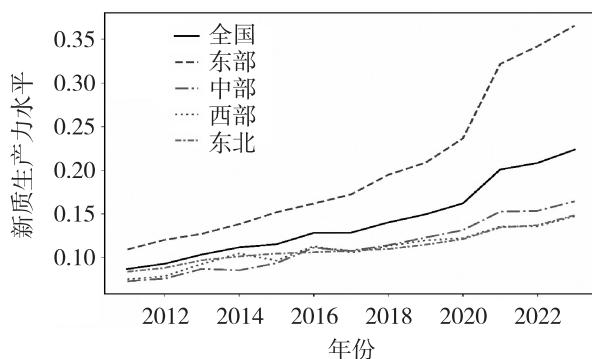


图 1 地区新质生产力发展趋势

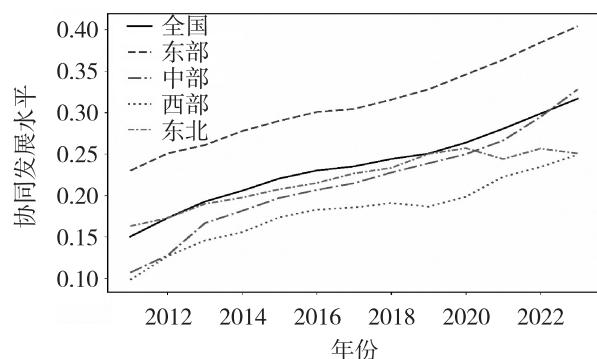


图 2 地区“数据要素×科技创新”水平趋势

四、实证结果分析

1. 空间自相关检验与空间计量模型选择

GNN 特征本质上是经过图神经网络学习得到的高维表征,已经在图结构中嵌入了空间与异质性信息,因而不需要进行空间相关性检验。新质生产力发展水平和“数据要素×科技创新”水平的全局 Moran's I 见表 5,均为正且通过了 1% 的显著性水平检验,表明二者均存在显著的空间自相关性,即新质生产力发展水平和“数据要素×科技创新”水平的空间分布并非随机的,而是表现出一定程度的空间集聚特征。因此,需要运用空间计量模型进行检验。

表 5 新质生产力发展水平和“数据要素×科技创新”水平的全局莫兰指数

年份	NQP			SYN			年份	NQP			SYN		
	<i>Moran'sI</i>	Z 值	P 值	<i>Moran'sI</i>	Z 值	P 值		<i>Moran'sI</i>	Z 值	P 值	<i>Moran'sI</i>	Z 值	P 值
2011	0.217	5.148	0.000	0.201	4.645	0.000	2018	0.183	4.346	0.000	0.168	4.238	0.000
2012	0.223	5.176	0.000	0.231	5.348	0.000	2019	0.178	4.243	0.000	0.157	3.939	0.000
2013	0.185	4.430	0.000	0.216	5.008	0.000	2020	0.201	4.659	0.000	0.160	4.014	0.000
2014	0.153	3.731	0.000	0.208	5.030	0.000	2021	0.168	4.116	0.000	0.158	4.023	0.000
2015	0.221	5.157	0.000	0.200	4.944	0.000	2022	0.164	4.038	0.000	0.174	4.315	0.000
2016	0.126	3.199	0.001	0.202	4.990	0.000	2023	0.155	3.826	0.000	0.159	3.904	0.000
2017	0.195	4.629	0.000	0.184	4.614	0.000							

在开展空间计量回归分析之前,需要检验模型的适用性与有效性,本文主要进行了 LM 检验、Hausman 检验、LR 检验和 Wald 检验(见表 6)。LM 检验结果表明样本数据存在空间效应,且空间滞后项与空间误差项均显著存在,可以初步判断适合构建空间杜宾模型;Hausman 检验统计量在 1% 的水平上显著,表明在空间杜宾模型的框架下,固定效应模型优于随机效应模型;LR 与 Wald 检验结果表明双向固定效应模型在分析过程中能够更有效地捕捉数据的复杂关系,空间杜宾模型无法简化为空间自回归模型,并优于空间误差模型。因此,本文最终选择采用包含双向固定效应的空间杜宾模型,并结合 GNN 特征构建 CNN-SDM 模型。

表 6 模型筛选检验结果

检验方法	统计值	检验方法	统计值
LM err	63.564 ***	LRsar	58.900 ****
Robust LM err	6.589 **	LR sem	60.560 ***
LM lag	74.565 ***	LR 固定效应	125.300 ***
Robust LM lag	17.590 ***	Waldsar	125.140 ***
Hausman	72.990 ***	Wald sem	1405.030 ***

注: *、**、*** 分别代表系数在 10%、5%、1% 的水平下显著,下表同。

2. 空间计量回归结果分析

在分析“数据要素×科技创新”的作用之前,本文分别以数据要素发展水平(*DAT*)和科技创新发展水平(*TEC*)及其交互项(*DAT*×*TEC*)为解释变量进行了 SDM 模型检验,估计结果见表 7 的列(1)、列(2)和

列(3)。 DAT 和 $W \times DAT$ 的系数、 TEC 和 $W \times TEC$ 的系数、 $DAT \times TEC$ 和 $W \times DAT \times TEC$ 的系数均显著为正,表明无论是数据要素发展还是科技创新发展,均对新质生产力发展具有显著的促进作用,且两者存在显著的协同效应(相互强化对方对新质生产力发展的促进作用);同时,上述影响均表现出显著的正向空间溢出效应,即数据要素发展和科技创新发展及其协同发展,不仅提升了本地的新质生产力发展水平,还促进了相邻地区的新质生产力发展水平提升。

基于地理经济距离嵌套矩阵的 SDM 模型和 GNN-SDM 模型检验结果见表 7 的列(4)和列(5)。 SYN 和 $W \times SYN$ 的系数均显著为正,表明“数据要素×科技创新”不仅能驱动本地区的新质生产力提升,还能够通过空间溢出效应带动相邻地区的新质生产力发展,进而实现区域联动的整体新质生产力进步。同时,空间依赖系数 rho 显著,进一步验证了模型中存在强烈的空间自相关性,有必要引入空间计量模型。由此,本文提出的假设 H1 和假设 H2 得到验证。

表 7 基准空间计量模型检验结果

变 量	NQP				
	(1)	(2)	(3)	(4)	(5)
DAT	0.629 *** (0.061)				
TEC		0.347 *** (0.050)			
$DAT \times TEC$			0.689 *** (0.075)		
SYN				0.468 *** (0.069)	0.393 *** (0.066)
$GNN1$					-0.377 *** (0.075)
$GNN2$					0.424 *** (0.068)
IND	-0.072 (0.079)	-0.050 (0.082)	-0.098 (0.082)	-0.035 (0.081)	-0.041 (0.074)
GOV	-0.001 (0.060)	-0.015 (0.068)	0.016 (0.067)	-0.049 (0.065)	-0.046 (0.059)
URB	-0.101 (0.096)	-0.158 (0.105)	0.009 (0.104)	-0.276 *** (0.104)	-0.186 ** (0.097)
INF	0.071 (0.127)	0.009 (0.130)	-0.016 (0.129)	0.059 (0.135)	0.056 (0.130)
FIN	-0.149 (0.282)	0.147 (0.281)	-0.153 (0.300)	0.500 * (0.274)	0.439 * (0.255)
$W \times DAT$	0.497 ** (0.212)				
$W \times TEC$		0.552 *** (0.138)			
$W \times DAT \times TEC$			0.694 *** (0.264)		
$W \times SYN$				0.422 ** (0.209)	0.474 ** (0.205)
$W \times GNN1$					0.368 *** (0.077)
$W \times GNN2$					-0.407 *** (0.070)
$W \times IND$	-0.056 (0.292)	-0.222 (0.325)	-0.106 (0.314)	-0.060 (0.317)	0.048 (0.294)
$W \times GOV$	-0.565 *** (0.207)	-0.814 *** (0.240)	-0.471 ** (0.221)	-0.848 *** (0.238)	-0.647 *** (0.232)
$W \times URB$	-0.473 * (0.260)	-0.827 *** (0.294)	-0.211 (0.263)	-0.772 ** (0.330)	-0.706 ** (0.306)
$W \times INF$	-0.032 (0.140)	0.077 (0.145)	0.009 (0.143)	0.060 (0.151)	0.044 (0.147)
$W \times FIN$	0.182 (0.803)	2.678 *** (0.841)	0.678 (0.827)	1.182 (0.845)	0.947 (0.794)
年份固定效应	控制	控制	控制	控制	控制
省份固定效应	控制	控制	控制	控制	控制
rho	0.589 *** (0.089)	0.336 *** (0.123)	0.491 *** (0.106)	0.592 *** (0.090)	0.582 *** (0.091)
样本量	390	390	390	390	390
R^2	0.642	0.572	0.591	0.563	0.638

注:括号内为标准误差,下表同。

此外, $GNN1$ 和 $GNN2$ 均表现出高度的显著性,且作用方向相反: $GNN1$ 显著抑制本地新质生产力发展,而 $GNN2$ 则表现出强烈的促进作用。这一发现有力地证实了 GNN 方法能够捕捉到传统计量模型可能忽略的非线性、异质性空间效应,从而揭示区域间更为丰富与复杂的关联机制。GNN 特征同样展现出显著的空间溢出效应, $GNN1$ 具有显著的正向空间溢出效应,而 $GNN2$ 则表现为负向溢出效应。此外,引入 GNN 特征后,模型的拟合优度也显著提高。上述检验结果表明, GNN-SDM 模型能够揭示传统方法难以捕捉的差异化空间溢出路径。

3. 稳健性检验

为确保实证分析结果的可靠性,借鉴徐瑾和单子强(2024)^[27]、郑威和陈辉(2023)^[28]的研究,进行以下稳健性检验:第一,更换空间权重矩阵。将基准模型中的地理经济距离嵌套矩阵 W 分别替换为地理距离矩阵 W_1 和经济距离矩阵 W_2 ,重新进行估计。第二,剔除直辖市样本。考虑到直辖市在资源配置与发展模式上具有较强的特殊性,将直辖市样本剔除后重新进行回归。第三,替换被解释变量。采用主成分分析法重新计算地区新质生产力发展水平($NQP1$),以其为被解释变量重新进行检验。第四,控制变量滞后处理。将控制变量进行滞后一期处理,重新进行模型回归。上述稳健性检验结果见表 8,除剔除直辖市样本检验中 $W \times SYN$ 的系数不显著外,其他结果均与基准模型基本一致,表明本文的分析结果具有较好的稳健性。

表 8 稳健性检验结果

变量	NQP		NQP		NQP	$NQP1$	NQP
	地理距离矩阵 W_1	经济距离矩阵 W_2	剔除直辖市样本	替换被解释变量	控制变量滞后一期		
SYN	0.390 *** (0.070)	0.413 *** (0.063)	0.325 *** (0.070)	3.779 *** (0.448)	0.413 *** (0.070)		
$GNN1$	-0.327 *** (0.069)	-0.414 *** (0.078)	-0.528 *** (0.089)	-2.030 *** (0.444)	-0.382 *** (0.075)		
$GNN2$	0.387 *** (0.063)	0.427 *** (0.069)	0.534 *** (0.080)	2.089 *** (0.400)	0.409 *** (0.069)		
$W \times SYN$	1.144 *** (0.306)	0.307 ** (0.133)	0.384 (0.239)	2.556 * (1.437)	0.476 ** (0.218)		
$W \times GNN1$	0.311 *** (0.070)	0.412 *** (0.079)	0.521 *** (0.091)	1.879 *** (0.456)	0.392 *** (0.076)		
$W \times GNN2$	-0.373 *** (0.065)	-0.401 *** (0.072)	-0.517 *** (0.082)	-2.009 *** (0.414)	-0.408 *** (0.071)		
样本量	390	390	338	390	360		
R^2	0.609	0.639	0.619	0.849	0.634		

注:所有模型均控制了控制变量和固定效应,限于篇幅,控制变量估计结果略,下表同。

4. 直接效应与间接效应分解

为更清晰地刻画解释变量对新质生产力的影响,本文进一步在三种空间权重矩阵下将 SYN 、 $GNN1$ 、 $GNN2$ 的总效应分解为直接效应与间接效应,结果见表 9。 SYN 的直接效应和间接效应均显著为正,且间接效应均远大于直接效应,表明“数据要素×科技创新”的区域联动对新质生产力发展至关重要。 $GNN1$ 的直接效应显著为负,间接效应显著为正,表明其刻画的空间结构具有反向激励的空间扩散特征,能够反映地区新质生产力发展中的内部阻碍,如结构性失衡或资源错配对本地新质生产力具有抑制作用,但对相邻地区的新质生产力发展可能产生促进作用。 $GNN2$ 的直接效应显著为正,间接效应显著为负,表明其刻画的空间结构具有“中心-外围”的极化特征,能够反映地区新质生产力发展中的“虹吸效应”,如

资源集聚促进本地新质生产力发展,同时导致相邻地区资源流出,进而抑制相邻地区的新质生产力发展。因此,对 $GNN1$ 值高的地区,应着力改善其内部结构与效率,消除阻碍新质生产力发展的内生因素;而对 $GNN2$ 值的高地区,应推动跨区联动,引导其充分释放外溢效应,促进新质生产力的区域联动发展。

表 9 直接效应与间接效应分解结果

变 量	NQP		
	地理距离矩阵 W_1	经济距离矩阵 W_2	地理经济距离嵌套矩阵 W_3
SYN	直接效应 0.391 *** (0.072)	0.490 *** (0.069)	0.447 *** (0.071)
	间接效应 1.140 *** (0.230)	1.221 *** (0.299)	1.677 *** (0.518)
$GNN1$	直接效应 -0.330 *** (0.066)	-0.392 *** (0.070)	-0.370 *** (0.070)
	间接效应 0.313 *** (0.068)	0.388 *** (0.077)	0.346 *** (0.078)
$GNN2$	直接效应 0.394 *** (0.061)	0.413 *** (0.063)	0.420 *** (0.063)
	间接效应 -0.379 *** (0.062)	-0.350 *** (0.072)	-0.376 *** (0.074)

5. 进一步讨论:“数据要素×科技创新”驱动本地新质生产力的异质性分析

本文进行以下 3 方面的异质性检验:一是地理区位异质性。根据图 1 和图 2,中部、西部以及东北地区发展状态相近,并与东部地区差异明显,故参考刘家民和马晓钰(2024)^[29]的研究,将全样本划分为东部地区和其他地区两组,分别进行检验。二是“数据要素×科技创新”水平异质性。根据 SYN 变量的均值,将样本划分为高耦合协调组和低耦合协调组,分别进行检验。三是新质生产力水平异质性。根据 NQP 变量的均值,将样本划分为高水平组和低水平组,分别进行检验。异质性分析结果见表 10。在各组样本中, SYN 的回归系数均在 1% 的水平上显著为正,但东部地区组的系数显著大于其他地区组、高耦合协调组的系数显著大于低耦合协调组、高水平组的系数显著大于低水平组,表明“数据要素×科技创新”对本地新质生产力发展的驱动作用在基础条件较优越的东部地区、“数据要素×科技创新”水平较高的地区、新质生产力领先的地区更强。因此,在加快发展新质生产力的过程中,必须秉持因地制宜、分类施策的原则。

表 10 异质性分析结果

变 量	NQP					
	地理区位异质性		耦合协调发展水平异质性		新质生产力水平异质性	
	东部地区	其他地区	高耦合协调组	低耦合协调组	高水平组	低水平组
	(1)	(2)	(3)	(4)	(5)	(6)
SYN	0.607 *** (0.137)	0.154 *** (0.028)	0.640 *** (0.141)	0.136 *** (0.033)	1.396 *** (0.236)	0.146 *** (0.027)
样本量	130	260	169	221	104	286
R^2	0.814	0.701	0.752	0.615	0.850	0.706
系数差异显著性	0.577 *** (0.087)		0.393 *** (0.071)		0.981 *** (0.119)	

五、结论与启示

本文采用我国 30 个省份 2011—2023 年的面板数据,以数据要素发展水平和科技创新发展水平的耦

合协调度衡量“数据要素×科技创新”水平,运用 GNN-SDM 模型检验“数据要素×科技创新”对地区新质生产力的影响及其空间溢出效应,研究发现:(1)新质生产力发展水平和“数据要素×科技创新”水平具有显著的空间自相关性,有必要采用空间计量模型来考察两者之间的关系。(2)采用 SDM 模型分析表明,数据要素发展和科技创新发展不仅能显著驱动本地新质生产力发展,还表现出显著的空间溢出效应,促进相邻地区的新质生产力发展;同时,数据要素与科技创新对新质生产力的驱动存在显著的协同作用(两者叠加产生更强的新质生产力驱动作用),该作用也具有显著的正向空间溢出效应。(3)采用 GNN-SDM 模型分析表明,“数据要素×科技创新”显著驱动了本地和邻近地区的新质生产力发展;同时, GNN1 刻画出“抑制本地、带动周边”的反向激励型空间扩散结构,而 GNN2 则刻画出“增强本地、抑制周边”的“中心-外围”型空间极化结构。因此,对于 GNN1 值较高的地区,应重点优化内部结构、提升资源配置效率和自主创新能力,打破对外部的依赖;而对于 GNN2 值较高的地区,应强化区域间的合作联动,避免形成资源过度集聚的“虹吸效应”。(4)异质性分析表明,“数据要素×科技创新”对本地新质生产力的驱动作用,在经济基础较好的东部地区、“数据要素×科技创新”水平较高的地区、新质生产力发展水平较高的地区更强。

基于本文研究,得到以下三点启示:第一,夯实数据要素市场基础,促进跨区域协同发展。加快建立规范的数据市场与产权制度,并实施“数据要素×”融合战略以促进数据要素与科技创新的跨区域共享,构建高效联动的区域发展格局,充分释放数据要素的乘数效应,为新质生产力发展提供强大动能。第二,构建多维空间识别模型,助力精准施策。加快构建一个融合经济基础、要素结构、创新网络与空间联动等多维度的识别模型,该模型的核心功能在于精准识别区域协同发展的异质性特征,并动态诊断发展瓶颈。基于模型输出的科学认知,决策者可优化区域创新布局与要素配置策略,制定前瞻性、靶向精准的差异化政策,高效推动区域新质生产力的协同发展。第三,因地制宜发展新质生产力,促进区域协同共进。政策制定应立足于各地区的异质性条件,实施精准匹配的分类指导。一方面,引导综合条件优越的“领先区”聚焦核心技术突破,发挥其创新策源与辐射带动作用;另一方面,支持“追赶区”结合自身资源禀赋,打造差异化比较优势。在此基础上,完善“领先区”与“追赶区”间的知识溢出、技术转移、产业协作机制,通过“结对共建”等模式形成高带低、强促弱的协同共进格局。

参考文献:

- [1] 王天雨,卢宇.数据要素赋能新质生产力发展与新业态培育的耦合关系、内在机制和发展建议[J].西南金融,2025(8):18-29.
- [2] 周文,许凌云.论新质生产力:内涵特征与重要着力点[J].改革,2023(10):1-13.
- [3] 冯永琦,林鳳峰.数据要素赋能新质生产力:理论逻辑与实践路径[J].经济学家,2024(5):15-24.
- [4] 徐翔,李帅臻,李涛.数据要素赋能新质生产力形成:逻辑、机理与路径[J].经济社会体制比较,2024(6):41-50.
- [5] 丁任重,李溪铭.新质生产力的理论基础、时代逻辑与实践路径[J].经济纵横,2024(4):1-11.
- [6] 梁昊光,黄伟.科技创新驱动新质生产力及其全球效应[J].财贸经济,2024,45(8):22-32.
- [7] 张斌,李亮.“数据要素×”驱动新质生产力:内在逻辑与实现路径[J].当代经济管理,2024,46(8):1-10.
- [8] 刘晨,李睿婕.科技创新促进文化新质生产力发展:理论分析与实证检验[J].中国科技论坛,2024(12):44-54.
- [9] 王文霞,贾永飞,孙童童.科技创新驱动新质生产力形成与发展机理研究——基于场域理论视角[J].经济与管理评论,2025,41(3):105-117.
- [10] 任宇新,贺正楚.“科技-产业-金融”融合互动与新质生产力发展路径[J].财经理论与实践,2025,46(3):67-74.
- [11] 姜汇泳,刘桂峰.数据要素赋能新质生产力的模型构建与实现机制[J].图书馆论坛,2025,45(10):21-29.

- [12] 陆扬,王育宝.数据要素市场化与新质生产力发展——基于双重机器学习的因果推断[J].城市问题,2024(7):80-90.
- [13] 姚博伦,邓云柯.数据要素集聚对企业新质生产力发展的影响研究[J].统计与信息论坛,2025,40(2):36-49.
- [14] 张小宇,张曼.数据要素共享与新质生产力——基于公共数据开放的经验证据[J].上海经济研究,2025(8):29-42.
- [15] 陈晔婷,何思源.数据要素价值化赋能数字新质生产力的效果及机制研究[J].科研管理,2025,46(10):21-30.
- [16] 徐建伟,李子文,韩晓.加快发展新质生产力:推进思路与政策建议[J].改革,2025(1):40-52.
- [17] ZHU D, LIU Y, YAO X, et al. Spatial regression graph convolutional neural networks: A deep learning paradigm for spatial multivariate distributions[J]. GeoInformatica, 2021, 26(4):1-32.
- [18] VELIKOVI P, CUCURULL G, CASANOVA A, et al. Graph attention networks[C]//Proceedings of the 6th International Conference on Learning Representations. Vancouver: Elsevier, 2018:1-12.
- [19] 蒋永穆.新质生产力理论对马克思主义生产力理论的继承和发展[J].经济学动态,2025(5):5-18.
- [20] 王珏,王荣基.新质生产力:指标构建与时空演进[J].西安财经大学学报,2024,37(1):31-47.
- [21] 张哲,李季刚,汤努尔·哈力克.中国新质生产力发展水平测度与时空演进[J].统计与决策,2024,40(9):18-23.
- [22] 潘宏亮,赵兰香,叶璐.我国数据要素发展水平的测度及时空演进研究[J].科学学研究,2025,43(1):205-216.
- [23] 陶长琪,丁煌.数据要素何以成为创新红利?——源于人力资本匹配的证据[J].中国软科学,2022(5):45-56.
- [24] 王洁洁,马晓君,范祎洁.数字经济与科技创新的协同效应对经济高质量发展的影响研究[J].统计与信息论坛,2024,39(4):46-62.
- [25] 张苏,朱媛.人口老龄化、数字化转型与新质生产力发展[J].北京工商大学学报(社会科学版),2024,39(3):28-39.
- [26] 邵帅,李欣,曹建华,等.中国雾霾污染治理的经济政策选择——基于空间溢出效应的视角[J].经济研究,2016,51(9):73-88.
- [27] 徐瑾,单子强.绿色技术创新对城市碳减排的影响研究——基于空间计量模型的实证分析[J].经济问题探索,2024(9):155-172.
- [28] 郑威,陈辉.数据要素市场化配置对城市经济韧性的影响:促进还是抑制?——基于数据交易平台设立的准自然实验[J].现代财经(天津财经大学学报),2023,43(12):78-92.
- [29] 刘家民,马晓钰.数智化创新政策如何推动企业新质生产力发展[J].西部论坛,2024,34(4):17-34.

The Spatial Spillover Effects of “Data Elements × Technological Innovation” on New Quality Productive Forces: An Empirical Test Based on a GNN-SDM Model

CHEN Yi-an^{1, 2}, YAN Yue¹

(1. School of Mathematics and Statistics, Chongqing Technology and Business University, Chongqing 400067, China;

2. Chongqing Key Laboratory of Economic and Social Applied Statistics, Chongqing 400067, China)

Summary: In the digital economy era, “data elements ×” is emerging as the core engine driving the development of new quality productive forces by unleashing the multiplier effect of data. It not only effectively empowers industrial transformation and upgrading but also serves as a key pillar for achieving high-quality economic development. In this process, data elements and technological innovation are deeply integrated, jointly constituting the dual-wheel drive for the growth of new quality productive forces. However, empirical research on the intrinsic mechanisms linking data elements, technological innovation, and new quality

productive forces remains scarce, particularly in identifying the pathways for their coordinated development and spatial diffusion.

This study employs provincial-level panel data from 2011 to 2023 to measure the coupling synergistic effects between data elements and technological innovation. Drawing on the methodology of Zhu et al. (2022), it utilizes graph neural networks (GNN) to extract fusion features of both elements. These features, along with their coupling synergistic effects, are embedded into a spatial Durbin model (SDM) to establish the GNN-SDM model for empirical analysis. The findings reveal that the coupling synergistic effect not only significantly promotes the development of new quality productive forces within the region but also drives improvements in surrounding areas through spatial spillover mechanisms. Heterogeneity analysis confirms significant regional variations in the synergistic driving effect of data elements and technological innovation, with this effect being particularly pronounced in eastern regions, areas with high synergy levels, and regions leading in new quality productive forces. This highlights the critical role of regional foundational conditions in constraining the realization of synergy dividends. Furthermore, intergroup differences in GNN feature variables validate the diversity of regional spatial interaction mechanisms. Moreover, GNN reveals two spatial diffusion mechanisms: GNN1 exhibits an inverse incentive structure where structural imbalances within a region inhibit its own development but can stimulate surrounding areas through spatial spillovers; GNN2 reflects a central-peripheral innovation polarization structure where intra-regional innovation linkage significantly boosts local new quality productive forces while generating a “siphon effect” on neighboring areas.

Compared to previous research, this paper makes two main contributions. First, from a synergy perspective, it proposes that data elements and technological innovation have a coupled-driving, spatially diffusing effect on new quality productive forces, thereby expanding the research perspective on new quality productive forces and providing empirical support and policy references for inter-provincial collaborative governance. Second, methodologically, by integrating GNN with SDM, it achieves precise characterization of complex spatial correlation structures and heterogeneous innovation diffusion pathways, enhancing the ability to identify the spatial spillover effects of synergistic development on new quality productive forces.

This study introduces graph neural networks as the core feature extraction tool. Leveraging their powerful structural learning capability, the model adaptively captures complex nonlinear interactions and spatial dependencies among regions, effectively extracting highly representative spatial embedding features. Building upon traditional spatial econometric models, this approach further enhances the fitting accuracy and generalization ability of the model for actual economic and geographical phenomena and enhances the interpretability of the model for spatial interaction mechanisms. This method provides a new analytical perspective and methodological support for deeply revealing the internal mechanism of regional coordinated development and identifying the transmission paths of cross-regional linkage effects.

Keywords: data element; technological innovation; new quality productive forces; coupling synergistic effect; GNN-SDM model

CLC number: F124; F49

Document code: A

Article ID: 1674-8131(2025)06-0041-14

(编辑:朱艳;刘仁芳)