

应用统计 与数据科学

Applied Statistics and Data Science



ART AND TECHNOLOGY PRESS INC.

(517 666 0904)

263 S KENWOOD ST 560

CASPER, WY 82601

Copyright © 2026 by ART AND TECHNOLOGY PRESS INC. (United States)

Complimentary Copy



ART AND TECHNOLOGY PRESS INC.
(United States)

编委会名单

主 编



王国长 (Guochang Wang), 暨南大学 (Jinan University),
邮箱: wanggc023@amss.ac.cn
研究兴趣: 函数型数据分析、时间序列、充分性降维、统计与机器学习
functional data analysis, time series, sufficiency dimension reduction, statistics and machine learning

执行副主编



张兴发 (Xingfa Zhang), 广州大学 (Guangzhou University),
邮箱: xingfazhang@gzhu.edu.cn
研究兴趣: 金融统计、环境统计、机器学习
financial statistics, environmental statistics, machine learning

编 委



胡志勇 (Zhiyong Hu), 广州大学 (Guangzhou University)
邮箱: zhyhu65@163.com
研究兴趣: 大数据, 人工智能, 财务与会计
big data, artificial intelligence, finance and accounting



雷田礼 (Tianli Lei), 深圳职业技术大学 (Shen Zhen Polytechnic University)
邮箱: ltl@szpu.edu.cn
研究兴趣: 数量经济、高职数学教育
auantitative economics, higher vocational mathematics education



舒连杰 (Lianjie Shu), 澳门大学 (University of Macau),
邮箱: ljshu@um.edu.mo
研究兴趣: 量化金融, 统计学习, 质量控制及管理
quantitative finance, statistical learning, quality control and management



王伟 (Wei Wang), 山东财经大学 (Shandong University of Finance and Economics)
邮箱: wangwei_0115@sdufe.edu.cn
研究兴趣: 统计机器学习、应用统计、计量经济
statistical machine learning, applied statistics, econometrics



杨凯 (Kai Yang), 长春工业大学 (Changchun University of Technology)
邮箱: yangkai@ccut.edu.cn
研究兴趣: 时间序列分析、高维数据分析、贝叶斯分析
time series analysis, high-dimensional data analysis, bayesian analysis



周彦 (Yan Zhou), 深圳大学 (Shenzhen University)
邮箱: zhouy1016@163.com
研究兴趣: 生物统计, 数据科学
Biostatistics, data science



李永明 (Yongming Li), 上饶师范学院 (Shangrao Normal University)
邮箱: lym1019@163.com
研究兴趣: 非参数统计, 金融统计, 极限理论与统计大样本性质, 小波方法
nonparametric statistics, financial statistics, limit theory and statistical large sample theory, wavelet method



刘照德 (Zhao de Liu), 广东财经大学 (Guangdong University of Finance & Economics)
邮箱: lzhaode@163.com
研究兴趣: 经济统计、多元统计、计量分析
economic statistics, multivariate statistics, econometric analysis



王小刚 (Xiaogang Wang), 北方民族大学 (North Minzu University)
邮箱: wongone@163.com
研究兴趣: 经济与社会统计
economic and social statistics



严晓东 (Xiaodong Yan), 西安交通大学 (Xi'an Jiaotong University)
邮箱: yanxiaodong@xjtu.edu.cn
研究兴趣: 统计机器学习、智能体统计学
statistical machine learning, agent statistics



朱柯 (Ke Zhu), 香港大学 (The University of Hong Kong)
邮箱: mazhuke@hku.hk
研究兴趣: 统计建模、金融时间序列分析、计量经济、金融大数据
statistical modeling, financial time series analysis, econometrics, financial big data

应用统计与数据科学

Applied Statistics and Data Science

第2卷 第4期 2026年4月刊

主办单位: ART AND TECHNOLOGY PRESS INC.

(United States)

承办单位: 广州统计师事务所

指导单位: 广东省现场统计学会、广州市统计学会

主 编: 王国长

执行副主编: 张兴发

编 委: 胡志勇、李永明、雷田礼、刘照德、
舒连杰、王小刚、王 纬、严晓东、
杨 凯、朱 柯、周 彦

地 址: 263 S KENWOOD ST 560

CASPER,WY 82601

网 址: <https://arttechpress.com>

本刊说明:

凡向本刊所投稿件, 全体作者需签署论文著作权转让声明书和论文发表承诺书, 声明、承诺及相关事项如下:

- 作者将论文的复制权、发行权、网络传播权、翻译权、汇编权、信息网络传播权、改编权等著作权在世界范围内免费转让给本刊。
- 论文不侵犯他人著作权和其他权利, 否则作者将承担由此产生的全部责任, 并赔偿由此给出版单位造成的全部损失。
- 论文署名作者享有该作品的完全著作权, 署名作者的身份真实。
- 论文未曾以任何形式公开发表过。
- 作者所投本刊稿件, 本刊编辑部拥有修改权。



管理统计 | MANAGEMENT STATISTICS

- 001 公益性与运营效率的平衡: 妇幼保健院绩效核算体系改革路径与效果分析 冯英娜
Balancing Public Welfare and Operational Efficiency: Reform Path and Effect Analysis of the Performance Accounting System in a Maternal and Child Health Hospital Feng Yingna

统计与教育教学 | STATISTICS AND EDUCATION TEACHING

- 005 科教融合与竞赛驱动的统计学综合实验教学改革与教学设计 周燕, 肖莉, 杨志程, 丁仕虹, 夏英俊
Integration of Scientific Research and Education with Competition-Driven Reform and Instructional Design in Comprehensive Statistics Experiment Teaching Zhou Yan, Xiao Li, Yang Zhicheng, Ding Shihong, Xia Yingjun
- 009 人工智能与数学竞赛双赋能《高等数学》教学改革研究 高会双
Research on the Teaching Reform of Advanced Mathematics Empowered by Artificial Intelligence and Mathematical Competitions Gao Huishuang
- 012 突发事件协同应急能力影响因素体系构建 李文博
Building a System of Factors Influencing Collaborative Emergency Response Capability for Emergencies Li Wenbo
- 019 OBE 理念下应用型会计人才“听学用”培养路径研究 刘欢娜
Research on the Training Path of "Understandable, Learnable, Applicable" for Application-Oriented Accounting Talents under the OBE Concept Liu Huanna

数据科学技术、方法与应用 | DATA SCIENCE TECHNOLOGY, METHODS AND APPLICATIONS

- 023 基于轻量化深度卷积神经网络的番茄叶片病害分类识别研究 崔鹏, 李杰, 汤岩春, 许名扬, 王晓松
Research on Classification and Recognition of Tomato Leaf Diseases Based on Lightweight Deep Convolutional Neural Networks Cui Peng, Li Jie, Tang Yanchun, Xu Mingyang, Wang Xiaosong
- 030 基于数据增强与 XGBoost 模型的信用卡欺诈检测算法 邓磊
Credit Card Fraud Detection Algorithm Based on Data Augmentation and XGBoost Deng Lei
- 037 新能源汽车协同效应与发展路径: 基于耦合协调与随机森林分析 王宇琪
Synergy Effects and Development Paths of New Energy Vehicles: Based on Coupling Coordination and Random Forest Analysis Wang Yuqi

统计方法及应用 | STATISTICAL METHODS AND APPLICATIONS

- 042 中国式全域旅游现代化指标评价体系构建研究——以江苏为例 刘广应, 马国婷, 林金官
Research on the Construction of a Chinese-style Modernization Indicator Evaluation System for All-for-One Tourism — A Case Study of Jiangsu Province Liu Guangying, Ma Guoting, Lin Jinguan

049	GARCH 型过程稳健控制设计及其在金融市场监管中的应用 Robust Control Design for GARCH-Type Processes and Its Application in Financial Market Monitoring	王志坚，麦智 Wang Zhijian, Mai Zhi
056	基于自适应权重的 MIDAS–GARCH 模型 MIDAS–GARCH Model Based on Adaptive Weights	刘培宇，许敏，尹兰江 Liu Peiyu, Xu Min, Yin Lanjiang
064	山西省二手交易市场现状分析 Analysis of the Current Situation of the Second-hand Trading Market in Shanxi Province	原缓缓，李雨航，史华杰，杨盼 Yuan Huanhuan, Li Yuhang, Shi Huajie, Yang Pan
069	体育赛事网络关注度与举办地旅游客流量的格兰杰因果检验——以 2025 年江苏省足球比赛（“苏超”）为例 The Granger Causality Test of the Network Attention of Sports Events and the Tourist Flow of the Host City — Taking the 2025 Jiangsu Football Match (' Suchao ') as an Example	王超，蔡捷 Wang Chao, Cai Jie

公益性与运营效率的平衡：妇幼保健院绩效核算体系改革路径与效果分析

冯英娜

佛山市三水区妇幼保健院，广东 韶关 528100

DOI:10.61369/ASDS.2026040001

摘 要： 在新医改与公立医院高质量发展的背景下，妇幼保健院面临坚守公益属性与提升运营效率的双重压力。本文以 S 妇幼保健院为案例，探讨其绩效核算体系改革路径。研究分析了医院如何通过构建以 DIP 病种分值、RBRVS 工作量点值与公益性 KPI 为核心的复合型绩效模型，有效平衡公益性与运营效率。实践证明，该改革在保障妇幼公共卫生任务落实、引导合理诊疗、控制医疗成本及调动医务人员积极性方面成效显著，为同类医疗机构在医保支付改革下的精细化管理提供了可借鉴的实践经验。

关 键 词： 公益性；运营效率；绩效改革；DIP（按病种分值付费）；RBRVS；妇幼保健院

Balancing Public Welfare and Operational Efficiency: Reform Path and Effect Analysis of the Performance Accounting System in a Maternal and Child Health Hospital

Feng Yingna

Maternal and Child Health Hospital of Sanshui District, Foshan City, Shaoguan, Guangdong 528100

Abstract： Under the background of new healthcare reforms and the high-quality development of public hospitals, maternal and child health hospitals face dual pressures of adhering to their public welfare nature and enhancing operational efficiency. Taking Hospital S, a maternal and child health hospital, as a case study, this paper explores the reform path of its performance accounting system. The research analyzes how the hospital effectively balances public welfare and operational efficiency by constructing a composite performance model centered on the Diagnosis-Intervention Packet (DIP) point system, the Resource-Based Relative Value Scale (RBRVS) workload point values, and public welfare Key Performance Indicators (KPIs). Practice has proven that this reform achieves remarkable results in ensuring the implementation of maternal and child public health tasks, guiding rational clinical practices, controlling medical costs, and motivating medical staff. It provides transferable practical experience for similar medical institutions in implementing refined management under healthcare payment reforms.

Keywords： public welfare; operational efficiency; performance reform; DIP (Diagnosis-Intervention Packet); RBRVS; maternal and child health hospital

引言

妇幼保健院是守护区域妇女儿童健康的核心力量，其独特的公共卫生职能决定了它必须在公益性与可持续运营之间寻求平衡。随着广东省按病种分值付费（DIP）改革的全面推行，医院传统的收入模式和运营逻辑被深刻重塑，既往以经济收入为导向的绩效核算体系已难以适应新发展要求。如何设计一套科学的“指挥棒”，既能激励效率，又能坚守公益，成为亟待解决的管理难题。本研究基于 S 妇幼保健院（以下简称 S 院）的改革实践，旨在系统剖析其绩效核算新体系的构建路径与实施效果，以期为推动同类机构的高质量发展提供务实参考与决策依据。

一、改革动因与理论框架：新形势下妇幼保健院的绩效困境与出路

当前，我国公立医院改革已进入以“高质量发展”为核心的新阶段，其内在要求是回归公益本源并提升运营效能。对于兼具临床与公共卫生职能的妇幼保健机构而言，这一挑战尤为严峻。一方面，《关于推动公立医院高质量发展的意见》等政策文件明确要求强化公益性；另一方面，广东省全面推行的按病种分值付费（DIP）改革，从根本上重塑了医院的补偿机制，使粗放式收入增长模式难以为继。在此背景下，S院原有的、以“收支结余”为核心的绩效核算体系弊端凸显：它客观上激励“创收”，与控费的DIP逻辑相悖，并导致轻视妇幼公卫项目、难以衡量技术劳务价值等内部不公。因此，构建一个能同时响应外部政策压力与内部激励需求的绩效新体系，不仅是实现医院可持续发展的管理命题，更是履行好政府赋予的公共卫生职责的必然选择。本章旨在系统梳理此次改革的内外动因，并奠定以“价值医疗”为导向，融合DIP激励与RBRVS量化工具的理论框架，为后续的路径分析与效果评估提供基础^[1]。

二、S院绩效核算新体系的构建路径

面对旧体系的积弊与外部环境的变革，S院以“公益导向、效率优先、优绩优酬、持续发展”为核心原则，启动了绩效核算体系的系统性重构。新体系摒弃了单一的经济导向，转而构建了一个融合工作量、成本控制、质量安全与公益属性的复合型核算模型，其核心架构如下。

（一）临床与医技科室：以价值医疗为核心的复合型核算模式

新体系下，科室绩效不再是简单的收入分成，而是由多重维度共同决定，其核心计算公式为：科室当月应发绩效工资 = (工作量绩效工资 + 成本控制绩效工资) × 医疗质量考评系数 × 管理KPI考评系数 × 风险系数 × 调节系数 ± 单项奖罚。该公式将分配与核心价值目标紧密捆绑。

1. 工作量绩效工资：从“创收”到“创值”的转变

工作量绩效的设计旨在科学衡量不同科室、不同岗位的技术劳务价值，淡化收入因素，引导资源配置优化^[2]。

门（急）诊工作量：创新性地引入了“门（急）诊工作量标准化系数”。该系数并非单一依据人次，而是综合考量药品处方量、检查检验单数量及其价值，通过加权计算（权重分别为0.2, 0.5, 0.3）得出一个综合值。此举更精准地反映了不同科室门诊工作的技术强度与复杂程度，避免了“只看人头不看内容”的粗放管理。

住院工作量：这是适应DIP改革的核心举措。可以将广东省DIP平台数据结果中的核心指标——病例组合指数（CMI）值，作为各科室住院工作量的标化系数（如下图）。CMI值越高，代表该科室收治患者的平均技术难度和资源消耗越高，其绩效奖励也相应提升。这直接引导临床科室愿意并积极收治病情更复杂、权

重更高的病种，与医保支付改革的导向完全一致^[3]。

表1：基于DIP-CMI的住院绩效激励导向

科室行为导向	对DIP结算的影响	对绩效核算的影响
收治病情更复杂、难度更高的病例	科室整体CMI值提升，DIP总分值增加	住院工作量绩效的核算系数提高，绩效工资增长
优化临床路径，合理控制费用	实际医疗费用低于DIP病组支付标准，形成“结余”	通过“成本控制绩效”获得额外奖励
规避重症，仅收治简单病例	科室整体CMI值下降，DIP总分值减少	住院工作量绩效的核算系数降低，绩效工资减少

手术与护理工作量：引入以资源为基础的相对价值比率（RBRVS）理论。对HIS系统中的每一项手术，依据其技术难度、风险程度和资源消耗综合设定点值。护理工作量则结合护理项目点值与护理床日数进行核算。这实现了“同工同酬、优劳优得”，精准衡量了医护人员的技术劳动价值。

2. 成本控制绩效工资：植入DIP下的运营效率基因

此项设计直面DIP预付制的挑战，公式为“（科室核算收入 - 科室核算支出）× 成本控制考评系数”。在DIP“结余留用”的机制下，科室只有主动优化临床路径、规范医疗行为、控制耗材与药品使用，降低“核算支出”，才能获得正向的绩效回报，这将成本控制的责任与利益直接下沉到科室，激发了全员参与运营管理的主动性^[4]。

3. 质量安全与公益性KPI：守护医院发展的生命线

前述两项绩效工资最终需要通过质量与公益的考核系数进行调节，设置“一票否决”项与浮动区间，确保发展不偏离轨道。

医疗质量考评系数：由医务科牵头，考核病历质量、抗菌药物使用强度、手术并发症发生率等核心医疗安全指标，增减幅度不超过绩效总额的5%。

管理KPI考评系数：此部分是平衡公益性的关键抓手。除了常规的运营效率指标，S院特别设置了公益性导向指标，如妇幼重大公卫项目（“两癌”筛查、艾梅乙阻断等）的完成率与质量、健康教育开展场次与效果、患者满意度、投诉处理及时率等，此系数增减幅度可达10%，强力引导科室行为符合公益定位。

（二）行政后勤部门：从“成本中心”到“价值伙伴”的探索

针对行政后勤部门工作难以量化、价值体现不直观的普遍难题，S院参考先进经验，启动了从“平均主义”向“价值贡献”的分配改革。其核心思路是建立“院科二级分配”体系^[5]。

院级分配（到部门）：计划借鉴“岗位价值评估”方法（如海氏评价法），从责任大小、知识技能、工作复杂性等维度，评估不同行政后勤部门的价值贡献，设定差异化的部门系数，打破“一刀切”的分配模式。

表2：行政后勤部门分类与系数设定示例

类别	部门特点	代表部门	建议系数区间
I类（核心管理）	专业性强，与临床融合深，责任风险高	医务科、护理部、信息科	1.4 - 1.5
II类（职能管理）	综合管理要求高，协调任务重	院办、党办、人力资源部、财务科	1.2 - 1.3
III类（保障服务）	提供基础运行保障，工作相对标准化	总务科、设备科、保卫科	1.0 - 1.1

部门二次分配（到个人）：将分配权部分下放至部门负责人。科室可参照临床 RBRVS 思路，对内部常规职责与临时性任务进行工作量赋值，或建立包含执行力、工作质量、服务态度等在内的 KPI 考核表，实现“千多干少、干好干坏”不一样，激发行政后勤人员服务临床一线的内生动力^[6]。

通过上述系统化的设计，S 院构建了一个能够响应 DIP 支付压力、体现技术劳务价值、并坚守公益底线的绩效核算新体系，为改革的落地实施奠定了坚实的制度基础。

三、改革实施的保障、成效与挑战分析

绩效核算体系的改革，是从理念蓝图走向管理实践的复杂过程。分析 S 院新绩效体系的实施保障措施，其核心举措与运营成效之间形成的有效耦合联动，并最终构建起一个从“方案设计”到“执行保障”，再到“成效评估与挑战反思”的完整价值闭环，是同类医疗机构的参考关键。

（一）系统实施的协同保障措施

为确保绩效改革平稳落地并有效运行，S 院构建了一个多维度、系统化的保障体系：

1. 组织与领导保障

成立了由院长担任组长，财务、医务、护理、人力资源、信息等核心职能部门负责人共同参与的“绩效改革领导小组”。该小组不仅负责方案的顶层设计与决策，更在推行过程中通过召开动员会、科室宣讲会及个别访谈等形式，全面解读维护公益性、调动积极性的改革理念，有效疏导并化解了科室的疑虑，为改革营造了良好的内部环境^[7]。同时，S 院借鉴了院科两级责任制的管理经验，通过签订综合目标管理责任书，将绩效目标层层分解，确保“一级管一级，层层抓落实”。

2. 指标体系整合与信息化支撑

新体系的复杂性对数据治理与信息化提出了极高要求。

①指标融合

将国家与省级卫生健康部门对妇幼保健机构的绩效考核要求（如辖区管理、服务提供、运行效率、持续发展、满意度评价等维度）内化到 S 院的管理 KPI 与公益性考评系数中。例如，将“辖区孕产妇系统管理率”、“婴儿死亡率”等公益性产出指标纳入科室管理 KPI 考核，权重向公共卫生服务倾斜。

②数据整合与系统升级

S 院投入专项资金，升级了医院信息系统（HIS）与绩效考核系统，重点实现了与广东省 DIP 平台的数据无缝对接。系统能够自动采集门急诊人次、手术操作、CMI 值、病种分值等关键数据，并按照预设的 RBRVS 点值表与标化系数模型进行自动化核算。这极大地提升了核算效率与准确性，避免了人为干预，确保了考核的公平、公正与透明。

③沟通反馈与政策协同

建立了常态化的“绩效考核沟通反馈机制”。绩效办公室定期向科室反馈绩效核算的初步结果，并允许科室在限定时间内进行数据核对与申述。这种开放的态度，既尊重了科室的知情权与参与权，也能及时发现了系统运行初期存在的数据抓取漏洞，实现了体系的持续优化。同时，S 院积极将考核结果与财政投入、绩效工资总量核定、评优评先等挂钩，强化了外部政策与内部激励的协同效应。

（二）改革成效的实证分析：公益与效率的协同显现

通过对改革前后（以 2022 年为基期，2023 年为观察期）关键数据的纵向对比，新体系的导向作用与积极成效逐步显现，具体如下表所示：

表 3：绩效改革关键指标对比分析表（2022vs2023）

评估维度	关键指标	改革前（2022 年）	改革后（2023 年）	变化趋势与解读
公益性体现	妇幼重大公卫项目完成率	95%	98.5%	显著提升，表明科室对公益任务的投入度增加。
	患者满意度	88.5%	91.8%	服务态度与就医体验改善，公益服务内涵深化。
运营效率提升	次均门诊费用（元）	285	278	在 DIP 成本考核与合理控制费用导向下，费用控制初见成效。
	全院 CMI 值	1.05	1.12	明确导向：科室更愿意收治疑难重症，体现技术价值。
	平均住院日（天）	6.5	6.0	临床路径优化，床位周转加快，效率提升。
	药占比（%）	28.5	26.1	成本控制意识增强，收入结构优化。
	DIP 医保基金结余合理率	-	显著改善	通过事前事中监控，医疗行为更规范。

深度分析：

1. 公益性的“主动回归”

改革并非以牺牲效率来保全公益，而是通过制度设计让公益行为“有价值”。当“公卫项目完成率”与“患者满意度”直接关联到科室和个人的绩效时，各科室从过去的“被动接受任务”转变为主动争取和优化服务，实现了公益性的内生驱动，这与妇幼保健机构绩效考核的初衷高度吻合^[7]。

2. 运营效率的“价值导向”

DIP - CMI 导向与成本控制绩效彻底改变了临床行为模式。以往可能被视为“负担”的高难度手术和复杂病例，现在因其能带来更高的 CMI 值和绩效回报而成为科室积极发展的方向。同

时，借助信息化系统对药占比、耗材比、次均费用等关键运营指标进行实时监测与预警，使运营效率的提升建立在全员、全流程的精细化管理基础上。

（三）面临的挑战与局限性反思

在取得显著成效的同时，改革进程中也暴露出一些亟待关注与解决的问题：

1. 数据治理的长期性与编码准确性：新体系高度依赖准确、及时的医疗与运营数据。初期，因诊断书写不规范、疾病分类编码与手术操作编码不准确等问题导致的数据失真时有发生，影响了核算结果的公信力。这要求我们必须持续加强病案首页质量管理与医务人员的数据素养培训。

2.“战略性亏损”科室的平衡难题：对于儿科、新生儿科等CMI值天然偏低但社会效益巨大的科室，尽管已通过门急诊标准化系数等进行调节，但其绩效水平与高技术难度外科科室的差距依然存在。如何更科学地评估其社会价值与健康产出（如降低婴儿死亡率）并给予合理补偿，是未来需要持续探索的课题^[8]。

3.行政后勤量化考核的深化困境：相较于临床科室，行政后勤的量化考核难度更大。虽然建立了二级分配框架，但岗位价值评估、工作量赋值的科学性与公平性仍需在实践磨合中不断完善。

4.潜在的“博弈”行为与监管需求：需要警惕在DIP和CMI指挥棒下，可能出现的“诊断升级”或“高套分组”等道德风险。这要求管理者必须借鉴“事前预警、事中监控、事后稽核”的智能监管体系，将强化医疗质量监管与智能审核作为绩效体系不可或缺的补充和纠偏机制^[9]。

综上所述，S院以公益性运营效率平衡为目标的绩效改革已取得阶段性成功，但其长期效果与可持续性，有赖于对上述挑战的持续应对与体系的动态优化。

四、结语

本研究通过对S妇幼保健院绩效核算体系改革的全面分析表明，公益性与运营效率并非不可兼得。通过建立以DIP病种分价值和RBRVS工作量为核心的绩效核算体系，医院成功实现了从“规模扩张”到“价值医疗”的转型发展。这一改革实践证实，将CMI值、公共卫生任务完成率等关键指标与绩效分配直接挂钩，能够有效引导医疗行为向提升技术服务价值和落实公益职能的方向转变，同时通过成本控制机制激发了全员参与精细化管理的主动性。

基于改革实践，建议医院管理者坚定改革决心，持续推进以价值为导向的绩效分配制度改革，并加强信息化建设作为支撑。同时，呼吁卫生行政部门完善配套政策，考虑建立针对儿科等公益性较强科室的专项绩效补偿机制，为公立医院改革创造更有利的政策环境。展望未来，绩效改革仍需在行政后勤考核、医联体绩效协同及长期成效评估等方面持续探索。

参考文献

- [1] 马青. 加强医院绩效核算及管理, 促进医院发展 [J]. 现代营销 (经营版), 2021, (04): 18-19. DOI: 10.19921/j.cnki.1009-2994.2021-02-0016-007.
- [2] 孙良琼. 医院绩效考核与会计成本核算体系的构建分析 [J]. 行政事业资产与财务, 2024, (22): 13-15.
- [3] 杨盛虹. 医院绩效考核与核算体系的现状及优化分析 [J]. 现代商贸工业, 2024, 45(16): 120-122. DOI: 10.19311/j.cnki.1672-3198.2024.16.044.
- [4] 司悦. 公立医院绩效核算与成本核算的结合研究 [J]. 财会学习, 2024, (03): 95-97.
- [5] 王红雁, 韩露, 姜丽萍. 公立医院绩效考核全成本核算体系设计 [J]. 中国市场, 2022, (17): 164-166. DOI: 10.13939/j.cnki.zgsc.2022.17.164.
- [6] 李冰. 有利于公益的三级公立医院绩效管理指标系统探索 [J]. 生产力研究, 2020, (09): 152-156+160. DOI: 10.19374/j.cnki.14-1145/f.2020.09.032.
- [7] 韩靖远, 李宗泽, 韩斌斌. 高质量发展背景下公立医院社会责任评价模型构建 [J]. 中国医院, 2024, 28(12): 36-39. DOI: 10.19660/j.issn.1671-0592.2024.12.07.
- [8] 王绒. 医院社会责任与可持续发展战略 [J]. 财经界, 2024, (26): 48-50. DOI: 10.19887/j.cnki.cn11-4098/f.2024.26.053.
- [9] 吕江. 公立医院高质量发展中的社会责任研究 [J]. 社会与公益, 2024, (09): 168-169.

科教融合与竞赛驱动的统计学综合实验教学改革 与教学设计

周燕, 肖莉, 杨志程, 丁仕虹, 夏英俊
华南农业大学 数学与信息学院, 广东 广州 510642
DOI:10.61369/ASDS.2026040002

摘要： 面对大数据时代对统计学人才创新能力与实践能力的更高要求，传统以验证性、封闭性为主的实验教学模式已显露出诸多不足。本研究依托省级教改课题，构建并实践了“科教融合、竞赛驱动”的统计学综合实验教学改革新模式。该模式以“创新能力培养”为核心，通过重构课程体系，将前沿科研课题、高水平学科竞赛与创新创业训练项目深度融合入教学全过程，构建“基础技能—综合实践—科研反哺—竞赛升华”四阶递进、双向贯通的育人体系。本文详细阐述了改革的教学理念、教学设计、具体实施路径，并紧密结合在全国大学生统计建模大赛、市场调查分析大赛及大创项目中产生的一系列高水平学生成果案例，检验了该模式在激发学生科研兴趣、培养综合数据分析能力、高质量教学成果产出等方面的显著成效。

关键词： 科教融合；竞赛驱动；实验教学；教学设计；创新能力

Integration of Scientific Research and Education with Competition-Driven Reform and Instructional Design in Comprehensive Statistics Experiment Teaching

Zhou Yan, Xiao Li, Yang Zhicheng, Ding Shihong, Xia Yingjun

College of Mathematics and Informatics, South China Agricultural University, Guangzhou, Guangdong 510642

Abstract : In response to the heightened demands for innovative and practical abilities among statistics professionals in the era of big data, the traditional experiment-based teaching model, which primarily emphasizes verification and closed-ended tasks, has revealed several shortcomings. Building upon a provincial-level educational reform project, this study proposes and implements a novel model for comprehensive statistics experiment teaching reform, characterized by the “integration of scientific research and education” and driven by academic competitions. Centered on fostering innovation capabilities, this model restructures the curriculum system by deeply incorporating cutting-edge research topics, high-level academic competitions, and innovation-entrepreneurship training programs into the entire teaching process. It establishes an educational framework featuring a “four-stage progressive system” — foundational skill development, comprehensive practice, research feedback, and competition elevation — with bidirectional integration. This paper elaborates on the teaching philosophy, instructional design, and specific implementation pathways of the reform. Furthermore, it closely aligns with a series of high-quality student achievements from national statistical modeling competitions, market research and analysis contests, and innovation training projects to validate the model's effectiveness in stimulating students' research interest, cultivating comprehensive data analysis skills, and generating high-quality teaching outcomes.

Keywords : integration of scientific research and education; competition-driven; experiment teaching; instructional design; innovation capability

基金项目：广东省高等教育教学改革项目“基于创新能力培养的统计学专业数据分析类实验课程改革探索与实践”（粤教高函[2024]9号516）、广东省质量工程项目“华南农业大学广州迪奥信息科技有限公司科产教融合实践教学基地”（粤教高函[2026]4号46）。

作者简介：

周燕，华南农业大学数学与信息学院，硕士，副教授，研究方向：数据挖掘、应用统计；

肖莉，华南农业大学数学与信息学院，博士，教授，研究方向：经济统计与统计决策；

杨志程，华南农业大学数学与信息学院，硕士，讲师，研究方向：应用统计及抽样调查；

丁仕虹，华南农业大学数学与信息学院，硕士，讲师，研究方向：应用统计与统计决策；

夏英俊，华南农业大学数学与信息学院，博士，讲师，研究方向：应用统计与统计决策。

引言

我们现在所处的是一个数据驱动决策的时代。统计学，一门以数据为基础的学科，其作用不仅限于简单的数据整理，而是一种能深入探寻规律的思维方式，对个人和政府决策具有重要影响。统计学作为从数据中提取信息、发现规律并进行预测推断的关键学科，其内涵与外延正在发生深刻变革。当今社会对统计学专业人才的培养需求，从传统的描述统计与参数估计，转为要求其具备从复杂现实场景中发现问题、收集与处理多源数据、构建与评估统计模型，最后将统计分析结果转化为商业决策或政策建议的综合能力，综合来说，是以数据科学为核心的复合型高阶创新能力^[1]。统计学专业的本科实验教学，正是培养这种能力的关键环节。通过实验教学把统计学的理论和方法，连接到真实的现实问题中。特别通过基于实际项目的、跨学科的动手训练，学生得以体验从数据获取、处理、建模到解读结果的全过程^[2]。

（一）国内统计学专业的实验教学现状

1. 教学内容滞后与方法脱节

实验内容多以验证经典统计方法（如假设检验、回归分析）为主，案例陈旧，未能及时纳入机器学习、文本挖掘、网络分析等数据科学技术，导致学生所学技能与行业实际应用存在脱节^[3]。

2. 教学模式单一与场景失真

传统的实验教学过程往往要求学生进行“数据给定、步骤固定、结果已知”的单一操作，并且，这种脱离真实、复杂、多变数据环境的教学，无法培养学生应对现实数据分析中数据质量参差、问题定义模糊等挑战的能力^[4]。

3. 创新能力激发不足与评价单一

实验教学中的考核往往多关注操作步骤的正确性与结果的准确性，忽视了对学生问题发现、方案设计、批判性思维和沟通表达等更高阶能力的评价，难以有效激发学生的创新能力^[5]。

4. 教学科研分离与资源割裂

统计学专任教师的科研前沿成果难以及时转化为教学养分，科研工作与本科实验教学常被视为两条平行线，而专业教学中发掘出来的优秀学生也缺乏高阶问题研究的孵化通道^[6]。

（二）实验教学改革思路

为破解这些困境，本教学团队以省级实验教改课题为依托，提出以“科教融合”与“竞赛驱动”综合实验教学改革的实现路径。打破课程实验、课程实习、毕业设计的传统实验课程的教学安排，以及教学、科研、学科竞赛三者彼此孤立的现状，通过系统性的教学设计，构建动态、开放、互促的实践教学体系。本文旨在阐述这一改革模式的理论框架、设计原则、实施路径，并以近年来指导学生取得的一系列具有代表性的竞赛获奖、科研论文及大创项目成果为教学案例，评估改革成效，为同类院校的统计学专业建设与教学改革提供可借鉴的教学经验。

一、“科教融合、竞赛驱动”教学模式的理论构建与教学设计

（一）理论构建

秉持“以研促学，以赛激创”的教学理念，本改革模式建立在“成果导向（OBE）”和“建构主义”理论上，构建了“四阶递进、双体贯通”的实验教学体系，强调以学生最终获得的能力产出为导向，通过创设真实或拟真的复杂学习任务，引导学生在主动探索和团队合作中，构建知识体系与发展创新能力^[7]。实现教学、科研、竞赛三者的良性循环与深度融合。

（二）四阶递进教学设计

1. 第一阶：基础技能层（课内实验）

（1）教学目标

确保学生熟练掌握进行现代数据分析所必需的编程工具、核心算法与基本流程。学生熟练运用 SPSS、Python、R 进行数据处理与建模，理解数据处理的基本流程，具备跟随实验教程进行基

本分析的能力。

（2）教学设计

对《统计学导论》、《数据挖掘》、《多元统计分析》等专业课程的实验内容进行重构。从教师科研项目、近期竞赛赛题或社会热点中抽取简化但真实的数据片段和基础方法作为教学案例。引入科研与竞赛中的作为教学案例。例如，在讲授文本分析时，引入某上市公司股吧评论的少量真实数据，让学生练习中文分词、去停用词和词频统计，为后续研究网络舆情埋下伏笔。

2. 第二阶：综合实践层（综合实验 / 实习课程）

（1）教学目标

培养学生综合运用多门课程知识解决复杂数据问题，独立完成一个从问题定义到报告呈现的完整数据分析项目的能力。

（2）教学设计

开设《统计学基础实习》、《专业拓展实习》等实践课程，以“选题 - 数据采集 - 清洗 - 探索 - 建模 - 解释 - 报告”完整数据分析流程为主线。选题库直接来源于教师科研子课题，历年学科

竞赛真题或近期经济社会热点问题。学生3-5人一组，在教师指导下，完成选题论证、方案设计、数据收集与处理、模型构建与优化、结果分析及撰写综合性研究报告的全过程。

3. 第三阶：科研反哺层（课外实践 / 大创项目）

（1）教学目标

引导学生接触学术前沿，培养初步的科研素养、批判性思维、独立探索能力和创新能力。学生初步掌握学术研究的基本范式，具备一定的理论联系实际能力和学术论文写作能力。

（2）教学设计

积极鼓励并指导在综合实验中表现突出的学生团队，将其选题进行深化、拓展，申报校级、省级和国家级“大学生创新创业训练计划”项目。教师将自身科研项目中的非核心但具有挑战性的任务（如指标体系的对比、新数据的尝试）转化为开放性问题，吸收学有余力的学生参与。学生在真实科研环境中，学习如何文献综述、提出假设、设计实验、处理数据、建立模型，并尝试撰写课题申报书与学术论文，实现科教深度融合。

4. 第四阶：竞赛驱动层（校外实践 / 学科竞赛）

（1）教学目标

学生具备快速学习、创新性解决问题、高效团队协作和出色表达沟通的能力，产出高水平应用成果，其成果获得行业及学界的认可。

（2）教学设计

将全国大学生统计建模大赛、市场调查分析大赛等教育部学科竞赛榜单赛事作为实验教学成果的校外实践教学平台。第二阶“综合实践”的优秀课程设计作品和第三阶“大创项目”的中期成果，经过进一步凝练和深化，作为参赛作品进行深度打磨。教学团队提供赛前集中培训、模型优化指导、论文撰写辅导等全过程支持。竞赛不仅检验了教学效果，其高强度的任务要求和明确的成果导向，也极大地激发了学生的内在动力和创新潜能。

（三）双向贯通教学机制

为确保这四个阶段有效联动，我们建立了“双向贯通”机制。一方面，教师的科研成果、竞赛的命题转化为第一、二阶的教学案例和项目题库，确保教学内容的前沿性。另一方面，学生在实验、竞赛中产生的见解、遇到的新问题，反馈给教师，促使教师就此开展更深入的方法论研究。这种双向流动，真正实现了教学相长与科教深度融合。

二、实验教学改革实施路径与教学案例

为具体说明上述改革模式如何运作，本节以“网络舆情与文本分析”这一贯穿多个优秀学生成果的研究方向为例，阐述其四阶递进的教学实施路径。

（一）基础技能训练（第一、二阶）

在《数据挖掘》和《统计学基础实习》等实验课程中，我们设立“社会感知与文本数据分析”教学模块。学生在此模块中，系统学习网络爬虫、中文文本预处理、词频统计、情感词典构建、主题模型（LDA）、基础情感分析等核心技能。课程作业

即要求选择一个热点事件，完成从数据采集到初步分析的完整流程。这为后续所有高阶研究奠定了通用的方法基础。

（二）问题深化与科研项目（第三阶）

基于课程中展现出的兴趣与能力，教师引导学生关注更具体的科研问题。例如，教师结合自身关于“企业融资约束”的研究课题，提出“网络舆情是否会以及如何影响中小企业融资？”这一子课题。学生以此为方向，在教师指导下申报了国家级大创项目《基于大数据文本分析的网络舆情对中小企业融资约束的影响研究》。在此阶段，学生需要深入学习更复杂的理论（如信息不对称理论、TOE 框架）和高级模型（如机器学习分类器、面板数据模型），开始尝试构建自己的分析框架。

（三）方法创新与学科竞赛（第四阶）

在大创项目研究基础上，学生团队对问题和方法进行持续迭代与创新，形成了多个具有参赛竞争力的作品。

作品 A（省级二等奖）：在市场调查与分析大赛中，团队作品《基于机器学习的网络舆情情感识别探究对中小企业融资约束的影响》则侧重于方法论的比较与优化，系统对比了不同机器学习算法在情感识别任务上的优劣，并将其应用于融资约束预测。

作品 B（省级一等奖）：在统计建模比赛中，学生团队将研究焦点细化，探讨了《基于大数据文本分析的网络舆情对农业中小企业融资约束的影响研究》。他们不仅运用了成熟的文本情感分析方法，还创新性地结合了农业领域的特色词典，并运用了因果推断模型，使研究更具针对性和深度^[8]。

作品 C（省级二等奖）：另一团队从“ESG（环境、社会 and 治理）”这一新兴视角切入，研究《ESG 绩效对农林牧渔业企业融资约束的影响分析——以企业声誉为中介变量》。该研究引入了结构方程模型（SEM）来验证复杂的中介路径，展现了学生处理多维构念和复杂模型的能力^[9]。

作品 D（国际会议论文）：学生将文本分析方法应用于文化产品领域，完成了《Analyzing and Predicting Online Public Opinion on Movies Based on Text Mining and Deep Learning》。该研究采用了 LSTM 深度学习模型进行情感趋势预测，并与传统 ARIMA 模型对比，展现了学生解决新领域问题的能力^[10]。

作品 E（省级三等奖）：统计建模大赛作品《生成式 AI 会成为学习助手还是“作弊神器”？——大学生使用意愿与行为调查》，采用分层 PPS 抽样进行问卷调查，综合运用多种统计模型，深入剖析了影响大学生使用生成式 AI 的多元因素。说明学生能够独立设计并执行一项从问卷设计到模型构建的完整社会调查研究^[11]。

所有这些成果，都源于一个共同的实验教学内容起点（第一、二阶），并在科研与竞赛的双重驱动下（第三、四阶），向不同的应用领域、理论深度和方法复杂度深化，最终开花结果。

三、总结

近几年，本团队使用改革模式在实验教学实践中取得了显著成效。一方面，学生通过“做中学、研中学、赛中学”，不仅掌

握了扎实的数据分析技能，更系统培养了从问题定义到学术写作的基础科研能力，能够独立解决复杂实际问题；另一方面，教学成果持续涌现，学生在国家级、省级学科竞赛及大创项目中累计获奖数十项，并发表多篇论文；教学过程同时反哺师资与课程建设，推动教师紧跟前沿，并将丰富的学生成果系统化整理为校本

特色案例库。我们将进一步优化“四阶递进”的实验教学体系，探索组建跨学科导师组，深化产教融合设立真实数据赛题，并利用在线平台构建开放共享的实验教学资源，持续深化创新型统计学人才的培养。

参考文献

[1] 大数据时代高校统计学专业现状分析与改革路径 [J]. 中国高等教育, 2019, (18): 58-60.

[2] 郭传好. 基于知识模块化和“一体三翼”教学法的管理统计学教学改革探索 [J]. 大学教育, 2023, (12): 75-77.

[3] 吴建祥. 新文科建设背景下统计学专业实验教学体系的构建与探索 [J]. 黑河学刊, 2025, (01): 70-78.

[4] 马丽娜. 《应用多元统计分析》课程实践教学改革研究 [J]. 办公自动化, 2024, 29(09): 35-37+7.

[5] 杜永军, 王小英, 张学翠. 新时代背景下经济管理类专业统计学课程教学改革 [J]. 高教学刊, 2023, 9(04): 122-125.

[6] 张楷欲, 邓丽娜, 柳晓晨. 智慧农业时代背景下高校生物统计与实验设计课程教学改革研究 [J]. 智慧农业导刊, 2026, 6(01): 137-140.

[7] 张扬, 解柠羽. 基于 OBE 理念的民族院校经管类本科专业统计学教学设计 [J]. 产业与科技论坛, 2022, 21(06): 127-128.

[8] 庄丽凤, 陈晓彤, 黄洁洁, 裴祺阳, 陈永学, 陈立宇, 周燕. 基于大数据文本分析的网络舆情对农业中小企业融资约束的影响研究 [J]. Applied Statistics and Data Science, 2025, 1(6) : 49-49.

[9] 庄丽凤, 陈晓彤, 黄洁洁, 裴祺阳, 陈永学, 陈立宇, 周燕. ESG 绩效对农林牧渔业企业融资约束的影响分析——以企业声誉为中介变量 [J]. Applied Statistics and Data Science, 2025, 1(7) : 40-45. 宋敬敬, 黄华存, 冯家勋. 新农科背景下提升农科学生创新实践能力的实验教学探索 [J]. 基因组学与应用生物学, 2023, 42(08): 888-894.

[10] Si-yao Huang, Hong-kai Zhang, Luo-shi Liang, Yuan-zhi Lin, Jia-hui Lin, and Yan Zhou. Analyzing and Predicting Online Public Opinion on Films Based on Text Mining and Deep Learning[C]. Proceedings of the 2025 2nd International Conference on Modeling, Natural Language Processing and Machine Learning. Association for Computing Machinery, New York, NY, USA, 261 - 267.

[11] 何思佳, 卜胤丹, 周燕. 生成式 AI 赋能下高校学生学术行为意愿与内在机制研究 [J]. Applied Statistics and Data Science, 2025, 1(6) : 55-55.

人工智能与数学竞赛双赋能《高等数学》教学改革研究

高会双

内蒙古民族大学 数学科学学院, 内蒙古 通辽 028000

DOI:10.61369/ASDS.2026040003

摘 要： 本文结合内蒙古民族大学教学实际情况，探索人工智能与数学竞赛双赋能的教学改革路径，从资源开发、流程设计、效果评估等方面构建适配民族院校的融合教学模式，以提升教师教学能力与学生数学核心素养，为同类高校教改提供参考。

关 键 词： 人工智能；数学竞赛；高等数学；教学改革

Research on the Teaching Reform of Advanced Mathematics Empowered by Artificial Intelligence and Mathematical Competitions

Gao Huishuang

College of Mathematics Science, Inner Mongolia Minzu University, Tongliao, Inner Mongolia 028000

Abstract： This paper combines the practical teaching reality of Inner Mongolia Minzu University to explore a teaching reform path with dual empowerment of artificial intelligence and mathematics competitions. It constructs an integrated teaching model suitable for universities for nationalities from the aspects of resource development, process design and effect evaluation, so as to improve teachers teaching ability and students' core mathematical literacy, and provide a reference for teaching reform in similar colleges and universities.

Keywords： artificial intelligence; mathematics competition; advanced mathematics; teaching reform

引言

《高等数学》作为高等院校理工科专业的核心公共基础课，是学生构建专业知识体系、培养逻辑推理能力与科学应用思维的重要基石，其教学质量直接关乎高等院校应用型、创新型人才培养的整体成效。在数智化教育全面推进的时代背景下，多生源背景院校《高等数学》教学的传统弊端日益凸显：教学模式仍以课堂单向讲授为主，师生互动不足，难以兼顾多生源背景院校学生数学基础参差不齐的个性化学习需求；教学内容偏重理论推导与公式记忆，实践应用环节薄弱，学生难以将抽象数学知识与专业学习、实际问题相结合；人工智能教学工具仅停留在课件演示、简单习题批改等浅层应用，未充分发挥其大数据分析、智能演示、个性化答疑的赋能价值；全国大学生数学竞赛等优质竞赛资源的思维训练价值未与课程教学有效衔接，竞赛成为少数学生的“课外拓展”，未能转化为推动全体学生数学能力提升的教学资源，形成“竞赛与教学两张皮”的现象。

全国大学生数学竞赛自2009年发起以来，已形成年度化、规范化的赛事体系，其高阶性、综合性、应用性的试题设计，为高等数学教学提供了优质的思维训练素材，是推动课程教学改革、提升学生数学应用能力的重要驱动力；而人工智能技术的快速发展，为我校高等数学教学提供了全新范式，智能动态演示、个性化学习推送、大数据学情分析等功能，能够突破传统教学的时空限制与形式局限，实现教学的精准化、个性化与高效化。基于此，本文以内蒙古民族大学为研究载体，开展人工智能与数学竞赛双赋能的《高等数学》教学改革研究，打破单一赋能的教学局限，构建“人工智能+数学竞赛+高等数学教学”的三维整合教学体系，推动内蒙古民族大学《高等数学》教学从“知识传授”向“能力培养”转变，为我校数学公共课教学改革注入新活力。

一、双赋能教学改革的研究背景与价值

国内学者围绕数学竞赛与高等数学教学融合、人工智能赋能高等数学教学展开了大量探索，形成了丰富的研究成果，但尚未构建起二者协同赋能的系统性教学框架。赛教融合领域，吕静等

(2025)以全国大学生数学竞赛为依托，发现数学竞赛有利于教师教学能力的培养和教学体系的改革，从变革教学内容、完善教学方法等方面提出了赛教融合策略^[1]。郭连红等(2024)提出融通一二课堂的“数学协会”教、学、做模式，数学竞赛有利于学生创新意识和教师实践能力的提升^[2]。付文娇(2022)认为数学竞

基金项目：内蒙古民族大学教育教学研究课题，“人工智能赋能下数学竞赛融入《高等数学》教学的实践与研究”（项目编号：JG2025006）；内蒙古自治区教育科学研究“十四五”规划课题，“专业融合模式下类比教学方法的研究与实践——以高等数学A下课程为例”（项目编号：NGJGH2024594）。

作者简介：高会双(1980-)，女，内蒙古民族大学数学科学学院副教授，主要从事算子代数和数值代数的研究。

赛有利于更新教师的教学理念,调整教育教学的方式方法,提升学生的学习兴趣^[3]。柴明硕(2020)、方正(2020)认为数学竞赛是破除当前公共课教学瓶颈,丰富课程教学内涵,纵深推进素质教育改革的一块“敲门砖”^[4、5],关丽红等(2017)认为数学竞赛有利于调动学生积极学习,提升学生学习热情,实现教学相长^[6]。

人工智能赋能教学领域,冷奇芳等(2025)依托“立德-授业-赋能-启思-实践”五层教学架构,利用大数据平台构建人工智能与高数课程融合课程,实现“五位一超”^[7],金迎迎(2024)对人工智能应用于高数教学资源库建设、实验指导、学习评价等方面进行案例实践,并对技术难点予以指出^[8],贾晓彤等(2025)对具体融合路径予以探究,为AI与高数融合实践提供案例借鉴^[9]。

但国内研究多为单一维度探索,尚未将AI技术与数学竞赛结合,形成协同赋能高等数学教学的体系化方案。

综上,当前高等数学教学改革亟需打破单一赋能的局限,实现人工智能与数学竞赛的协同赋能,内蒙古民族大学需要结合自身学生基础、教学资源等实际情况,探索适配的双赋能教学模式,破解教学痛点,提升教学质量。

二、人工智能与数学竞赛双赋能《高等数学》教学的核心实施路径

双赋能教学改革的核心是实现人工智能、数学竞赛与《高等数学》教学的深度融合,而非简单的元素叠加,本文从教学资源开发、全流程教学实施、教学效果评估三个维度设计实施路径,同时明确路径推进中的重点、难点,并融入创新设计,确保改革的可操作性与实效性。

(一) 双赋能教学资源的开发与适配

教学资源是双赋能教学改革的基础,研究重点在于实现竞赛资源与高等数学教学的深度适配、AI资源与学生认知的精准匹配,构建体系化、本土化的“竞赛案例+AI工具”融合教学资源库,研究难点则是平衡数学竞赛案例的专业性与《高等数学》教学的基础性,避免案例难度过高导致学生产生畏难情绪,或难度过低无法达到思维训练的效果。

在数学竞赛资源开发方面,紧扣《高等数学》极限、导数、定积分、微分方程等各章节核心知识点,以全国大学生数学竞赛(非专业组)真题为核心素材,结合我校学生的数学基础进行本土化改编,简化复杂推导过程,保留核心思维训练点。例如在定积分应用章节,选取竞赛中平面图形面积计算、简单优化问题的真题进行简化,作为课堂例题与课后练习;在微分方程章节,拆解竞赛案例中高数知识的应用逻辑,转化为贴合教学进度的教学素材,让竞赛资源服务于课堂教学,实现“竞赛为教学奠基,教学为竞赛赋能”。

在人工智能辅助资源开发方面,筛选适配的数学类AI工具(智能演示、个性化答疑、大数据分析类),开发三大核心资源:一是动态演示资源,选取竞赛案例的解题过程,利用AI工具制作

动态演示课件,将抽象的公式推导、定理应用具象化,帮助学生理解解题思路;二是分层练习资源,利用AI工具的个性化生成功能,打造“基础层(竞赛入门题型)-提升层(竞赛真题改编题)”的阶梯式练习体系,匹配不同基础学生的学习需求;三是答疑知识库,借助AI工具的智能答疑功能,整理竞赛与高数融合知识点的常见疑问及标准化解答,为学生自主学习提供支撑。同时,依托学习通等智能教学平台,将各类资源进行整合,实现预习、学习、练习、答疑的一体化对接。

(二) 双赋能全流程课堂教学的设计与实施

打破传统“满堂灌”的教学模式,研究重点在于将人工智能与数学竞赛融入课前、课中、课后全教学流程,设计衔接顺畅、互动性强的实施路径,研究难点在于引导民族院校学生熟练使用人工智能工具开展自主学习,解决部分学生工具使用基础薄弱的问题,让AI工具真正成为学生学习的“助推器”。本文的创新设计在于摒弃零散的元素加入,让人工智能、数学竞赛与高等数学教学围绕知识点形成协同效应——人工智能作为技术载体,为竞赛内容融入提供精准化支撑;数学竞赛作为实践场景,为人工智能技术应用提供落地方向,二者共同服务于高等数学教学目标。

课前:AI赋能精准预习,竞赛案例激发兴趣依托人工智能教学平台,向学生推送高数知识点预习材料、微课视频及简化的数学竞赛案例片段,设置针对性的预习思考题,引导学生标注知识疑问点与竞赛案例的探究点。AI平台自动收集学生的预习数据,通过大数据分析梳理学生的知识薄弱点,教师根据平台反馈精准调整课堂教学重点,实现“以学定教”;同时以趣味性的竞赛案例导入预习环节,打破传统预习的枯燥性,激发学生的学习兴趣与探究欲望,让学生带着问题进入课堂。针对学生AI工具使用基础薄弱的问题,课前设置简单的AI工具操作指引微课,降低学生使用门槛。

课中:AI演示+竞赛拆解,互动探究培养能力采用“基础知识讲解-AI动态演示-竞赛案例拆解-小组互动探究”的教学模式,层层递进,实现知识传授与能力培养的有机结合。首先,教师紧扣教学大纲,讲解高等数学基础知识点与基本解题方法,夯实学生的知识基础;其次,利用AI工具动态演示竞赛案例中公式、定理的应用过程,将抽象的数学知识具象化,帮助学生理解竞赛案例与课堂知识点的内在联系;再次,教师拆解竞赛案例,提炼其中蕴含的高数核心思维方法(数形结合、分类讨论、转化与化归等),引导学生掌握“从竞赛到基础”的知识迁移能力;最后,组织学生开展小组讨论,尝试运用所学知识点解决简化的竞赛案例,教师巡视指导,收集学生的典型解题思路,进行针对性点评与讲解,强化学生的合作探究能力与数学思维。

课后:AI个性化辅导,竞赛任务巩固提升依托人工智能平台实现课后教学的精准化与个性化,同时借助竞赛任务强化知识应用,延伸学习时空。一是个性化答疑与错题解析,AI平台根据学生的课堂表现与练习情况,推送个性化的错题解析,针对学生的知识薄弱点匹配对应的练习题,实现“千人千练”;二是竞赛小组协作任务,安排两人一组完成简单的竞赛题目,培养学生的协作解题能力,让学生在实践中深化对知识点的理解;三是常态化

答疑直播，每周开展一次 AI 智能答疑直播，针对性解决学生在双赋能学习中遇到的问题，同时持续指导学生熟练使用 AI 工具开展自主探究学习。

（三）双赋能教学效果的综合评估与动态优化

教学效果评估是保障双赋能教学改革实效性的关键，研究重点在于建立科学的综合评估体系，同时构建动态优化机制，创新设计在于摒弃固化不变的教学模式，基于评估结果形成“评估－反馈－优化－实施”的闭环教学体系，让双赋能教学模式始终适配我校的教学实际。在评估体系构建方面，采用定量＋定性的综合评估方式，确保评估结果的全面性与客观性。定量评估以自然班为实验对象，设置实验班（双赋能教学）与对照班（常规教学），对比两班的高等数学单元测试和期末成绩，重点分析融入竞赛思路的应用题、综合题得分情况，直观反映学生数学应用能力的提升效果；同时统计实验班学生的 AI 工具使用熟练度、数学竞赛参与率，量化双赋能教学的实施效果。定性评估通过问卷调查了解实验班学生对双赋能教学模式的接受意愿、学习体验与收获；通过师生座谈会收集教师在教学实施过程中的问题与建议；以教学日志的形式记录教学过程，梳理教学中的亮点与不足，包

括 AI 工具的应用效果、竞赛案例的难度适配性、教学环节的衔接顺畅度等。

在动态优化机制构建方面，基于定量与定性评估结果，及时开展教学优化调整：针对竞赛案例难度不当的问题，重新筛选、改编竞赛资源，确保难度贴合学生认知水平；针对学生 AI 工具使用不熟练的问题，增加工具操作指导课时，制作简易操作手册；针对教学环节衔接不畅的问题，优化课前、课中、课后的教学流程，强化各环节的衔接设计。每月开展一次教学反思会，总结教学经验，解决教学痛点，让双赋能教学模式在实践中持续优化，提升其适用性与可持续性。

人工智能与数学竞赛双赋能的《高等数学》教学改革，是顺应数智化教育发展趋势、破解民族院校传统高等数学教学痛点的有效路径，也是提升高等数学教学质量、培养学生数学核心素养的重要举措。本文结合内蒙古民族大学的教学实际，从教学资源开发、全流程教学实施、教学效果评估三个核心维度，设计了双赋能教学改革的实施路径，同时将研究重点、难点与创新点融入路径设计，构建了“技术＋竞赛＋课程”的三维整合教学体系，实现了人工智能与数学竞赛的协同赋能，而非简单的元素叠加。

参考文献

[1] 吕静,刘文英,钟太勇,等.基于以赛促教模式的高职数学课程教学改革实践——以全国大学生数学建模竞赛为例[J].中国教育技术装备,2025,(14):66-69.

[2] 郭连红,金迎迎.赛教融合模式在高等数学课程教学改革中的实践研究[J].牡丹江教育学院学报,2024,(10):58-62.

[3] 付文娇.“以赛促教、以赛促学”教学模式在高等数学教学中的应用[J].数学学习与研究,2022,(05):20-22.

[4] 柴明硕.大学生数学竞赛在大学数学公共课教学中的作用[J].现代职业教育,2020,(07):164-165.

[5] 方正.浅谈大学生数学竞赛在大学数学公共课教学中的作用[J].科教文汇(上旬刊),2020,(01):61-62.

[6] 关丽红,赵亚男.浅谈高等数学教学过程中大学生数学竞赛能力的培养[J].长春大学学报,2017,27(10):84-87.

[7] 冷奇芳,帅燕.人工智能与高职高等数学课程融合创新教学模式的研究[J].太原城市职业技术学院报,2025,(01):87-89.

[8] 金迎迎.人工智能在“高等数学”教学中的应用探究[J].科技风,2024,(01):135-137.

[9] 贾晓彤,王利岩,赵晓丽,等.人工智能赋能高等数学课程教学模式的探索和实践[J].创新教育研究,2025,13(7):230-233.

突发事件协同应急能力影响因素体系构建

李文博

济南大学 管理科学与工程学院, 山东 济南 250000

DOI:10.61369/ASDS.2026040004

摘 要 : 突发事件频发、环境复杂和主体多元化,使得突发事件的应对过程是一个涉及多要素、多主体间相互作用的复杂动态过程。随着多主体参与的应急联动在突发事件管理中的应用,该机制的科学性和有效性被反复检验。然而,作为应急联动的核心协同应急能力影响因素体系构建问题目前还没有获得充分的理论支持,因此,协同应急能力影响因素体系构建迫在眉睫。本研究以协同治理理论为基础,剖析相关概念,从多主体协同视角构建突发事件协同应急能力影响因素体系。以组织协同关系作为中介变量,选取技术条件、组织资源、社会环境、复杂性、以及协同应急能力五个潜变量,运用 AMOS 和 SPSS 为工具平台构建结构方程模型,基于问卷数据进行实证分析。研究表明:技术条件,组织资源,社会环境,复杂性因素,组织协同关系,是影响多主体协同应急能力的重要因素;技术条件,组织资源,社会环境会对组织协同关系造成显著的正向影响从而间接正向作用于协同应急能力;同时,技术条件与组织资源会对复杂性因素产生显著的负向影响。

关 键 词 : 应急管理;多主体协同应急;结构方程

Building a System of Factors Influencing Collaborative Emergency Response Capability for Emergencies

Li Wenbo

School of Management Science and Engineering, University of Jinan, Jinan, Shandong 250000

Abstract : The frequency of emergencies, the complexity of the environment, and the diversity of subjects make the response process to emergencies a complex, dynamic process involving multiple elements and interactions between multiple subjects. With the application of multi-subject emergency linkage in emergency management, the scientificity and effectiveness of this mechanism have been repeatedly tested. However, the problem of constructing a system of influencing factors for collaborative emergency response capacity, which is the core of emergency linkage, has not yet received sufficient theoretical support, and therefore the construction of a system of influencing factors for collaborative emergency response capacity is urgent. Based on the theory of collaborative governance, the research analyzes the relevant concepts and constructs a system of influencing factors for collaborative emergency response capacity from the perspective of multi-organizational collaboration. Five latent variables, namely technical conditions, organizational resources, social environment, complexity, and collaborative emergency response capacity are selected as mediating variables, and a structural equation model is constructed using AMOS and SPSS as the tool platform. Empirical analysis is conducted based on questionnaire data. The study shows that technical conditions, organizational resources, social environment, complexity, and organizational synergy are the most important factors influencing the ability of multi-organizational synergy; technical conditions, organizational resources, and social environment have a significant positive effect on organizational synergy and thus indirectly have a positive effect on synergy; meanwhile, technical conditions and organizational resources have a significant negative effect on complexity.

Keywords : emergency management; multi-subject collaborative emergency response; structural equation modelling

引言

近年来,国内外各类突发事件频发,我国也频繁遭受风雹、干旱、洪涝袭扰,尤其2021年7.20特大暴雨以及突发性公共卫生事件

不仅造成国家巨大的经济损失和人员伤亡，还会引起次生和衍生灾害，进而引发一系列多米诺骨牌灾害链的发生。突发事件频发、环境复杂和主体多元化，使得整个应对过程是一个涉及多要素、多主体间相互作用的复杂动态过程。突发事件下，以政府为主导的应急管理体系，发挥多元主体联动作用，并根据已有资源整合探寻有效的应急能力建设路径，是实现《“十四五”国家应急体系规划》^[1]全面加强应急救援力量建设，大幅提升应急管理法治水平、科技信息化水平和综合保障能力，实现现代化相适应的中国特色大国应急体系的有力保证。

随着多主体参与的应急联动在突发事件管理中的应用，该机制的科学性和有效性被反复检验。然而，作为应急联动的核心——协同应急能力影响因素体系构建问题目前还没有获得充分的理论支持，而这一领域的研究又十分复杂，需要系统深入地开展相关理论研究。因此本研究将在多主体协同视角下，探索影响应急能力的因素，并且构建相应体系。

一、应急能力以及影响因素理论模型的构建

本研究以协同应急能力以及影响因素为研究对象，共选取了六个潜在变量，分别是技术条件、组织资源、社会环境、复杂性、组织协同关系、以及协同应急能力。本研究的理论概念模型利用 AMOS 软件进行构建，如图 1 所示：

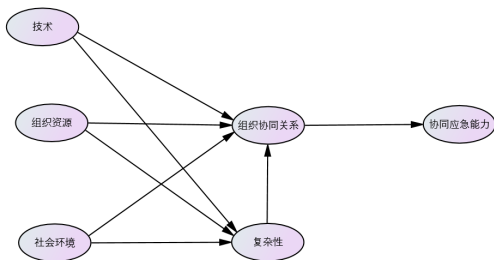


图 1：理论概念模型

（一）研究假设

1. 组织协同关系假设

从技术条件来看 N.Kapucu（2008）通过对 2004 年在美国佛罗里达发生的龙卷风应急事件的案例分析，相关技术的有效采用能够提高应急协同的效率^[2]。同时，樊博（2015）等人以食品药品安全联合监管为实证背景跨部门协同能力的影响因素研究中发现技术因素与信息沟通共享存在正相关关系^[3]。从组织资源方面来看，Chen（2008）等人研究表明，在部门应急管理中如果政府拥有的应急资源数量越多，种类越丰富，共享程度越高，同时人员和资源能够较为合理分配，则部门间越容易协同合作，从而提升了应急协同能力^[4]。钟开斌（2020）认为应急资源是突发事件治理过程中所需资源的集合，应急资源是应急管理中的支撑性要素，陈瑞认为应急资源的充足性将影响应急协同效率^[5]。李艳飞（2022）等人在对于地方政府应急能力影响因素及其提升路径探究中认为要建设常态化应急能力，短期内需增加应急资源投入^[6]。从社会环境方面来看，吴国斌、张凯（2011）在对湖北省高速公路的研究中通过对六个观测变量进行研究证明应急法律法规能够对于协同应急能力的产生影响^[7]。王力立（2015）通过对伙伴能力、政府回应性、公共服务质量等环境性因素的研究发现伙伴能力、政府回应性和公共服务质量这三个因子除了各自都会对协同行为产生正向影响之外，还会通过互相依赖因子的部分中介效应

对协同行为产生一定正向影响^[8]。因此论文做出如下假设：

H1：技术条件与组织协同关系呈正相关

H2：组织资源与组织协同关系呈正相关

H3：社会环境与组织协同关系呈正相关

2. 复杂性假设

在实际的应急管理中，突发事件的复杂性，需求的复杂性，任务的复杂性均会给协同应急管理带来的压力。Chen（2008）将灾害生命周期方法引入应急管理之中，提出了包括灾害前阶段、灾害阶段、恢复阶段的应急协同生命周期的框架，认为应急工作任务的复杂性、信息传递效率、应急指挥的有效性、应急资源的充足性、组织协同关系等影响了应急协同的效率^[4]。L.K.Comfort（2005）等认为，突发事件的状态将会对跨部门应急协同能力产生重要的影响^[9]。樊博（2015）等人以食品药品安全联合监管为实证背景跨部门协同能力的影响因素研究中发现事件因素与协同能力存在负相关关系影响^[3]。Kapucu（2006）认为专业技术人员的参与，可以为突发事件的解决提供可行性的技术方案，分析事件可能造成的后果，从而缩减事件波及的范围，减少事件的不确定^[10]。山少男、段霞（2022）通过元分析方法梳理了前人研究中关于公共危机多元主体协同治理行为的影响因素并进行定量综合分析，研究发现内部因素与外部因素皆对公共危机多元主体协同治理行为产生正向影响，其中关系性因素、感知性因素、效能性因素、复杂性因素与技术性因素是重要影响因素^[11]。因此论文做出以下研究假设：

H4：技术条件与复杂性呈负相关

H5：组织资源与复杂性呈负相关

H6：社会环境与复杂性呈负相关

H7：复杂性与组织协同关系呈负相关

3. 协同应急能力假设

在突发事件协作中，组织协同关系显得尤为重要，若缺乏一种成熟、稳定的信任关系，则很难实现主体之间的协作。L.K.Comfort（1990）提出应急组织冲突对应急协同效率具有影响。L.K.Comfort（2001）进一步归纳出应急协同效率的因素包括：应急信息的有效传递、事件的动态变化、应急目标的一致性、组织文化差异、应急风险的共享程度^[12]。王力立（2015）等人在地方政府网络治理协同行为实证研究中发现地方政府在组建治理网络、选择合作伙伴、实现成功合作的过程中，要特别注重

与合作伙伴形成高度的信任关系^[8]。S.E.Robinson (2006) 通过对卡特琳娜飓风应急事件的分析,提出了影响应急协同效率的因素:应急规范和标准的差异、突发公共事件应急阶段差异、组织在事前形成的关系^[13]。Gonzalez 等在提出了相关的概念性框架模型后,通过对鹿特丹港危机应对的个案研究,得出应急信息沟通共享的质量越高,即信息的精准度、及时性、一致性程度越高,越促进协同合作的完成^[14]。基于对上述文献的研究,论文做出如下假设:

H8: 组织协同关系与协同应急能力呈正相关

(二) 变量设计与测量

本研究选取结构方程模型作为主要工具来分析在突发事件下,多主体协同应急能力影响因素,并根据涉及的结构方程模型的要求来进行变量的选取,进行问卷设计。在本次研究中,选取技术条件,组织资源,社会环境为自变量组织协同关系,复杂性因素为中介变量,协同应急能力为因变量。

在研究中,技术条件分基础设施水平和大数据平台两个观测维度。组织资源分应急物资,应急资金,应急人员,应急设备四个观测维度。社会环境分为应急法规,公共服务质量,经济发展水平三个观测维度。复杂性分为事物复杂性,需求复杂性,任务复杂性三个观测维度。组织协同关系分为信任关系,协同沟通,组织目标三个观测维度。协同应急能力分为协同演练预案,协同联合预警监测、资源协同保障、信息协同共享、救援指挥、设施重建恢复、后期评估七个观测维度,故论文的研究假设模型如图2所示:

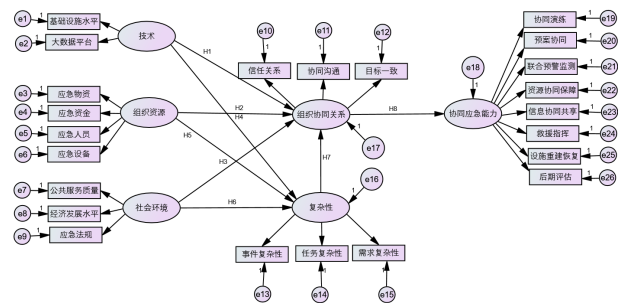


图2: 协同应急能力影响因素研究假设模型图

1. 技术条件的测量指标

参考李艳飞 (2022)^[6], 唐玉青 (2022)^[15], 姚乐野、李明 (2018)^[16], 山少男 (2022)^[11] 等学者文献中的量表论文选取基础设施水平、大数据平台作为技术条件的观测变量,其中具体测量题项如表1所示:

表1: 技术条件测量量表

测量维度	编号	题项	参考文献
技术条件	JS1	基础设施建设水平能够满足公共危机事件发生时的使用	李艳飞 (2022); 唐玉青 (2022); 姚乐野、李明 (2018); 彭媛、张玲 (2022)
	DS1	各应急组织的大数据平台功能完整,能够实现实时数据监测和分析,数据可视化,数据挖掘和分析,预测模型建立和应用	
	DS2	各应急组织的大数据平台数据收集和及时处理有效	

2. 组织资源的测量指标

参考邓云峰 (2005)^[17], 樊博 (2014)^[3], 山少男 (2022)^[11], 张永颂 (2010)^[18] 樊博、杨文婷、尹鹏程 (2017)^[19] 等学者文献中的成熟量表,论文选取应急物资、应急资源、应急人员、应急设备四个观测变量,具体题项如表2所示:

表2: 组织资源测量量表

测量维度	编号	题项	参考文献
应急物资	WZ1	储备的应急物资的种类和数量在面对公共危机事件时能够满足了公共危机事件的需要	
	WZ2	储备应急物资能够及时有效送往受灾现场	
应急资源	ZY1	在公共危机事件发生时有充足应急资金能够应对公共危机事件	樊博 (2014); 山少男 (2022); 邓云峰 (2005); 张永颂 (2010)
	ZY2	在公共危机事件发生时应急资金能够流畅无阻	
应急人员	RY1	应急人员的培训和教育对应急响应效果良好	
	RY2	应急现场专业技术人员能够及时有效的参与现场救援	
应急设备	SB1	运送至应急救援现场的应急设备功能能够满足应急活动的使用	
	SB2	应急设备故障能够及时解决	

3. 社会环境的测量指标

参考王力立 (2015)^[8], 山少男 (2022)^[11], 聂爽 (2017)^[20], 李尧远、曹蓉 (2015)^[21] 等学者的文献研究,论文选取公共服务质量、应急法规、经济发展水平三个观测变量,具体题项如表3所示:

表3: 社会环境测量量表

测量维度	编号	题项	参考文献
公共服务质量	GF1	公共服务的响应迅速	
	GF2	公共服务质量能够满足公共要求	
社会环境	ZD1	相关应急制度建设程度完整,能够满足应对公共危机事件的需要	王力立 (2015); 山少男 (2022); 聂爽 (2017); 李尧远、曹蓉 (2015)
	ZD2	在公共危机事件发生时,应急组织能够依靠相关法律法规,合法的参与应急管理活动中	
	ZD3	在公共危机事件处置过程中应急组织间相应的法律责任划分明确	
经济发展水平	JJ1	您所在地区的经济发展水平	

4. 组织协同关系的测量指标

参考王力立 (2015)^[8], Bharosa N (2010)^[22], 吴国斌、张凯 (2015)^[7], 本论文选取信任关系、协同沟通、组织目标、为组织协同关系的测量指标。具体题项如表4所示:

表4: 组织协同关系测量量表

测量维度	编号	题项	参考文献
信任关系	XR1	应急组织或其他参与组织很信任,愿意共享资源或强化自身行为来完善合作	王力立 (2015); Bharosa (2010); 吴国斌, 张凯 (2011); 山少男 (2022)
	XR2	应急组织之间能够进行有效的动态反馈与及时对话,并且能够获取全部信息	
协同沟通	XG1	在公共危机事件发生时,应急组织的组织目标一致	
组织目标	ZM1		

5. 复杂性因素的测量指标
汪渊（2017）^[23] 等学者的研究，论文选取事物复杂性、需求复杂性、任务复杂性三个观测变量，具体测量题项如表5所示：

表5：复杂性测量量表

测量维度	编号	题项	参考文献
事物复杂性 需求复杂性	SF1	在公共危机事件发生时事件往往比较复杂,突发性强,不确定性高,影响范围广	樊博 (2014); 山少男 (2022); 王力立 (2015); 汪渊 (2017) 等
	XF1	公共危机事件处置过程中的需求往往存在很多不确定性,难以准确把握	
	XF2	公共危机事件处置过程中的需求往往需要经过多方协商才能得到满足	
任务复杂性	RF1	公共危机事件处置过程中需要协调的人员数量多,时间比较紧迫,资源的限制程度高,任务复杂性高	

6. 协同应急能力的测量指标
参考文献，论文选取了协同演练预案，协同联合预警监测，资源协同保障，信息协同共享，救援指挥，设施重建恢复，后期评估八个观测变量，具体测量题项如表6所示：

表6：协同应急能力测量量表

测量维度	编号	题项	参考文献
协同 应急 能力	协同演练	YL1 各应急组织在危机事件发生前,能够有效地进行预案演练和应急演练,以提高应急响应能力和配合能力	曹惠民 (2015); 赖先进 (2015); 邓云峰 (2004); 闪淳昌 (2007); 齐春泽 (2019) 等
	预案协同	FX1 您认为应急组织能准确识别突发事件的种类,并制定各类突发事件的应急预案是否符合实际	
		FX1 您认为在制定应急预案时,应急组织能够充分考虑其他组织的应急预案规划,与其他单位应急预案衔接性高是否符合实际	
	联合监测预警	XY1 应急组织能够实时监测风险信息、快速分析处理风险信息是否符合实际	
		XY2 您认为应急部门在危机事件发生时,能够及时地向社会公众发布相关信息,以及有效地应对公众的关切和需求是否符合实际	
	资源协同共享	ZX1 您认为各应急组织的应急资源协同配合能力高,能够在紧急情况下完成应急资源分配和调配是否符合实际	
	信息协同共享	XX1 您认为应急组织能够有效地共享信息,并协调好各自的行动,以达到最佳的应急效果是否符合实际	
	救援指挥	JZ1 您认为应急组织在危机事件发生时,能够及时向上级部门报告相关情况,并接受上级部门的指挥和调度是否符合实际	
		JZ2 您认为应急组织在危机事件发生时能够快速处理安全事故并最大限度保证人民的生命与财产安全的能力是否符合实际	
	设施重建恢复	HF1 您认为在公共危机事件发生后中能够全面落实灾后恢复重建的主体责任,统筹规划安排灾后恢复重建的建设项目是否符合实际	
		HF2 您认为在公共危机事件发生后中能够加大财政资金的统筹支持力度,进行灾后设施重建与恢复是否符合实际	
	后期评估	PG1 您认为应急组织在危机事件发生后,能够及时进行事故调查和总结,以及及时修订和完善应急预案和应急措施是否符合实际	

二、实证分析

本论文的问卷主体部分包括36个测量题项。问卷基于李克特五级量表法进行设计，结合研究内容加以修改与完善。为确保问卷数据质量，本论文选取政府部门，企业，非盈利性组织，社会公众为研究对象，在全国范围内进行问卷调查。被访问者基于对于多主体协同应急的了解程度，该问卷并做出相应选择，具体问卷设计见附录。2025年4月2日至5月5日，共收集到210份问卷，剔除回答时间太短、答案相同的问卷数据，共有200份有效的问卷，有效率达95%。按照样本数量是题目的5–10倍的原则，本次调查问卷的数量达到了所需标准。

（一）样本数据描述性分析

样本的描述性统计分析主要针对调查对象的总体样本特点，如调查对象的所属机构、学历、地域等。具体如表7所示：

表7：样本特征分布描述

变量	选项	频率	百分比
参与主体	政府机构	53	27%
	非盈利性组织	27	14%
	企业	53	27%
	个人	48	24%
	其他	19	10%
年龄	18岁以下	7	4%
	18~25	127	64%
	26~30	35	18%
	31~40	21	11%
	40以上	10	5%
受教育情况	博士	11	6%
	硕士	35	18%
	本科	141	71%
	本科以下	13	7%
地区	城市	118	59%
	城镇	48	24%
	农村	34	17%

由表7可知从受调查者的基本信息上来看,受访者来自政府机构和企业的总人数相同为53人,占比27%,来自非盈利性组织27人占比14% 剩下67人为社会公众占比34%。受访者年龄大多数为18岁至25岁,127人,占比64%。在受访者中,学历为本科的人数最多为141占71%,学历在本科以下的人数最少为13人占比7%。在受访者中59%的人来自城市,24%的人来自城镇,17%的人来自农村。因此,从以上的研究结果看,论文的样本结构比较合理,有较好的代表性,可以对后续的信度、效度进行分析。

(二) 数据分析

1. 问卷信度检验

在本次研究中主要的因素均通过量表的形式进行测量,因此,对于测量结果的数据质量进行检验是保证后续分析具有重要意义的前提。首先,通过 Cronbach a 值信度检验方法分析各个维度的内部一致性。Cronbach a 取值范围在0到1之间,检验结果数值越高,信度越高。一般认为信度系数在0.6以下,则认为信度不可信,需要重新设计问卷或尝试重新收集数据。在本次分析中,信度分析结果如图表8所示,多主体协同应急能力总体以及各个维度信度系数均在0.7以上,因此说明本次研究所使用的量表具有良好的内部一致性,信度很好。

表8: 量表信度分析

变量	Cronbach a	项数	Cronbach a	标准
技术条件	0.879	3	0.943	Cronbach a>0.7
组织资源	0.919	7		
社会环境	0.901	6		
复杂性因素	0.914	4		
组织协同关系	0.872	3		
协同应急能力	0.948	12		

2. 探索性因子分析(EFA 分析)

探索性因子分析属于效度检验,效度即有效性,是指测量的有效性程度,就是指一个测验的准确性与有用性。本研究利用 KMO 和 Bartlett 检验来进行探索性因子分析,分析认为 KMO 中检验系数大于0.5, Bartlett 检验显著性 Sig 小于0.001问卷才有结构效度,才能进行因子分析。具体分析情况如表9所示:

表9: 量表信效度分析

变量	KMO 值	KMO	Sig.
技术条件	0.736	0.954	0
组织资源	0.929		
社会环境	0.900		
复杂性因素	0.825		
组织协同关系	0.735		
协同应急能力	0.947		

由表9可以看出每个潜变量的 KMO 值均在0.6以上,总体的 KMO 值为0.954,且 Bartlett 检验显著性均为0.000。综上,问卷数据具有一定的可靠性,可用来进行下一步的验证性因子分析。

3. 验证性因子分析(CFA 分析)

(1) 模型适配度检验

根据表10的模型适配度检验结果可以看出, χ^2/df (卡方自由比)=1.813,在1-3的范围内, RMSEA(误差均方根)=0.064,在<0.08的范围内。另外的 GFI、AGFI、IFL、TLI 以及 CFI 的检验结果均达到了0.8以上的水平。因此综合本次的结果可以说明,样本数据具有良好的模型适配度。

表10: 模型适配度检验

拟合指标	CMIN/DF	RMSEA	CFI	IFL	TLI	GFI	AGFI
参考标准	1.813	0.064	0.858	0.845	0.86	0.8	0.8
实测结果	1-3	<0.08	>=0.8	>=0.8	>=0.8	>=0.8	>=0.8
拟合判断	是	是	是	是	是	是	是

(2) 收敛效度和组合信度检验

表11: 协同应急能力影响因素量表各个维度收敛效度和组合信度检验

	路径关系	Estimate	AVE	CR
DS2	<---	技术	0.8	
DS1	<---	技术	0.773	0.5148
JC1	<---	技术	0.554	
ZJ1	<---	组织资源	0.606	
ZJ2	<---	组织资源	0.504	
RY2	<---	组织资源	0.564	0.441
SB1	<---	组织资源	0.583	
WZ2	<---	组织资源	0.607	
WZ1	<---	组织资源	0.622	
JJ1	<---	社会环境	0.449	
ZD3	<---	社会环境	0.753	
ZD2	<---	社会环境	0.675	0.4251
ZD1	<---	社会环境	0.722	
GF2	<---	社会环境	0.631	
GF1	<---	社会环境	0.638	
XF2	<---	复杂性	0.775	
XF1	<---	复杂性	0.711	0.5804
SF1	<---	复杂性	0.77	
RF1	<---	复杂性	0.789	
XR1	<---	组织协同关系	0.694	0.4104
XG1	<---	组织协同关系	0.583	
ZM1	<---	组织协同关系	0.64	
YL1	<---	协同应急能力	0.597	
FX1	<---	协同应急能力	0.678	0.4339
FX2	<---	协同应急能力	0.561	
XY1	<---	协同应急能力	0.666	
XY2	<---	协同应急能力	0.685	
ZX1	<---	协同应急能力	0.663	
XX1	<---	协同应急能力	0.739	
JZ1	<---	协同应急能力	0.709	
JZ2	<---	协同应急能力	0.669	
HF1	<---	协同应急能力	0.577	
HF2	<---	协同应急能力	0.651	
PG1	<---	协同应急能力	0.686	

在样本数据具有良好的模型适配度的前提条件下,将进一步检验量表的各个维度的收敛效度(AVE)和组合信度(CR)。检验流程通过建立的 SEM 模型计算出各个测量题项在对应维度的标准化因子载荷。然后通过 AVE 和 CR 的计算公式计算出,各个维度的收敛效度值和组合信度值,根据标准 AVE 值最低要求达到0.36,CR 值最低要求达到0.7,才能说明具有良好的收敛效度和组合信度。根据表11的分析结果可以看出,在本次量表效度检验

中，各个维度 AVE 值均达到了 0.36 以上，CR 值均达到了 0.7 以上，综合可以说明各个维度均具有良好的收敛效度和组合信度。

（三）模型假设检验

1. 模型运行

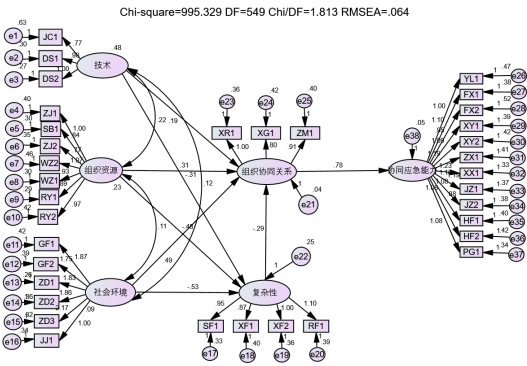


图 3：模型运行图

在进行了以上数据分析，认为所得数据通过了信效度检验，故可以导入至模型中进行分析。本研究在结构方程模型的基础上

对技术，组织资源，社会环境，组织协同关系，复杂性以及协同应急能力六个变量之间的关系进行分析，利用 AMOS 软件进行模型运算，结果如图 3 所示：

2. 路径分析

根据表 12 的分析结果可以看出，在本次研究的路径假设关系检验中，社会环境对于复杂性的影响作用不明显 ($B=-0.159, p>0.05$)，因此假设 H6 不成立。技术显著的负向复杂性 ($B=-0.231, p<0.05$) 因此假设 H4 成立。组织资源对于复杂性有显著的负向作用 ($B=-0.429, p<0.05$)，因此假设 H5 成立。技术显著的正向影响组织协同关系 ($B=0.207, p<0.05$)，因此假设 H1 成立。社会环境显著的正向影响组织协同关系 ($B=0.234, p<0.05$) 因此假设 H3 成立。复杂性显著的负向影响组织协同关系 ($B=-0.333, p<0.001$) 因此假设 H7 成立。组织资源显著的正向影响组织协同关系 ($B=0.304, p<0.05$) 因此假设 H2 成立。组织协同关系显著的正向影响协同应急能力 ($B=0.893, p<0.001$)，因此假设 H8 成立。

表 12：协同应急能力影响因素 SEM 路径关系检验结果

路径关系	std.	Unstd.	S.E.	C.R.	P	Label
复杂性 <--- 社会环境	-0.159	-0.393	0.347	-1.134	0.257	H6
复杂性 <--- 技术	-0.231	-0.245	0.124	-1.973	0.049	H4
复杂性 <--- 组织资源	-0.429	-0.653	0.279	-2.34	0.019	H5
组织协同关系 <--- 技术	0.207	0.175	0.082	2.135	0.033	H1
组织协同关系 <--- 社会环境	0.234	0.461	0.218	2.113	0.035	H3
组织协同关系 <--- 复杂性	-0.333	-0.265	0.073	-3.623	***	H7
组织协同关系 <--- 组织资源	0.304	0.368	0.182	2.024	0.043	H2
协同应急能力 <--- 组织协同关系	0.893	0.786	0.109	7.178	***	H8

本研究假设具体通过情况如下表 13 所示：

表 13：研究结果

假设编号	假设内容	通过
H1:	技术条件与组织协同关系呈正相关	通过
H2:	组织资源与组织协同关系呈正相关	通过
H3:	社会环境与组织协同关系呈正相关	通过
H4:	技术条件与复杂性呈负相关	通过
H5:	组织资源与复杂性呈负相关。	通过
H6:	社会环境与复杂性呈负相关	不通过
H7:	复杂性与组织协同关系呈负相关	通过
H8:	组织协同关系与协同应急能力呈正相关	通过

同应急能力影响因素。但很少有人关注协同应急能力影响因素之间的关系，在查阅了大量的文献之后，本论文结合其他学者关于协同行为影响因素以及应急能力影响因素的研究，提出了突发事件协同应急能力影响因素体系理论模型，并通过实证分析验证该模型的有效性。运用 SPSS 和 AMOS 软件进行模型分析最终得出以下结论：第一、技术条件对于组织协同关系具有显著的正向作用。第二、组织资源显著的正向影响组织协同关系。第三、社会环境对于组织协同关系具有显著的正向影响。第四、技术对于复杂性具有显著的负向影响。第五、组织资源对于复杂性具有显著的负向影响。第六、复杂性对于组织协同关系存在显著的负向影响。第七、组织协同关系对于协同应急能力具有显著的正向影响。

三、结束语

关于突发事件应急管理的研究，目前大多数学者都是从突发事件的三个时间节点即事件发生前，发生中，发生后，来探索协

参考文献

[1] 国务院关于印发“十四五”国家应急体系规划的通知 [J]. 中华人民共和国国务院公报, 2022, 1761(06): 30-48.
[2] Kapucu N. Collaborative emergency management: better community organising, better public preparedness and response[J]. Disasters, 2008, 32(2): 239-262.
[3] 潘潇，樊博. 应急管理中跨部门协同能力的影响因素研究——以食品药品安全联合监管为实证背景 [J]. 软科学, 2014, 28(02): 52-55+60.

- [4]Chen R, Sharman R, Rao H R, et al. Coordination in emergency response management[J]. Communications of the ACM, 2008, 51(5): 66-73.
- [5] 钟开斌. 国家应急管理体系: 框架构建、演进历程与完善策略 [J]. 改革, 2020, 316(06): 5-18.
- [6] 李艳飞, 李苏梅, 王亦虹等. 地方政府应急能力影响因素及其提升路径探究——基于31省面板数据的实证研究 [J]. 中国安全生产科学技术, 2022, 18(04): 47-53.
- [7] 吴国斌, 张凯. 多主体应急协同效率影响因素实证研究——以湖北省高速公路为例 [J]. 工程研究 - 跨学科视野中的工程, 2011, 3(02): 164-173.
- [8] 王力立, 刘波, 姚引良. 地方政府网络治理协同行为实证研究 [J]. 北京理工大学学报 (社会科学版), 2015, 17(01): 53-61.
- [9]Comfort L K, Ko K, Zagorecki A. Coordination in rapidly evolving disaster response systems: The role of information[C]//Agent-Based Simulation: From Modeling Methodologies to Real-World Applications: Post-Proceedings of the Third International Workshop on Agent-Based Approaches in Economic and Social Complex Systems 2004. Springer Tokyo, 2005: 208-219.
- [10] Kapucu N. Interagency communication networks during emergencies: Boundary spanners in multiagency coordination[J]. The American review of public administration, 2006, 36(2): 207-225.
- [11] 山少男, 段霞. 复杂性视角下公共危机多元主体协同治理行为的影响因素与行动路径——基于元分析与模糊集 QCA 的双重分析 [J]. 公共管理与政策评论, 2022, 11(01): 104-119.
- [12]Comfort L K, Sungu Y, Johnson D, et al. Complex systems in crisis: Anticipation and resilience in dynamic environments[J]. Journal of contingencies and crisis management, 2001, 9(3): 144-158.
- [13]Robinson S E, Berrett B, Stone K. The development of collaboration of response to Hurricane Katrina in the Dallas area[J]. Public Works Management & Policy, 2006, 10(4): 315-327.
- [14]Gonzalez R A, Bharosa N. A framework linking information quality dimensions and coordination challenges during interagency crisis response[C]//2009 42nd Hawaii International Conference on System Sciences. 2009: 1-10.
- [15] 唐玉青. 大数据时代公共突发事件中政府的跨域协同治理 [J]. 江汉论坛, 2022, 527(05): 60-65.
- [16] 姚乐野, 李明, 曹杰. 基于 Multi-Agent System 的应急管理多元主体信息互动机制初探 [J]. 情报资料工作, 2018, 222(03): 44-50.
- [17] 邓云峰, 郑双忠, 刘功智等. 城市应急能力评估体系研究 [J]. 中国安全生产科学技术, 2005(06): 33-36.
- [18] 张永领. 基于应急管理阶段的应急能力评价指标体系研究 [J]. 工业安全与环保, 2010, 36(09): 32-34.
- [19] 樊博, 杨文婷, 尹鹏程. 应急资源协同联动信息系统构建的影响机理——基于 IS 成功模型的分析 [J]. 系统管理学报, 2017, 26(05): 801-808.
- [20] 聂爽. 城市协同应急准备能力的影响因素研究 [D]. 上海交通大学, 2017.
- [21] 李尧远, 曹蓉. 我国应急管理研究十年 (2004-2013): 成绩、问题与未来取向 [J]. 中国行政管理, 2015, 355(01): 83-87.
- [22]Bharosa N, Lee J K, Janssen M. Challenges and obstacles in sharing and coordinating information during multi-agency disaster response: Propositions from field exercises[J]. Information systems frontiers, 2010, 12: 49-65.
- [23] 汪渊. 重庆市江北区政府与社会组织间关系的影响因素研究 [D]. 西南大学, 2017.
- [24] 曹惠民, 黄炜能. 地方政府应急管理评估指标体系探讨 [J]. 广州大学学报 (社会科学版), 2015, 14(12): 60-66.
- [25] 赖先进. 论城市公共危机协同治理能力的构建与优化 [J]. 中共浙江省委党校学报, 2015, 31(01): 60-66.
- [26] 闪淳昌. 加强应急预案体系建设 提高应对突发事件和风险的能力 [J]. 现代职业安全, 2007, 71(07): 82-83.
- [27] 齐春泽, 代文锋. 基于云模型的城市灾害应急能力评价 [J]. 统计与决策, 2019, 35(04): 41-45.

OBE 理念下应用型会计人才“听学用”培养路径研究

刘欢娜

广东白云学院 会计学院, 广东 广州 510550

DOI:10.61369/ASDS.2026040005

摘 要 : 新时代背景下社会对会计人才的专业素养有了更高的要求, 企业更需要能够将会计专业知识应用于实践的复合型人才, 培养高质量应用型会计人才是本科高校的必然选择。然而, 应用型人才的培养困囿于传统的教育方式, 现实困境包括教学内容与行业需求脱节、教学方式缺乏互动性、学习资源匮乏、实习机会有限、缺乏自主学习空间等, 这直接导致会计专业学生“听不懂、学不会、用不上”的难题。本研究以 OBE 理念为指导, 围绕让学生“听得懂、学得会、用得上”这一核心目标, 厘清会计人才培养内在逻辑的应然之境, 提出强化实践教学、创新教学方法、加强资源建设、深化校企合作、减轻学生负担等使然之策, 以破解应用型本科教育高质量发展面临的实然之困, 构建一套符合 OBE 理念、适应市场需求的应用型会计人才“听学用”培养路径。

关 键 词 : 本科高校; 应用型会计人才; OBE 理念; “听学用”

Research on the Training Path of "Understandable, Learnable, Applicable" for Application-Oriented Accounting Talents under the OBE Concept

Liu Huanna

College of Accounting, Guangdong Baiyun University, Guangzhou, Guangdong 510550

Abstract : Against the background of the new era, society has put forward higher requirements for the professional literacy of accounting talents. Enterprises are more in need of interdisciplinary talents who can apply professional accounting knowledge to practice, and cultivating high-quality application-oriented accounting talents is an inevitable choice for undergraduate universities. However, the cultivation of application-oriented talents is confined by traditional educational methods. The practical dilemmas include the disconnection between teaching content and industry needs, the lack of interactivity in teaching methods, the shortage of learning resources, limited internship opportunities, and the lack of independent learning space. These problems directly lead to the difficulty that accounting majors "cannot understand, cannot learn, and cannot apply" what they have learned. Guided by the OBE concept, this study focuses on the core goal of enabling students to "understandable, learnable and applicable", clarifies the ideal state of the internal logic of accounting talent training, and puts forward practical strategies such as strengthening practical teaching, innovating teaching methods, enhancing resource construction, deepening school-enterprise cooperation, and reducing students' burden, so as to solve the practical difficulties faced by the high-quality development of application-oriented undergraduate education and construct a "Understandable, Learnable, Applicable" training path for application-oriented accounting talents that conforms to the OBE concept and meets market demands.

Keywords : undergraduate universities; application-oriented accounting talents; OBE concept; "understandable, learnable and applicable"

引言

2024年11月14日, 教育部、人力资源社会保障部召开了2025届全国普通高校毕业生就业创业工作会, 对高等教育应用型人才培养提出了新的要求。应用型人才培养离不开成果导向教育(Outcome-Based Education, 简称OBE)。OBE作为一种先进的教育理念, 倡导以学生为中心并关注学生的全面发展^[1]。然而, 在实践中OBE的实施往往面临着诸多挑战, 许多学生面临着“听不懂、学不会、用不上”的困境, 造成这个难题的原因之一是教学与实际需求脱节^[2]。另外, 随着信息化的深入发展, 会计行业经历着前所未有的转变, 企业对会计人才的需求不再是传统的账务处理与报表编制, 而是具备扎实理论基础、卓越实践能力并且能够灵活应对复杂经济环境的应用型会计人才^[3]。

作者简介: 刘欢娜, 广东白云学院会计学院, 硕士研究生, 会计师, 研究方向为会计、企业经济、高等教育。

本研究旨在探讨如何在 OBE 理念下构建一个使学生“听得懂、学得会、用得上”的本科高校应用型会计人才培养路径。首先明确培养目标,分析 OBE 理念下“听学用”会计人才培养的应然之境,接着分析当前会计人才培养的实然之困,最后提出“听学用”会计人才培养的使然之策,即针对存在的问题提出具体的改进建议。本研究期望能够为本科高校应用型会计人才的培养提供新的思路,促进 OBE 理念在会计教育领域的深入实施,为培养更多符合社会需求的人才贡献力量。

一、OBE 理念下“听学用”会计人才培养的应然之境

在理想的 OBE 理念下,“听得懂、学得会、用得上”不仅是会计人才培养的核心目标,也是教育实践的应然之境。

(一) 让学生“听得懂”:构建清晰易懂的知识体系

在 OBE 理念下,本科高校应用型会计人才的培养应达到的理想状态之一是让学生“听得懂”。“听得懂”是基础,它要求教师构建一个清晰易懂的知识体系,如果学生连老师课堂上所讲解的知识都听不懂,那么他们将很难学会并将其应用于实际工作中。传统的教学模式过于注重理论的深度,导致学生难以理解,为了实现“听得懂”这一目标,首先教师必须精心筛选会计课程内容,去除冗余过时的知识,保留前沿的理论及实践技能,做到教学内容精简且逻辑清晰。同时教师应采用更加生动的教学方法,使复杂的会计理论变得生动易懂,这要求教师要具备丰富的实践经验,用浅显的语言解释知识并帮助学生建立起理论基础。教学课程的设计需要遵循由浅入深、循序渐进的原则,确保学生能够充分吸收每个知识点,学生只有在课堂上“听得懂”,才能够学以致用^[4]。

(二) 让学生“学得会”:注重实践能力的培养与提升

本科高校应用型会计人才培养的理想状态之二是让学生“学得会”。“学得会”是关键,学生只有熟练掌握所学知识才能解决实践中的问题。学校可以提供丰富的实践机会,比如,设计更多的实践课程、实训项目,让学生把所学内容真正变成自己的本领。学校还可以提供个性化的学习方案,根据学生的兴趣为其量身定制学习计划,确保学生都能有效掌握会计技能。此外,教师可采用创新型人才培养模式,注重培养学生的实践能力以及创新思维,鼓励他们主动探索新的解决方案^[5]。教师也应持续完善教学方法,确保学生不仅“听得懂”,而且能够“学得会”。

(三) 让学生“用得上”:实现教育与职业的无缝对接

在 OBE 理念下,本科高校应用型会计人才的培养应达到的理想状态之三是让学生能够将所学知识应用在实际工作中,即“用得上”,“用得上”是终极目标。即使学生听懂了课堂上的知识并且学会了如何应用这些知识,但如果他们无法在实际工作中使用这些知识来解决问题或者创造价值,那么前面的学习和努力都将失去意义。

要想让学生所学知识“用得上”,那么教育内容不仅要涵盖会计理论,还要紧密结合行业实际需求,这样才能确保学生实习后能够迅速适应工作环境、拥有解决问题的能力。特别是在智能财会人才培养大趋势下,必须要提高会计专业学生的新技术应用能

力^[6]。学校设置的课程应与行业需求相结合,引入最新的行业知识,也可以设置职业规划与就业指导课程,为学生提供全面的职业规划服务,包括行业分析、岗位介绍、简历撰写、面试技巧等,以帮助学生顺利步入职场。在教学过程中,教师也应该引导学生关注市场需求变化,帮助他们了解未来的就业方向,教师还应该建立有效的反馈机制,定期收集学生、用人单位的意见,对教学内容、教学方法进行及时优化,以满足学生的学习需求以及市场的用人需求,真正让学生所学知识“用得上”。

总之,在 OBE 理念下,“听得懂、学得会、用得上”是会计人才培养的三个必然目标,它们之间构成了一种辩证统一的关系。其中,“听得懂”是基础,学生只有听懂了才能进一步的掌握知识,“学得会”是关键,学生只有学会了才能真正应用所学知识,而“用得上”则是终极目标,只有将所学知识用在职业中并取得成果,前面的学习和努力才有意义,否则,达不到“用得上”的效果,“听得懂”、“学得会”终将是徒劳。只有三个目标全部达到才能实现以“学生为中心”的教育理念,才能培养出符合社会需求的高素质会计人才。

二、OBE 理念下“听学用”会计人才培养的实然之困

尽管 OBE 理念在理论上为本科高校应用型会计人才培养指明了方向,但在实际操作过程中却往往未能充分实现这一目标,反而变相增加了学生任务,现实困境主要包括:

(一) 教学内容与行业需求脱节

在 OBE 理念下“听学用”会计人才的培养面临诸多实际困难,其中最突出的困难便是实践教学环节薄弱,教学内容与市场需求脱节。许多高校在实施 OBE 过程中往往忽视了实践教学的重要性,致使无法用所学内容和自有知识储备解决实际问题,直接导致所学知识“用不上”的困境。部分本科院校在学习设置上仍侧重理论内容的讲解,致使学生所学内容与岗位需求存在差距。另外,OBE 要求教学内容紧密对接行业需求,但现实中许多高校的课程设置与市场需求脱节或滞后于市场变化,教学内容更新不及时且无法跟上行业发展的要求,使得学生的知识体系陈旧,难以满足实际工作需要以及就业市场的需求。因此,如何加强实践教学环节,使教学内容与市场需求紧密结合,成为 OBE 理念下会计人才培养亟待解决的问题。

(二) 教学方式缺乏互动性

在 OBE 理念下,教学方法缺乏互动性也是培养会计人才所面临的问题之一。虽然 OBE 提倡灵活多样的教学方法,但在大多

数高校中，由于课堂规模较大，教师难以关注到每个学生的学习需求，难以实现与学生的互动，缺乏趣味性的教学方式不利于学生深入理解。此外，传统的课堂教学往往缺乏足够的案例分析、模拟实训等实践性教学环节，这使得学生只是被动接受，难以将所学理论知识转化为实际操作技能，无法实现“听得懂、学得会、用得上”。因此，如何丰富教学方法、提高教学的互动性，是OBE理念下每个教育者都值得思考的问题。

（三）学习资源难以满足学生个性化需求

会计人才培养的又一困境是有限的学习资源难以满足学生个性化的学习需求。一些高校在OBE实施的过程中忽视了学习资源的多样性，这导致学生在学习时难以获得适合自己的学习材料，不能根据自己的兴趣有效地学习。加之教学资源的限制，教师通常无法为每个学生提供单独的指导服务，学生在学习中遭遇难题也无法及时得到帮助，这限制了学生的学习效果和学习的主动性。因此，如何丰富学习资源以满足学生的个性化学习需求，成为OBE理念下会计人才培养不容忽视的问题。

（四）学生实习实训机会有限

在OBE理念下，本科高校“听学用”会计人才培养的实然之困还包括校企合作深入不够。由于资源的限制，一些学校无法提供足够的实践教学，导致学生的实习机会有限，缺少实际操作经验会影响其职业技能的发展。许多高校在实施OBE时往往拘泥于表面，缺少应有的深度及广度。这种浅层次的合作使学生难以在实际工作中应用所学知识，无法提高解决实际问题的能力，从而直接导致学生“用不上”的困境。此外，由于企业参与度不高，学生难以了解行业发展趋势以及职业前景，这严重影响了学生的就业竞争力。因此，如何深化校企合作，为学生提供更多的实习实训机会，成为OBE理念下会计人才培养亟待解决的问题之一。

（五）学生缺乏自主学习空间

OBE理念下本科高校应用型会计人才的培养面临诸多实际困难之一便是学生负担过重，缺乏自主学习空间。当前在实施OBE过程中，部分高校为了强调成果导向并且达到各种评估标准，往往是通过增加课程难度和作业量来实现的，这大大增加了学生的负担，这种教学模式不仅难以激发学习主动性，反而可能抑制他们的创造性，也使得学生难以在课后和工作中有效地运用所学知识，这大大降低了学习效果，也背离了OBE的初衷。因此，如何在保证教学质量的同时，减轻学生的负担，提供足够的自主学习空间，成为OBE理念下会计人才培养亟待解决的问题。

三、OBE理念下“听学用”会计人才培养的使然之策

针对上述困境，我们提出以下策略，以期推动OBE理念下“听学用”会计人才培养的有效实施：

（一）强化实践教学

为了确保OBE理念下“听学用”应然之境目标的实现，关键步骤是加强实践教学环节。许多高校在实施OBE时忽略了课程设置与实际工作的联系，导致学生所学知识难以在实际工作中应

用，造成学生所学知识“用不上”的难题。为了解决这一问题，民办本科应紧密贴合市场需求加强与企业行业的对接，了解市场对会计人才需求的变化，在此基础上引入新的会计准则、升级课程设置、优化教学内容，增加与实际操作紧密相关的教学。同时高校还可以增加实验实训设施投入，扩大校企合作范围，积极与企业合作共同培养人才。此外，高校还可以通过强化就业指导服务、鼓励学生参与科研项目及社会实践活动，来锻炼学生的创新能力、培养学生的职业素养，从而提升学生的综合能力。

（二）创新教学方式

为了让学生“听得懂、学得会、用得上”，创新教学方式、增强课堂互动性也是关键措施之一。许多高校仍旧采用传统的讲授式教学模式，教学过程缺乏互动性，为了解决这一困境，教师可以创新教学方法^[7]。比如，采用项目式学习、翻转课堂、角色扮演、案例分析、模拟实训等，通过增加实践环节比重来增强教学的互动性。另外教师还可以借助技术手段，通过建立有效的反馈机制来完善教学，为学生提供个性化的辅导服务，确保每位同学都能按照个人需求进行学习，以最大程度地提升教学效果。通过这些方法，可以在有效提升学生学习体验的同时使他们“听得懂、学得会、用得上”。

（三）加强资源建设

加强课程资源建设能够充分满足学生个性化的学习需求。在实施OBE过程中，许多高校虽然提供了一定的学习资源，但其往往缺乏针对性，难以满足不同学生的学习需求。为了确保学生能够“听得懂、学得会、用得上”，高校应加大学习资源投入，建设丰富多样的学习资源体系，比如在线课程、电子书籍、案例库、实验平台等，为学生提供多元化的选择，同时可以组建数字化会计教学团队，将现代技术有机融入课堂教学^[8]。高校也要注意学习资源的更新优化，确保相关资源适应行业发展，此外，学校还应当拥有国际化视野，为有出国意愿或者有意在外企工作的学生增设相应国际化学习资源，比如语言沟通、财务报表翻译等相关课程^[9]。

（四）深化校企合作

为了实现OBE理念下应用型会计人才的有效培养，针对学生实习实训机会有限的困境，深化校企合作、拓展实训渠道是关键策略之一。深化校企合作同强化实践教学不可分离，虽然许多高校在实施OBE过程中反复强调与企业的合作，但往往工作只停留在表面层次，缺乏与企业的深度交流。为了真正实现以学生为中心的教育目标，高校应积极与企业建立长期稳定的合作关系，共同制定人才培养方案、教学计划，将企业的真正需求设置于课程之中，从而使人才培养质量符合行业企业的需求。同时，高校还可以开展深度校企合作教学模式^[10]，比如，邀请企业导师分享最新行业动态，使学生更加了解发展趋势以及就业前景，这能有效促成“听学用”目标的达成。

（五）减轻学生负担

尽管许多高校在开展OBE时强调成果导向，但实施过程中往往增加了学生的学业压力。针对学生学习任务重、缺乏自主学习空间的困境，教育者可以通过优化课程设计、增加实践环节等

措施优化不必要的理论内容，减少学生任务量，让学生所学的知识更加贴近实际工作需求。学校也可以建立有效的反馈机制，了解他们在学习过程中遇到的困难，帮助他们更好地吸收知识，避免学生压力过大及相互间的过度竞争。只有合理安排作业量、控制任务难度、脱离应试教育、避免过度竞争、按需提供个性化服务，才能够有效减轻学生的负担，使他们能够在轻松愉快的氛围中学习，真正做到“听得懂、学得会、用得上”。

OBE 理念下“听学用”会计人才培养策略的有效实施需要高校、教师和学生三方面的共同努力，但牵头力量依然是学校和教师。通过强化实践教学、创新教学方法、加强资源建设、深化校企合作、减轻学生负担等路径可以解决本科高校应用型会计人才

培养所面临的问题，推动 OBE 理念在会计教育领域的有效实施。

四、结束语

OBE 理念在本科高校应用型会计人才培养中的实践已是不争的事实。本文通过对 OBE 理念下“听学用”会计人才培养的应然之境、实然之困以及使然之策进行了系统的分析，为会计教育改革提供了有益参考。未来，OBE 理念在会计人才培养中的应用将越来越广泛，但其成功并非一蹴而就，而是需要教育者不断地实践、改进。作为教育者，我们应继续秉持以学生为中心的教育理念，不断探索符合时代需求的新型会计人才培养模式。

参考文献

[1] 周建平. OBE 理念下应用型本科教育课程改革的几点反思 [J]. 黑龙江高教研究, 2023, 41(12):149-154.

[2] 郭斌, 刘杨南旺. “校企合作校地融合”视角下应用型本科高校人才培养模式创新研究 [J]. 皮革科学与工程, 2023, 33(05):110.

[3] 夏会, 程平. “三链”融合驱动下高校财会人才大数据能力培养 [J]. 财会月刊, 2023, 44(14):105-109.

[4] 叶炼. 要让学生听得懂学得会得上 [J]. 数理化解题研究, 2021, (04):2-3.

[5] 程平, 陈锐. 互联网+MPAcc 人才培养长效机制探析——基于“三链”融合视角 [J]. 财会月刊, 2023, 44(19):120-124.

[6] 孙倩, 宋在科. 应用型本科高校智能财会人才培养路径研究 [J]. 财会通讯, 2022, (11):167-171.

[7] 王蕾, 葛军. 应用型本科院校会计专业课程思政建设研究——基于数智化背景 [J]. 财会通讯, 2024, (05):166-171.

[8] 杨忠海, 李瑛玫, 李潭, 等. 数字化时代《会计学》课程教学改革实践与思考 [J]. 会计之友, 2024, (06):90-96.

[9] 周晓苏. 关于国际化会计人才培养的思考 [J]. 会计之友, 2024, (22):153-155.

[10] 东晓莉. 基于 OBE 理念的营销实践教学模式改革与探索——以宝鸡文理学院市场营销专业为例 [J]. 新经济, 2022, (02):103-107.

基于轻量化深度卷积神经网络的番茄叶片病害分类识别研究

崔鹏¹, 李杰¹, 汤岩春^{1, 2}, 许名扬³, 王晓松^{1*}

1. 山东工商学院 管理科学与工程学院, 山东 烟台 264005

2. 山东黄金矿业(玲珑)有限公司, 山东 招远 265400

3. 黑龙江省依安县信访局, 黑龙江 依安 161500

DOI:10.61369/ASDS.2026040006

摘 要 : 由于传统深度卷积神经网络(CNN)模型参数冗余高、计算资源消耗大,难以满足农业无人机及边缘计算设备的实时推理需求,本研究系统分析和验证 MobileNetV3、ShuffleNetV2、GhostNet 与 EfficientViT 等典型轻量化模型在番茄叶片病害细粒度分类任务中的性能表现。针对包含 10 类病害特征的番茄叶片统一数据集,从分类精度、计算复杂度(FLOPs)、推理延迟及模型可解释性等多维度进行量化评估。实验结果表明, EfficientViT 模型在综合性能上表现最优, F1-score 值达到 0.95, 展现出对复杂纹理病斑的卓越特征提取能力。而 ShuffleNetV2 则凭借最小的 FLOPs (0.15G) 与最高的推理吞吐量 (5625 images/s), 成为实时检测场景的首选方案。GhostNet 虽然参数量极致压缩,但在处理复杂病害特征时存在一定的精度瓶颈。Grad-CAM 可视化分析进一步证实, EfficientViT 能够更精准地聚焦于病灶区域, 具有更强的抗背景干扰能力。本研究为智慧农业装备中的边缘端模型选型提供了科学的理论依据与数据支撑。

关 键 词 : 番茄叶片病害; 轻量化模型; 细粒度分类; 复杂病害特征

Research on Classification and Recognition of Tomato Leaf Diseases Based on Lightweight Deep Convolutional Neural Networks

Cui Peng¹, Li Jie¹, Tang Yanchun^{1, 2}, Xu Mingyang³, Wang Xiaosong^{1*}

1. School of Management Science and Engineering, Shandong Technology and Business University, Yantai, Shandong 264005

2. Shandong Gold Mining (Linglong) Co., Ltd., Zhaoyuan, Shandong 265400

3. Yi'an Bureau of Letters and Visits, Yi'an, Heilongjiang 161500

Abstract : Because traditional deep convolutional neural network (CNN) models have high parameter redundancy and consume a large amount of computing resources, it is difficult to meet the real-time inference requirements of agricultural drones and edge computing devices. This study systematically analyzes and verifies the performance of typical lightweight models, including MobileNetV3, ShuffleNetV2, GhostNet, and EfficientViT, in the fine-grained classification task of tomato leaf diseases. Focusing on a unified tomato leaf dataset containing 10 types of disease characteristics, quantitative evaluations were conducted from multiple dimensions, such as classification accuracy, computational complexity (FLOPs), inference latency, and model interpretability. The experimental results show that the EfficientViT model achieves the best comprehensive performance, with an F1-score reaching 0.95, demonstrating superior feature extraction capabilities for complex texture lesions. Meanwhile, ShuffleNetV2, with the lowest FLOPs (0.15G) and the highest inference throughput (5625 images/s), emerges as the preferred solution for real-time detection scenarios. Although GhostNet highly compresses the number of parameters, it encounters certain accuracy bottlenecks when processing complex disease features. Grad-CAM visualization analysis further confirms that EfficientViT can more accurately focus on the lesion areas, possessing stronger resistance to background interference. This study provides a scientific theoretical basis and data support for edge model selection in smart agricultural equipment.

Keywords : tomato leaf diseases; lightweight models; fine-grained classification; complex disease feature

基金项目: 2025 省级大学生创新创业训练项目一视析果丰——基于 Vision Transformer 与数据增强的苹果叶片病害分类识别研究 (S202511688012)。

作者简介:

崔鹏 (2006-), 男, 汉族, 山东青岛人, 本科在读, 研究方向: 信息管理与信息系统;

李杰 (1997-), 男, 汉族, 山东泰安人, 研究生在读, 研究方向: 工业工程;

汤岩春 (1990-), 男, 汉族, 山东烟台人, 研究生在读, 山东黄金矿业(玲珑)有限公司地质工程师, 研究方向: 矿产地质勘查和矿山工程管理;

许名扬 (2001-), 女, 汉族, 黑龙江依安人, 本科, 研究方向: 信息管理与信息系统。

通讯作者: 王晓松, 山东工商学院管理科学与工程学院, 教授, 硕士生导师, 研究方向为大数据分析与应用。

引言

番茄作为全球重要的经济作物，其叶片病害的早期精准识别对保障产量与品质至关重要。传统人工诊断受主观经验与时效性的限制，难以适配现代化农业设施实现精准植保的目标。以 ResNet、DenseNet 为代表的深度卷积神经网络虽然在图像识别领域取得了显著成果，但受限于庞大的参数规模与高额的计算负载，在农业无人机和手持检测仪等资源有限的终端上进行部署是个难题，且难以实现实时推理。

近年来，以 MobileNet、ShuffleNet 为代表的轻量化网络架构，以及融合 Transformer 机制的 EfficientViT 等模型，通过深度可分离卷积、通道混洗及注意力机制等技术，在保证一定精度的前提下显著降低计算复杂度。然而，现有研究多聚焦于通用图像分类或单一模型的改进，缺乏针对番茄病害这一特定细粒度分类任务的系统性横向对比研究，尤其是不同轻量化架构对局部病斑纹理与全局色彩特征的适配性还不明确。

因此，本研究通过构建统一的实验基准，从模型精度、运算效率、收敛特性及可解释性等维度，对四类主流轻量化模型进行全方位评估，以其为农业智能装备的算法选型提供客观、量化的参考依据。

一、相关研究

（一）农作物病害识别技术

1. 传统机器学习方法

早期病害识别主要依赖传统图像处理技术，通过人工设计的颜色、纹理及形状特征（如 LBP、GLCM）结合支持向量机（SVM）或 K 近邻（KNN）分类器进行判别^[1]。然而，该类方法严重依赖专家先验知识进行特征工程，特征鲁棒性较差，难以适应非结构化农业环境中的复杂光照变化与背景干扰^[2]。随着深度学习技术的发展，卷积神经网络（CNN）凭借端到端的特征自学习能力，显著提升了识别精度，逐渐成为当前研究的主流范式^[3, 4]。

2. 深度学习方法

深度学习技术在计算机视觉领域发展迅猛，基于深度神经网络的农作物病害识别方法取得突破性进展^[5, 6]。卷积神经网络通过多层非线性变换，能够从原始图像中自动学习具有强判别力的特征表示，显著减少对人工特征设计的依赖^[7]。大量研究表明，深度学习模型在多种作物叶片（如水稻、玉米、小麦、苹果、葡萄等）病害识别任务中均取得优于传统方法的识别性能^[8, 9]。通过端到端训练方式，模型能够有效捕获病斑的颜色变化、纹理结构及空间分布特征，大幅提高对细粒度病害类别的区分能力^[10]。然而，随着网络结构不断加深（如，ResNet、DenseNet 等），模型参数量和计算复杂度显著增加，模型在农业无人机、手持终端等资源受限场景中的部署面临巨大挑战^[11, 12]。

（二）轻量化深度学习模型

1. MobileNet 模型

MobileNet 系列模型通过引入深度可分离卷积，将标准卷积分解为深度卷积和逐点卷积，从而在保持精度的同时显著减少计算量和参数数量。MobileNetV3 在此基础上进一步引入注意力机制（SE 模块）和高效激活函数（h-swish），在模型精度、推理速度与计算复杂度之间实现较好的平衡^[5]。由于其结构紧凑、效率较高，MobileNet 系列已成为轻量化网络研究中的典型代表，被广泛应用于移动端病害识别任务中。

2. ShuffleNet 模型

针对组卷积中的信息流通不畅问题，ShuffleNet 系列模型

提出了通道分割和通道混洗机制，有效减少卷积运算中的冗余计算，同时降低内存访问成本。ShuffleNetV2 在结构设计过程中进一步考虑实际硬件加速特性，在相同 FLOPs 条件下实现更高的推理速度^[13]。该类模型在保证较低计算复杂度的同时，能够维持较强的特征表达能力，特别适用于对实时性要求较高的检测场景。

3. GhostNet（GhostModule）模型

GhostNet 通过提出 GhostModule 的概念，利用廉价的线性变换生成冗余特征图（Ghostfeatures），从而减少对传统卷积操作的依赖。该方法在保持模型基本性能的同时，实现参数量和计算复杂度的极致压缩。尽管 GhostNet 在极低算力条件下具有显著的应用优势，但在处理复杂背景下的细粒度病害特征时，特征表达能力可能受到一定限制^[14]。

4. EfficientNet/EfficientViT 模型

EfficientNet 系列模型通过复合缩放策略，在网络深度、宽度和输入分辨率之间实现统一优化，后续研究引入 Fused-MBConv 等结构以提升训练效率^[15, 16]。近年来，VisionTransformer（ViT）结构逐渐被引入轻量化模型设计中。EfficientViT 将卷积操作的局部感知能力与自注意力机制的全局建模能力相结合，在保持较低计算复杂度的同时增强模型对上下文信息的捕获能力，在复杂纹理和细粒度分类任务中表现出良好潜力^[17]。

（三）轻量化模型在农业中的应用现状

随着农业智能化与信息化水平的不断提升，轻量化深度学习模型在作物病害识别、生长监测和智能农业装备等领域得到广泛关注^[14, 18]。研究表明，轻量化模型在保证一定识别精度的前提下，能够有效满足农业现场对实时性、低功耗和易部署性的严苛要求^[4, 5]。在农作物病害识别任务中，轻量化模型已被成功应用于多种作物叶片图像分类研究，并在移动终端和嵌入式设备上取得初步应用成效。然而，现有研究多集中于单一模型的改进或应用，不同轻量化架构（如，纯卷积架构与混合架构）在番茄病害细粒度分类任务中的性能差异尚缺乏系统比较，各种轻量化模型网络结构特性与病斑特征识别效果之间的内在联系有待进一步研究。因此，本研究在相同实验条件下，对多种典型轻量化模型在番茄叶片病害分类任务中的性能进行系统对比分析，以期

智能系统中边缘端模型的选型提供科学的实验依据和参考^[19, 20]。

二、材料与方法

（一）数据集构建

1. 数据来源与分布

本研究数据集由公开植物病害图库与田间实采图像整合而成，涵盖10个类别，包括细菌性斑点、早期疫病、晚期疫病、叶霉病、黄化曲叶病毒及健康叶片等，样本总量约1.9万张，分辨率归一化为 224×224 ，各类别样本的具体数量及占比分布如图1所示。数据集呈现显著的不均衡性，其中样本量最高的番茄黄化曲叶病毒（约4000张）与最低的番茄花叶病毒（约500张）之间存在约8倍的极值差异。针对数据集的这种分布不均衡特征，实施标准化数据增强策略，包括随机旋转（ $\pm 30^\circ$ ）、镜像翻转、HSV色彩空间扰动及随机裁剪，以提升模型在不同光照与姿态下的泛化能力。同时，为减轻这一分布特征对模型训练的负面影响，在训练阶段引入加权损失函数（WeightedLoss）进行修正，最终将数据按8:1:1的比例划分为训练集、验证集和测试集。

Distribution of Tomato Leaf Disease Categories in Dataset

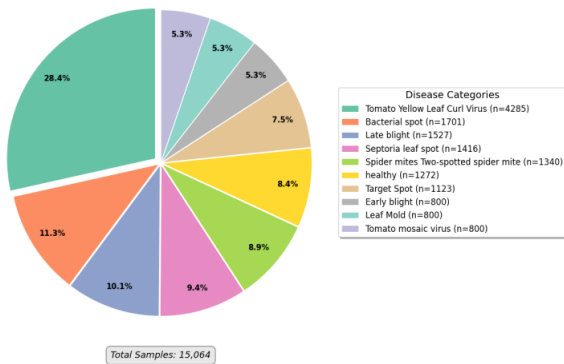


图1：番茄叶片病害数据集各类别样本数量分布

2. 数据增强

为提升模型泛化能力并抑制过拟合风险，本研究在训练阶段实施了标准化的数据增强策略，具体效果如图2所示。增强方法主要包括：(1) 旋转增强，随机旋转 $\pm 30^\circ$ 以模拟叶片的自然生长姿态；(2) 翻转增强，执行水平与垂直翻转，赋予模型对叶片方向的不变性；(3) 色彩扰动，随机调整图像亮度、饱和度与对比度，以模拟田间复杂多变的光照条件；(4) 随机裁剪，在 $0.8\sim 1.0$ 的比例范围内随机裁剪并重缩放至目标尺寸，增强模型对病灶尺度变化的鲁棒性。上述策略既严格保持图像语义一致性，又显著扩展训练样本多样性空间。

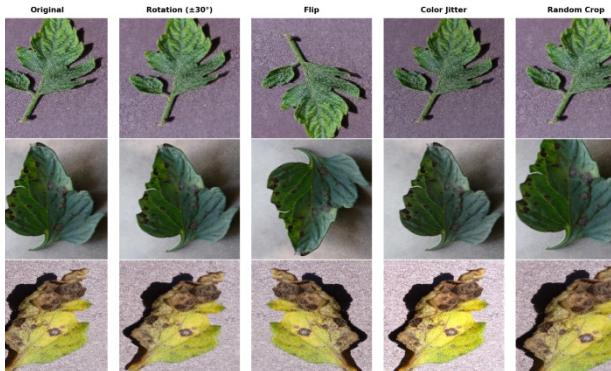


图2：数据增强策略可视化效果对比

（二）对比模型结构

本研究选取EfficientViT、GhostNet、ShuffleNetV2以及MobileNetV3四种具有代表性的轻量化网络架构进行对比。这四类模型在设计理念上各有侧重，分别代表当前轻量化设计的主流方向（如，CNN架构优化、混合架构设计等）。为了更直观地阐述各模型的结构特性与应用潜力，本研究从核心机制、设计优势及适用场景三个维度对上述模型进行总结对比，对比结果如表1所示。

表1：不同轻量化模型在测试集上的识别性能对比

模型	核心机制	设计机制	适用场景
MobileNetV3 ^[5]	深度可分离卷积；SE注意力模块 +h-swish激活	在模型大小、精度与速度之间实现良好的平衡	通用的移动设备设备及常规嵌入式应用
ShuffleNetV2 ^[13]	通道分割；通道混洗	有效降低计算冗余；推理速度显著快于同级模型	对实时性要求极高、需快速推理的检测任务
GhostNet ^[14]	Ghost模块；线性变换生成特征图	在保持精度前提下极大减少计算开销；实现极致的参数压缩	存储空间与算力极其有限的边缘设备
EfficientViT ^[17]	融合CNN与Transformer架构； 优化缩放策略	兼具局部特征提取与全局感知能力；在复杂纹理识别任务中表现优异	对识别精度要求较高、需处理复杂病斑纹理的场景

（三）实验环境与超参数设置

实验依托TensorFlow/Keras深度学习框架，硬件环境配备NVIDIA RTX4060 GPU。训练过程采用Adam优化器，初始学习率设定为0.001，Batchsize为32，迭代轮次（Epoch）为30。损失函数选用交叉熵损失（Cross-Entropy Loss）。为保证实验结果的可比性与可复现性，本文所有性能指标（Accuracy、

Precision、Recall、F1-score及Loss）均来源于同一轮实验配置下模型在测试集上的最终评估结果，全文数值统一保留两位小数。

（四）评价指标

为全面、客观地评价不同轻量化模型在番茄叶片病害识别任务中的分类性能，评价从整体准确性与类别判别能力两个层面出

发，选取准确率（Accuracy）、精确率（Precision）、召回率（Recall）和 F1-score 作为主要评价指标。其中，Accuracy 用于衡量模型整体分类正确程度，而 Precision、Recall 与 F1-score 则更关注模型在类别层面的判别性能，能够有效反映模型在类别不平衡条件下的综合表现。准确率（Accuracy）定义为模型预测正确的样本数占总样本数的比例，其计算公式为：

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

精确率（Precision）用于衡量模型预测为某一类别的样本中，实际属于该类别的比例，反映模型预测结果的可靠性，其计算公式为：

$$Precision = \frac{TP}{TP + FP}$$

召回率（Recall）用于衡量某一类别样本中被模型正确识别的比例，反映模型对目标类别的覆盖能力，其计算公式为：

$$Recall = \frac{TP}{TP + FN}$$

F1-score 是 Precision 与 Recall 的调和平均值，综合考虑模型的查准率与查全率，在类别分布不平衡的情况下具有更强的评价稳定性，其计算公式为：

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

其中，TP 表示被正确预测为正类的样本数量，TN 表示被正确预测为负类的样本数量，FP 表示被错误预测为正类的样本数

量，FN 表示被错误预测为负类的样本数量。在多分类任务中，本文对各类别分别计算 Precision、Recall 与 F1-score，并采用宏平均（Macro-average）方式对各类别指标进行综合统计，以保证不同病害类别在整体评价中的权重一致，从而更客观地反映模型

$$F1_{macro} = \frac{1}{C} \sum_{i=1}^C F1_i$$

其中，C 表示类别总数， $F1_i$ 表示第 i 类的 F1-score。

三、实验结果与分析

（一）性能分析

为全面评估各模型在番茄叶片病害分类任务中的识别性能，本文选取准确率（Accuracy）、精确率（Precision）、召回率（Recall）和 F1-score 作为主要评价指标。

实验结果如表 2 所示，四种轻量化模型在测试集上均表现出较好的识别能力。EfficientViT 在各项指标上均位居首位，其准确率达到 96.5%，F1-score 达到 0.95，显著高于其他卷积神经网络模型，证明其结构在复杂病斑纹理与细粒度特征建模方面的优越性。MobileNetV3 与 ShuffleNetV2 性能相近，准确率分别为 93.8% 和 92.1%，在计算复杂度较低的情况下维持了稳定的分类效果。GhostNet 尽管在模型规模上极具优势，但其准确率（90.6%）和召回率（88.6%）在四类模型中相对较低，表明其在极简结构下对复杂特征的捕获能力存在一定局限。

表 2：不同轻量化模型在测试集上的性能对比

模型	Accuracy(%)	Precision(%)	Recall(%)	Loss	F1-score
MobileNetV3	93.8	93.5	90.1	0.47	0.90
ShuffleNetV2	92.1	91.8	89.2	0.29	0.88
GhostNet	90.6	89.8	88.6	0.43	0.89
EfficientViT	96.5	95.8	95.2	0.14	0.95

（二）模型轻量化效率分析

在农业应用场景中，模型的推理效率和资源消耗同样具有重要意义。为分析不同模型在训练与测试过程中的稳定性与收敛特性，实验对各模型在测试集上的准确率与损失随训练轮次变化情况进行了对比，如图 3 和图 4 所示。

从测试集准确率变化曲线可以看出，EfficientViT 在训练初期即可快速达到较高准确率，并在后续训练过程中保持较小波动，显示出良好的收敛速度与稳定性。MobileNetV3 在前期训练阶段存在一定波动，但在训练中后期逐渐趋于稳定。

ShuffleNetV2 和 GhostNet 的收敛过程相对平缓，整体准确率略低于 EfficientViT。

损失曲线分析结果与准确率曲线趋势基本一致，EfficientViT 的测试集损失值整体最低，且变化幅度较小，进一步验证了其较强的泛化能力。

尽管 EfficientViT 在 FLOPs 上略高于其他轻量化模型，如表 3 所示，但其显著提升的分类性能和稳定性在高精度农业识别场景中具有更高综合价值。

表 3：不同轻量化模型在测试集上的效率对比

模型	F1-Score ↑	FLOPs(G) ↓	推理速度 (images/s) ↑	参数量 (M) ↓	训练时间 (s) ↓
MobileNetV3	0.92	0.23	2800.04	4.21	36.21
ShuffleNetV2	0.90	0.15	5625.89	1.26	19.82
GhostNet	0.89	0.15	2671.71	3.91	38.60
EfficientViT	0.95	0.41	1393.88	4.02	60.05

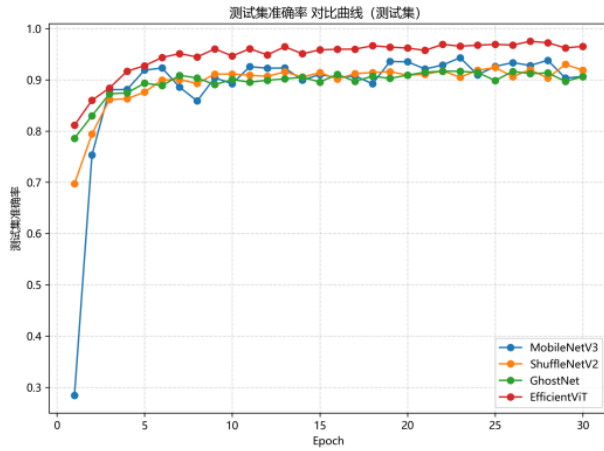


图3：不同模型测试集准确率随 Epoch 变化曲线

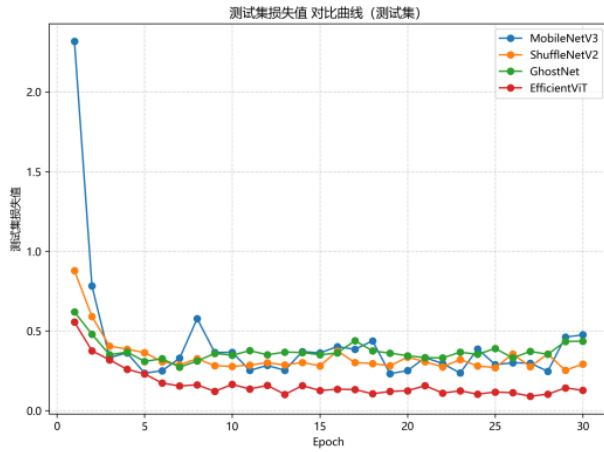


图4：不同模型测试集损失值随 Epoch 变化曲线

（三）混淆矩阵分析

为进一步分析模型在不同病害类别上的分类表现，本文对四种模型在测试集上的预测结果绘制了混淆矩阵。混淆矩阵能够直观反映各类别的正确分类情况及误分类分布特征。

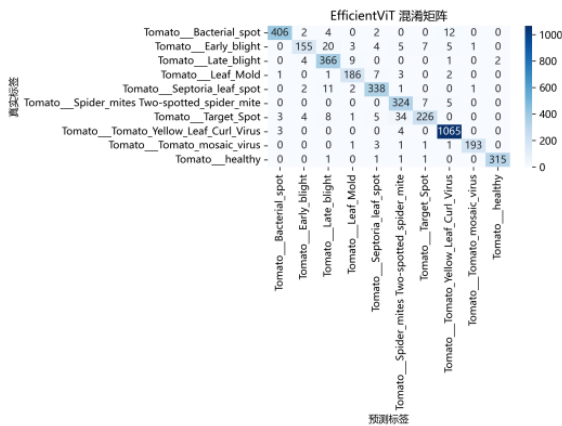


图5：EfficientViT 模型混淆矩阵

混淆矩阵（图5-8）进一步揭示模型在细粒度病害区分上的性能差异。分析显示，所有模型在特征显著且样本量充足的类别（如“健康叶片”和“番茄黄化曲叶病毒”）上均保持极高的识别准确率。然而，在区分“早疫病”与“晚疫病”这一难点上，模型表现出现显著分化。由于这两类病害的病斑均呈现褐色坏死

状，且纹理特征极具相似性，MobileNetV3和 GhostNet 在这两类之间出现了较高的相互误判率。这表明，部分极度轻量化的网络结构在大幅压缩计算量的同时，牺牲对图像高频细节特征的捕获能力，导致其难以提取出区分相似病斑的关键细微纹理。

值得注意的是，在样本量最少的“番茄花叶病毒”类别上，EfficientViT 表现出更强的均衡性，误判数量明显少于其他模型，这说明融合自注意力机制的模型能够更好地学习到小样本类别的判别性特征，减弱了数据集分布不均衡带来的负面影响。

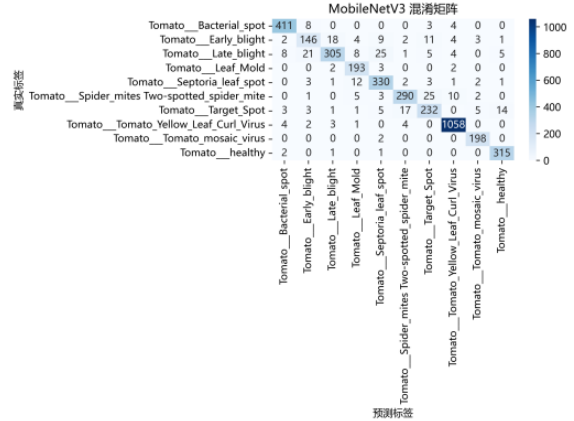


图6：MobileNetV3 模型混淆矩阵

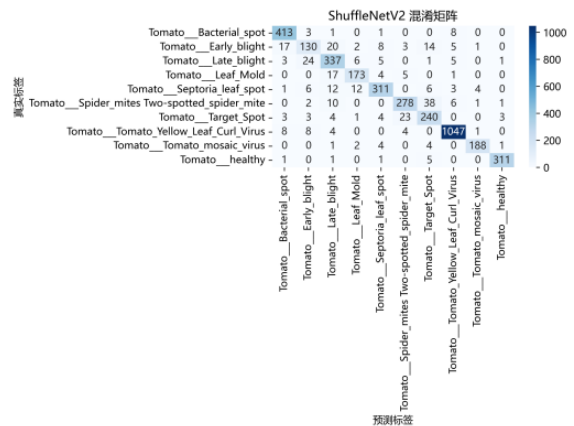


图7：ShuffleNetV2 模型混淆矩阵

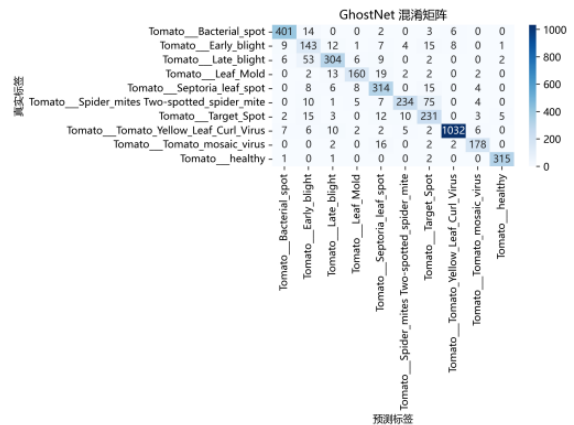


图8：GhostNet 模型混淆矩阵

（四）激活图 Grad-CAM 分析

为增强模型决策过程的可解释性，本文采用 Grad-CAM 方法

对不同模型的分类结果进行可视化分析。通过对网络最后一层卷积特征图进行梯度加权并映射至输入图像空间,生成类别相关的显著性热力图,从而观察模型在分类过程中关注的关键区域。

从 Grad-CAM 可视化结果可知, EfficientViT 模型能够较为准确地聚焦于叶片病斑区域和关键纹理结构,而对背景区域的响应较弱,表明其分类决策主要基于病害相关特征。相比而言, MobileNetV3、ShuffleNetV2 和 GhostNet 在部分样本中存在关注区域分散或背景干扰现象,这在一定程度上影响了其分类性能。

Grad-CAM 可视化是观察模型性能和特点的依据之一。如表 4 所示, EfficientViT 能够精准聚焦于叶片的病斑区域和关键边缘纹理,对背景噪声的响应极弱。相比之下, ShuffleNetV2 和 GhostNet 在处理部分复杂背景样本时,注意力区域存在一定程度的分散或偏移,这解释了其在细粒度分类精度上略逊于 EfficientViT 的原因。

表 4: 四种轻量化模型的 Grad-CAM 热力图可视化对比

MobileNetV3	GhostNet	ShuffleNetV2	EfficientViT
			

(五) 综合分析

综合实验结果分析可得, EfficientViT 适用于追求高精度识别的专业诊断设备, ShuffleNetV2 是农业无人机等实时检测终端的最佳选择, MobileNetV3 提供了性能与资源的良好折中, 而 GhostNet 则适用于对存储空间有极致要求的微型传感器节点。

四、结论与未来工作

(一) 结论

通过量化实验深刻揭示不同轻量化架构在番茄病害识别中的

性能差异及其适用边界。实验结果表明, EfficientViT 所采用的 CNN 与 Transformer 混合架构有效突破纯卷积网络在复杂背景下细粒度特征提取的瓶颈, 以 96.5% 的精度确立其在高标准诊断场景中的标杆地位。与此同时, 验证了 ShuffleNetV2 在极低算力约束下实现实时推理的不可替代性。研究不仅证明各模型在“精度-效率”权衡中的具体位置, 更为差异化的农业智能装备研发提供精准的算法选型策略与理论支撑。

(二) 未来工作展望

基于现有研究成果, 未来将从数据生态、多模态感知及系统集成等维度进行深入研究, 推动技术落地。

复杂田间环境下的鲁棒感知是未来研究需要攻克的难点之一。研究如何拓展采集光照剧变与叶片遮挡下的实景数据, 利用域自适应技术攻克“受控-实景”域偏移难题, 显著提升模型在非结构化巡检场景中的泛化边界。

多模态数据融合与早期预警则是研究的另一个重点问题。突破单一视觉局限, 融合 RGB、高光谱及环境传感数据, 构建“环境-表型”双驱动感知模型, 大幅提升对潜伏期微小病斑的检测灵敏度。

“诊-决-治”闭环系统集成是研究的最终目标。将识别模型嵌入农业 IoT 系统, 联动无人机变量施药平台; 基于病害热力图实现定点、定量的精准作业, 构建“监测-决策-执行”的智慧农业全流程闭环。

参考文献

[1] BILLAH M M, SULTANA A, AFTAB S R, et al. Leaf disease detection using convolutional neural networks: A proposed model using tomato plant leaves[J]. Neural Computing and Applications, 2024: 1-11.

[2] SHARMA R, NAAZ S, VAIDYA P. Harnessing deep learning with AlexNet for tomato leaf disease detection in the Indian Himalayan terrain[J]. Journal of Electrical and Computer Engineering, 2025, 2025(1): 2807347.

[3] GUHA S P, AMIN A B, ARPA S, et al. A real-time application-based convolutional neural network approach for tomato leaf disease classification[J]. Array, 2023, 19: 100313.

[4] 王晨. 基于轻量级卷积神经网络的番茄病虫害分类研究 [D]. 长春: 吉林农业大学, 2023.

[5] HE Z, TONG M. LT-YOLO: A lightweight network for detecting tomato leaf diseases[J]. Computers, Materials & Continua, 2025, 82(3): 4301-4317.

[6] ABDULLAH A, AMRAN A G, TAHMID A M S, et al. A deep-learning-based model for the detection of diseased tomato leaves[J]. Agronomy, 2024, 14(7): 1593.

[7] NI S, JIA Y, ZHU M, et al. An improved ShuffleNetV2 method based on ensemble self-distillation for tomato leaf diseases recognition[J]. Frontiers in Plant Science, 2025, 15: 1521008.

[8] 杨进进, 张文慧, 王哲. 基于改进的 ResNet18 模型识别番茄叶片多种病害 [J]. 现代计算机, 2024, 30(9): 30-34.

[9] 侯文慧, 龚昌智, 曹文昊, 等. 基于超分辨率增强与改进 YOLOv8 的番茄叶片病害检测 [J]. 农业工程学报, 2025, 41(16): 211-220.

[10] YAN C, LI H. CAPNet: Tomato leaf disease detection network based on adaptive feature fusion and convolutional enhancement[J]. Multimedia Systems, 2025, 31(2): 178.

- [11] JIAPING J, SHUFEI L, CHEN Q, et al. A tomato disease identification method based on leaf image automatic labeling algorithm and improved YOLOv5 model[J]. Journal of the Science of Food and Agriculture, 2023, 103(14): 7070–7082.
- [12] SUN Y, NING L, ZHAO B, et al. Tomato leaf disease classification by combining EfficientNetV2 and a Swin Transformer[J]. Applied Sciences, 2024, 14(17): 7472.
- [13] 宋国柱, 黄文静, 崔帅帅, 等. 基于改进 YOLOv8n 的轻量化番茄叶片小目标病害识别方法 [J]. 农业工程学报, 2025, 41(10): 232–242.
- [14] SLADOJEVIC S, ARSENOVIC M, ANDERLA A, et al. Deep neural networks based recognition of plant diseases by leaf image classification[J]. Computational Intelligence and Neuroscience, 2016(6): 1–11.
- [15] CHEN M, WANG C, LIU C, et al. Tomato leaf disease detection method based on improved YOLOv8n[J]. Scientific Reports, 2025, 15(1): 25837.
- [16] ZAHID U, NAJAH A, MONA J, et al. EffiMob-Net: A deep learning-based hybrid model for detection and identification of tomato diseases using leaf images[J]. Agriculture, 2023, 13(3): 737.
- [17] ALAMPALLY S, CHIRANJEEVI M. A smart solution for tomato leaf disease classification by modified recurrent neural network with severity computation[J]. Cybernetics and Systems, 2024, 55(2): 409–449.
- [18] PREETI B, SHARMA R J, KETAN K. TomConv: An improved CNN model for diagnosis of diseases in tomato plant leaves[J]. Procedia Computer Science, 2023, 218: 1825–1833.
- [19] KUMAR V, KAUR M, MITTAL V, et al. TLD-FDL: Tomato leaf disease classification using end-to-end fusion of deep learning models[J]. SN Computer Science, 2025, 6(8): 933.
- [20] 许悦, 陈琳. 基于改进 YOLO v8 的番茄叶片病害检测算法 [J]. 江苏农业科学, 2024, 52(17): 192–200.

基于数据增强与 XGBoost 模型的信用卡欺诈检测算法

邓磊

合肥工业大学 数学科学学院, 安徽 合肥 230009

DOI:10.61369/ASDS.2026040007

摘 要 : 信用卡欺诈行为对经济、金融等诸多领域产生严重不良影响, 如何有效预防、识别信用卡欺诈行为具有重要理论和实际价值。另一方面, 来自银行匿名的真实数据集, 高度不平衡, 正常数据远远大于欺诈数据。本文首先基于 CGAN 和 Isolation Forest 对数据加强, 构建贝叶斯优化器的 XGBoost 模型, 用于检测信用卡欺诈, 称为 CGAN-IF-Bayes-XGB 算法。具体来说使用 CGAN 生成新的、真实的数据来补充原始数据集, Isolation Forest 则用于计算每条样本的“异常性”评分, 作为新的特征注入, 以此来增加样本中的有效信息量, 为 XGBoost 模型更好地识别诈骗交易做好铺垫。在 CGAN-IF-Bayes-XGB 算法中使用了贝叶斯优化器来对 XGBoost 中的超参数进行优化, 确保 XGBoost 识别效果达到最佳。其次, 本文把 CGAN-IF-Bayes-XGB 算法与其他常用的机器学习算法(例如 KNN 和 Light GBM)进行了比较, 比较结果表明: CGAN-IF-Bayes-XGB 算法具有更高的准确率、真阳性率、真阴性率和马太相关系数, 这使其成为检测信用卡欺诈的有效方法。

关 键 词 : CGAN 模型; XGBoost 模型; Isolation Forest 模型; 贝叶斯优化器

Credit Card Fraud Detection Algorithm Based on Data Augmentation and XGBoost

Deng Lei

School of Mathematical Sciences, Hefei University of Technology, Hefei, Anhui 230009

Abstract : Credit card fraud detection is challenged by severe class imbalance in real-world banking datasets. This study proposes a hybrid framework, termed CGAN-IF-Bayes-XGB, which integrates Conditional Generative Adversarial Networks (CGAN), Isolation Forest (IF), Bayesian Optimization, and XGBoost for fraud detection. CGAN is used to generate synthetic fraudulent samples to alleviate data imbalance, while IF provides anomaly scores as auxiliary features to enhance discriminative representation. Bayesian Optimization is applied to automatically tune XGBoost hyperparameters for optimal classification performance. Experimental results on a real-world anonymized banking dataset show that the proposed model outperforms baseline methods, including KNN and LightGBM, in terms of accuracy, TPR, TNR, and MCC, demonstrating its effectiveness and robustness in credit card fraud detection.

Keywords : CGAN model; XGBoost model; Isolation Forest model; Bayesian optimization

引言

随着电子支付和在线交易的快速普及, 信用卡支付成为全球范围内主要的非现金支付方式。然而, 信用卡的广泛应用也使其欺诈问题日益突出, 这也成为金融机构面临的重大风险来源^[1]。相关研究和行业报告表明, 全球每年因信用卡欺诈造成的经济损失高达数百亿美元, 且呈现出逐年上升的趋势^[2]。在中国, 信用卡日益普及、金融电子化加速发展, 信用卡累计发行量已突破数亿张, 全国范围内交易金额庞大, 市场规模巨大, 这也给金融风险管理带来严峻挑战^[3]。

在上述背景下, 欺诈性的信用卡交易行为如何被准确识别出来已经成为风险控制领域的核心研究课题。相较于传统的规则匹配或人工审核, 数据驱动的自动化欺诈检测能够在大规模交易数据中实时识别潜在的风险交易, 这种方法已被金融机构广泛应用^[4]。然而, 在实际应用中, 欺诈的交易行为往往有着隐蔽性强、模式多变以及与正常交易高度相似等特点, 这使得欺诈检测面临较大挑战^[1]。

基于上述分析, 本文提出了一种基于条件生成对抗网络(CGAN)的数据增强、加入 *Isolation Tree* 异常特征评分以及使用贝叶斯优化 *XGBoost* 模型的新型信用卡欺诈检测方法(CGAN-IF-Bayes-XGB)。该方法从数据层、特征层和模型层三个层面对传统欺诈检测流程进行改进: 首先, 利用 CGAN 学习真实欺诈交易的分布特征, 生成高质量的欺诈样本以缓解交易数据高度不平衡的问题^[5, 6]; 其次, 使用 *Isolation Tree* 对 CGAN 生成的欺诈样本进行质量评估与筛选, 将高质量的生成数据与原始训练集进行融合^[6]; 同时, 基于 *Isolation*

Tree 的异常检测对交易行为进行建模,并将其输出的异常评分作为辅助特征引入后续模型^[7-9],以增强对欺诈交易的判别能力;最后,采用贝叶斯优化器对 *XGBoost* 分类模型中关键的超参数进行自动搜索,以提升模型整体性能和稳定性^[10、11]。

针对信用卡交易数据高度不平衡和欺诈样本特征难以表达等问题,本文的主要创新性体现在如下几个方面:一是提出了一种融合条件生成和异常检测的 CGAN-IF-Bayes-XGB 集成方法。将 CGAN 用于欺诈数据增强、*Isolation Tree* 用于异常性特征注入,并结合贝叶斯优化的 *XGBoost* 分类器,形成一个完整有效的欺诈检测方法。二是使用 CGAN 生成高质量的欺诈样本以缓解数据不平衡问题。与传统过采样方法相比,CGAN 能够学习真实欺诈样本的分布,从而生成更自然、更加多样的新欺诈样本。三是引入 *Isolation Tree* 异常评分作为附加特征以增强区分能力^[7、8]。对每条交易计算异常得分,并作为新特征融合进训练数据,使分类器能够更好地捕获潜在欺诈行为的异常性,提高识别准确率。四是采用贝叶斯优化自动搜索 *XGBoost* 的最优超参数。相比于网格搜索和随机搜索,贝叶斯优化不仅搜索效率更高,而且能更稳定地找到性能最优的超参数组合。

一、文献综述

传统的信用卡欺诈检测方法主要包括基于规则的系统、统计方法以及机器学习(Machine Learning, ML)方法。基于规则的系统依赖预先定义的规则集,主要来源于业务经验和专家知识,无法处理新型的欺诈模式,误报率较高,且规则维护依赖人工,系统灵活性不足^[4]。统计方法对交易数据的统计指标进行建模,如均值、中位数、标准差及聚类分析等,用于识别异常交易。这类方法通常基于分布假设,在处理复杂、特征维度较高的信用卡交易数据时适用性有限^[4]。随着数据规模和计算能力的提升,机器学习方法被广泛应用于信用卡欺诈检测,通过从历史数据中学习非线性特征来提高检测性能。然而,在高度不平衡的数据环境下,传统机器学习模型仍易受到类别分布不均衡的影响,导致难以识别欺诈性交易^[12]。

在检测信用卡欺诈问题中主要挑战是数据分布不平衡,即与合法交易相比,欺诈交易非常少,这可能会导致检测模型有偏差。国内外许多研究人员都尝试使用机器学习和深度学习算法来检测信用卡欺诈。他们为了解决数据不平衡的问题广泛使用 SMOTE 算法来确保训练数据集中正样本和负样本的平衡表示^[13],他们概述了各种采样技术及其不同的方法,还研究了不平衡的数据如何影响学习算法的性能,从而导致不准确、结果不正确以及 F1 评估、召回率和精确率分数降低。

近年来,机器学习方法被广泛应用于异常检测中,并在金融欺诈、网络安全和工业故障诊断等领域取得了很大成就。Nassif 等人对机器学习异常检测方法进行了系统综述,指出在高维、类别高度不平衡的数据场景下,无监督异常检测方法具有显著优势。其中,*Isolation Tree* 由于无需依赖数据分布、计算效率高、能够有效识别异常样本以及能够输出连续型异常评分而被广泛应用,其生成的异常评分还可作为辅助特征引入后续分类模型,从而提升整体检测性能^[1]。

近期,Choi et al. 将 CGAN 模型应用于信用卡数据集上生成新的交易数据^[14]。CGAN 会生成一个信用卡数据集,该数据集可以在不透露实际数据集的情况下用于训练。他们的研究结果表

明,深度学习模型可以应用于信用卡交易数据生成领域。

Lee 和 Par 认为,在处理大量数据集时,深度学习模型的性能优于机器学习模型。作者认为,以往在关于解决不平衡类挑战的研究中,由于过度拟合和数据丢失而存在局限性。为了克服这些问题,他们利用 GAN 和他们提出的模型来处理不平衡的类。然后,他们使用随机森林分类评估模型的检测能力,并基于 GAN 进行数据增强。结果表明,他们提出的模型优于不考虑不平衡类的模型。

总的来说,现有研究在信用卡欺诈异常检测中取得了一定成果,但欺诈检测仍面临诸多挑战,例如交易数据中有效信息密度较低,模型在不同时间段下的稳定性和泛化能力不足、在大规模数据环境下的计算效率不足。针对上述问题,相关研究普遍认为,通过融合数据生成、异常检测与监督分类的多阶段建模策略,可在一定程度上缓解数据不平衡和特征表达受限等问题^[15、16]。

二、模型与方法

本文提出了一种融合数据生成、异常检测与监督分类的信用卡欺诈检测方法,称之为 CGAN-IF-Bayes-XGB。该方法首先使用 CGAN 生成欺诈的交易样本^[5、6],并通过 *Isolation Tree(If)* 筛选出高质量的欺诈数据,以缓解样本类分布不平衡的问题。然后,将筛选后的合成数据与原始训练集数据混合,利用 IF 计算交易异常评分,并将其作为新特征引入数据集,以增强特征表达能力^[6]。最后,在训练集中训练 *XGBoost* 分类模型,并使用贝叶斯优化器对模型的超参数进行调节,通过 AUC 指标选取最优超参数作为最终的欺诈检测模型^[10、17]。整体流程如图 1 所示:

(一) 条件生成对抗网络(CGAN)

传统 GAN 由生成器和判别器构成,通过博弈进行联合优化:生成器 G:生成与真实数据分布近似的数据。判别器 D:区分数据是来自真实数据集还是生成器。其目标函数可表示为^[6]:

$$\underset{G}{\text{MIN}} \underset{D}{\text{MAX}} V(D, G) = E_{x \sim p_{\text{data}}} [\log D(x)] + E_{z \sim p_z} [\log (1 - D(G(z)))] \quad (1)$$

其中: $x \sim p_{\text{data}}(x)$ 表示真实样本 x 服从于真实样本 $p_{\text{data}}(x)$ 分布,即

从原始信用卡交易样本集中随机采样得到的交易样本； $z \sim p_z$ 是从正态分布中采样的随机向量； $G(z)$ 是生成器生成的数据； $D(x)$ 是判别器判断 x 是来自真实样本的概率。

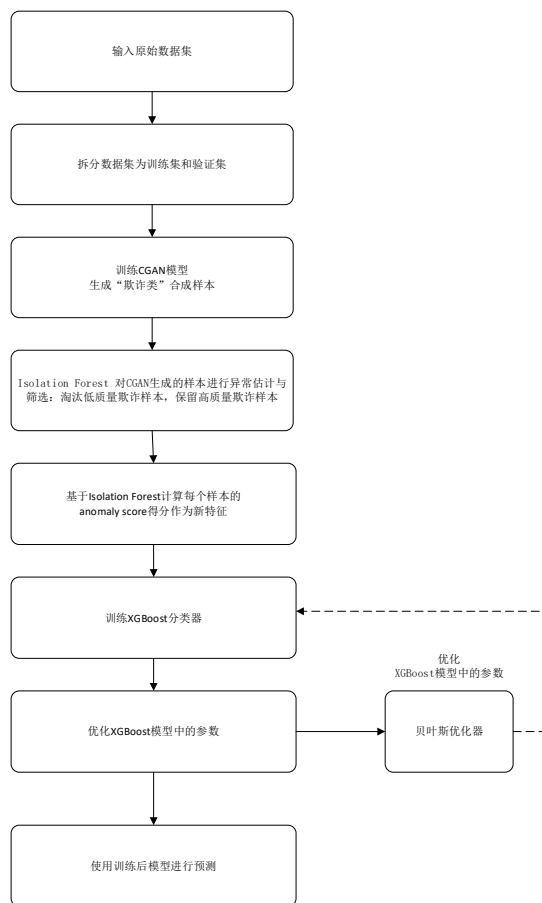


图1：基于 CGAN-Isolation Forest-XGBoost 的信用卡欺诈检测模型流程图

CGAN 是将条件变量 y （如类别、标签、特征）同时引入生成器和判别器，生成特定类型的样本^[18, 19]。比如在信用卡欺诈场景中，将 $y=0$ 代表正常交易， $y=1$ 代表欺诈交易，作为条件变量控制生成样本的类型^[20-22]。CGAN 的目标函数如下^[23]：

$$\min_G \max_D V(D, G) = E_{x \sim p_{data}} [\log D(x|y)] + E_{z \sim p_z} [\log (1 - D(G(z|y)))] \quad (2)$$

在生成器中，将 z 与条件 y 一起拼接，输入网络： $G(z|y) = G([z, y])$ 。

在判别器中，将生成/真实样本 x 与条件 y 一起输入： $G(x|y) = G([x, y])$ 。

在模型训练过程中，CGAN 通过交替优化生成器与判别器，使二者在对抗中逐步收敛。判别器在给定条件标签的约束下学习区分真实交易样本与生成样本，而生成器则在随机向量与条件信息的共同驱动下，不断调整参数以生成更接近真实欺诈样本分布的样本。随着训练的进行，生成器能够逐步识别欺诈交易的潜在特征，从而生成满足条件约束且分布一致性较高的欺诈样本，为后续异常筛选与分类建模提供可靠的数据基础^[19]。

在本文提出的 CGAN-IF-Bayes-XGB 框架中，CGAN 主要用于生成条件为 $y=1$ 的欺诈交易样本，以缓解原始数据集中缺少欺诈交易样本的问题。

（二）Isolation Forest 异常检测模型

虽然 CGAN 能够生成与真实样本分布相似的样本，但其生成样本的质量仍存在不确定性，部分样本可能偏离真实分布且引入噪声^[6]。为此，本文进一步引入 *Isolation Tree* (IF) 对 CGAN 生成数据进行筛选，并使用 IF 对交易样本进行异常性建模，将得到的异常性评分加入数据集中，从而增强模型对欺诈行为的区分能力^[7, 8]。

Isolation Tree 是一种基于树结构的无监督异常检测方法，其核心思想是：异常样本更容易被随机划分过程“孤立”，即在较少的分割步骤中即可与其他样本区分^[7]。与传统依赖距离或密度度量的方法不同，*Isolation Tree* 采取一种完全不同的方式：通过随机划分特征空间对样本进行分离，从而根据一个样本被隔离所需的步骤多少来判断其是否异常^[7]。

在每棵 *Isolation Tree* 中，对于一个样本 x ，其路径长度 $h(x)$ 被定义为：样本 x 从树根节点走到叶子节点（即被孤立）所需的切分次数。通常情况下，异常样本位于稀疏区域，往往在前几次切分中就被单独隔离出去，路径长度 $h(x)$ 较小^[1]。而正常点路径长：它们被大量其他点包围，需要较多次随机切分才能完全隔离，路径长度 $h(x)$ 较大^[7]。由于算法构造了多棵树，因此对每个样本 x 的路径长度会取平均值：

$$E[h(x)] = \frac{1}{t} \sum_{i=1}^t h_i(x) \quad (3)$$

其中 t 是森林中的树数， $h(x)$ 是第 i 棵树中样本 x 的路径长度。

为了便于不同规模下的比较，*iForest* 引入了路径长度的归一化项：调和期望路径长度：对于包含 n 个样本的数据集，树的平均路径长度期望为^[7]：

$$C(n) = 2H(n-1) - \frac{2(n-1)}{n} \quad (4)$$

最终异常评分定义为^[7]：

$$S(x, n) = 2 \frac{E[h(x)]}{C(n)} \quad (5)$$

当 $E[h(x)]$ 很小（即容易孤立）时， $S(x, n) \rightarrow 1$ ，代表高度异常；

当 $E[h(x)]$ 很大（即难以孤立）时， $S(x, n) \rightarrow 0$ ，代表正常点^[7]。

（三）XGBoost 模型

XGBoost 是一种基于梯度提升（Gradient Boosting）的集成学习模型，其基本学习器为 CART 回归树^[24]。从统计学习角度看，*XGBoost* 通过构建加性模型，对交易特征与欺诈标签之间的复杂非线性关系进行建模，从而实现条件分布 $P(Y=1|X=x)$ 的有效逼近^[25, 26]。

设交易样本定义为 (X, Y) ，其中 $x \in \mathbb{R}^d$ 表示包含了 d 维特征的交易特征向量， $Y \in \{0, 1\}$ 为对应的类别标签，分别表示正常交易与欺诈交易，对于给定的训练样本集 $D = \{(x_i, y_i)\}_{i=1}^n$ ，*XGBoost* 模型通过构建多颗回归树组成的加型模型，对样本的预测输出表示为：

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i), f_k \in F \quad (6)$$

其中 F 表示 CART 回归树构成的函数空间, K 为基学习器的数量。

在模型训练过程中, $XGBoost$ 通过最小化经验损失与正则化项之和来学习模型参数, 其目标函数定义为:

$$Obj = \sum_{i=1}^n L(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k) \quad (7)$$

其中 $L(\cdot)$ 为损失函数, 用于衡量预测值与真实标签之间的误差, $\Omega(\cdot)$ 为正则化项, 用以约束模型复杂度并防止过拟合。在本文中, 基学习器采用平方误差损失函数:

$$L(y_i, \hat{y}_i) = (y_i - \hat{y}_i)^2 \quad (8)$$

对应的正则化项定义为:

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T \omega_j^2 \quad (9)$$

其中 T 表示回归树的叶节点数量, ω_j 为第 j 个叶节点的输出权重, γ 与 λ 为正则化系数, 用于控制模型复杂度^[24, 25]。

在优化过程中, $XGBoost$ 利用损失函数关于预测值的一阶与二阶梯度信息, 对目标函数进行二阶泰勒近似, 并在每一轮迭代中生成一棵新的回归树, 从而在保证计算效率的同时提升模型的稳定性与泛化能力^[24, 27]。

在本文提出的 CGAN - IF - Bayes - XGB 框架中, *Isolation Tree* 模型为每一笔交易样本计算异常评分, 记为 S_i , 该评分用于刻画样本在特征空间中的异常程度。最终, 每个交易数据的输入特征向量被扩展表示为: $\tilde{x}_i = [x_i, s_i]$, 即将异常评分作为附加统计特征与原始交易特征共同输入 $XGBoost$ 分类模型。需要强调的是异常评分仅作为辅助特征参与建模, 从而增强模型对潜在欺诈行为的区分能力, 而非直接作为分类标签。

(四) 使用贝叶斯优化器对 XGBoost 中参数进行优化

在 $XGBoost$ 模型中, 我们希望找到一组最优超参数 $\theta = (\theta_1, \theta_2, \dots, \theta_n)$, 使得模型在验证集上的性能指标——AUC (Area Under the ROC Curve) 达到最优^[10, 28]。这个优化问题可表示为:

$$\theta^* = \arg \min_{\theta \in D} f(\theta) \quad (10)$$

其中, $f(\theta)$ 为验证误差 (目标函数), 通常是无法解析的“黑盒”; D 是超参数空间; $\theta = \{n_estimators, max_depth, min_child_weight, learning_rate, \gamma, \lambda\}$ 表示 $XGBoost$ 模型的超参数集合。

从统计建模角度看, 贝叶斯优化通过构建目标函数在超参数空间中的概率代理模型, 对模型性能的不确定性进行刻画。在给定期前 t 次迭代中已观测到的超参数 - 性能对:

$$D_{tr} = \{(\theta_i, f(\theta_i))\}_{i=1}^t \quad (11)$$

贝叶斯优化器通过估计目标函数的后验分布:

$$P(f(\theta) | \theta, D_{tr}) \quad (12)$$

来预测不同超参数配置下模型性能的潜在表现, 并结合采集函数在探索未知区域与利用已有优良参数之间权衡, 从而选择下一组最具潜力的超参数进行评估^[28, 29]。

本文采用树结构的 Parzen 估计器 (Tree-structured Parzen Estimator, TPE) 作为代理模型。TPE 通过分别对性能较优与较差的参数区域建立概率模型, 在有限计算预算下优先采样潜在最优区域, 从而显著提高超参数搜索效率^[11, 30]。该方法已被证明在信用评分与欺诈检测等任务中具有良好的稳定性与实用性^[30]。

贝叶斯优化 $XGBoost$ 的工作流程如下图 2 所示:

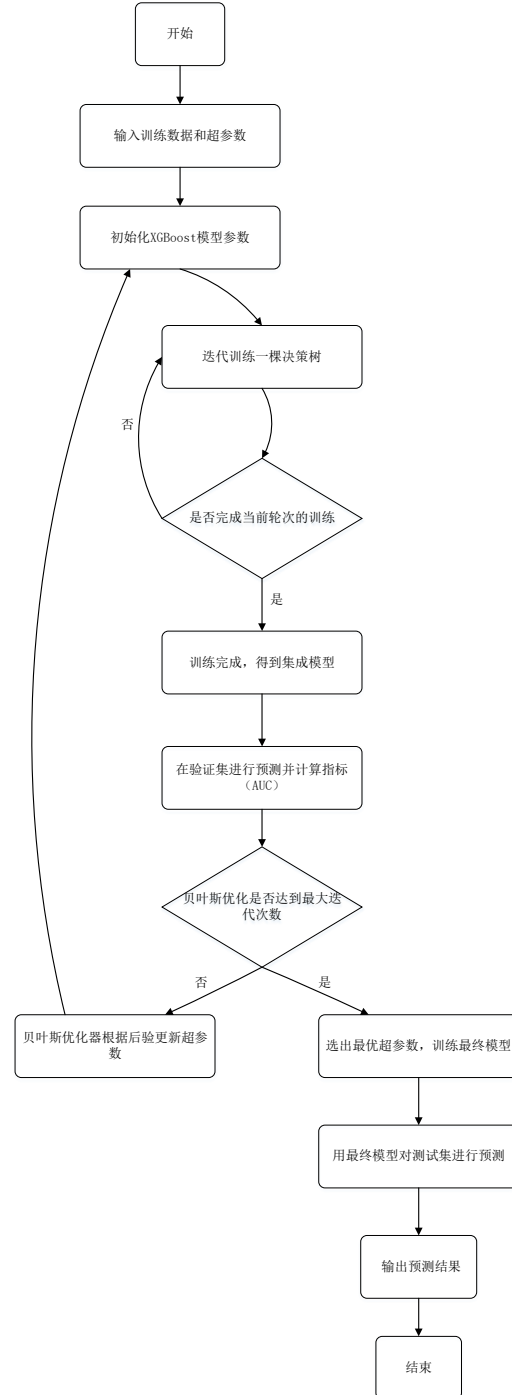


图 2: 贝叶斯优化 XGBoost 流程图

三、数据集说明

验证所提出模型在信用卡欺诈检测任务中的有效性与优越性，本文选用 Worldline 公司与比利时布鲁塞尔自由大学（Universit  Libre de Bruxelles, ULB）机器学习小组在大数据挖掘与欺诈检测研究合作中公开的数据集进行实验分析。该数据集已在 Kaggle 平台公开发布，被广泛用于信用卡欺诈检测相关研究。

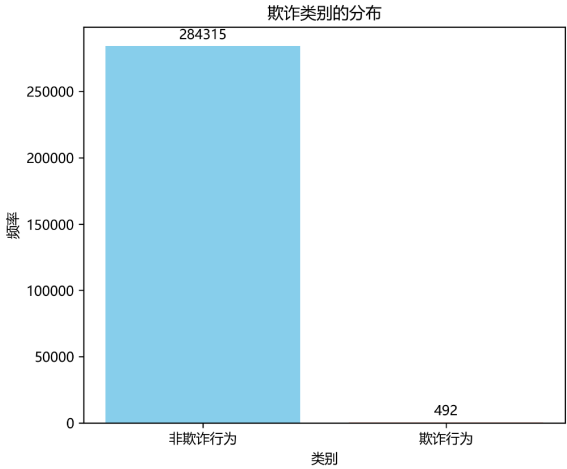


图3：欺诈类数据分布直方图

该数据集包含连续两天内记录的信用卡交易数据，共计 284,807 笔交易，其中欺诈交易仅 492 笔，占比约为 0.17%，具有典型的信用卡欺诈数据特征，数据集类别分布如图3所示。

数据集的特征变量由28个经过主成分分析（Principal Component Analysis, PCA）处理后的匿名特征以及交易金额（Amount）组成。由于交易金额特征与其他经 PCA 处理后的特征在数值尺度上存在显著差异，本文首先对交易金额进行归一化处理，以消除不同特征量级差异对模型训练的影响。

在完成数据预处理后，按照 8:2 的比例将数据集划分为训练集和测试集，其中训练集用于模型训练与参数优化，测试集用于评估模型的泛化性能^[26]。

四、讨论与实验结果

（一）数据集处理

为评估 CGAN+IF 方法在数据增强与样本筛选方面的有效性，本文基于训练集对生成结果进行分析。鉴于原始数据集中包含较多特征，为便于可视化比较与分布分析，本文选取具有代表性的关键特征进行实验验证。

在特征选择方面，本文利用 XGBoost 模型中提供的 `get_fscore` 函数计算各特征在模型训练过程中的重要性得分^[31]，然后将它们的特征分数按从大到小进行排序，如下图4所示，可以看出不同特征对模型预测结果的贡献存在明显差异，其中 V14 和 V29 的贡献最大。

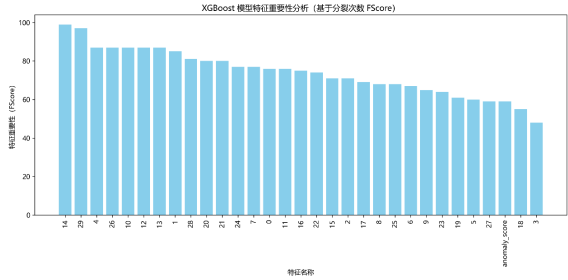


图4：特征重要性排序图

基于上述结果，本文选取特征 V29 和 V14 这两个特征对 CGAN 生成的欺诈样本与原始欺诈样本的分布情况进行对比分析，如下图5所示。可以观察到，生成的欺诈样本主要分布在原始欺诈样本的高密度区域，表明 CGAN 能够较好地学习真实欺诈数据的潜在分布特征。同时 *Isolation Tree* 模型对生成样本进行了进一步筛选，成功剔除了部分偏离真实分布的异常样本，从而提升了合成数据的整体质量。

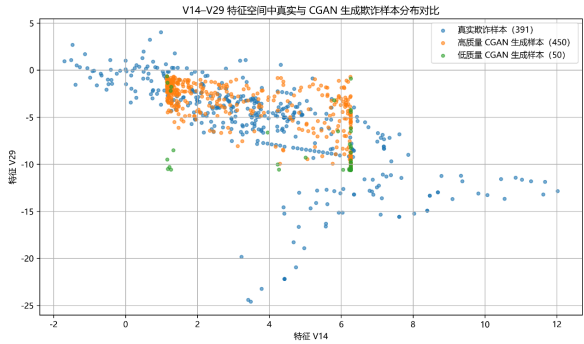


图5：生成欺诈数据与原始欺诈数据分布对比图

此外，为分析 CGAN 的训练稳定性，表1给出了不同训练轮次（Epoch）下判别器与生成器损失函数的变化情况。可以看出，随着训练轮次的增加，判别器在真实样本与生成样本上的损失值整体呈下降趋势并逐渐趋于稳定，说明判别器训练过程较为平稳。生成器损失在训练初期存在一定波动，但随后逐步收敛，表明 CGAN 在训练过程中逐渐学习到真实欺诈样本的分布特征，整体训练过程未出现明显的模式崩溃（Mode Collapse）现象。

表1：CGAN 训练过程中生成器与判别器损失变化

训练轮次 (Epoch)	判别器损失（真实 样本）	判别器损失（生 成样本）	生成器损失
0	0.9634	0.6085	0.8477
500	0.6138	0.6396	0.7517
1000	0.5935	0.6226	0.7883
1500	0.5775	0.5826	0.8333
2000	0.5568	0.5506	0.8905

（二）模型性能评估

1. 贝叶斯优化器与 XGBoost 模型设计

本文将训练集按照 8:2 的比例进一步划分为训练集与验证集^[26]，其中 80% 的数据来训练 XGBoost 模型，20% 的数据用于模型验证，并采用 AUC（Area Under the ROC Curve）作为性能评价指标，以衡量模型在类别不平衡条件下的分类能力^[17]。贝叶

斯优化器以验证集 AUC 作为目标函数, 迭代搜索 30 次, 并最终选择 AUC 值最高时对应的 *XGBoost* 超参数组合作为最终模型^[10、17]。

由下图 6 可以看出, 随着贝叶斯优化迭代次数的增加, 模型性能逐步提升并趋于稳定, 最终获得的 *XGBoost* 模型在验证集上的 AUC 值达到 0.99 以上, 表明贝叶斯优化策略能够有效提升 *XGBoost* 模型在信用卡欺诈检测中的判别能力。

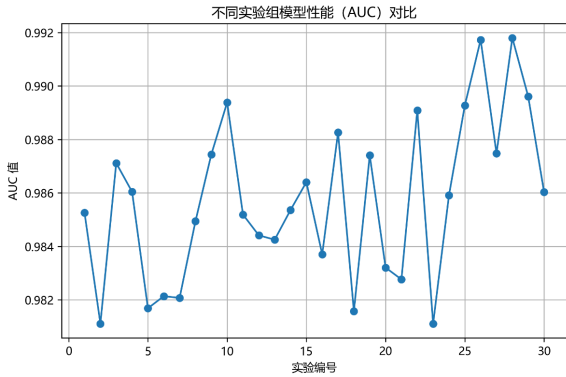


图 6: 在贝叶斯优化器下 XGBoost 的 ACU 得分

2. 模型评价指标

使用 *XGBoost* 模型对每一笔交易输出一个预测得分 (score), 该得分可理解为交易属于欺诈类别的置信度或风险水平^[32], 因此, 在实际应用中需要确定一个合适的分类阈值, 以判断当预测得分高于该阈值时是否将交易判定为欺诈^[32], 为了获得满足特定性能要求的最优阈值, 通常需要对阈值进行微调 and 移动, 并综合评估不同阈值下模型的分类性能^[33], 本文通过多种评价指标对不同阈值进行比较分析, 最终确定最优决策阈值。相关评价指标包括 MCC、ACC、TPR 和 TNR 等^[17、33]。模型性能的评估指标定义如下:

$$TNR = \frac{TN}{TN + FR} \quad (13)$$

$$FPR = \frac{FP}{TN + FP} \quad (14)$$

$$TNR = \frac{TN}{TN + FP} \quad (15)$$

$$TPR = \frac{TP}{TP + FN} \quad (16)$$

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad (17)$$

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP) \times (TP + FN) \times (TN + FP) \times (TN + FN)}} \quad (18)$$

其中 TP 定义了真阳性, 表示模型准确预测了欺诈标签。FN 表示假阴性, 表示模型未能准确预测欺诈标签。FP 表示误报, 表示模型错误地预测了欺诈标签。TN 代表真阴性, 表示模型对非欺诈标签的准确预测。TPR 是真阳性率, 表示准确预测的欺诈样本的比例。TNR 是真阴性率, 表示准确预测的非欺诈样本的比例。FPR 是假阳性率, 表示将非欺诈样本预测成欺诈样本的比例。马修斯相关系数 (Matthews Correlation Coefficient, MCC) 用于衡量模型预测结果与真实标签之间的整体相关性, 其取值范围为

[-1,1]。其中, 1 表示完全正确的预测, 0 表示与随机预测效果相当, -1 表示完全错误的预测^[33]。

图 7 展示了在不同阈值条件下 ACC、MCC、TPR 和 TNR 的变化情况。从图中可以看出, 当阈值设置为 0.4 时, 各项指标均取得了较为平衡且优良的表现, 因此本文将该阈值作为最终模型的判定阈值。

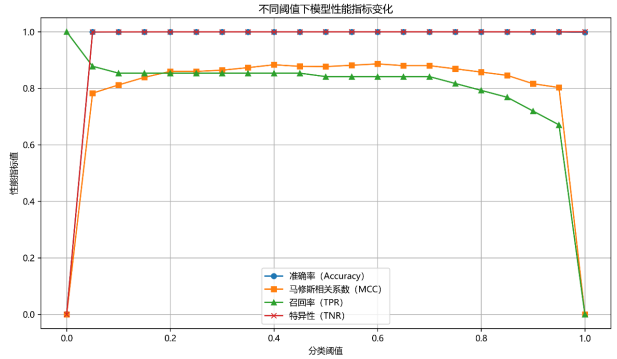


图 7: 不同阈值下的 ACC、MCC、TPR 和 TNR 得分

3. 不同模型的效果对比

在实验的最后, 我们将 CGAN-IF-Bayes-XGB 与相关方法进行性能对比, 如下表 2 所示, 我们可以看出综合四个指标 CGAN-IF-Bayes-XGB 模型在对信用卡欺诈识别上有着显著的优势。

表 2: CGAN-IF-Bayes-XGB 与对比方法的分类性能比较

模型	准确率	召回率	特异性	马修斯系数
CGAN-IF-Bayes-XGB	0.99960	0.86585	0.99980	0.88333
KNN	0.96914	0.88353	0.97113	0.59032
XGBoost	0.96963	0.85294	0.99954	0.81003
Random Forest	0.95832	0.80253	0.99894	0.69024
AE-based Clustering	0.98902	0.81632	0.98933	0.30585

(三) 模型讨论与应用

根据实验结果我们可以得出 CGAN-IF-Bayes-XGB 模型在处理不平衡数据集时有着更大的优势, 同时与传统机器学习模型以及深度学习模型相比在处理信用卡欺诈数据集时优势显著, 与 XGBoost 相比在具有代表性的 ACC 指标上提升了 3.09%, MCC 指标上提升了 9.05%, 与 KNN 相比在 ACC 指标上提升了 3.14%, 在 MCC 指标提升了 49.64%, 这表示在引入 CGAN 和 Isolation Tree 模型后对 XGBoost 的性能有着较大的提升, 但模型仍然存在着训练复杂度较高的问题以及 TPR 指标不如 KNN 的现象, 未来我们将从优化模型复杂度方面入手, 旨在开发出性能更好的 CGAN-IF-Bayes-XGB 模型, 让其应用在更广的风控场景。

五、结束语

本文针对银行信用卡交易中类别高度不平衡和欺诈样本信息不足的问题, 提出了一种融合数据生成、异常检测与监督分类的 CGAN-IF-Bayes-XGB 欺诈检测方法。该方法通过 CGAN 扩充少数类欺诈样本, 并结合 Isolation Forest 提取异常评分作为辅助特征, 有效增强了训练数据的判别信息表达; 同时, 引入贝叶

斯优化对 XGBoost 模型的超参数进行自动调优, 并通过阈值搜索确定最优决策边界。实验结果表明, 在真实银行匿名交易数据集上, 所提出方法在多项评价指标上均显著优于 XGBoost、KNN 等对比模型, 尤其在不平衡数据场景下表现出更强的稳定性和综合判别能力。本文的研究结果表明, 将条件生成模型、异常检测机

制与梯度提升分类器进行有机融合, 是缓解信用卡欺诈检测中数据不平衡与特征表达不足问题的一种有效途径。未来研究将进一步关注模型计算复杂度的优化, 并在不同来源的银行风控数据集上验证其泛化能力。

参考文献

- [1] Nassif, A.B.Talib., M.A.Nasir., Q.Dakalbab., F.M. Machine Learning for Anomaly Detection: A Systematic Review[J]. IEEE Access, 2021, 9, :pp.78658 - 78700.
- [2] THE NILSON REPORT. Global Card Fraud Losses at \$33 Billion[R]. GlobeNewswire, 2026-01.
- [3] 人民网. 信用卡诈骗案一年5万起 银行损失数百亿仅追回16.5亿元 [N]. 2016-08.
- [4] SUBUDHI S., PANIGRAHI S., DASH S.R. ARMS: Automated rules management system for fraud detection[EB/OL]. (2020-02)[2026-01-31]. <http://arxiv.org/abs/2002.06075>.
- [5] LEE J., PARK K. GAN-based imbalanced data intrusion detection system[J]. Personal and Ubiquitous Computing, 2021, 25: 121-128.
- [6] GOODFELLOW I.J., POUGET-ABADIE J., MIRZA M., XU B., WARDE-FARLEY D., OZAIR S., COURVILLE A., BENGIO Y. Generative Adversarial Nets[C]// GHAHRAMANI Z., WELLING M., CORTES C., LAWRENCE N.D., WEINBERGER K.Q., eds. Advances in Neural Information Processing Systems 27. Montreal, Canada: Neural Information Processing Systems Foundation; Curran Associates, Inc., 2014: 2672-2680.
- [7] Liu F.T., Ting K.M., Zhou Z. Isolation-based anomaly detection ACM Trans[J]. Knowl. Discov. Data (TKDD), 2012, 6 (1): 1-39.
- [8] Li, X., Liang, D., Li, X. et al. Quality monitoring of real-time PPP service using isolation forest-based residual anomaly detection[J]. GPS Solutions, 2024, 28(3): 118.
- [9] LIU F.T., TING K.M., ZHOU Z. Isolation Forest[J]. IEEE Transactions on Knowledge and Data Engineering, 2008, 20(8): 413-422.
- [10] SNOEK J., LAROCHELLE H., ADAMS R.P. Practical Bayesian Optimization of Machine Learning Algorithms[C]// LECUN Y., BENJIOU Y., HINTON G., eds. Advances in Neural Information Processing Systems 25. Lake Tahoe, NV, USA: Neural Information Processing Systems Foundation; Curran Associates, Inc., 2012: 2951-2959.
- [11] BERGSTRÄ J., YAMINS D., COX D.D. Making a Science of Model Search: Hyperparameter Optimization in Hundreds of Dimensions for Vision Architectures[C]// GOMEZ F., TUKEY J., EDWARDS R., eds. Proceedings of the 30th International Conference on Machine Learning (ICML 2013). Atlanta, GA, USA: JMLR.org, 2013: 115-123.
- [12] CHEN S., GOO Y.-J.J., SHEN Z.-D. A Hybrid Approach of Stepwise Regression, Logistic Regression, Support Vector Machine, and Decision Tree for Forecasting Fraudulent Financial Statements[J]. The Scientific World Journal, 2014, 11: 1-9.
- [13] ALAMRI M., YKHLEF M. Survey of Credit Card Anomaly and Fraud Detection Using Sampling Techniques[J]. Electronics, 2022, 11: 4003.
- [14] CHOI E., BISWAL S., MALIN B., DUKE J., STEWART W.F., SUN J. Generating multi-label discrete patient records using generative adversarial networks[C]// SMITH J., WANG L., eds. Proceedings of the Machine Learning for Healthcare Conference. Boston: Springer, 2017: 286-305.
- [15] FENG X., ZHAO K., ZENG H. Combining GANs with anomaly detection for generating high-quality minority samples[J]. Knowledge-Based Systems, 2020, 193: 105408.
- [16] STRELCENIA E., PRAKONWIT S. A survey on GAN techniques for data augmentation to address imbalanced data issues in credit card fraud detection[J]. Machine Learning and Knowledge Extraction, 2023, 5(1): 304-329.
- [17] FAWCETT T. An introduction to ROC analysis[J]. Pattern Recognition Letters, 2006, 27(8): 861-874.
- [18] SUH S., LEE H., LUKOWICZ P., LEE Y.O. CEGAN: Classification Enhancement Generative Adversarial Networks for unraveling data imbalance problems[J]. Neural Networks, 2021, 133: 69-86.
- [19] ZAREAPOOR M., SHAMSOLOMOALI P., YANG J. Oversampling adversarial network for class-imbalanced fault diagnosis[J]. Mechanical Systems and Signal Processing, 2021, 149: 107175.
- [20] MULLICK Sankha Subhra S., DATTA Shounak, DAS Swagatam. Generative adversarial minority oversampling[C]// Zhou Jie, Feng Shuiwang, eds. Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, Republic of Korea: IEEE, 2019: 16951704.
- [21] MIRZA M., OSINDERO S. Conditional Generative Adversarial Nets[J]. Computer Science, 2014: 2672-2680.
- [22] BA HUNG. Improving detection of credit card fraudulent transactions using generative adversarial networks[J]. arXiv, 2019, 1907.03355.
- [23] SOMORJIT L., VERMA M. Variants of generative adversarial networks for credit card fraud detection[C]// KAR NIRMALYA, SAHA ASHIM, DEB SUMAN, eds. Trends in Computational Intelligence, Security and Internet of Things: Third International Conference, ICCISIoT 2020, Proceedings. Cham, Switzerland: Springer, 2020: 133-143.
- [24] CHEN T., GUESTRIN C. XGBoost: A scalable tree boosting system[C]// KRISHNAPURAM B., SHAH M., SMOLA A.J., AGGARWAL C.C., SHEN D., RASTOGI R., eds. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco, CA, USA: ACM, 2016: 785-794.
- [25] 李航. 统计学习方法 [M]. 北京: 清华大学出版社, 2012.
- [26] HASTIE T., TIBSHIRANI R., FRIEDMAN J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction[M]. 2nd ed. New York, NY, USA: Springer, 2009.
- [27] LIANG W., LUO S., ZHAO G., WU H. Predicting hard rock pillar stability using GBDT, XGBoost, and xgboost algorithms[J]. Mathematics, 2020, 8(765): 1-12.
- [28] 王刚., 李忠伟., 任明明., 李鹏. 机器学习 - 贝叶斯和优化方法 [M]. 2版. 北京: 机械工业出版社, 2021.
- [29] WU J., CHEN X.Y., ZHANG H., XIONG L.D., LEI H., DENG S.H. Hyperparameter Optimization for Machine Learning Models Based on Bayesian Optimization[J]. Journal of Electronic Science and Technology, 2019, 17(1): 26-40.
- [30] SUN X., XU W., WANG X. Bayesian Optimization for Hyperparameter Tuning in XGBoost-based Credit Scoring Model[C]// LI H., ZHANG Y., eds. Proceedings of the 2020 International Conference on Artificial Intelligence and Big Data (ICAIBD). Chengdu, China: IEEE, 2020: 393-398.
- [31] CHEN T., GUESTRIN C. XGBoost: A scalable tree boosting system[C]// KRISHNAPURAM B., SHAH M., SMOLA A.J., AGGARWAL C.C., SHEN D., RASTOGI R., eds. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco, CA, USA: ACM, 2016: 785-794.
- [32] FRIEDMAN J.H. Greedy Function Approximation: A Gradient Boosting Machine[J]. Annals of Statistics, 2001, 29(5): 1189-1232.
- [33] HE H., GARCIA E.A. Learning from imbalanced data[J]. IEEE Transactions on Knowledge and Data Engineering, 2009, 21(9): 1263-1284.

新能源汽车协同效应与发展路径：基于耦合协调与随机森林分析

王宇琪

上海对外经贸大学 统计与数据科学学院，上海 201620

DOI:10.61369/ASDS.2026040008

摘要： 随着全球能源转型与环境保护需求的加剧，新能源汽车产业已成为推动社会可持续发展的重要引擎。本研究首先构建新能源汽车—生态环境系统耦合协调度模型，量化二者协调程度及时间演变规律；进而，构建“环境质量—能源结构—经济水平—技术创新”四维度预测框架，结合随机森林与 SHAP 可解释性方法识别关键变量的作用路径。研究发现，新能源汽车发展与生态环境改善存在显著协同效应，且清洁能源基础设施与经济技术水平是核心驱动力。本研究为制定差异化的产业政策、优化基础设施布局与激励机制提供了实证依据与决策参考。

关键词： 新能源汽车；耦合协调度模型；随机森林；SHAP

Synergy Effects and Development Paths of New Energy Vehicles: Based on Coupling Coordination and Random Forest Analysis

Wang Yuqi

School of Statistics and Data Science, Shanghai University of International Business and Economics, Shanghai 201620

Abstract： With the intensifying global energy transition and environmental protection demands, the new energy vehicle industry has become a vital engine driving sustainable social development. This study first constructs a coupling coordination degree model for the new energy vehicle–ecological environment system, quantifying their coordination level and temporal evolution patterns. Subsequently, a four-dimensional predictive framework encompassing "environmental quality–energy structure–economic level–technological innovation" is established. Combining random forest and SHAP explainability methods, the study identifies the causal pathways of key variables. Findings reveal a significant synergistic effect between new energy vehicle development and ecological environment improvement, with clean energy infrastructure and economic–technological level serving as core drivers. This research provides empirical evidence and decision-making references for formulating differentiated industrial policies, optimizing infrastructure layout, and refining incentive mechanisms.

Keywords： new energy vehicles; coupling coordination degree model; random forest; SHAP

引言

近年来，随着全球能源危机和环境问题的日益突出，新能源和节能领域获得广泛关注。中国政府将新能源汽车作为战略性新兴产业的重点发展方向，并制定了一系列政策措施来推动其发展。其中，《节能与新能源汽车技术路线图2.0》和《新能源汽车产业发展规划（2021—2035年）》提出力争到2035年纯电动汽车成为新销售车辆主流。

以习近平新时代中国特色社会主义思想为指引，围绕创新、协调、绿色、开放、共享的新发展理念。我国新能源汽车发展在核心技术创新、能源供给适配、空气污染协同治理、与经济市场深度融合等方面仍存在短板，亟待进一步强化。从现有研究来看，学界多聚焦单一政策或技术路径的直接影响，缺乏对“技术—能源—环境—经济”多维度机制分析，系统进行新能源汽车保有量和生态环境协同演进的长期影响和预测分析。

基于上述背景，本文首先利用耦合协调度模型分析新能源汽车系统与生态环境系统的协同演进特征及驱动因素；继而基于随机森林模型对新能源汽车保有量进行多维度预测，并结合 SHAP 值解释方法量化各变量的边际贡献；最后提出针对性政策建议，以期为我国新能源汽车与生态环境的协同发展与可持续路径提供决策参考。

一、文献综述

近年来,新能源汽车(NEV)与生态环境的发展关系成为多学科交叉研究的热点。徐宗煌(2025)^[1]通过构建新能源汽车与生态环境评价指标体系,并结合运用博弈论组合赋权模型、模拟退火模型等模型对2010—2021年中国新能源汽车与生态环境耦合协调发展水平进行分析。潘苏楠(2019)基于生态系统视角对新能源汽车可持续发展进行综合评价以及影响因素分析^[2],通过构建障碍度模型识别影响我国新能源汽车产业发展的障碍因素有行业发展环境可持续化等。张锦程等人(2022)通过构建新能源汽车创新生态系统的“结构-种群-环境”演化分析框架^[3],发现我国新能源汽车产业政策变迁与创新生态系统演变是一个协同演化的过程。

在新能源汽车预测领域,多维度框架逐渐取代单一经济驱动模型。王达等人(2020)^[4],通过随机森林模型提升系统性风险识别的准确性,并利用SHAP值揭示关键风险驱动因素,结果表明随机森林模型对系统性风险的识别显著优于传统的逻辑回归模型,采用Shapley值和作用力图展示不同特征变量的边际贡献。

SHAP值可以量化变量对预测结果的贡献度,能够提高机器学习模型的可解释性(Joseph, 2019; Bracke et al., 2019)^[5,6]。陈小昆等人(2022)引入SHAP框架增强模型可解释性^[7],综合四种模型特征重要性排序结果。罗妍等人(2022)将4种机器学习模型进行比较^[8],并引入SHAP模型用于提供个性化评估和解释。

目前已经有许多学者对新能源汽车方面进行研究,但有关生态环境系统与新能源汽车系统的协调发展以及保有量预测较为单一,故本研究还有一定的探索空间。本研究通过构建两大系统九大指标对协调关系进行分析,进而识别系统重要指标,给予一定的政策建议。同时本文通过构建随机森林模型对新能源汽车保有量进行预测,并引入SHAP值增加模型的可解释性。

二、理论方法介绍

(一) 耦合模型介绍

在物理学中,耦合^[9]是指两个(或两个以上的)系统或运动形式通过各种相互作用而彼此影响的现象。其中耦合度就是衡量系统间相互作用程度的大小,如果系统间相互配合则为良性耦合。协调则反映系统间共同发展的有序状态。协调度是度量系统之间或系统内部要素之间在发展过程中彼此和谐一致的程度,体现了系统由无序走向有序的趋势^[10]。耦合协调度模型就是反映多个系统之间的协同发展水平,常用于评估多个系统的综合效益问题^[11]。本文基于新能源汽车-生态环境系统构建耦合协调度模型,通过非线性耦合关系捕捉系统间动态平衡。两系统表现出耦合和协调发展的关系如下图所示。

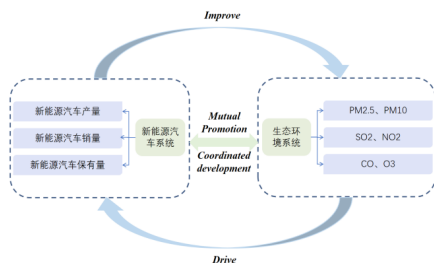


图1：新能源汽车与环境的耦合协调机制

(二) 耦合协调度计算方法

1. 数据标准化处理

为了消除量纲差异,需要对新能源汽车系统内指标 x_i 与生态环境系统内指标 y_i 进行标准化处理,处理方式如下:

$$x'_i = \begin{cases} \frac{x_i - \min\{x_i\}}{\max\{x_i\} - \min\{x_i\}} & (\text{正向指标}) \\ \frac{\max\{x_i\} - x_i}{\max\{x_i\} - \min\{x_i\}} & (\text{负向指标}) \end{cases} \quad (1)$$

$$y'_i = \begin{cases} \frac{y_i - \min\{y_i\}}{\max\{y_i\} - \min\{y_i\}} & (\text{正向指标}) \\ \frac{\max\{y_i\} - y_i}{\max\{y_i\} - \min\{y_i\}} & (\text{负向指标}) \end{cases} \quad (2)$$

2. 确定组合权重及子系统综合评价值

参照徐宗煌(2025年)^[1]关于新能源耦合协调发展及驱动因素的研究,本文使用博弈论组合赋权。这是通过优化模型平衡主客观权重差异,寻求使组合权重与各方法权重偏差最小的均衡解。本文综合主客观权重,采取阶关系分析法(ORA)以及熵权法^[12]确定组合权重大小。其中ORA^[13]是一种基于解析层次的主观加权方法。客观权重采取熵权法,是根据每个指标的相对变化对整个系统的影响来计算指标权重系数^[14,15]。具体步骤如下。

(1) 构建权重矩阵:

设主观权重向量为 w^{ORA} ,客观权重向量为 $w^{Entropy}$,则组合权重表示为:

$$w = \alpha \cdot w^{ORA} + \beta \cdot w^{Entropy} \quad (3)$$

其中 $\alpha + \beta = 1$, $\alpha, \beta \geq 0$ 。

(2) 目标函数最小化:

$$\min \left(\alpha^2 \cdot \|w^{ORA}\|^2 + \beta^2 \cdot \|w^{Entropy}\|^2 \right), s.t. \alpha + \beta = 1. \quad (4)$$

(3) 求解最优系数:

$$\alpha = \frac{\|w^{Entropy}\|^2}{\|w^{ORA}\|^2 + \|w^{Entropy}\|^2}, \beta = 1 - \alpha. \quad (5)$$

(4) 组合权重计算:

$$w_i = \alpha \cdot w_i^{ORA} + \beta \cdot w_i^{Entropy} \quad (i = 1, 2, \dots, n). \quad (6)$$

(5) 系统综合评价值计算:

$$U_k = \sum_{i=1}^n w_i x'_i. \quad (7)$$

其中, $U_k(k=1,2)$ 分别代表新能源汽车系统和生态环境系统。

3. 耦合度(C)、综合发展指数(T)及协调度(D)计算

$$C = \frac{2\sqrt{U_1 \cdot U_2}}{U_1 + U_2}, T = \alpha U_1 + \beta U_2, D = \sqrt{C \cdot T}. \quad (8)$$

其中, $\alpha + \beta = 1$, $D \in [0,1]$ 。

(三) 随机森林模型

随机森林(Random Forest)是一种基于决策树的集成学习算

法，由 Leo Breiman 和 Adele Cutler 于 2001 年^[16]提出。它通过构建多棵决策树，在树的每个节点通过选择最优的分裂特征不停地进行分类，直到达到建树的停止条件，显著降低单棵决策树过拟合的风险，通过降低方差来提高模型的泛化能力。同时，随机森林对噪声和缺失数据表现出强容忍度。

本文对比多元回归、XGBoost 等众多预测模型算法后，选取随机森林模型对新能源汽车保有量进行预测。由于“汽车保有量”数据具有明显的时间依赖性，因此在自变量中纳入汽车保有量的滞后一期观测值，以捕捉时间序列的动态特征。

在建模前，模型使用 StandardScaler 进行列向标准化消除量纲差异^[17]。数据分割采用分层抽样策略，基于时间先后将数据集按 7:3 划分为训练集与测试集，确保样本分布一致性。设定因变量 z 为“汽车保有量”，自变量 x 为除“汽车保有量”外的其他变量以及汽车保有量滞后值 z_{t-1} 。使用下式表示：

$$X = \{x_1, x_2, \dots, x_{13}, z_{t-1}\}, z = \text{汽车保有量} \quad (9)$$

为了消除不同变量量纲带来的影响，对自变量 x 使用标准化转换，其中 μ 表示均值， σ 表示标准差：

$$X_{std} = \frac{X - \mu}{\sigma} \quad (10)$$

文本通过构建多棵决策树并集成其预测结果来降低过拟合风险。其预测结果为所有树的均值。其中 $T_k(x)$ 表示第 k 棵树的预测值， K 为树的总数：

$$\hat{z} = \frac{1}{K} \sum_{k=1}^K T_k(x) \quad (11)$$

三、实证分析

（一）耦合协调度指标与描述性统计

根据新能源汽车、生态环境两系统的作用原理以及指标数据可获得性、可比性、整体性、代表性等构建如下表各系统指标构建，数据来源于 CSMAR 数据库以及中国汽车工业协会的中国新能源汽车产量和销售数据（月）（2022—2005 年）、中国各省空气质量 API 数据（2022—2014 年）。其中新能源汽车保有量即为新能源汽车销量累加数据。

表 1：新能源汽车 - 生态环境系统指标构建		
体系	主要指标	符号说明
新能源汽车系统	新能源汽车产量（辆）	x_1
	新能源汽车销量（辆）	x_2
	新能源汽车保有量（辆）	x_3
	PM2.5 排放量（ $\mu\text{g}/\text{m}^3$ ）	y_1
生态环境系统	PM10 排放量（ $\mu\text{g}/\text{m}^3$ ）	y_2
	SO ₂ 排放量（ $\mu\text{g}/\text{m}^3$ ）	y_3
	NO ₂ 排放量（ $\mu\text{g}/\text{m}^3$ ）	y_4
	CO 排放量（ mg/m^3 ）	y_5
	O ₃ 排放量（ $\mu\text{g}/\text{m}^3$ ）	y_6

在此基础上对新能源汽车 - 生态环境系统指标做描述性统计与分析。

表 2：新能源汽车 - 生态环境系统指标描述性统计表

指标	最小值	最大值	平均值	标准差
新能源汽车产量	10012	795000	197281.19	188323.74
新能源汽车销量	11828	814000	193286.21	185523.60
新能源汽车保有量	25698	21010057	6751372.44	5305978.05
PM2.5 排放	15.68	107.58	41.22	17.62
PM10 排放	31.77	156.14	73.38	25.41
SO ₂ 排放	7.13	62.94	16.95	10.63
CO 排放	0.53	1.74	0.89	0.27
NO ₂ 排放	13.73	50.32	27.88	8.04
O ₃ 排放	47.25	131.19	90.26	22.27

结果显示，新能源汽车产量、销量、保有量的均值分别为 197281.19 辆、193286.21 辆和 6751372.44 辆，标准差均处于较高水平，样本间离散程度较大；大气污染物中 NO₂ 排放均值为 27.88，范围在 13.73—50.32 之间，PM2.5、PM10、SO₂、CO、O₃ 排放均值分别为 41.22、73.38、16.95、0.89、90.26，整体排放水平波动相对平稳。

（二）耦合协调模型结果

1. 协调度时间演变分析

新能源汽车与生态环境两系统间的耦合协调关系呈现显著的规律性特征。散点整体沿一条向右上方倾斜的 Loess 拟合曲线分布，清晰地揭示出耦合度（C）与协调度（D）之间存在稳定的正向关联。这意味着，在观测期内（2014—2022 年），两系统间相互作用的耦合度越高，其协同发展的有序性与协调度也倾向于更高。

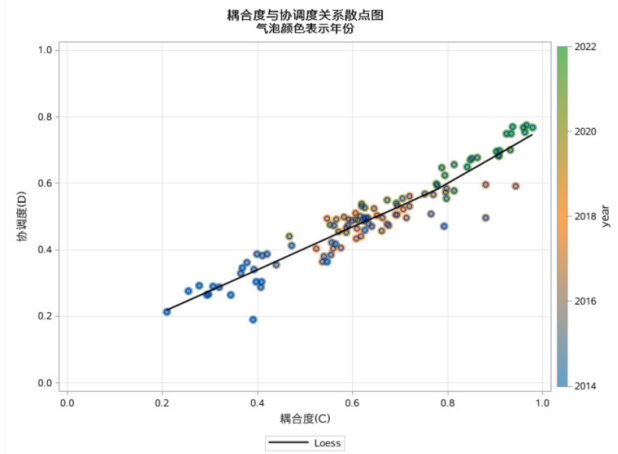


图 2：耦合度与协调度关系散点图

2. 协调等级分析

在全部样本中，协调等级呈现“中间大、两头小”的纺锤形分布。处于“良好协调”与“初步协调”等级的样本数量最为集中，合计占比超过 71%，表明我国多数地区的新能源汽车发展与生态环境治理已初步建立起良性互动关系，迈入了协调发展的轨道。然而，代表最高协同水平的“优质协调”样本仅有 7 个，占比不足 6.5%，凸显出当前协调发展的质量仍有较大提升空间，系统间的深度融合与高水平均衡尚未成为普遍现象。同时，仍有部分的样本处于“轻度失调”及“严重失调”状态，尤其是存在 1 个“严重失调”样本，这表明在部分区域或特定时期，两系统的发展仍存在显著冲突或脱节现象。

3. 驱动因素重要性分析

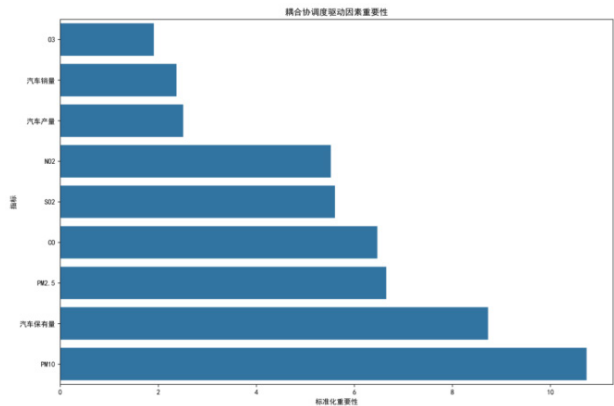


图3：耦合协调度驱动因素重要性排序

从图3中可以看出不同指标对耦合协调度的驱动作用存在显著差异，具体分析如下：

(1) PM10、PM2.5、CO、SO₂、NO₂、O₃总体来说对于耦合协调度的驱动具有重要作用。它们都是衡量空气污染的重要指标，与生态环境系统具有直接的关系。政府应当制定更加全面的制度，加大对造成污染气体、污染物企业的监管力度，加大各地政府的执行力度。

(2) 汽车保有量对于耦合协调度的驱动具有重要作用。汽车的保有量在一定程度上反映了市场上公众的接受程度，接受度越高表明公众更加倾向于选择新能源汽车而不是油车，在很大程度上改善生态环境系统。同时汽车产量、销量也与保有量具有相关性。

(三) 随机森林指标构建

在探索新能源汽车保有量的预测模型时，参照党的十八届五中全会提出的创新、协调、绿色、开放、共享的新发展理念，以及刘耀迪（2023）^[18]纳入空气质量、能源等因素对新能源汽车建模，本文选取了我国2014—2022年空气质量、清洁能源、经济GDP和科技专利这四类月度时间序列关键指标，来源于CSMAR数据库、国家统计局。

(四) 随机森林模型结果

模型优化采用网格搜索（Grid Search）与5折交叉验证^[19]，重点调整了决策树数量（n_estimators）和单棵树的最大深度（max_depth）两类参数。特征重要性通过袋外数据（OOB）计算，衡量各变量在分裂节点时对MSE的减少贡献。随机森林的优势在于可以处理高维数据和非线性关系、能够降低共线性干扰。

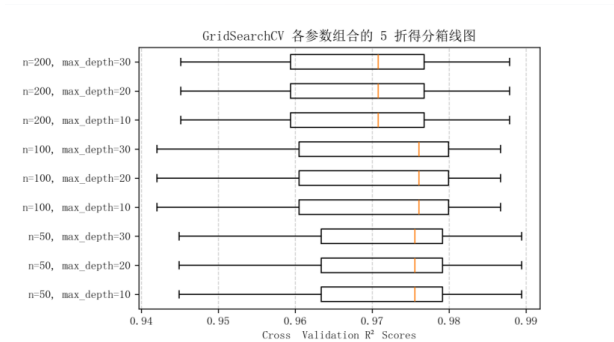


图4：五折得分箱线图

对每个参数组合的5折得分做箱线图，图4展示各折之间的波动情况。从箱线图可以看到，n=50，max_depth=10的组合不仅拥有最高的中位（约0.99），其IQR也最小（约0.945），说明在各折验证中的表现既好又稳定；相比之下，n=200，max_depth=30的组合中位最小，且波动范围更大，显现出较差的鲁棒性。

训练后本文网络搜索最终确定最佳参数为：

$$\{n_estimators:50,max_depth:10\}.$$

使用最佳参数训练模型后，判断训练集结果没有过拟合或欠拟合后，本文用Python中的predict函数进行预测得到随机森林结果，对测试集预测与真实值对比曲线如下图5。

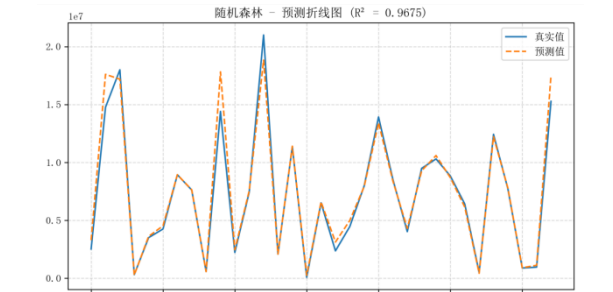


图5：随机森林测试集预测与真实值对比曲线

表3：模型检验

模型	R^2	MSE	MAE
随机森林	0.97	981635972704.18	516793.39

由表3可知，随机森林该模型取得最高的 R^2 （0.96753），表明其在捕捉自变量非线性关系及变量间交互作用方面表现较好。

本文参考Lundberg等（2020）提出的Tree SHAP算法^[20]，来解释特征值重要性。通过计算满足局部精度、缺失性与一致性的Shapley值：

$$\phi_i(f, x) = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(M-|S|-1)!}{M!} [f_x(S \cup \{i\}) - f_x(S)] \quad (12)$$

M=14为特征总数，S为特征子集， f_x 为模型预测函数。通过平均绝对SHAP值量化各变量对新能源汽车保有量的全局影响强度。

根据图6的SHAP分析结果，特征重要性排序显示保有量滞后一期指标具有显著的主导地位。其次，能源供给指标（太阳能/核能发电量）对预测结果相对较为重要，其高值区间与汽车保有量增长呈正相关。宏观经济指标（国民总收入、人均GDP）作为次要因子，呈现线性驱动趋势，而绿色技术专利指标出现“申请量>授权量”的反常影响层级，或反映政策申报周期与市场转化延迟。环境污染物（PM2.5、O₃）与火力、水力发电量处于特征权重末端，表明当前模型更加侧重清洁能源转型与经济增长的协同效应，对传统污染因子的敏感性较低。

为了进一步探究各特征对模型预测结果的具体影响方向及分布情况，我们绘制了SHAP蜂群图（图7）。图中每个点代表一个样本在该特征上的SHAP值，横轴表示SHAP值的大小，颜色则反映了特征值的高低。

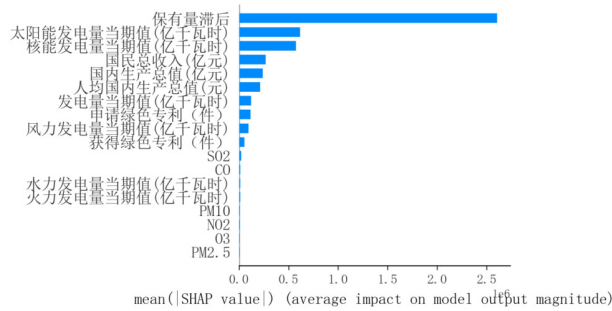


图6：特征重要性

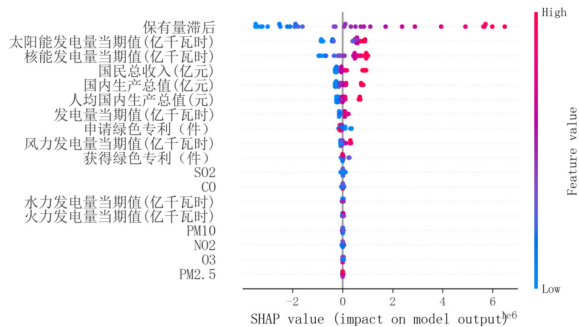


图7：SHAP 蜂群图

根据 SHAP 分析结果，社会驱动因素可划分为三个层级：

- (1) 滞后保有量 - 能源供给主导层 ($\text{SHAP} \geq 0.5 \times 10^6$)。滞后保有量与太阳能发电量、核能发电量构成核心驱动因素，其贡献度的占比达总解释力的 48.5%。
- (2) 经济 - 技术协同层 ($0.25 \times 10^6 \leq \text{SHAP} < 0.5 \times 10^6$)。国民总收入 (0.4×10^6) 与人均 GDP (0.3×10^6)、申请绿色专利 (0.25×10^6) 呈现显著经济阈值效应。
- (3) 环境边缘层 ($\text{SHAP} < 0.25 \times 10^6$)。空气污染物 ($\text{PM}_{2.5} = 0.2 \times 10^6$, $\text{PM}_{10} = 0.18 \times 10^6$) 的影响强度不足经济因素的 50%。其正向作用暗示，地方政府可能将污染治理压力转化为新能源汽车推广政策，而非直接引导消费者决策。

参考文献

- [1] Xu Z H. Coupling Coordination Development and Driving Factors of New Energy Vehicles and Ecological Environment in China[J]. Wuhan University Journal of Natural Sciences, 2025, 30(1): 79–90.
- [2] 潘苏楠, 李北伟, 聂洪光. 我国新能源汽车产业可持续发展综合评价及制约因素分析——基于创新生态系统视角[J]. 科技管理研究, 2019, 39(22): 41–47.
- [3] 张锦程, 方卫华. 政策变迁视角下创新生态系统演化研究——以新能源汽车产业为例[J]. 科技管理研究, 2022, 42(11): 173–182.
- [4] 王达, 周映雪. 随机森林模型在宏观审慎监管中的应用——基于 18 个国家数据的实证研究[J]. 国际金融研究, 2020, (11): 45–54.
- [5] Joseph A. Shapley regressions: A framework for statistical inference on machine learning models[J]. SSRN Electronic Journal, 2019.
- [6] Bracke P, Datta A, Jung C, et al. Machine learning explainability in finance: an application to default risk analysis[J]. Bank of England, 2019.
- [7] 陈小昆, 左航旭, 廖彬, 等. 融合 XGBoost 与 SHAP 的冠心病预测及其特征分析模型[J]. 计算机应用研究, 2022, 39(06): 1796–1804.
- [8] 罗妍, 王枫, 叶文玲. 基于 XGBoost 和 SHAP 的急性肾损伤可解释预测模型[J]. 电子与信息学报, 2022, 44(01): 27–38.
- [9] Vefie L. The penguin dictionary of Physics[J]. Beijing: Foreign Language Press, 1996: 92–93.
- [10] 吴跃明, 张翼, 王勤耕, 等. 论环境 - 经济系统协调度[J]. 环境污染与防治, 2001(1): 25–29.
- [11] 李虹, 袁颖超, 王娜. 区域绿色金融与生态环境耦合协调发展评价[J]. 统计与决策, 2019, 35(08): 161–164.
- [12] 王奕森, 夏树涛. 集成学习之随机森林算法综述[J]. 信息通信技术, 2018, 12(01): 49–55.
- [13] 徐慧, 蔡慧, 张伟, 等. 国家科技安全风险指标体系和模型的构建[J]. 信息科学, 2023, 41(12): 165–173+182.
- [14] 朱喜安, 魏国栋. 熵值法中无量纲化方法优良标准的探讨[J]. 统计与决策, 2015, (02): 12–15.
- [15] 王帮俊, 崔林玉, 王智乐, 等. 绿色金融试点政策对新能源企业技术创新的影响研究[J]. 南京财经大学学报, 2024, (02): 67–77.
- [16] Breiman L. Random forests[J]. Machine learning, 2001, 45: 5–32.
- [17] 方匡南, 吴见彬, 朱建平, 等. 随机森林方法研究综述[J]. 统计与信息论坛, 2011, 26(03): 32–38.
- [18] 刘耀迪, 肖钰惠, 展晶晶. 基于灰色预测模型的长三角地区新能源汽车的市场保有量预测研究[J]. 内燃机与配件, 2023(12): 107–109.
- [19] 董倩, 孙娜娜, 李伟. 基于网络搜索数据的房地产价格预测[J]. 统计研究, 2014, 31(10): 81–88.
- [20] Lundberg S M, Erion G G, Lee S I. Consistent individualized feature attribution for tree ensembles[J]. Statistics, 2018.

四、结束语

本研究通过对新能源汽车与生态环境的耦合协调发展、相关驱动因素及其可持续发展路径进行综合探究，得出如下核心结论：

一方面，新能源汽车系统、生态环境系统及其两者的耦合协调水平在考察期内呈现出持续改善的总体趋势。研究发现，在众多影响耦合协调度的指标中，PM10、新能源汽车保有量及 PM2.5 的驱动效应最为显著，同时汽车产销量、PM2.5 与 PM10 在各自系统内所占权重较高，凸显了产业发展与环境质量关键因子在二者协同演进中的重要作用。

另一方面，对新能源汽车可持续发展的预测研究表明，滞后保有量、太阳能与核能发电量是最具影响力的前瞻性特征，这证实了清洁能源基础设施的完善是产业长远发展的基石。同时，宏观经济水平与绿色专利活动也是关键的驱动力量，表明居民购买力、整体市场规模以及技术创新能力共同构成拉动市场增长、提升产品竞争力与市场信心的重要支撑。

基于上述结论，本研究从环境、能源、经济与技术四个维度提出以下政策建议：第一，应加强对重点地区传统燃油车的排放监管，并借助科普宣传提升公众环保意识与社会接受度；第二，需在清洁能源富集地区优先建设“绿电直供”充电网络，并通过电价机制激励清洁电力消纳；第三，针对经济发达地区，可优化“购置税减免 - 充电补贴”等组合型经济激励，以充分释放消费潜力；第四，应持续深化产学研合作，加大对技术创新尤其是中小企业关键攻关的支持力度，通过结构性改革推动全产业链的效能提升与成本优化。

综上所述，推动新能源汽车产业的健康发展，需要构建一个涵盖严格环境约束、稳固清洁能源底座、有效经济激励与持续技术创新在内的协同政策体系。未来研究可进一步探索多区域差异化政策的动态效应，以及智能网联、车网互动等新兴技术对“人 - 车 - 环境”系统深度融合的深远影响。

中国式全域旅游现代化指标评价体系构建研究 ——以江苏为例

刘广应*, 马国婷, 林金官

南京审计大学 统计与数据科学学院, 江苏 南京 211815

DOI:10.61369/ASDS.2026040009

摘 要 : 旅游业现代化是中国式现代化的一个重要组成部分。分析国家“十四五”旅游业发展规划指导方针与中国式全域旅游现代化的理论内涵, 将中国式全域旅游现代化分为四个一级指标: 需求多元化、供给品质化、区域协调化和成果共享化。将游客主观评价指标纳入到中国式全域旅游现代化指标评价体系。利用主成分分析方法, 对指标体系进行分类赋权, 构造了中国式全域旅游现代化评价指标。通过该指标体系对江苏省2000年至2022年数据进行实证分析。研究发现, 品质供给化指数和区域协调化指数得分相对需求多元化指数、成果共享化指数得分较高, 表明江苏省旅游供给充分, 全域旅游协调较好; 需求多元指数化和成果共享化指数, 在疫情期间得分有所下降, 相对得分偏低, 表明旅游需求和成果共享还有待进一步挖掘和提升。研究丰富了构建中国式全域旅游现代化指标体系的理论和方法, 为找准全域旅游现代化着力点并优化相关政策提供重要借鉴。

关 键 词 : 全域旅游现代化; 主观满意度; 中国式现代化; 文本数据; 网络爬虫

Research on the Construction of a Chinese-style Modernization Indicator Evaluation System for All-for-One Tourism — A Case Study of Jiangsu Province

Liu Guangying*, Ma Guoting, Lin Jinguan

School of Statistics and Data Science, Nanjing Audit University, Nanjing, Jiangsu 211815

Abstract : Tourism modernization is a crucial component of Chinese-style modernization. By examining the guiding principles of the national "14th Five-Year Plan" for tourism development and the theoretical connotations of Chinese-style all-for-one tourism modernization, this study proposes a framework that divides Chinese-style all-for-one tourism modernization into four first-level indicators: diversified demand, quality supply, coordinated regional development, and shared outcomes. To construct a comprehensive evaluation system for Chinese-style all-for-one tourism modernization, this study employs web-crawling technology to collect tourist review data and utilizes text analysis methods to extract high-quality subjective evaluation data from tourists. These subjective evaluation indicators are then integrated into the evaluation system for Chinese-style all-for-one tourism modernization. Principal component analysis is applied to classify and weight the indicators, thereby constructing an evaluation index for Chinese-style all-for-one tourism modernization. An empirical analysis was conducted using this index system on data from Jiangsu Province from 2000 to 2022. The results indicate that the indices for quality supply and coordinated regional development scored relatively higher than those for diversified demand and shared outcomes. This suggests that Jiangsu Province has a sufficient tourism supply and that all-for-one tourism is well-coordinated. The indices for diversified demand and shared outcomes experienced a decline during the pandemic period and scored relatively lower, indicating that there is room for further exploration and enhancement of tourism demand and shared outcomes. This study enriches the theoretical and methodological approaches for constructing an evaluation system for Chinese-style all-for-one tourism modernization and provides important references for identifying the key points of all-for-one tourism modernization and optimizing relevant policies.

Keywords : all-for-one tourism modernization; subjective satisfaction; Chinese-style modernization; textual data; web crawling

基金项目: 国家社会科学基金重大项目“中国式现代化的统计监测评价问题研究”(23&ZD036); 江苏省高等学校自然科学研究重大项目“复杂结构金融数据的统计建模及其在金融风险应用”(25KJA110002)。

作者简介:

马国婷, 硕士研究生, 研究方向为经济社会统计与金融统计;

林金官, 教授, 博士生导师, 研究方向为统计建模与经济社会统计。

通讯作者: 刘广应, 教授, 博士生导师, 研究方向为统计建模与金融统计。

引言

2015年5月,国家旅游局首次提出“全域旅游”这个新概念。从此,中国旅游业进入了一个全新的发展阶段。相较于传统旅游模式,全域旅游的一个突出点在于整个区域旅游资源的统一。这一概念提出后,更能反映我国旅游的中国特色,为旅游业可持续发展创造了全新的环境。“十三五”期间,作为满足人民对美好生活向往的重要产业,旅游业得到了极大发展。同时,旅游业在这段时间内巩固了自身作为经济战略性支柱产业的地位,也成为了践行“绿水青山就是金山银山”生态文明建设的重要行业。在“十四五”时期,中国特色社会主义现代化建设新征程中,国家提出了到2035年要将旅游业发展成为经济社会发展、提升国人幸福生活以及宣传中国文化的核心和重要手段,构建符合国情和中国新时代特点的全域旅游现代化指标评价体系已到了刻不容缓的程度。构建中国式全域旅游现代化评价体系既可以对区域内全域旅游发展现状以及潜在发展状况作出客观评价,同时,还可以从对策分析的角度深入剖析,从而更好地作出有说服力、有指导意义的对策支撑,为全域旅游的不断发展提供动力源泉。本文主要探究和构建符合中国国情的全域旅游现代化指标评价体系。

一、文献综述

（一）旅游领域中国式现代化的理论内涵

中国特色社会主义现代化模式即中国式现代化,成为广大学者研究的新热点,广大学者就中国式现代化的理论意蕴进行了相关的研究,论述了中国特色社会主义现代化的特殊性和作用性。

中国式现代化主要有五个特征,在旅游领域也有其具体表现:(1)人口规模巨大。中国人口规模巨大基本国情决定了中国式现代化的艰巨性和复杂性。中国式现代化要注重处理好人口规模巨大现状下经济和社会稳定等关系,不能照搬西方小国寡民式现代化模式^[1]。中国式现代化在旅游市场体现了中国潜在巨大的旅游市场需求,旅游市场规模巨大,旅游产业发展将促使人民大众生活水平的提升,促进人民消费理念转变,旅游业的发展可以更好地提供更多的就业机会,给社会带来更多经济收益,促进社会健康发展。(2)全体人民共同富裕。中国式现代化以实现全体人民的共同富裕为本质目标,与西方少数人富裕的现代化有本质区别^[2]。共同富裕贯穿于中国式现代化的全过程,是中国式现代化区别于其他现代化模式的重要特征^[3]。旅游业的发展将推动自然环境优美但工业欠发达地区的经济发展,提高该地区旅游等相关行业从业人员的收入水平,促进地区发展,推进共同富裕。(3)物质文明和精神文明相协调。中国式现代化不仅讲究物质财富增长、而且讲究物质文明、精神文明全面发展。为防止出现物质富强但精神贫瘠的中国式现代化,强调精神物质共同建设对于理解中国式现代化发展具有重要意义^[4]。旅游开阔人们的视野、丰富人们物质文明,更重要的是有利于人类心灵的治愈,使得精神上得到放松和愉悦,推进旅游领域的中国式现代化,对于推进人类物质文明和精神文明建设具有重要意义。(4)人与自然和谐共生。中国式现代化摒弃了西方“先污染、后治理”的发展模式,强调生态文明建设与经济社会发展的同步推进。中国式现代化必须坚持绿色发展理念,实现经济发展与生态保护的良性互动^[5、6]。旅游业天然以绿色为背景,发展旅游业与自然美景和生态平衡密不可分。发展旅游业,注重绿色生态保护,将促进人与自然的和谐共生。(5)走和平发展道路。中国式现代化打破了国强必霸的惯常大国

崛起路径,走和平发展道路。中国式现代化通过参与全球治理、推动构建人类命运共同体等为世界和平与发展作出贡献^[7]。旅游业为不同国家人民提供了一个交流平台,推动中国旅游业的发展,在促进各国之间的文化交流、使不同国家能够互相包容,推动世界和平与发展的过程将发挥重要作用。

（二）全域旅游的理论内涵

1. 全域旅游的基本内涵

全域旅游是一种以旅游业为引领,推动区域经济发​​展的理念与模式。它将旅游业作为整体产业的重要组成部分,紧密融入社会经济发展的全局。从全方位、全角度出发,优化资源配置,实现区域内旅游资源的合理整合、产业间的融合发展以及社会大众的共建共享。通过将旅游业打造为经济发展的重要引擎,以旅游资源富集地区为目的,从空间和产业层面优化生产要素的配置,提高其利用效率,从而以旅游产业为统领,持续提升区域竞争力,实现区域协调发展^[8]。全域旅游也指以旅游产业为主导或者引导在旅游资源富集地区的空间和产业上实现对旅游资源进行合理高效的配置要素生产要素的一种过程,充分利用旅游业统筹区域经济发展,持续提升区域竞争力^[9]。全域旅游超越了以传统旅游和小旅游为范畴的内容,将一个地区整体功能作为一个完整有目的地进行开发建设的一种旅游发展理念和模式^[10]。

2. 全域旅游的实践意义

相对于传统旅游发展模式中的分散、孤立、割裂、条块化、僵化,全域旅游注重全局,具有重要的战略性作用与应用价值。主要体现在:(1)全域旅游注重推动旅游产业与其他产业融合发展,提升旅游产业的产业综合竞争力。(2)全域旅游的发展,可以促进区域经济协调发展,推动新型城镇化与乡村振兴战略的落实,推动全域旅游可持续发展;(3)全域旅游不仅对旅游精准扶贫有着重大意义,而且有助于推动贫困区经济社会的全面发展^[11、12]。(4)全域旅游除了旅游发展方式的转变,更是区域经济社会发展方式的转变,推动区域经济协调发展、旅游产业综合竞争力提升和社会文明程度提升^[10]。

（三）全域旅游评价指标与测度方法

对于旅游领域的评价,学术界主要从以下几个方面展开了相

关研究。在全域旅游未提出之前，国内对地区旅游的评价测度主要是围绕旅游可持续发展、新发展理念、旅游竞争力等方面展开。例如，崔凤军从生态环境类指标、旅游经济类指标、社会文化类指标以及旅游支持系统指标出发，提出了区域旅游可持续发展评价指标体系^[13]；Ke采用层次分析法构造旅游可持续发展评价体系^[14]。十八届五中全会后，习近平总书记提出了创新、协调、绿色、开放、共享的新发展理念后，出现了一些文献从新发展理念角度构建旅游评价指标。吕雁琴等基于创新、协调、绿色、开放、共享五大发展理念构建了省域旅游业发展水平综合评价指标体系^[15]。孙晓等构建了包括创新、协调、绿色、开放、共享5个维度的旅游产业高质量发展指标体系^[16]。魏谨瑞和崔浩萌基于网络搜索数据，构造了区域旅游指数^[17]。此外，还有一些文献讨论旅游与其他产业的融合问题。常家玲等构造了中国省域乡村旅游综合评价指标体系，并分析了城乡融合与乡村旅游之间的耦合协调关系^[18]。

上述文献主要反映了区域的旅游发展情况，对地区综合全域旅游的反映存在一定局限性。在全域旅游的概念提出后，学者开始对（全域）旅游进入更深层次的探索。丰晓旭等在总结分析全域旅游内涵的基础上，从产业域、空间域和管理域三个维度选取一些有代表性的评价指标，采用全局主成分分析和熵值法相结合的方式，对我国31个省（市、区）的全域旅游发展水平进行了试评价^[19]。王旭科等从生态环境、公共服务、发展效益、旅游竞争力、融合发展及保障体系等维度构建了全域旅游发展水平评价指标体系，使用聚类模型及因子分析对指标设置权重^[20]。罗新颖从全域旅游推力因素、全域旅游引力因素以及全域旅游支持因素角度出发，使用主成分分析技术对山东省全域旅游发展水平进行评价研究^[21]。

（四）综合评述与问题的提出

上述文献对对地区旅游发展评价体系进行了构建，对全域旅游的评价也进行了初步研究，取得了一定的进展，未见文献对全域旅游从中国式现代化视角进行综合评价。一方面，现有文献在构建旅游评价体系中，主要依据客观数据，没有充分利用游客评价等主观数据，少数文献采用了主观数据，主要是专家评审以及问卷调查等数据^[22]，这类数据一般数据量不大，也很难充分客观全面反映真实的情况。而中国式现代化的物质文明和精神文明相协调的本质，要求指数评价需要综合考虑人民是否满意等主观数据。此外，主观数据更能体现中国式现代化人民满意的程度。另一方面，现有旅游评价指标体系未能与中国特色充分结合^[2、4、23]。传统构建的指标评价体系，不能很好地反映中国式全域旅游现代化的发展程度。

因此，本文在现有研究的基础上，采用客观数据与主观数据相结合方法，对全域旅游从中国式现代化视角进行综合评价研究。本文按照“十四五”旅游业发展规划将中国式全域旅游现代化发展体系分为四个一级指标，尝试构建一套新的符合中国特色的全域旅游现代化评价体系。利用爬取游客评价文本大数据等主观数据，对其进行文本分析，获得高质量主观数据，克服传统主观数据收集中的回收信效度低、数量少等问题^[24]，并将该指标数

据拉入综合评价指数体系。

二、中国式全域旅游现代化的统计含义及评价体系构建

（一）中国式全域旅游现代化的理论内涵

中国式全域旅游现代化应当既包含中国式现代化的特征，也具备全域旅游的内涵^[24]。本文突出全域旅游和中国特色，将其理论内涵界定为区域旅游业作为整体并协同经济可持续发展与中国式现代化特征的良性互动与结合。中国式全域旅游现代化主要体现在以下四个方面。第一，基于人口规模的需求多元性。人口规模巨大决定了旅游市场的潜力，也反映了消费者需求的多元性。第二，共同富裕导向下的区域协调性。借助全域旅游促进区域内不同子单元的协调发展，缩小贫富差距，实现“先富带后富”的富民效应。第三，物质精神协调的供给品质性。打造承载物质文明与精神文明兼备的旅游产品，创造经济价值，也能够传承中华优秀传统文化，富裕游客精神水平。第四，人与自然和谐共生的成果共享性。坚持生态优先、绿色优先的发展理念，确保实现绿水青山的同时，让人民共享发展成果。

表1：中国式全域旅游现代化评价体系

一级指标	二级指标	三级指标	正负性
需求多元化	客源结构	国际接待人数（人，X1）	+
		人均访区域旅游次数（次，X2）	+
	游客行为	区域国内游客人均支出（元，X3）	+
	需求规模	旅客铁路周转量（亿人公里，X4）	+
	市场关注度	日均5A 景区百度搜索次数（次，X5）	+
供给品质化	生态环境质量	绿化覆盖率（%，X6）	+
		人均公园绿地面积（m ² ，X7）	+
	旅游产品质量	游客评价（分，X8）	+
	基础设施保障	每万人拥有公厕数（座，X9）	+
		每万人拥有公交数量（辆，X10）	+
区域协调化	经济协调	人均旅游收入基尼系数（X11）	-
	设施均衡	单位旅游星级酒店数量基尼系数（X12）	-
	环境协调	PM2.5地表浓度基尼系数（X13）	-
	资源均衡	单位5A 级景区个数基尼系数（X14）	-
成果共享化	公共服务共享	每万人拥有医师数（人，X15）	+
		每千万人有用博物馆数（座，X16）	+
	民生与就业共享	年末城镇登记失业率（%，X17）	-
	生态共享	污水处理率（%，X18）	+
	经济与收入共享	旅游收入占 GDP 比重（%，X19）	+
	创新成果共享	研究与试验发展经费支出占地区生产总值比重（%，X20）	+

（二）中国式全域旅游现代化的评价体系

《国务院关于印发“十四五”旅游业发展规划的通知》为我

国“十四五”时期的旅游发展指明了目标的方向。规划中指出：“到2035年，旅游需求大众化、旅游供给品质化、旅游发展协调化、旅游成果共享化特征更加凸显，优质旅游供给更加丰富，旅游业基本建成世界旅游强国，为社会主义现代化建设作出积极作用”。基于对我国全域旅游现代化基本内涵以及统计意义上内涵的理解，并结合我国旅游发展现状，将发展规划中的发展目标转化为评价体系的一级指标，使评价体系具有科学性与政策导向性特征。通过评价指标体系来反映我国全域旅游发展是否满足国家战略的要求，为政策制定与调整提供科学依据。具体的指标体系见表1。

1. 需求多元化指数

基于中国式特征，本文将需求多元化指数构建工作分客源构成、游客行为、需求规模以及市场关注度这四个方面分层构建。以国际接待游客量与国内人均访区域旅游次数反映客源构成二级指标，通过国际接待游客量反映国际化游客的规模，体现区域旅游对国际市场的吸引力；以国内人均访区域旅游次数为指标，衡量国内游客对区域旅游的重复消费意愿和忠诚度，反映国内游客的需求差异性。以区域国内游客人均消费额，刻画游客在消费中出行、住宿、购物等消费能力，体现旅游需求层次的游客行为二级指标。以旅客铁路运输总量来体现旅游需求在铁路运量运输上的需求规模二级指标。市场关注度由日5A景物游客百度搜索量作为其表征特征，用于度量优质资源的市场价值吸引与信息知晓度。通过各层次指标耦合于数，反映旅游市场需求的多元化发展规模大小。

2. 供给品质化指数

供给品质化指数着眼于旅游产业供给侧，本文从生态环境质量、旅游产品质量和基础设施保障三个方面确定评价指标。生态环境质量方面选择绿化覆盖率和人均公园绿地面积作为指标测度自然生态环境的供给；游客满意度能够直观体现旅游产品的质量状况，综合评价住宿、景区、旅游服务等各个环节，因此利用游客评价作为三级指标来度量旅游产品质量二级指标；每万人口拥有的公共厕所数量和公交车数量用于测算基础公共服务设施和交通的保障能力。

3. 区域协调化指数

协调化指数更多面向旅游协调发展这一指标，本文从经济协调、设施均衡、环境协调和资源均衡四个角度度量区域协调化一级指标。选取人均旅游收入基尼系数、单位旅游星级饭店数量基尼系数、PM2.5地表浓度基尼系数及单位5A级景区个数基尼系数4个三级指标测度区域协调化。通过人均旅游收入基尼系数测算区域旅游经济收益平衡性，实现旅游红利的普惠共享；基于单位旅游星级饭店数量基尼系数测度反映高等级住宿资源分布的均衡性，防范旅游服务设施二元化发展风险；通过PM2.5地表浓度基尼系数测度区域空气质量的不均衡性，实现旅游全域均质生态环境打造；基于单位5A级景区个数基尼系数测算优质旅游资源空间分布均衡性。该指数从多均衡性指标维度揭示区域旅游结构性不平衡。

4. 成果共享化指数

成果共享化指数遵循“人民共享发展成果”的核心理念，从公共服务、民生就业、生态治理、经济红利、创新驱动五个维度

构建评价体系。公共服务共享以每万人拥有医师数与博物馆数为指标，衡量医疗与文化资源的普惠程度，体现居民与游客共同享用的服务质量；民生与就业共享以年末城镇登记失业率为参照，反映旅游产业对就业的拉动效应；生态共享以污水处理率为核心指标，测度生态治理成效，保障环境红利的全民共享；经济与收入共享以旅游收入占GDP比重为依据，衡量旅游产业的经济贡献度，体现发展成果的全民共享属性；把研究与试验发展经费支出占GDP比重作为创新成果共享。

（三）权重和目标值的确定

本文采用主成分分析法直接对全域旅游现代化指标评价体系三级指标进行处理，对其进行赋权。主成分分析方法的核心是通过数据的线性变换，找到使数据方差最大化的坐标轴^[21]。

为测度区域旅游现代化程度，需对各指标的现代化目标值做出测定。本文目标值确定遵循以下3个原则。第一，中国式特点体现。以北京市2022年（人均GDP达中等发达国家之际）实现值，而非其他资本主义国家实现值作为目标值，更加符合中国社会主义国情以及中国式特征。第二，客观性度量。国际之间的数据由相似国情国家的数值做度量：国际接待人数以国际为基准。德日两国与中国有相似的社会结构以及紧密的经济合作交流，因此以其国际接待人数作为目标值。第三，科学性原则。所有目标值均依照科学的方法进行确定。在确定所有指标的权重及其对应目标值后，结果如表2所示。

表2：中国式全域旅游现代化指数各指标权重及其目标值

三级指标	目标值	权重	三级指标	目标值	权重
X1	89.4	3.92%	X11	0.8	5.76%
X2	0.51	1.74%	X12	0.8	5.83%
X3	1945.19	4.70%	X13	0.8	5.22%
X4	875.27	5.87%	X14	0.8	5.98%
X5	2.07	4.74%	X15	36	5.82%
X6	41	1.90%	X16	46.85	5.91%
X7	16	6.01%	X17	1.08	5.68%
X8	4.8	4.86%	X18	100	5.68%
X9	4.75	3.05%	X19	17.58	5.80%
X10	19	5.62%	X20	4.34	8.91%

（四）中国式全域旅游现代化发展阶段进程测度

本文采取测度不同指标的实现程度来测度历年中国式全域旅游现代化发展程度。在测度不同的指标时，由于存在的指标并非全都是极大型指标，因此，本文需对所有的指标进行极大化处理。对其进行测度存在以下两种情况。

1. 极大型或极小型指标的测度分数。首先对极小型指标进行极大化，设 M_i 为指标向量 X_i 的实际值，记 $M_{i*} = \max\{M_i\} - M_i$ ，将 M_{i*} 的值重新赋给 M_i ，再对其发展阶段（分数）进行测算。设 T_i 为指标向量极大化后的目标值，则指标向量的测度分数 S_i ：

$$S_i = \begin{cases} \frac{M_i}{T_i} \times 100, & \text{若 } \frac{M_i}{T_i} < 1 \\ 100, & \text{若 } \frac{M_i}{T_i} \geq 1 \end{cases}$$

2. 中间型指标的测度分数。设 X_i 为中间型指标, 若 X_i 存在适度区间 $[n_1, n_2]$, 其上下界限为 $[m_1, m_2]$, 那么指标向量的测度分数 S_i :

$$S_i = \begin{cases} 0, & \text{若 } M_i \notin [m_1, m_2] \\ \frac{1}{(n_1 - m_1)^2} M_i^2 - \frac{2n_1}{(n_1 - m_1)^2} M_i + \frac{m_1^2 - 2n_1 m_1}{(n_1 - m_1)^2} \times 100, & \text{若 } M_i \in [m_1, n_1] \\ 100\%, & \text{若 } M_i \in [n_1, n_2] \\ \frac{1}{(m_2 - n_2)^2} M_i^2 - \frac{2n_2}{(m_2 - n_2)^2} M_i + \frac{m_2^2 - 2n_2 m_2}{(m_2 - n_2)^2} \times 100, & \text{若 } M_i \in [n_2, m_2] \end{cases}$$

三、江苏省中国式全域旅游现代化指数的测度与分析

(一) 数据来源与数据处理

本文收集了江苏省2000年至2022年共23年数据。除百度搜索指数、游客评分等特色数据外, 其他数据来源于国家统计局年鉴与CEIC中国经济数据库。针对缺失数据利用插值法进行填补。百度搜索指数的数据来源于百度搜索引擎历年对关键词的搜索统计, 通过对每个景点逐年记录其百度搜索指数, 反映其市场关注度。游客评分的数据则来源于去哪儿网, 基于游客对景点的历年文本评价。通过软件(Octopus)爬取游客对不同景点的文本评价数据, 并进行数据清洗处理。在运用主成分分析方法计算权重时, 为剔除疫情对旅游业的潜在影响, 选取的输入数据范围为2000年至2018年。

(二) 游客评价文本数据的建模分析

表3给出了部分游客评价文本数据示例, 也给出了本文基于Logistic模型的情感得分。针对评价文本数据我们采用如下步骤得到文本的情感得分。(1) 对网络爬取收集到的评价文本数据进行数据清洗, 去除缺少评价文本、时间或景点名称的数据, 并清除数据文本中无意义字符, 仅保留中文、英文、数字及标点。(2) 使用jieba库对存储评价文本进行分词。导入停用词表, 过滤掉常见无实际语义的停用词, 如“的”、“是”、“了”等, 以提升情感分析效果。(3) 识别出评价文本中的情感词(分为积极、消极两类), 否定词以及程度副词。遍历所有情感词, 查看当前情感词前面是否有否定词或程度副词。若无否定词, 则情感词 $\times 1$; 若存在否定词或多个否定词, 则 $\times (-1)^n$, n 为否定词个数。同样地, 若情感词前存在程度副词, 就在当前情感词前乘以程度副词的程度等级。(4) 加权计算评价得分。对当前文本数据的最终得分 y 利用如下公式计算:

$$y = \frac{\sum_{i=1}^m (-1)^{n_i} \cdot c_i}{\sum_{i=1}^m (-1)^{n_i} \cdot c_i - \sum_{j=1}^w (-1)^{n_j} \cdot c_j},$$

公式中, n_i 表示第 i 个情感词前面的否定词个数, c_i 表示第 i 个情感词前面程度副词的程度等级。不难看出, 公式分子为积极情感词的总得分, 分母为积极情感词与消极情感词的得分之和。将 y 代入 Logistic 函数中, 设置阈值进行分类, 得到最终得分 S :

$$s = \begin{cases} 1, & \text{Logistic}(y) < 0 \\ 3, & \text{Logistic}(y) = 0 \\ 5, & \text{Logistic}(y) > 0 \end{cases}$$

由于景点太多, 本文选取了所有5A级景区作为代表, 江苏省共有26家5A级景区。按照如上方法, 对江苏省26家5A级景区的23969条游客评价数据进行清洗整合后, 共获得22762条有效样本数据, 数据涵盖2014年至2022年。对文本利用上述每条评论得分, 最后对每个景区的情感得分进行平均得到该景区游客评分。

表3: 游客评价数据与模型评分结果示例

年份	景点名称	评价内容	得分
2015	中华恐龙园	简运期间开放的东西太少, 有点不尽兴, 可惜了。	1
2016	夫子庙秦淮风光带	环境优美, 细节完美, 水上风景如画, 值得一看, 有蚊子, 注意不要被咬到	5
2017	云龙湖	景色宜人空气清新那啥湖远不如海	3
2018	溱湖国家湿地公园	生态保持的不错, 环境很好。玩的很开心。	5
2020	中山陵景区	非常值得一去, 了解那些人那些事	5
2021	周庄	值得推荐。纯玩的人。旅游公司也很好。	5
2022	鼋头渚	夜景真的好看, 这灯光绝了绝了, 这门票值	5

(三) 江苏省中国式全域旅游现代化指数测度结果与分析

对江苏省23年中国式全域旅游现代化指数进行测度, 逐年总指数与各一级指标结果如图1与表4所示。总体看来, 2000年至2022年间, 总指数由52.61增长至82.03, 整体呈现长期上升趋势, 江苏省中国式全域旅游现代化水平持续提高。

表4: 江苏省2000-2022年历年中国式全域旅游现代化进程得分

年份	区域协调	成果共享	需求多元	供给品质	进程得分
2000	76.00	43.16	24.75	70.33	52.61
2001	76.00	42.79	26.19	63.56	51.33
2002	76.00	34.05	27.89	63.93	48.73
2003	76.00	36.86	28.80	65.02	50.13
2004	76.00	43.04	33.13	67.38	53.69
2005	76.00	55.96	36.80	71.99	59.95
2006	76.00	53.70	46.27	75.93	61.99
2007	72.90	58.27	44.79	78.20	63.05
2008	74.04	60.39	53.81	80.66	66.47
2009	73.55	62.96	55.79	80.05	67.53
2010	78.34	66.22	61.48	79.93	70.93
2011	78.82	67.82	67.19	80.19	72.85
2012	83.03	72.45	72.43	79.70	76.41
2013	83.28	75.12	66.48	81.87	76.62
2014	83.91	76.87	70.33	84.89	78.82
2015	85.66	78.27	72.25	87.54	80.68
2016	88.21	80.12	75.91	88.31	82.84
2017	87.92	81.41	80.16	89.84	84.44
2018	88.11	83.39	83.10	90.31	85.89
2019	88.18	86.16	80.31	92.06	86.66
2020	87.03	80.35	62.40	94.39	81.12
2021	89.08	83.63	66.24	94.58	83.57
2022	88.09	82.69	61.67	94.41	82.03

1. 中国式全域旅游现代化水平的总体特征

本文对江苏省中国式全域旅游现代化水平和时间进行最小二乘回归建模,得出可决系数 $R^2 = 0.9180$,拟合效果优越。图2给出了江苏省中国式全域旅游现代化水平和时间的散点图和回归拟合直线。根据拟合回归直线预测结果显示,拟合直线稳步增长,预测未来也会持续保持上升趋势,因此江苏省中国式全域旅游现代化具有一定的可持续性和成长性。结果显示,如果整个旅游市场大环境没有变化情况下,在2027年中,江苏省中国式全域旅游现代化指数值达到100,也即中国式全域旅游现代化得以实现。但考虑到2019年和2020年疫情带来的冲击,全域旅游现代化水平在2020年有所下降,若全域旅游现代化水平增长与疫情前保持相同速度,则2032年中江苏省即将实现中国式全域旅游现代化。江苏省是我国华东地区的大省之一,经济发展水平在全国名列前茅,旅游资源丰富且交通便捷,江苏省有望在中国式全域旅游现代化进程中发挥示范引领作用。

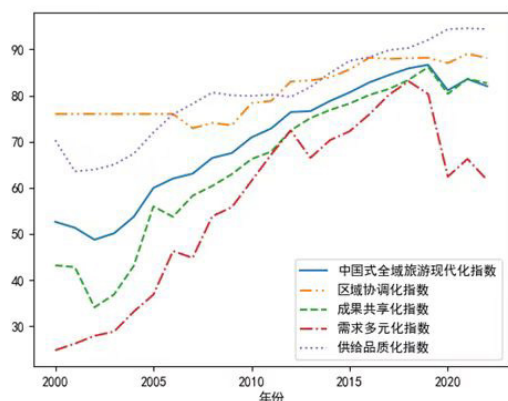


图1：2000—2022年江苏省中国式全域旅游现代化指数及其一级指标时序图

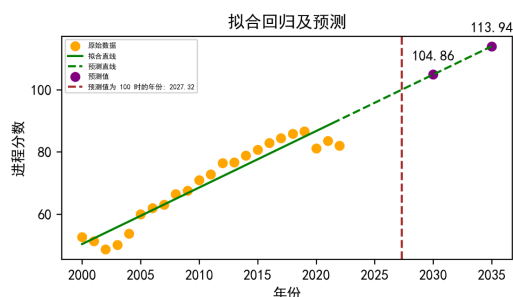


图2：江苏省中国式全域旅游现代化指数的进程分数散点图与拟合回归直线

2. 中国式全域旅游现代化水平的维度特征

本文对各个一级指标水平进行分别测算,以探究和分析中国式全域旅游现代化在发展过程中的优势和相对不足。图1包含了一级指标时序图,表4也给出了不同年份各一级指标取值结果。从图1可以明显看出,需求多元化指数折线波动明显,在2020年前呈现明显上升趋势。这归因于2020年前,随着中国式现代化发展,我国市场活力好,因此其能够随消费升级稳步增加,在全域旅游现代化方面也表现显著。需求多元化在2020年后出现一个较大的下降,主要在于国内新冠疫情期间,一方面,出行管理相对严格,另一方面,人们出于健康安全等因素,也主动减少了出行和旅

游,进而全域旅游的需求出现了较大的下降。在疫情后期,旅游需求也逐步开始复苏,市场需求得以释放和增长。可以看出,相对于区域协调化、供给品质化、成果共享化三个维度,需求多元化指数值相对较低,也表明需求多元维度未来还需奋力追赶。

区域协调化、供给品质化、成果共享化维度方面。供给品质化指数在2001年后,一直呈现增长趋势,即使在疫情期间旅游需求得到影响,该指数也在增长,表明江苏省地方政府和旅游市场主体一直在积极推动供给,推动地区旅游发展。此外,2016年全域旅游政策启动与2020年疫情倒逼,也共同推动了江苏省旅游供给的发展。区域协调指数总体相对处在较高水平,在2007年降低之后,呈现缓慢上升趋势,在疫情期间有所下降,但下降幅度很小,表明江苏省在旅游协调方面表现稳健,区域协调能力和表现也在稳步提升。成果共享化指数一直处于上升状态,同时与中国式全域旅游现代化指数的图线高度相似,这也验证了在我国中国式现代化进程中,成果全体人民共同享是一个核心主题。

3. 未来发展路径分析

为了更好更快地推动中国式现代化,推动中国式全域旅游现代化,还需从需求多元化、区域协调化、供给品质化、成果共享化等方面共同发力。

需求多元化发展路径:(1)积极推动旅游市场,满足个性化需求。随着人们生活水平的提高和消费观念的转变,旅游需求将更加多样化和个性化。全域旅游发展需进一步细分市场,针对不同年龄、性别、兴趣爱好、消费能力的游客群体,开发更多定制化、小众化的旅游产品和服务,推动中国式全域旅游现代化。(2)推动消费升级,提升旅游需求。推动旅游消费从传统的观光游览向深度体验、休闲度假、文化探索、健康养生等多元化方向发展和升级。(3)密切关注旅游市场动态,及时调整旅游产品和服务供给,注重从供给侧推动需求的多元和发展。例如,随着“银发经济”的崛起,开发适合老年人的慢节奏、康养型旅游产品;2025年由江苏省体育局和各设区的市人民政府主办的江苏省城市足球联赛(简称“苏超”),带动了举办城市的旅游、餐饮等产业,表明品质供给可以激发需求,越来越旺盛需求也进一步提升的供给的品质,形成良性互动。

区域协调化发展路径:(1)加强区域旅游规划的统筹协调,打破行政区划限制,形成优势互补、协同发展的全域旅游格局。(2)加强不同地区之间的旅游合作与交流,实现资源共享、客源互送、线路互联。例如,“苏超”联赛期间,举办比赛城市文旅部门积极实施球迷旅游优惠措施,比赛城市双方交通部门为球迷出行实施便捷出行等等,这些区域协调措施,大大激活了比赛举办地旅游市场,区域旅游一体化发展,助力全域旅游现代化。(3)促进乡村旅游与城市旅游的协同发展,通过发展乡村旅游带动农村经济发展,缩小城乡差距,推动城乡融合。同时,城市旅游要为乡村旅游提供市场、资金、技术等方面的支持,实现城乡旅游良性互动。这也保证了区域协调化指数与成果共享化指数拥有持续增加的能力。

供给品质化发展路径:(1)加强旅游从业人员培训,提高服务意识和专业技能,为游客提供更加优质、高效、贴心的服务。同时,完善旅游投诉处理机制,提高投诉结案率,提升游客满意

度。(2) 加大对旅游基础设施的投入,完善交通、住宿、餐饮、卫生等配套设施,提升旅游目的地的可进入性和舒适度。(3) 促进旅游与文化、农业、体育、康养等产业的深度融合,打造更多复合型旅游产品和服务,提升旅游供给的品质和内涵,推动产业融合创新。例如,2025年江苏创设“苏超”足球联赛,大大增加了品质供给,激活了比赛举办地城市的旅游市场。

成果共享化发展路径:(1) 通过发展全域旅游,带动更多地区和人群参与旅游产业,分享旅游发展带来的经济红利,促进区域协调发展和共同富裕。例如,推动乡村旅游发展,让农民通过参与旅游经营、提供特色农产品等方式增加收入。(2) 加强旅游公共服务设施建设,提高公共服务的供给质量和均等化水平,提升公共服务水平,让更多的游客和当地居民受益。例如,可以完善旅游交通网络、建设旅游集散中心、提升旅游信息服务水平等,提高成果共享。

四、结束语与研究结论

本文基于中国式现代化理论内涵及全域旅游的特点,结合国家“十四五”旅游业发展规划的指导方针,界定了中国式全域旅游现代化的理论内涵,并在此基础上,构建包含区域协调化、供给品质化、需求多元化、成果共享化这四个一级指标的中国式全域旅游现代化指标评价体系。为了减少传统指标体系只利用宏观经济指标数据等客观数据,很难充分描述自然人的满意度,本文爬取了游客评价文本数据,利用情感建模得出游客评价得分,将该得分纳入到中国式全域旅游现代化指标评价体系中。对2000年—2022年江苏省中国式全域旅游现代化水平进行测度。研究结

参考文献

- [1] 赵西君,郭剑锋,季小妹,等.中国式现代化指标体系构建与实证研究[J].社会科学家,2023(07):34-40+53.
- [2] 任保平,张倩.构建科学合理的中国式现代化的评价指标体系[J].学术界,2022(06):33-42.
- [3] 徐振宇,周智翔,孔新兵,等.我国共同富裕评价指标体系及测度——基于省级行政区与区域层面的探讨[J].统计研究,2024,41(03):3-17.
- [4] 马晓河,周婉冰.中国式现代化:评价指标体系构建及统计测度[J].贵州社会科学,2023(08):105-115.
- [5] 李柏文,郭凌.中国式现代化情境下的旅游高质量发展理论研究[J].旅游学刊,2024,39(01):13-15.
- [6] Hoang H T T, Truong Q H, Nguyen A T, et al. Multicriteria evaluation of tourism potential in the central highlands of vietnam: Combining geographic information system (GIS), analytic hierarchy process (AHP) and principal component analysis (PCA)[J]. Sustainability, 2018, 10(9): 3097.
- [7] 蒋永穆,李想,唐永.中国式现代化评价指标体系的构建[J].改革,2022(12):22-35.
- [8] 戴学锋.全域旅游:实现旅游引领全面深化改革的重要手段[J].旅游学刊,2016,31(09):20-22.
- [9] 杨振之.全域旅游的内涵及其发展阶段[J].旅游学刊,2016,31(12):1-3.
- [10] 张辉,岳燕祥.全域旅游的理性思考[J].旅游学刊,2016,31(09):15-17.
- [11] 郭文.旅游赋能中国式现代化发展的时代诉求与内涵诠释[J].四川师范大学学报(社会科学版),2024,51(03):31-38+200.
- [12] 李君轶,高慧君.信息化视角下的全域旅游[J].旅游学刊,2016,31(09):24-26.
- [13] 崔凤军,许峰,何佳梅.区域旅游可持续发展评价指标体系的初步研究[J].旅游学刊,1999(4):42-45.
- [14] Ke C. A Research of AHP Approach in Evaluating Sustainable Tourism Development[C]// Proceedings of the 3rd International Conference on Information Management and Management Science. 2020: 93-97.
- [15] 吕雁琴,陈静,邱康权.省域旅游业发展水平、耦合协调与时空演变[J].统计与决策,2022,38(18):114-118.
- [16] 孙晓,刘力钢,寅克武,等.旅游产业高质量发展水平测度和区域差异分析[J].统计与决策,2022,38(19):92-97.
- [17] 魏瑾瑞,崔浩萌.基于网络搜索数据的区域旅游指数及其微观动态:以西安为例[J].系统科学与数学,2018,38(02):177-194.
- [18] 常家玲,张杉,苗红,等.中国省域城乡融合与乡村旅游耦合协调关系时空格局及驱动机制研究[J].中国农业资源与区划,2024,45(12):214-228.
- [19] 丰晓旭,夏杰长.中国全域旅游发展水平评价及其空间特征[J].经济地理,2018,38(4):183-192..
- [20] 王旭科,刘文静,李华.全域旅游发展水平评价指标体系构建与实证[J].统计与决策,2019,35(24):51-54.
- [21] 罗新颖.山东省全域旅游发展水平评价研究[J].国土与自然资源研究,2021(02):92-94.
- [22] 刘洋,赵彦云,张鑫钰.中国旅游统计体系的现状、问题与现代化改革[J].统计与信息论坛,2023,38(9):17-31.
- [23] 吴开军.旅游大数据研究热点及特征探析——基于国外文献的分析[J].统计与信息论坛,2019,34(4):105-113.
- [24] 徐金海,夏杰长.中国式现代化视域下的现代旅游业体系构建[J].社会科学家,2023,(8):47-51.

果发现:这23年来,以江苏省为例的中国式全域旅游现代化水平呈现稳步增加趋势,目前完成度已超80%。品质供给化和区域协调化一级指标得分相对需求多元化、成果共享化两个一级指标得分较高,表明江苏省旅游供给充分,全域旅游协调较好。需求多元化和成果共享化得分,在疫情期间得分下降,相对得分偏低,表明旅游需求还有待进一步挖掘和激发,成果共享也需进一步深入。

基于上述研究发现,可以得到如下政策启示。(1) 推动全域旅游产业升级。推动人工智能、区块链等前沿技术在旅游产业的合理应用,完善全域旅游数字基建,以保证旅游供给品质化。制定专项财政支持与税收优惠,引导传统景区向“文旅+”模式转型,培育智慧旅游、绿色低碳旅游等新兴业态。(2) 完善旅游行业风险防控体系和风险预警体系。完善“互联网+全域旅游”应急管理机制,提升旅游突发事件、旅游重大公共卫生事件、自然灾害等预防和应急处理水平,提升旅游市场主体抵御风险能力。加强旅游市场主体风险防范能力建设,指导旅游企业重视旅游应急管理投入,建立旅游企业内部风险防范防控体系。加强旅游从业人员和游客风险管理教育,普及旅游应急及自救互助知识,增强全社会旅游风险防范意识和应对突发事件能力,筑牢旅游行业风险防范的“防火墙”。(3) 强化物质文明与精神文明相协调的认知,培育本地化与特色化的旅游产品,加大对乡村旅游、红色旅游、民俗文化旅游等本地化旅游资源的开发与扶持力度,鼓励各地挖掘自身特色文化与自然资源,打造具有地域特色的旅游产品与项目。(4) 借鉴“苏超”等活动的品质供给激活需求的成功经验,通过创新,增加品质供给,通过品质供给创作和激发新的旅游需求,供给和需求的良性互动。

GARCH 型过程稳健控制设计及其在金融市场监控中的应用

王志坚, 麦智

广东财经大学统计与数据科学学院, 广东 广州 510320

DOI:10.61369/ASDS.2026040010

摘要 : 传统 GARCH 型质量控制图的中心限与控制上下限对离群值敏感, 常会导致监测结果与实际不符。本研究采用 HL-Shamos 稳健组合估计量与 Huber 权函数对其进行稳健改进, 构建出稳健 GARCH 型控制图。数值模拟结果发现: 第一, 在不同样本量、不同污染率下, 稳健 GARCH 型控制图与删除离群值的传统控制图监测效果接近, 而与含离群值的传统控制图相差较大。第二, 随着污染率增加, 传统控制图异方差控制限间距逐渐增大, 而稳健控制图几乎不变或变化轻微。第三, 在大样本、高污染率下, 传统控制图异方差控制限近似成直线型且远离控制中心限, 而稳健控制图却改变不大, 表现出较强的抗差性和稳健性。实证分析结果进一步验证了稳健 GARCH 型控制图的可行性和有效性。研究结论可用于离群值与异方差并存的时间序列稳健建模及统计过程控制。

关键词 : GARCH 型控制图; HL-Shamos 组合估计量; Huber 权函数; 稳健估计

Robust Control Design for GARCH-Type Processes and Its Application in Financial Market Monitoring

Wang Zhijian, Mai Zhi

School of Statistics and Data Science, Guangdong University of Finance & Economics, Guangzhou, Guangdong 510320

Abstract : The central limit and control upper & lower limits of the traditional GARCH quality control chart are sensitive to outliers, which often leads to the inconsistency between the monitoring results and the actual situation. In this study, the HL-Shamos robust combination estimator and Huber weight function are used to improve it robustly, and a robust GARCH control chart is constructed. The results of numerical simulation show that: first, under different sample sizes and different contamination rates, the monitoring effect of the robust GARCH control chart is close to that of the traditional control chart with outliers deleted, but is significantly different from that of the traditional control chart with outliers. Second, with the increase of contamination rate, the heteroscedasticity control limit interval of traditional control chart increases gradually, while the robust control chart is almost unchanged or changed slightly. Third, under large samples and high contamination rate, the heteroscedasticity control limit of traditional control chart is approximately linear and far from the control center limit, while the robust control chart has little change, showing strong robustness and robustness. The empirical analysis results further verify the feasibility and effectiveness of the robust GARCH control chart. The research findings can be applied to robust statistical process control for time series featuring concurrent outliers and heteroskedasticity.

Keywords : GARCH control chart; HL-Shamos combination estimator; Huber weight function; Robust estimation

引言

统计过程控制 (简称 SPC) 是应用统计技术对过程中的各个阶段进行评估和监测, 建立并保持过程处于可接受且稳定的水平, 从而保证产品与服务符合规定要求的一种质量管理技术。常规控制图应用的基本假定是观察数据独立同分布 (iid), 但在大数据时代背景下, 时间序列数据常出现复杂情况, 如非正态性、离群值、自相关性和异方差性并存等特征, 导致常规控制图有时不能很好发挥作用, 因此有必要进一步研究。

基金项目: 本文获得广东省基础与应用基础研究基金项目“多变量自相关过程稳健控制设计及评价方法研究”(项目编号: 2026A1515011461)资助。

作者简介:

王志坚 (1982.12-), 男, 江西余干人, 博士, 广东财经大学副教授, 研究方向: 稳健统计、时间序列; 邮箱: 32680822@qq.com;

麦智 (2002.09-), 男, 广东佛山人, 广东财经大学硕士研究生, 研究方向: 稳健统计、时间序列; 邮箱: 1910287603@qq.com。

一、文献综述

对于序列仅出现自相关情形,目前学术界有两种研究路径:一是修正控制图,即针对受控过程的自相关性,对常规控制图控制限加以调整,其中包括改进的 EWMA 控制图和 ARMA 控制图等; Ajadi JO^[1] 设计出了稳健多变量 EWMA 分散控制图; Koyuncu, N^[2] 研究了稳健修正的 R 型控制图; Cabana Elisa 等^[3] 基于稳健重加权收缩估计量提出了稳健多变量控制图设计原理; Yadpirun 等^[4] 针对自相关过程提出了改进型指数加权移动平均控制图并给出了平均运行长度计算方法。二是转向残差控制图,因为当序列值拟合适当的时间序列模型后,其残差是独立的,可通过残差控制图达到对质量特性的监控目的,建立了关于残差 EWMA 和 CUSUM 控制图; 张立健等^[5] 构建了 AR-ARCH 型残差控制图; Sermad Abbas 等^[6] 构建了局部线性时间序列均值稳健残差控制图; Salah Mohamed 等^[7] 研究了伽马回归模型下残差控制图的剖面监测; Li Xue 等^[8]、Ramy M 等^[9] 分别提出了不同的残差控制图设计方法并将其应用于不同的领域监控。

当序列出现离群值(非正态性)与自相关混合情形时,多采用基于模型拟合残差控制图方法,但该方法严重依赖于模型的估计精度,当模型精度不够时,残差并非独立同分布导致监测失效。目前解决这类问题的思路主要是构建稳健时间序列模型,确保得到的残差无自相关性,符合残差控制图使用条件,从而构建稳健残差控制图。Liu 等^[10] 研究了在非正态分布下,用样本中位数来构建稳健 EWMA 控制图,该稳健控制图在检测均值漂移方面比传统控制图更优。Derya Karagöz^[11] 对非正态分布下的非对称控制图控制限进行稳健设计。Ishaq 等^[12] 对污染过程及违背基本假定的双控制图进行了稳健改进。Liang 等^[13] 将空间符号协方差矩阵和最大范数应用于指数加权移动平均方案,提出了一种多变量协方差二期控制图。Nasir Abbas^[14] 构建了稳健 S2 控制图。Kao Shih Chou^[15] 基于平方 A 估计量构建了稳健标准差控制图。Muhammad Arslan 等^[16] 基于稳健 L 协方差矩阵构建霍特林 T2 控制图,用于在存在异常值的情况下监测高维非正态多元数据。

当序列仅有异方差特征(波动簇聚性)情形时,通过文献梳理不难发现,此类问题的研究相对比较滞后。主要有: Severin^[17]、Sehipper^[18]、夏远强^[19] 先后提出应用受控过程条件标准差取代无条件标准差来构造控制图控制限方法,后面许多学者均是采用该方法。Sabahno Hamed^[20] 用变参数控制图监测存在自相关的多变量变异系数。当出现离群值(非正态性)、自相关与异方差三种混合情形时,目前国内外对这方面的研究进展不大。主要有:王志坚^[21] 通过构建稳健 ARMA 模型和 GARCH 模型得到了稳健残差控制图。庄芳等^[22] 采用绝对偏差中位数对 EWMA 方差控制图进行稳健设计,效果良好。苏拥英,王志坚^[23] 研究了常规过程控制图的敏感性分析及稳健设计。Salih O 等^[24] 对因素选择控制图的最优稳健设计进行了研究。Atef F. Hashem 等^[25] 研究了一种新的基于变异系数的距离加权控制图,用于监测核事件数据。

本研究在前人研究基础上,首先对 GARCH 型质量控制图进

行了敏感性分析,包括:第一, GARCH 型一般控制图,即无自相关性但有异方差性的控制图;第二, GARCH 型残差控制图,即既存在自相关性又存在异方差性的控制图。其次,结合 HL-Shamos 稳健组合估计量与 Huber 权函数对 GARCH 型控制图控制中心、异方差控制限进行稳健改进,构建了稳健 GARCH 型控制图,并通过模拟与实证分析检验稳健 GARCH 型控制图的可行性与有效性,从而较好的解决了在离群值(非正态性)、自相关与异方差三种情形并存时的传统过程控制失效问题。研究结论可为大数据时代的过程控制方法提供新思路。

二、传统 GARCH 型控制图的敏感性分析及相关稳健估计量

(一) 无自相关有异方差情形

假设 $\{y_t\}$ 为无自相关的异方差序列,其表达式如下:

$$y_t = m + \varepsilon_t \quad (1)$$

其中, m 为序列均值; ε_t 为相互独立的随机扰动项。(1)式残差可表示为:

$$\varepsilon_t = y_t - m$$

由前文假设可知,序列 $\{y_t\}$ 无自相关性,故残差序列 $\{\varepsilon_t\}$ 相互独立; $\{y_t\}$ 具有异方差性,故残差 $\{\varepsilon_t\}$ 存在 GARCH 效应。因此,基于过程 $\{y_t\}$ 的控制图,其控制限可以表示为:

$$\begin{cases} \text{控制上限} & \text{UCL} = m + 3\sigma_t \\ \text{中心限} & \text{CL} = m \\ \text{控制下限} & \text{LCL} = m - 3\sigma_t \end{cases} \quad (2)$$

显然, m 与时间无关, σ_t 与时间有关,因此由(2)式确定的控制中心为恒定,但控制限是随时间而变化的异方差控制限,故得到 GARCH 型一般控制图。显然 m 与 σ_t 均对离群值敏感^[21],故 GARCH 型一般控制图具有不稳健性。

(二) 自相关与异方差并存情形

假设 $\{y_t\}$ 为自相关与异方差并存的序列,其表达式如下:

$$y_t = m_t + \varepsilon_t = E[y_t | F_{t-1}]$$

其中, m_t 为在给定前期信息集 F_{t-1} 条件下 y_t 的条件均值; ε_t 为相互独立的随机扰动项。由定义可知, $\{y_t\}$ 具有异方差性,故残差 $\{\varepsilon_t\}$ 有 GARCH 效应。则下式成立:

$$y_t \in (\hat{m}_t - 3\sigma_t, \hat{m}_t + 3\sigma_t)$$

因此,基于过程 $\{y_t\}$ 的控制图,其控制限可以表示为:

$$\begin{cases} \text{控制上限} & \text{UCL} = \hat{m}_t + 3\sigma_t \\ \text{中心限} & \text{CL} = \hat{m}_t \\ \text{控制下限} & \text{LCL} = \hat{m}_t - 3\sigma_t \end{cases} \quad (3)$$

显然, $\hat{m}_t$ 和 σ_t 与时间有关,因此由(3)式确定的三条控制限均不再恒定,而是随时间变化而变化,控制限的不恒定性会导致对异常点界定带来困难。在此,将(3)式变形如下:

$$y_t - \hat{m}_t = \varepsilon_t \in (-3\sigma_t, 3\sigma_t)$$

可知在 t 时刻, 残差 ε_t 落在 $(-3\sigma_t, 3\sigma_t)$ 内的概率也是99.73%。故对原序列监测与对残差的监测效果一致。故可采用对残差序列的监测去替代对原序列监控, 即为残差控制图, 其控制限表达式为:

$$\begin{cases} \text{控制上限} & \text{UCL} = 3\sigma_t \\ \text{中心限} & \text{CL} = 0 \\ \text{控制下限} & \text{LCL} = -3\sigma_t \end{cases}$$

由残差控制图控制限表达式可知, 中心限是恒定, 只有控制上、下限是波动, 显然这比之前的三条控制限均是波动要改善很多。针对控制上、下限具有波动性特点, 可以采用 GARCH 模型去刻画之, 得到 GARCH 型残差控制图。GARCH 型残差控制图需要先建立序列 ARMA 模型, 再求出异方差控制限。而 GARCH 型一般控制图只需求出序列均值, 再求出异方差控制限, 显然 GARCH 型残差控制图比 GARCH 型一般控制图更为复杂, 由于两者构建异方差控制限的思路一致, 因此下文主要围绕 GARCH 型残差控制图展开。

在受控过程无离群值时, 用异方差构造的控制限显然比同方差构造控制限更科学, 且能更好地捕捉到受控过程的异常现象。但随着 SPC 技术在质量管理和经济金融等领域的广泛使用, 诸如疫情的发生、金融危机、观测仪器的偶然故障、人为的观测差错等均会导致序列中异常情况发生。但由于 GARCH 型控制图异方差控制限 $(-3\sigma_t, 3\sigma_t)$ 对离群值敏感, 缺乏稳健性, 易导致上、下控制限间距扩大, 引起“漏报警”现象, 从而出现了在对异常现象监控过程中反而被异常点所干扰的困境。鉴于此, 本研究采用 σ_t 的稳健估计技术来解决该问题, 即用稳健标准差代替无条件标准差来构造控制限, 从而设计出稳健的 GARCH 型质量控制图。

(三) 相关稳健估计量

最小二乘法的基本原理是残差平方和函数最小, 但该函数不稳健, 单个的离群值能将其变得很大; 而最小一乘估计如 $\rho(x) = |x|$, 虽能减弱异常值对参数估计的影响, 但因其不可导而使其应用受限。对此, Huber (1964) 给出了一般意义上的 ρ 函数, 即对于如下样本回归模型:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_k X_{ik} + e_i = X_i' \beta + e_i$$

这时目标函数是

$$\sum \rho(e_i) = \sum \rho(Y_i - X_i' \beta) \quad (4)$$

β 的 M 估计是极小化 (4) 式。对 (4) 式关于参数 β 求偏导数, 并令其等于零, 则得到如下方程:

$$\sum \psi(Y_i - X_i' \beta) X_i' = 0 \quad (5)$$

其中 $\psi(\cdot)$ 为 $\rho(\cdot)$ 导函数, 为保证所估参数与因变量观测单位无关, 需下式成立:

$$\hat{\beta}(X_i, aY_i) = a\hat{\beta}(X_i, Y_i)$$

上式表示 M 估计满足尺度同变性, α 为任意常数。经尺度同

变换后 (5) 变为:

$$\sum_{i=1}^n \psi\left(\frac{Y_i - X_i' \beta}{s}\right) X_i' = \sum_{i=1}^n \frac{\psi(e_i / s)}{e_i / s} X_i' \cdot e_i / s = 0 \quad (6)$$

记 $\omega_i = \frac{\psi(e_i / s)}{e_i / s}$ 为回归估计方程的权重, 即方程 (6) 可以写成

$$\sum \omega_i X_i' e_i = 0$$

显然上式即变成了一个加权最小二乘估计问题, 而且权重与回归系数通过残差来相互依赖; 适当选择目标函数 ρ 与权重函数 $\omega(e)$, 使残差大的异常点权重变小, 从而减弱离群值对参数估计的影响。M 估计法的函数很多, 本研究选择 Huber 权函数, 其表达式如下:

$$\omega(\varepsilon_i) = \begin{cases} 1 & |\varepsilon_i| \leq k \cdot \sigma \\ k \cdot \sigma / |\varepsilon_i| & |\varepsilon_i| > k \cdot \sigma \end{cases} \quad (7)$$

(7) 式中常数 k 称为阈值 (也叫细调参数), 在实际中, k 值的选择要依赖于对样本中有多大比例离群值的先验信息基础上, 经试验比较, 本研究取 k=2.2。 σ 为序列的标准差, ε_i 为残差。

本研究借鉴 Hodges & Lehmann (1963) 提出的均值稳健估计量, 表达式为:

$$\text{median}\left(\frac{X_i + X_j}{2}\right) \quad i, j = 1, 2, \dots, n$$

而后, Chanseok Park 等 (2022) [26] 将其推广为如下三种形式, 分别简称为 HL1、HL2 及 HL3:

$$HL1 = \text{median}_{i < j} \left(\frac{X_i + X_j}{2} \right),$$

$$HL2 = \text{median}_{i \leq j} \left(\frac{X_i + X_j}{2} \right),$$

$$HL3 = \text{median}_{\forall (i, j)} \left(\frac{X_i + X_j}{2} \right)$$

在此, 经比较后本研究选 HL3。

而标准差 σ 的稳健估计量选取 Shamos 估计量的无偏估计, 表达式为:

$$\text{Shamos} = \frac{\text{median}_{i < j} (|X_i - X_j|)}{\sqrt{2} \Phi^{-1}(3/4)} \approx 1.048358 \cdot \text{median}_{i < j} (|X_i - X_j|)$$

由于 Huber 权函数中的标准差 σ 对离群值敏感, 在此, 我们将 Huber 权函数进行稳健改进, 得到改进的 Huber 权函数, 本研究将其记为:

$$\omega^R(\varepsilon_i) = \begin{cases} 1 & |\varepsilon_i| \leq k \cdot \text{Shamos} \\ k \cdot \text{Shamos} / |\varepsilon_i| & |\varepsilon_i| > k \cdot \text{Shamos} \end{cases} \quad (8)$$

(8) 式即为 (7) 式的稳健版本, 因此, 对于构建 GARCH 型一般控制图, 采用 HL3 作为控制中心 (均值) m 的稳健估计量, 用原序列减去稳健均值得到残差, 再通过残差结合改进的 Huber 权函数算出原序列权重, 最后将该权重乘原序列后计算出稳健异方差波动率。

对于 GARCH 型残差控制图, 先建立稳健 ARMA 模型并通过模型算出残差, 确保残差服从独立同分布, 再通过残差结合改进的 Huber 权函数算出原序列权重, 最后将该权重乘以原序列后计算出稳健异方差波动率, 进而构建稳健控制限。稳健 GARCH 型控制图设计原理如图 1 所示。

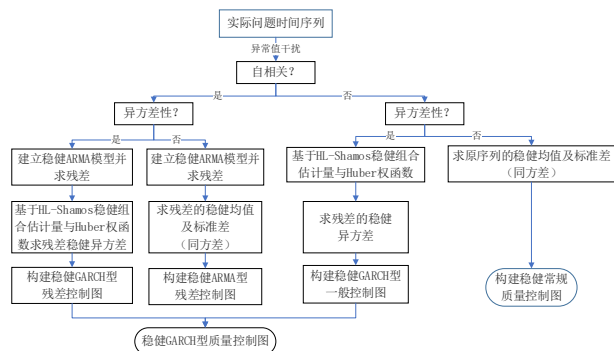


图 1: GARCH 型过程稳健设计原理

三、模拟研究

为验证本研究提出的稳健 GARCH 型控制图可行性与有效性, 在此模拟一组 GARCH (0,2) 系列, 其中 $\alpha_1 = 0.01$, $\alpha_2 = 0.9$ 。选取样本量分别为: 50、100、200、500, 污染率分别为 1%、5%、10%。目的是从小样本、大样本, 低污染率、高污染率全方位考察稳健效果。需要说明的是, 根据研究需要, 样本量为 500 时, 只考虑污染率为 10% 情形, 采用污染分布产生离群值, 表达式如下:

$$P = (1 - \varepsilon)N(\mu_0, \sigma^2) + \varepsilon U(\mu, \sigma^2)$$

其中 ε 表示污染率。污染数据来自于均匀分布, 通过随机抽取。

在此, 定义污染前的波动率称为原始波动率, 被污染后的波动率称为含离群值波动率, 稳健改进后的含离群值波动率称为稳健波动率, 此处波动率图即为控制图的上限。不同样本量、不同污染率下的模拟结果如图 2- 图 5 所示。需要说明的是, 由于控制图关于控制中心对称, 在此, 模拟部分只给出控制图上限, 目的是便于比较稳健改进前后的效果。

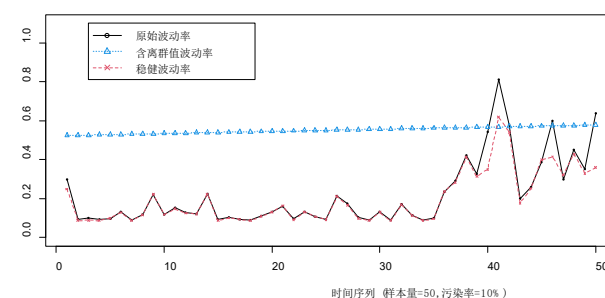
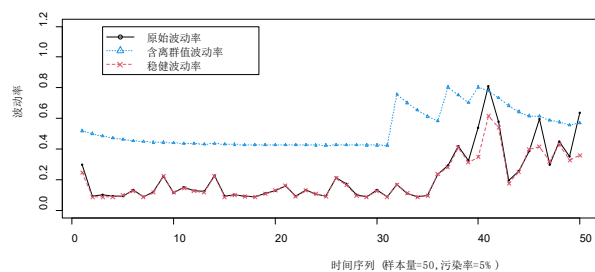
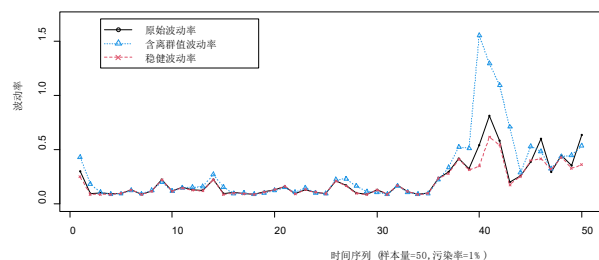


图 2: 样本量为 50, 污染率为 1%、5%、10% 的各波动率图

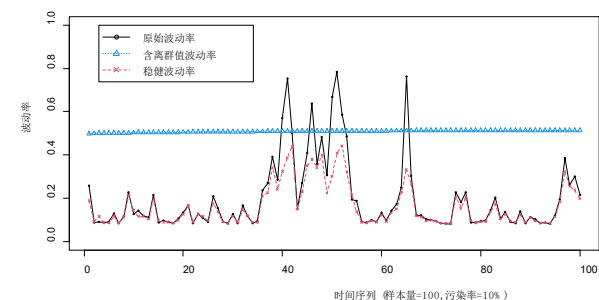
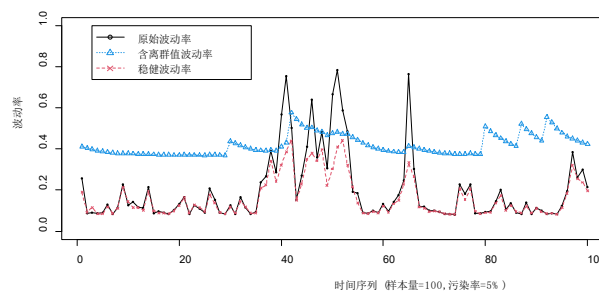
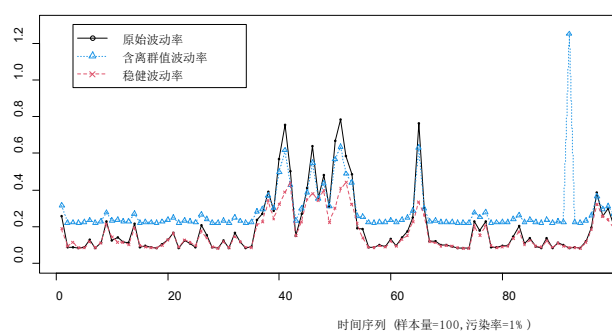


图 3: 样本量为 100, 污染率为 1%、5%、10% 的各波动率图

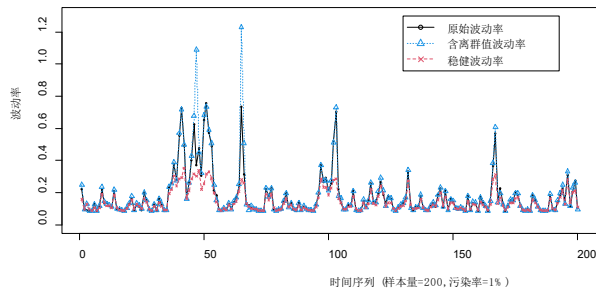


图4：样本量为200，污染率为1%、5%、10%的各波动率图

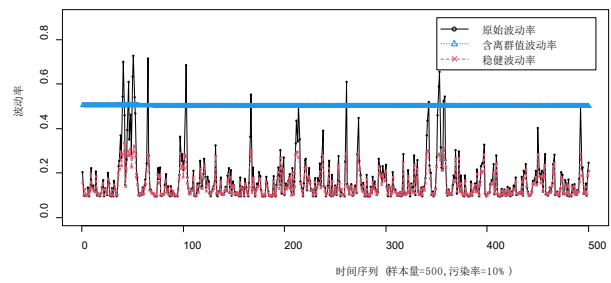


图5：样本量为500，污染率为10%的波动率图

从图2-图5的模拟结果可以得出以下三个结论：第一，在不同样本量、不同污染率下，原始波动率与稳健波动率差别不大，而与含离群值波动率相差较大。第二，随着污染率的增加，含离群值波动率偏离原始波动率越来越大，控制限间距被拉得越来越开，而稳健波动率几乎不变或变化轻微。第三，在大样本、高污染率下，原始波动率形状已近似成为一条直线且偏离控制中心较大，导致监控能力越来越差，甚至说已不再具备监控效果（在样本量为200、500，污染率为10%图中足已说明），而稳健波动率却改变不大，表现出较强的抗差性和稳健性。

四、实证分析

为了进一步检验稳健 GARCH 型控制图的可行性和有效性，下面以一个实际金融市场例子为例，选取脸书（Facebook）公司股票作为研究对象，研究期间为：2020年1月6日至2021年8月20日，共411个样本数据，数据来源于 yahoo 财经网站。股票前6天后6天各指标基本情况如表1所示。

表1：脸书公司股票2020-01-06至2021-08-20前6后6天情况

日期	开票价	最高价	最低价	收盘价	成交量	调整价
2020-01-06	206.70	212.78	206.52	212.60	17058900	212.60
2020-01-07	212.82	214.58	211.75	213.06	14912400	213.06
2020-01-08	213.00	216.24	212.61	215.22	13475000	215.22
2020-01-09	217.54	218.38	216.28	218.30	12642800	218.30
2020-01-10	219.20	219.88	217.42	218.06	12119400	218.06
2020-01-13	219.60	221.97	219.21	221.91	14463400	221.91
2021-08-13	362.97	364.66	361.48	363.18	7192500	363.18
2021-08-16	362.52	366.95	357.38	366.56	11222100	366.56
2021-08-17	362.98	364.67	355.07	358.45	12328700	358.45
2021-08-18	356.25	359.86	355.15	355.45	10042400	355.45
2021-08-19	351.35	357.57	351.00	355.12	12004500	355.12
2021-08-20	354.60	359.99	353.31	359.37	8693500	359.37

在各变量中选取收盘价的调整价，通过调整价算出脸书股票的对数收益率，计算公式如下：

$$r_t = \ln(P_t - P_{t-1}) * 100$$

其中， P_t 表示第t天的调整价， r_t 表示第t天的对数收益率，用R语言进行统计分析。首先对调整价及收益率做一个探索性数据分析图（如图6）及基本统计分析（如表2）来直观了解脸书公司股票调整价及收益率。

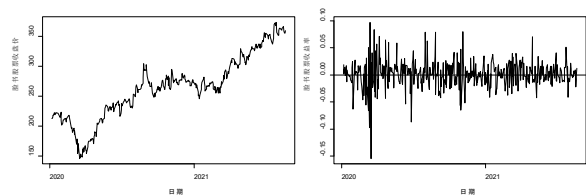


图6：脸书股票收盘价（左）及收益率（右）时序图

图6显示，收盘价在2020年上半年急速下降后一路上升，

中间虽偶有滑落现象，但不影响总体上涨趋势。而对于收益率来说，很明显可以看出有一些较为突出的点，至于这些点是否为统

计学上定义的离群点，需要通过统计检验来界定。

表2：脸书股票收益率基本统计量

最小值	最大值	均值	中位数	标准差	偏度	峰度
-0.154	0.0974	0.001321	0.00149	0.025461	-0.375313	4.688253

从收益率基本统计量看（表2），收益率没有呈现正态分布特征，而是呈现左偏分布，且有高峰厚尾的现象。进一步来观察脸书股票收益率的直方图及 Q-Q 图，如图7所示：

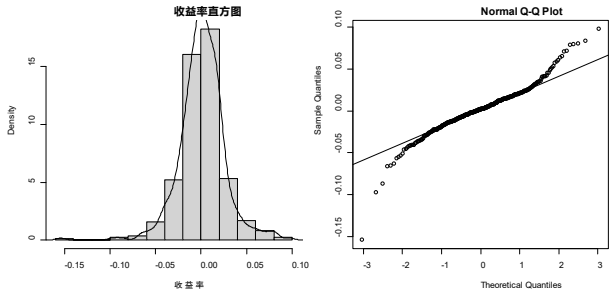


图7：脸书股票收益率直方图及 Q-Q 图

图7显示，收益率直方图及 Q-Q 图均已偏离正态分布。进行 Jarque Bera Test 正态性检验，得到检验统计量 X-squared = 392.18, p-value < 2.2e-16, 可知收益率不是正态分布，结合前面收益率时序图，我们初步判定：由于离群值的存在，导致收益率偏离正态分布。接下来建立 ARMA 模型，通过对收益率序列进

行平稳性检验，检验结果为平稳，再看收益率自相关和偏自相关图（如图8）：

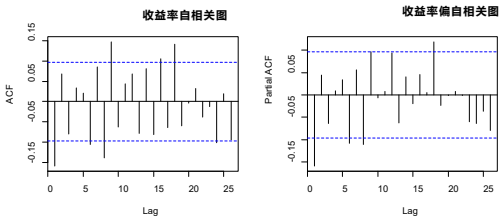


图8：脸书股票收益率自相关和偏自相关图

表3：ARMA 模型阶数的确定

	模型	AR(p)	MA(q)	ARMA(p,q)
ACF		拖尾	q 阶截尾	拖尾
PACF		P 阶截尾	拖尾	拖尾

结合图8再根据表3中 ARMA 模型的定阶准则，很难清晰地确定 ARMA 模型的阶数，因此需要比较2阶及以下各模型情况，如表4：

表4：ARMA 各模型显著性情况

模型	ARMA(1,0)	ARMA(0,1)	ARMA(1,1)	ARMA(1,2)	ARMA(2,1)	ARMA(2,2)
系数显著性	全显著	全显著	不全显著	全显著	全显著	全不显著
AIC	-1856.86	-1856	-1856.21	-1851.14	-1852.13	-1851.98

在保证系数显著的前提下，再比较模型的 AIC 值，AIC 值越小越好，最后选择 ARMA(1,0) 模型。建立 ARMA(1,0) 模型后再求出模型的残差，对残差序列进行 ARCH 效应检验，检验结果如表5所示：

表5：残差序列的 ARCH 效应检验结果

ARCH LM-test; Null hypothesis: no ARCH effects
data: resid*2
Chi-squared = 49.451, df = 12, p-value = 1.744e-06

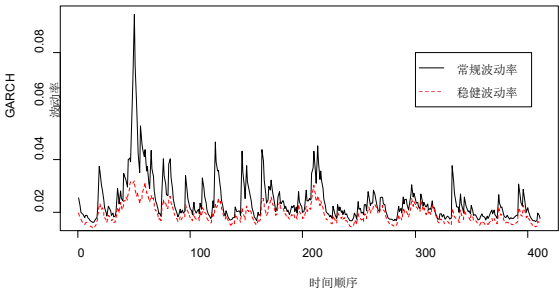


图9：脸书股票收益率序列的常规波动率与稳健波动率比较

表5显示：检验结果 p-value = 1.744e-06, 拒绝原假设，说明残差序列具有显著的 ARCH 效应，因此可以对收益率序列求出波动率，再建立 GARCH 型控制图。

图9为收益率序列的常规波动率与稳健波动率，图中显示，稳

健波动率比常规波动率波动幅度明显要小，说明稳健波动率受离群值影响较小。

结合常规 GARCH 型控制图（图10左）与稳健 GARCH 型控制图（图10右）对离群值监测效果来看，受离群值影响，常规控制图异方差控制限波动幅度参差不齐，而稳健控制图异方差控制限相对平稳。以序号48、49的点为例，该两点为离群点，但在常规控制图中由于离群点将异方差控制限间距拉大，导致该两个点在控制限内，使得原本为离群点而被误判为正常点，“掩盖现象”发生。

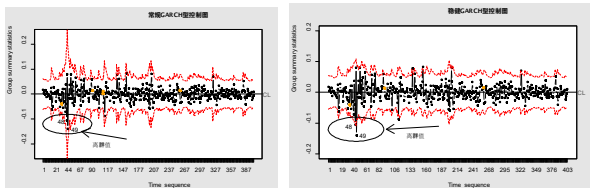


图10：常规 GARCH 型控制图（左）与稳健 GARCH 型控制图（右）

在此我们将文献^[27]提出的稳健时间序列离群值检测法检测结果作为参照，发现常规 GARCH 型控制图对离群值检测相对正确率为44.44%，而稳健 GARCH 型控制图对离群值检测相对正确率为77.78%，稳健 GARCH 型控制图监测正确率显著高于常规 GARCH 型控制图（表6）。

表6：离群值监测正确率比较											
	稳健时间序列离群值诊断法					常规 GARCH 型控制图					稳健 GARCH 型控制图
离群值序号	47	48	49	55	121	121	145	163	332	18	48 49 55 95 121 145 149 163 332
相对正确率	--					44.44%					77.78%

由于稳健 GARCH 型控制图采用了稳健控制限，进而消除或减弱了离群值对控制限的影响，从而较好的解决了在用 GARCH 型控制图对离群值监测过程中反而被离群值所干扰的困境。

五、结语

当受控过程出现离群值（非正态性）、自相关与异方差情形，

参考文献

[1]Ajadi JO, Zwetsloot IM, Tsui K-L. A New Robust Multivariate EWMA Dispersion Control Chart for Individual Observations[J]. Mathematics. 2021,9(9) :1–18.

[2]Koyuncu, N., Karagöz, D. Designing robust modified R control charts for asymmetric distributions under ranked set and median ranked set sampling[J]. Computational Statistics, 2021,36:1093 – 1121.

[3]Cabana Elisa ,Lillo Rosa E.. Robust multivariate control chart based on shrinkage for individual observations[J]. Journal of Quality Technology ,2022, 54(4): 415–440.

[4]Yadpirun Supharakonsakun; Yupaporn Areepong; Korakoch Silpakob.Explicit ARL Computational for a Modified EWMA Control Chart in Autocorrelated Statistical Process Control Models [J] Computer Modeling in Engineering & Sciences.2025. 145(1): 699–720 .

[5]张力健，杨继平，张秋菊.AR-GARCH 型残差控制图及其应用研究 [J]. 中国管理科学 ,2006, 14(1):15–19.

[6]Sermad Abbas ； Roland Fried. Robust control charts for the mean of a locally linear time series[J]. Journal of Statistical Computation and Simulation,2020,90(15):1–25.

[7]Salah Mohamed,Engy Mohamed,Shereen Abdel Latif. Profile monitoring of residuals control charts under gamma regression model[J]. Eastern-European Journal of Enterprise Technologies,2022,6(4):23–31.

[8]Li Xue; Qiuyu Wang; Congkai Li; Lisheng An.Economic design of residuals MEWMA control chart with variable sampling intervals and sample size [J] Communications in Statistics – Simulation and Computation. 2025.54(2) : 467–488.

[9]Ramy M. Khalifa; Soumaya Yacout; Samuel Bassetto; Yasser Shaban.Condition monitoring and warning of a belt drive system based on a logical analysis of data regression-based residual control chart [J] Structural Health Monitoring. 2025. 24(3): 1657–1673.

[10]Lin Y C, Chao Y C, Chen C H. Robustness of the EWMA median control chart to non-normality[J]. International Journal of Industrial & Systems Engineering, 2017, 25(1):35–45.

[11]Derya Karagöz. Asymmetric control limits for range chart with simple robust estimator under the non-normal distributed process[J].Mathematical Sciences, 2018,12(4) :249–262.

[12]Ishaq Adeyanju Raji, Muhammad Riaz, Nasir Abbas. Robust dual-CUSUM control charts for contaminated processes[J]. Communications in Statistics–Simulation and Computation, 2019, 48(7):2177–2190.

[13]Liang, Xiang, Pu, Li, Jin. A robust multivariate sign control chart for detecting shifts in covariance matrix under the elliptical directions distributions[J]. Quality Technology & Quantitative Management, 2019, 16(1): 113–127.

[14]Nasir Abbas. A robust S2 control chart with Tukey’s and MAD outlier detectors[J].Quality and Reliability Engineering International, 2020, 36(1):403–413.

[15]Kao Shih Chou. A robust standard deviation control chart based on square A estimator[J]. Quality and Reliability Engineering International, 2022, 38(5):2715–2730.

[16]Muhammad Arslan; Usman Shahzad; Ali Yeganeh; Huiming Zhu; JC Malela Majika; Shakeel Ahmad.A Robust L - Comoments Covariance Matrix - Based Hotelling’s T2s{T^{2}} Control Chart for Monitoring High - Dimensional Non - Normal Multivariate Data in the Presence of Outliers [J] Quality and Reliability Engineering International. 2025, 41(7): 3308–3317.

[17]Severin T., Schmid W.. Monitoring changes in GARCH models[J]. Allgemeines Statistisches Archiv, 1999, 83 :281–307.

[18]Schipper S.,Schmid W.. Control charts for GARCH process[J]. Nonlinear Analysis,2001, 47:2049–2060.

[19]夏远强，韩文秀，何民.GARCH 型相关过程及其质量控制图. 天津大学学报 [J]. 2003, 36(0):92–95.

[20]Sabahno Hamed, Celano Giovanni. Monitoring the multivariate coefficient of variation in presence of autocorrelation with variable parameters control charts[J]. Quality Technology & Quantitative Management, 2023, 20(2):184–210.

[21]王志坚. 稳健残差控制图的构建及在金融时序中的应用 [J]. 数理统计与管理 , 2017, 36(05):930–942.

[22]庄芳，郁淼淼，吴纯杰. 带参数估计的 EWMA 方差控制图稳健性分析 [J]. 系统科学与数学 , 2018, 38(01):101–118.

[23]苏拥英，王志坚. 常规过程控制图的敏感性分析及稳健设计 [J]. 统计与决策 , 2021, 37(10):160–164.

[24]Salih O. Duffuaa, Ahmed M. Ghaithan, Ahmed M. Attia. A mathematical model for the optimal robust design of cause selecting control charts[J]. European Journal of Industrial Engineering, 2022, 16(2):169–193.

[25]Atef F. Hashem; Najam ul Hassan; Maqbool Hussain Sial; Syed Muhammad Muslim Raza.A novel coefficient of variation based distance weighted control chart for the monitoring of nuclear event data [J] Journal of Radiation Research and Applied Sciences..2025, 18(3) :101743–10174.

[26]Chanseok Park, Haewon Kim & Min Wang . Investigation of finite-sample properties of robust location and scale estimators[J].Communications in Statistics – Simulation and Computation, 2022, 51(5):2619–2645.

[27]王志坚，王斌会. 稳健改进的 AO 型异常点检测法在金融时序中的应用 [J]. 数理统计与管理 . 2016, 35(2):369–380.

研究者通常构建 GARCH 型一般控制图或 GARCH 型残差控制图对受控过程进行监测，而忽略了离群值本身对控制中心及异方差控制限的影响，结果往往导致监测效果不理想，对此类问题的解决即是本研究的创新之处。文章分析了 GARCH 型一般控制图及 GARCH 型残差控制图的敏感性，并基于 HL–Shamos 稳健组合估计量与 Huber 权函数方法对以上两类控制图的控制中心及控制限进行稳健改进，构建了稳健 GARCH 型控制图。研究进一步发现：在构建稳健 GARCH 型控制图时，将控制中心及控制限两者看成有机整体设计出的稳健组合估计量效果优于单独构建的稳健估计量。

基于自适应权重的 MIDAS-GARCH 模型

刘培宇, 许敏, 尹兰江*

广州大学 经济与统计学院, 广东 广州 510006

DOI:10.61369/ASDS.2026040011

摘 要 : 针对传统 MIDAS-GARCH 模型中固定权重函数难以有效捕捉高频数据微观结构噪声与时变特征的局限, 本文提出自适应权重优化的 MIDAS-GARCH 模型。通过引入基于高频数据交易特征与波动动态调整权重的函数, 克服传统方法权重分配模式固定的局限, 实现高频信息的高效融合。理论层面论证了模型参数估计量的一致性与渐近正态性, 并采用极大似然估计 (QMLE) 实现参数求解。模拟研究和实证研究验证了模型在不同数据生成机制下的高估计精度与稳健性。

关 键 词 : 混频数据; 自适应权重; MIDAS-GARCH 模型; 数值模拟; 波动率预测

MIDAS-GARCH Model Based on Adaptive Weights

Liu Peiyu, Xu Min, Yin Lanjiang*

School of Economics and Statistics, Guangzhou University, Guangzhou, Guangdong 510006

Abstract : This paper proposes an adaptive weight MIDAS-GARCH model to address the limitations of traditional fixed-weight specifications, which fail to effectively capture the microstructure noise and time-varying characteristics of high-frequency data. By introducing a weighting function that dynamically adjusts based on trading characteristics and volatility, the proposed model overcomes the rigidity of fixed weight allocation schemes and achieves efficient fusion of high-frequency information. Theoretically, we establish the consistency and asymptotic normality of the parameter estimators derived via Quasi-Maximum Likelihood Estimation (QMLE). Simulation and empirical studies validate the model's superior estimation accuracy and robustness across various data generation mechanisms.

Keywords : mixed frequency data; adaptive weights; MIDAS-GARCH model; numerical simulation; volatility forecasting

引言

波动率是衡量金融资产价格变动剧烈程度的核心指标, 直接关联风险暴露、投资组合收益及衍生品定价, 是金融市场风险管理与资产配置的关键考量。传统波动率度量多基于低频数据 (如宏观政策), 依赖 ARCH、GARCH 等模型虽能捕捉波动集群等特征^[1-3], 却难以充分整合高频信息 (如日内价量)。近年来, 高频数据 (分钟级、秒级交易记录) 的可得性提升, 为波动率短期动态提供了更精细的视图。实证研究表明, 将高频信息纳入模型可显著增强预测能力 (如 Zhang (2005)^[4]、Barndorff (2009)^[5] 等方法)。然而, 现有研究多聚焦单一频率, 较少从混频视角探索高低频信息的协同价值。因此, 高效融合高低频数据以优化波动率预测的准确性与稳健性, 成为金融时间序列分析领域的核心挑战。

为突破传统时间序列分析在处理高低频混合数据时的频率壁垒^[6-8], Engle (2013) 等人提出了混频数据抽样 (MIDAS) 模型, 并将其与 GARCH 模型结合形成 MIDAS-GARCH 模型。该模型将波动率分解为短期波动成分和长期趋势成分, 其基本设定如下:

$$r_{i,t} = E(r_{i,t} | F_{i-1,t}) + \sigma_{i,t}, \quad (1)$$

$$\sigma_{i,t} = \sqrt{\tau_t \times g_{i,t}} \varepsilon_{i,t}, \varepsilon_{i,t} | F_{i-1,t} \sim N(0,1), \quad (2)$$

$$g_{i,t} = (1 - \alpha - \beta) + \alpha \frac{\sigma_{i,t}^2}{\tau_t} + \beta g_{i-1,t}, \quad (3)$$

作者简介:

刘培宇, 广州大学经济与统计学院本科生;

许敏, 广州大学经济与统计学院硕士研究生。

通讯作者: 尹兰江, 广州大学经济与统计学院硕士研究生。

$$\tau_t = m + \theta \sum_{k=1}^K \phi_k(\omega_1, \omega_2) X_{t-k} \quad (4)$$

其中， $r_{i,t}$ 为第 t 日的收益率， $g_{i,t}$ 为短期波动成分，遵循日度 GARCH(1,1) 过程； τ_t 为长期波动成分，通过低频变量的加权滞后项捕捉长期趋势。该模型通过引入混频数据采样机制，有效捕捉了不同频率数据间的动态关系，已成为金融波动性分析的重要工具。目前，其应用场景已从国外的股市、加密货币市场，拓展至国内的股市波动、金融状况指数测度及原油期货分析等领域。

在上述 MIDAS-GARCH 框架中，权重函数（即公式中的权重部分）的选择决定了高频滞后项的分布形态，是模型构建的核心。现有研究多采用固定形式的参数化权重函数，常见的类型包括：等权重函数、Beta 函数族（含非零 Beta）、Almon 多项式族（含指数 Almon）^[9] 以及 Legendre 权重等。这些传统权重函数虽然通过参数化结构实现了时间尺度的一致性，能稳健提取长期波动趋势，但其权重分配模式相对僵化，难以适配金融市场非线性、非对称的复杂动态特征。特别是面对具有微观结构噪音和日内剧烈波动的高频数据时，固定权重往往难以根据实时交易特征进行动态调整，从而限制了模型对短期市场变化的捕捉能力。

针对现有研究在处理高频数据微观结构信息挖掘不足及权重分配僵化的问题，本文主要在以下两个方面进行了改进与创新：

第一，构建了基于交易特征的自适应权重函数。借鉴 Feng（2023）^[10] 在高频数据线性回归中的数据驱动思想，本文设计了一种能够根据日内即时波动和交易特性动态调整的权重机制。该函数突破了传统 MIDAS 方法仅依赖时间序列排列或固定函数形式的局限，显著提升了从高频数据中剥离有效信息的效率。

第二，建立了引入自适应权重的 MIDAS-GARCH 模型。在此基础上，本文对经典 GARCH 模型进行了结构性拓展，构建了自适应权重 MIDAS-GARCH 模型。该模型实现了高频微观数据与低频宏观数据的深度耦合，能够更敏锐地捕捉高频交易数据的微观动态演变。实证结果表明，相较于传统固定权重模型，该模型在强化预测精准度和提升模型整体稳健性方面均具有显著优势。

一、模型及估计

（一）自适应权重原理

高频金融数据因非平稳性而呈现波动模式的时变特性，使传统固定权重方法难以适应。为此，本章基于 Feng（2023）提出的框架，通过数据驱动方式动态调整权重，以提升信息提炼效率。

对于平稳过程 $\{y_t, X_t\}$ ，当 $X_t^T X_t$ 的所有特征值均为正，可以证明一结论：对于固定样本量 n ，估计量 $\hat{\beta}$ 的方差迹 $tr[\text{var}(\hat{\beta})]$ 的下界可以表示为：

$$\frac{\text{var}(y_t | X_t) (p+1)^2}{E(X_t^T X_t) n} (1 + o_p(1)), \quad (5)$$

对标准的混频数据抽样模型进行简化处理。模型的基本形式为：

$$y_t = \beta_0 + \beta_1 \sum_{k=1}^m x_{t1,k/m} \omega_k + \beta_2 \sum_{k=1}^m x_{t2,k/m} \omega_k + \cdots + \beta_p \sum_{k=1}^m x_{tp,k/m} \omega_k + \varepsilon_{t,k/m}, \quad (6)$$

通过定义加权聚合后的低频变量 $x_{tq}^* = \sum_{k=1}^m x_{tq,k/m} \omega_k$ ($q=1,2,\dots,p$) 和 $\varepsilon_t^* = \sum_{k=1}^m \varepsilon_{t,k/m} \omega_k$ ，可以将复杂的混频模型简化为一个标准的低频回归模型：

$$y_t = \beta_0 + \beta_1 \sum_{k=1}^m x_{t1,k/m} \omega_k + \beta_2 \sum_{k=1}^m x_{t2,k/m} \omega_k + \cdots + \beta_p \sum_{k=1}^m x_{tp,k/m} \omega_k + \varepsilon_{t,k/m} \quad (7)$$

这一简化的关键在于，它将权重选择的问题从高频数据整合的框架，转化为了一个如何优化低频模型参数估计效率的问题。

权重的选择基于一个核心的计量经济学原理——提升参数估计的效率。理论分析表明，参数估计量 $\hat{\beta}$ 的方差迹 $tr[\text{var}(\hat{\beta})]$ 存在一个下界。通过最小化这个下界，可以获得更精确、更稳定的参数估计。

最终，权重优化的目标被转化为求解以下优化问题：

$$\omega = \arg \min_{\omega \in \mathcal{R}} \frac{\omega^T \hat{\Sigma}_\varepsilon \omega}{\omega^T \hat{\Sigma}_X \omega} \quad (8)$$

其中： $\hat{\Sigma}_X = \frac{1}{n} \sum_{t=1}^n X_t^T X_t$ 是高频解释变量样本协方差矩阵的估计， $\hat{\Sigma}_\varepsilon = \frac{1}{n} \sum_{t=1}^n \varepsilon_{t\omega_{t-1}}^T(u) \varepsilon_{t\omega_{t-1}}(u)$ 是误差项协方差矩阵的估计。

（二）自适应 MIDAS-GARCH 模型

1. 模型构建

在混合数据采样（MIDAS）模型的构建过程中，如何有效弥合高频变量与低频变量之间的频率差异，始终是一个关键挑战。传统 MIDAS 模型多依赖于固定函数形式或预设权重方案来聚合高频数据，这类方法虽简化了模型设定，但由于其权重结构缺乏对数据生成过程的动态响应，往往难以充分捕捉高频数据中所蕴含的丰富信息，从而在一定程度上制约了模型的估计精度与样本外预测表现。为突破这一局限，本研究借鉴 Feng（2023）提出的高频数据聚合框架，提出一种基于自适应权重机制的 MIDAS 建模新框架。该方法摒弃传统预设权重的思路，转而通过数据驱动的方式动态学习高低频数据间的最优关联结构，实现权重分配的自动优化。这一改进不仅增强了模型对复杂经济与金融动态的适应能力，也有望显著提升参数估计的准确性及预测的稳健性，为处理混频数据提供更为灵活与有效的建模途径。构建模型如下：

$$\begin{cases} y_t = \beta_0 + \sum_{q=1}^p \beta_q \sum_{k=1}^m \omega_k x_{tq,k/m} + \varepsilon_t, \\ \omega = \arg \min_{\omega \in \mathcal{R}} \frac{\omega^T \hat{\Sigma}_\varepsilon \omega}{\omega^T \hat{\Sigma}_X \omega}, \\ \hat{\Sigma}_X = \frac{1}{n} \sum_{t=1}^n X_t^T X_t, \\ \hat{\Sigma}_\varepsilon = \frac{1}{n} \sum_{t=1}^n \hat{\varepsilon}_t^T \hat{\varepsilon}_t. \end{cases} \quad (9)$$

其中, y_t 是低频被解释变量, β_i 为模型系数, ε_t 为模型误差项, m 表示混频数据倍差, 权重向量 $\omega = (\omega_1, \omega_2, \dots, \omega_m)'$, ω_k 是自适应权重, 满足 $\omega_k > 0$ 且 $\sum_{k=1}^m \omega_k = 1$, $k = 1, 2, \dots, m$ 。

2. 基于自适应权重的 MIDAS-GARCH 模型构建

为了充分挖掘高频交易数据在预测未来波动率方面的价值, 本文借鉴普通 MIDAS 及 MIDAS-GARCH 模型的设计思路, 在传统 GARCH 模型框架中引入了一种新的权重函数。该函数旨在提取高频交易数据中的关键信息, 并将其作为解释变量纳入模型, 具体构建形式如下:

$$\begin{cases} y_t = \sqrt{h_t} \varepsilon_t, \\ \omega = \arg \min_{\omega \in R} \frac{\omega^T \hat{\Sigma}_\varepsilon \omega}{\omega^T \hat{\Sigma}_x \omega}, \\ h_t = u + \sum_{i=1}^q \beta_i h_{t-i} + \sum_{j=1}^p \alpha_j \left(\sum_{k=1}^m \omega_{t-j,k} y_{t-j,k/m}^2 \right), \\ \hat{\Sigma}_x = \frac{1}{n} \sum_{t=1}^n X_t^T X_t, \\ \hat{\Sigma}_\varepsilon = \frac{1}{n} \sum_{t=1}^n \hat{\varepsilon}_t^T \hat{\varepsilon}_t. \end{cases} \quad (10)$$

其中, y_t 表示日间收益率, $\{y_{t,k/m}\}$ 是日内高频收益率序列在第 t 天第 k 个观测时刻的值, h_t 为收益率的条件方差。模型包含未知参数 u 、 β_i 和 α_j , 通常假定为非负常数, 模型的滞后阶数 p 和 q 可通过 AIC、BIC 等信息准则来确定。 ε_t 为模型误差项, m 表示混频数据倍差, 权重向量 $\omega = (\omega_1, \omega_2, \dots, \omega_m)'$, ω_j 是自适应权重, 满足 $\omega_j > 0$ $\sum_{j=1}^m \omega_j = 1$ 且, $j = 1, 2, \dots, m$ 。 $\{x_{t,j,k/m}\}$ 是第 j 个高频解释变量在时间 t 的第 k 个观测值, $\hat{\varepsilon}_t$ 为模型误差项的估计量, $\hat{\Sigma}_x$ 是高频解释变量的样本协方差矩阵, $\hat{\Sigma}_\varepsilon$ 是误差项协方差矩阵。

$\{x_{t,j,k/m}\}$ 是与高频收益序列 $\{y_{t,k/m}\}$ 相伴而生的高频数据。通过构建自适应权重函数, 波动率模型将高频数据 $x_{t,k/m}$ 、 $y_{t,j/m}$ 与低频数据 y_t 结合, 利用多源混频信息来预测未来的波动率。为简化计算过程, 本文仅选取一个高频解释变量 $\{x_{t,k/m}\}$ 计算权重。金融市场高频交易环境会生成多种高频实时交易数据如高频成交量、高频收益率等, 这些均可作为 $\{x_{t,k/m}\}$ 潜在的候选变量, 鉴于在条件异方差模型中, 平方收益率能够高效地捕捉波动率的动态变化, 因此, 本文将 $\{y_{t,k/m}^2\}$ 序列选定为 $\{x_{t,k/m}\}$, 以期在波动率预测中发挥更加精准和有效的作用。

(三) 参数估计

本文借鉴 Han 和 Park (2012)^[11] 关于普通 GARCH 模型的参数估计理论框架证明推导过程, 选择拟极大似然估计法对基于自适应权重的 MIDAS-GARCH 模型进行参数估计。

1. 拟极大似然估计

本文参考普通 GARCH 模型参数估计的流程, 针对所提出的自适应权重 MIDAS-GARCH 模型, 采用拟极大似然估计进行参数估计, 模型对数似然密度函数为:

$$Ln(\theta) = -\frac{1}{2} \sum_{t=1}^n \left[\log(h_t) + \frac{y_t^2}{h_t} \right]. \quad (11)$$

其中, h_t 是模型的条件方差, $\theta = (u, \alpha_1, \dots, \alpha_p, \beta_1, \dots, \beta_q, \omega)$ 是参数向量, 记真实参数为 θ_0 , $\hat{\theta}_n = \operatorname{argmax}_{\theta \in \Theta} Ln(\theta)$ 是参数的估计量, Θ 表示完整参数空间。从形式上看, 自适应权重 MIDAS-GARCH 模型的似然函数与传统 GARCH 模型的似然函数极为相近。然而, 实际上, 模型中的 h_t 包含了复杂的参数结构。因此, 为了获得估计量 $\hat{\theta}_n$ 的大样本性质, 需要引入如下条件:

① 模型参数应满足: $0 < \underline{k} < \bar{k}$, $\underline{k} \leq \min(u, \alpha_1, \alpha_2, \dots, \alpha_p, \beta_1, \beta_2, \dots, \beta_q) \leq \bar{k}$, $\beta_1 + \beta_2 + \beta_3 + \dots + \beta_q = \rho < 1$, $q\underline{k} < \rho$, $0 < \lambda < \bar{\lambda}$, 其中 $\underline{k}$ 、 $\bar{k}$ 、 ρ 和 $\bar{\lambda}$ 为已知常数;

② 高频收益率时间序列 $y_{t,j/m}$ 平稳遍历, 并满足 $E(y_{t,j/m}^2) < \infty$;

③ 独立同分布的误差项 ε_t 为平稳遍历并满足 $E(\varepsilon_t | F_{t-1}) = 0$, $D(\varepsilon_t | F_{t-1}) = 1$, 其中 F_{t-1} 为 $t-1$ 时刻的信息集;

④ 随机序 $\{y_{t,j/m} | F_{t-1}\}$ 列为非退化分布;

以上设定与普通 GARCH 模型中的假设相似, 不仅在逻辑上自洽, 更在数学严谨性层面构建了参数估计的底层框架: 首先, 通过约束参数空间 Θ 的紧性, 条件①不仅利用 Bolzano-Weierstrass 定理确保了似然函数全局最优解的存在性, 还通过对截断值的限定从根本上保障了条件方差函数的非负性, 从而有效避免了数值迭代过程中可能出现的逻辑失效; 在此基础上, 条件②作为计量经济学中不可或缺的识别性假设, 确保了参数与分布映射的唯一性, 为证明估计量的一致性扫清了障碍; 进一步而言, 针对高频收益率序列施加的平稳遍历性假设(条件③), 与实际金融市场中典型的“波动率聚集”及“均值回归”特征高度契合, 这使得大数定律与中心极限定理在统计推断中的应用具备了理论合法性; 最后, 由于条件④与标准 GARCH 模型假设保持了高度的内在一致性, 这确保了模型在处理不同频率数据时均能保有拟极大似然估计(QMLE)的渐近正态性。综上所述, 这些约束条件并非孤立存在, 而是环环相扣, 共同为后续参数估计的一致性与统计稳健性奠定了坚实的理论基石。

2. 一致估计量:

当拟极大似然估计量 $\hat{\theta}_n$ 满足条件①至④时, 容易证明: 存在常数 $\delta > 0$ 使得 $E(|\varepsilon_0^{t+\delta}|) < \infty$, $E(\varepsilon_0^2) = 1$, $\lim_{t \rightarrow \infty} t^{-u} P\{\varepsilon_0^2 \leq t\} = 0$, $u \geq 0$, 那么拟极大似然估计量 $\hat{\theta}_n$ 为真实参数 θ 的一致估计量, 即:

$$\hat{\theta}_n \xrightarrow{p} \theta, \text{ as } n \rightarrow \infty. \quad (12)$$

3. 渐进正态性:

若条件①-④成立, 且存在某个常数 $\delta > 0$, 使得 $E(y_{t,j/m}^{t+\delta}) < \infty$, 依据 Han 和 Kristense 的结论, 结合极大似然估计的渐近正态性, 可以证明: 那么拟极大似然估计量 $\hat{\theta}_n$ 满足:

$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{d} N(0, B_0^{-1} A_0 (B_0^{-1})^T), n \rightarrow \infty. \quad (13)$$

其中,

$$A_0 = \operatorname{cov}(\ell'_0(\theta)), B_0 = E(\ell'_0(\theta)) \quad \ell'_0(\theta) = -\frac{1}{2} \left[\frac{h'_0(\theta)}{h_0(\theta)} - \frac{y_0^2 h'_0(\theta)}{h_0^2(\theta)} \right].$$

二、模拟研究

(一) Sieve-Bootstrap 方法

本文借鉴 Buhlman 提出的 Sieve-Bootstrap 方法^[12]，以适应数据间的相依结构，并估计渐近协方差。

Sieve-Bootstrap 方法利用误差项的独立同分布属性，借助迭代程序来估算协方差矩阵，其具体实施流程如下：

①用数据 $\{y_t\}$ 、高频率列 $\{x_{t,k/m}\}$ 、 $\{y_{t,j/m}\}$ 得到参数估计 $\hat{\theta}_l$ 和残差 $\hat{\varepsilon}_t$ ；

②对 $\hat{\varepsilon}_t$ 抽样，结合 $\hat{\theta}_{l-1}$ 和高频率列算出低频收益率 $\{\tilde{y}_t\}$ ，再估计新参数 $\hat{\theta}_l$ ；

③重复步骤 2 生成参数序列 $\{\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_n\}$ ，据此计算样本协方差矩阵，作为估计量的渐近协方差阵。

(二) 生成模拟数值

本文构建了两组具有不同特征的模拟数据以开展相关检验。首组模拟数据的生成方式如下：

$$\begin{cases} y_t^* = 0.6y_{t-1}^* + 0.3y_{t-2}^* + \varepsilon_{1t}, \\ \hat{\Sigma}_x = \frac{1}{n} \sum_{t=1}^n X_t^T X_t, X_t = \begin{pmatrix} 1 & x_{t,1/m} \\ 1 & x_{t,2/m} \\ \vdots & \vdots \\ 1 & x_{t,m/m} \end{pmatrix}, \\ \omega = \arg \min_{\omega \in R} \frac{\omega^T \hat{\Sigma}_x \omega}{\omega^T \hat{\Sigma}_x \omega}, \omega = (\omega_1, \omega_2 \cdots \omega_m)', \\ h_t = 0.2 + 0.6h_{t-1} + 0.3 \left(\sum_{k=1}^m \omega_k y_{t-1,k/m}^2 \right), y_t = h_t^{1/2} \eta_{1t}. \end{cases} \quad (14)$$

其中， $\{y_t^*\}$ 代表高频收益率，首个等式用 AR(2) 生成高频收益率数据，误差项 $\varepsilon_{1t} \sim N(0, 0.02)$ 。考虑到实际交易场景，将参数 m 设为 240，对应每分钟产生的高频交易数据量。第二、三个等式为自适应权重函数，解释变量 $x_{t,k/m}$ 构建方式如下：

$$x_{t,k/m} = 0.3 + 5 \cdot |y_{t,k/m}| + \varepsilon_{1t}^* \quad (15)$$

其中 $\varepsilon_{1t}^* \sim N(0, 0.01)$ 。

对于第二组模拟数据，其构建方式如下所示：

$$\begin{cases} y_t^* = 0.6y_{t-1}^* + \varepsilon_{2t} + 0.3\varepsilon_{2t-1}, \\ \hat{\Sigma}_x = \frac{1}{n} \sum_{t=1}^n X_t^T X_t, X_t = \begin{pmatrix} 1 & x_{t,1/m} \\ 1 & x_{t,2/m} \\ \vdots & \vdots \\ 1 & x_{t,m/m} \end{pmatrix}, \\ \omega = \arg \min_{\omega \in R} \frac{\omega^T \hat{\Sigma}_x \omega}{\omega^T \hat{\Sigma}_x \omega}, \omega = (\omega_1, \omega_2 \cdots \omega_m)', \\ h_t = 0.1 + 0.5h_{t-1} + 0.3 \left(\sum_{k=1}^m \omega_k y_{t-1,k/m}^2 \right), y_t = h_t^{1/2} \eta_{2t}. \end{cases} \quad (16)$$

其中，误差项 $\varepsilon_{2t} \sim N(0, 0.03)$ ， $\eta_{2t} \sim \frac{t(10)}{\sqrt{1.25}}$ 。 $t(10)$ 是自由度为 10

的 t 分布，为使 η_{2t} 方差为 1， η_{2t} 除以 $\sqrt{1.25}$ ，在第二组模拟数据中，用于自适应权重函数的解释变量 $x_{t,k/m}$ 生成方式为：

$$x_{t,k/m} = 1 + 0.5e^{|y_{t,k/m}|} + \varepsilon_{2t}^* \quad (17)$$

其中，误差项满足 $\varepsilon_{2t}^* \sim N(0, 0.01)$ 。

对于上述两种模拟数据产生机制，本文考虑不同情形：收益率数据的生成机制、条件异方 h_t 差的参数设置、日收益率误差项的分布、伴随高频率列 $\{x_{t,k/m}\}$ 的生成方式。

(三) 参数估计结果分析

数值模拟中，设定不同样本量评估其对参数估计的影响，用梯度下降法结合 MATLAB 实现估计，通过偏差 (BIAS，越小越准)、标准误差 (SE，越低越稳)、标准差 (SD，越小越稳定) 三个指标，综合分析参数估计的统计性质与有效性，最终得到参数估计结果。如表 1、表 2。

表 1：误差项服从正态分布时的参数估计结果
Tab.1: Parameter Estimation Results with Normally Distributed Error Terms

交易日 n	指标	$\mu=0.2$	$\alpha=0.3$	$\beta=0.6$
n=100	BIAS	0.000794	-0.009495	-0.009163
	SE	0.002234	0.027992	0.058677
	SD	0.001670	0.021664	0.066401
n=300	BIAS	0.001026	0.033723	0.027165
	SE	0.001363	0.056969	0.175007
	SD	0.000884	0.035465	0.203329
n=500	BIAS	0.000356	0.023342	0.045427
	SE	0.002130	0.018145	0.030406
	SD	0.002639	0.018792	0.035033

表 1 详尽展示了在误差项服从标准正态分布 $N(0,1)$ 的基准情形下，模型参数的数值模拟估计结果。表首行明确标注了条件方差模型中各项参数的理论真实值，随后的各行则依次呈现了在不同模拟试验次数 n (样本容量) 设定下，模型捕捉真实参数能力的各项评估指标。

从估计精度的绝对与相对维度来看：

数值实验结果显示，在不同样本规模下，参数估计值的最大绝对偏差波动范围为 0.001026 至 0.045427，而最小绝对偏差则低至 0.000356 至 0.009495。若进一步消除参数量纲的影响，观察相对偏差可以发现：其最大值分布在 0.513% 至 11.241% 之间，最小值则仅为 0.178% 至 3.165%。这一极窄的偏差区间有力地证明了 Sieve-Bootstrap 方法在估计模型参数时具备极高的一致性，即便在有限样本约束下，该方法依然展现出极强的抗干扰能力与数值稳定性，能够精准锁定参数的真实分布区域。

从统计推断的稳健性维度来看：

尤为关键的一点在于，实验观察到的模拟标准差 (SD) 与通过理论公式计算得到的平均标准误差 (SE) 在数值上高度趋同。在计量经济学中，这种 SD 与 SE 的近乎重合并非巧合，而是渐近理论有效性的关键判准：

SD 反映了估计量在多次重复实验中的实证离散程度；

SE 则代表了基于本文所推导的渐近协方差矩阵所计算出的理论精度。

两者的一致性不仅深度验证了本文所给出的渐近协方差表达式在数学推导上的准确性，更揭示了 Sieve-Bootstrap 方法在捕捉样本波动性特征时的优异表现。这意味着该方法不仅能给出准确的点估计，更能提供可靠的区间估计与假设检验基础，证明了其在复杂金融时间序列分析中的理论价值与应用可靠性。

表2: 误差项服从 t 分布时的参数估计结果
Tab.2: Parameter Estimation Results with Student's t -distributed Error Terms

交易日 n	指标	$\mu=0.1$	$\sigma=0.3$	$\beta=0.5$
$n=100$	BIAS	-0.001079	-0.014079	-0.009359
	SE	0.001141	0.017181	0.061844
	SD	0.002057	0.028018	0.071254
$n=300$	BIAS	-0.000159	-0.006573	-0.025573
	SE	0.001034	0.057272	0.176025
	SD	0.000655	0.033149	0.201649
$n=500$	BIAS	-0.000139	-0.010620	-0.056298
	SE	0.001580	0.015979	0.028521
	SD	0.001322	0.012784	0.033747

表2进一步详细展示了当误差项服从特定非正态分布时，模型参数估计的数值模拟表现。该表结构与表1保持一致：表首行明确列示了条件方差模型中各参数的理论真值，随后的各行则记录了在不同样本容量 n 设定下，通过 Sieve-Bootstrap 方法得到的各项统计评估指标。

从偏差分布与估计精度来看：

实验结果表明，在不同的样本规模约束下，模型参数估计值的绝对偏差处于 0.001079 至 0.056298 这一较低区间，对应的最大相对偏差被严格限制在 1.08% 至 11.26% 之间；而最小绝对偏差仅为 0.000139 至 0.009359，对应的最小相对偏差低至 0.139% 至 2.19%。

尽管在特定小样本情形下，估计结果呈现出轻微的系统性偏差，但这在有限样本统计推断中属于预期的理论范畴。关键在于，随着样本容量 n 的逐步扩大，估计偏差表现出明显的收敛迹象。这一发现有力地证实了 Sieve-Bootstrap 方法在处理非标准分布误差项时具备优异的参数捕捉能力，即便在有限样本条件下，依然能维持极佳的估计准确性与性能稳定性。

从协方差矩阵的估计有效性来看：

尤为值得关注的是，实验所得的模拟标准差（SD）与基于理论推导得到的标准误差（SE）在数值上高度契合。这种 SD 与 SE 的一致性（Coherence）提供了双重理论证据：一方面，它深度验证了本文所构建的渐近协方差表达式在复杂分布下的推导正确性；另一方面，它表明该方法能够精确测算参数估计的不确定性。这种对参数协方差矩阵的有效估计，不仅提升了模型在非正态环境下的统计推断效率，也为后续的假设检验与风险测度提供了稳健的量化基础。

三、实证研究

为说明自适应权重函数对高频数据有效信息的提取优势，验证基于该权重的 MIDAS-GARCH 模型相较于传统模型的预测精

度优越性，本实证选用上证综合指数 2019 年 1 月 2 日至 2024 年 12 月 2 日（1434 个交易日）的 30 分钟高频收盘价数据（共 11472 个观测值）为研究样本。数据划分为训练集（2019 年 1 月 2 日至 2024 年 7 月 4 日，1334 个交易日，用于模型训练与参数估计）和测试集（2024 年 7 月 5 日至 2024 年 12 月 2 日，100 个数据，用于样本外预测验证），兼顾训练样本量与模型泛化性能评估。选取标准 GARCH、GARCH-RV、传统 MIDAS-GARCH 及自适应权重 MIDAS-GARCH 模型进行对比^[13]，分析其波动率分析与预测表现。

（一）平稳性检验

金融数据回归分析中，平稳性检验是必要步骤（需满足时间序列平稳性假设，避免伪回归、提升预测准确性）。建模前对上证综合指数收益率序列进行 ADF 单位根检验，结果显示 ADF 值为 -10.602（ $P < 0.01$ ），表明序列显著平稳、无单位根，可开展后续模型实证分析。

（二）ARCH 效应检验

ARCH 效应检验用于检测时间序列波动聚集现象，是判断是否适用 ARCH/GARCH 模型（尤其金融领域捕捉异方差）的关键。上证综合指数收益率序列虽平稳，但存在显著波动集聚效应（大幅波动后随大幅波动、小幅波动后随小幅波动），提示可能存在条件异方差，后续将通过自相关函数（ACF）和偏自相关函数（PACF）进一步验证。

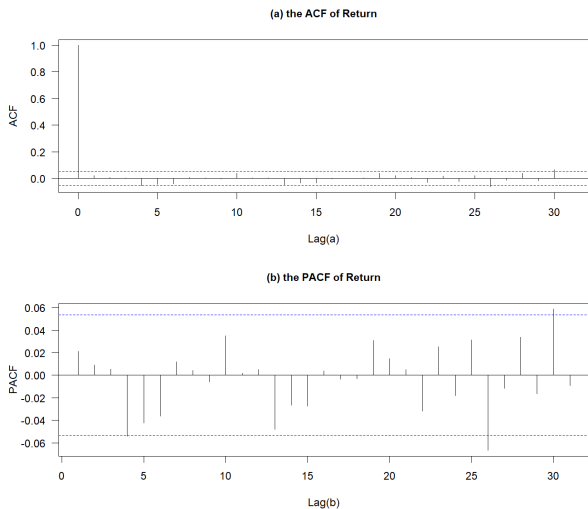


图1：收益率序列的 ACF 图和 PACF 图

Fig.1: ACF Plot and PACF Plot of the Yield Rate Series

图1显示，上证综合指数收益率序列的自相关和偏自相关系数多在置信区间内波动，自相关性弱（近乎无显著自相关性），因此均值方程可采用“常数项 + 随机扰动项”形式，符合 GARCH 模型设定要求。

尽管收益率序列弱相关，但需检验其平方序列自相关性（GARCH 模型核心假设为方差序列具有序列相关性，收益率平方自相关性是重要体现）。图2显示，收益率平方序列的 ACF 图具有显著自相关性持续性，验证了方差序列的序列相关性，满足 GARCH 模型适用的关键条件。

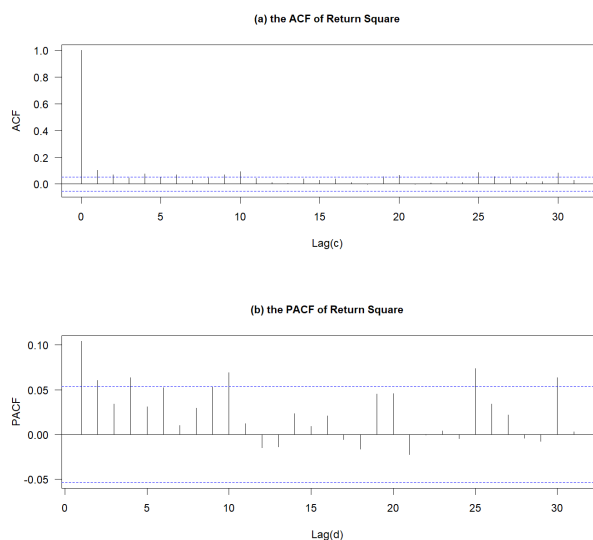


图2：收益率平方序列的 ACF 图和 PACF 图

Fig.2: ACF and PACF Plots of the Squared Return Series

本文通过 LM 检验验证 ARCH 效应（原假设为残差无 ARCH 效应），结果显示检验统计量为 42.371，P 值远小于 1%，故推翻原假设，证实收益率序列存在显著 ARCH 效应，表现为波动聚集特征（高 / 低波动时期集中出现）。综上，上证综合指数收益率序列适合用 GARCH 模型拟合。

（三）MIDAS 模型下不同权重对比

为验证自适应权重的优越性，本文构建 MIDAS 模型并对比不同权重函数。鉴于传统 MIDAS 忽视因变量动态特性，而金融波动率常具自回归特征，本文采用引入 ADL 项的 ADL-MIDAS 模型以有效捕捉该特性。模型表示为：

$$y_t = \alpha + \sum_{i=1}^p \rho_i y_{t-i} + \sum_{j=1}^J \beta_j \sum_{k=1}^m \omega_j x_{tj,k/m} + \varepsilon_{t0} \quad (18)$$

其中， $\{y_t\}$ 为上证综合指数日收益率， $\{x_{t-j,k/m}\}$ 为上证综合指数每 30 分钟高频收益率， $\sum_{i=1}^p \rho_i y_{t-i}$ 表明低频数据 $\{y_t\}$ 的自回归项。

本节实证基于 ADL-MIDAS 模型展开。为了验证本文构建的自适应权重函数的有效性，将其与第二章详述的多种经典权重形式进行对比分析。具体而言，选取了 Beta 权重（含非零形式）、Almon 权重（含指数形式）、无约束 MIDAS 多项式（U-MIDAS）、Step 权重及 Legendre 权重分别代入模型进行回归测试。

表3：不同权重下样本外预测结果比较

Tab.3: Comparison of Out-of-Sample Prediction Results Under Different Weighting Methods

权重	v	MSFE	DMSFE
Beta	1.62519	2.64124	0.02008
非零 Beta	1.62814	2.65085	0.02302
Almon	1.63615	2.67700	0.02113
指数 Almon	1.63633	2.67758	0.02146
U-MIADS	1.62657	2.64574	0.02073
Step	1.63220	2.66406	0.01993
Legendre	1.64185	2.69566	0.02101
自适应权重	1.53472	2.35536	0.01942

如表3所示，本文选取了三种指标来评估模型性能。首先，

利用均方根误差（RMSE）来衡量预测结果的整体精确性，其计算方式为预测偏差平方均值的开方；其次，采用均方预测误差（MSFE）来刻画模型的整体准确度，即预测误差的平方均值；最后，考虑到时间维度的影响，引入了折扣均方预测误差（DMSFE）。该指标是对 MSFE 的改进，通过设定折扣因子对近期误差赋予更高权重，从而侧重于考察模型对近期市场动态的预测能力。具体公式如下：

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (\hat{y}_t - y_t)^2}, \quad (19)$$

$$MSFE = \frac{1}{n} \sum_{t=1}^n (\hat{y}_t - y_t)^2, \quad (20)$$

$$DMSFE = \frac{1}{n} \sum_{t=1}^n \lambda^{n-t} (\hat{y}_t - y_t)^2. \quad (21)$$

其中， $\hat{y}_t$ 为预测值， y_t 为真实值， n 为样本数量， λ 为折扣因子，通常 $0 < \lambda < 1$ ，本实证中模型为了平衡近期和远期误差，既重视近期的关键信息，又不完全忽视远期的影响，取常见折扣因子经验值 $\lambda = 0.85$ 。

表3显示，自适应权重函数在 RMSE（1.5347）、MSFE（2.3554）及 DMSFE（0.0194）三项指标上均达到最优。相较于表现次优的 Beta 权重，其 RMSE 与 MSFE 分别下降了 5.6% 和 10.8%，有力证实了动态权重机制在提升预测精度与稳定性方面的优势。

在传统权重对比中：标准 Beta 权重表现尚可，而非零 Beta 因稀疏约束导致部分有效信息损失，误差微增；Almon 类（含指数形式）与 Legendre 权重受限于固定多项式或基函数设定，难以适应复杂时序，后者更是位列末席；U-MIDAS 处于中游水平；Step 权重虽在折扣误差（DMSFE）上逼近自适应权重，但整体精度差距依然显著，说明分段静态分配无法替代动态自适应。

综上所述，传统权重因固定数学形式的局限，难以充分应对上证指数波动率的复杂性。相比之下，自适应权重凭借数据驱动的灵活性，能更精准地捕捉动态滞后效应，从而显著降低预测偏差，增强了 ADL-MIDAS 模型的实证适应性。

（四）基于自适应权重的 MIDAS-GARCH 模型及其对比模型

1. 基准模型与各模型参数估计结果

为验证本文提出的自适应权重 MIDAS-GARCH 模型的有效性，本章选取了三类经典模型构建对照组。首先，采用满足参数平稳性约束即 $a_1, b_1 \geq 0, a_1 + b_1 < 1$ 的 GARCH(1,1) 作为标准低频基准。其次，引入 GARCH-RV 模型，该模型通过对第 t 日内 m 个高频收益率的平方进行简单加总（即 $RV_t = \sum_{k=1}^m y_{t,k/m}^2$ ），实现了高频信息的等权重聚合。最后，纳入传统 MIDAS-GARCH 模型，其特点在于利用预设的权重函数对高频数据实施差异化的加权处理。通过将上述模型与自适应权重模型进行横向实证对比，旨在甄别出最高效的高频信息提取路径，从而提升波动率预测的精度

与模型的动态适应性。

各模型基于训练集的参数估计结果为：

标准 GARCH 模型：

$$h_t = 0.051418 + 0.844815h_{t-1} + 0.113214y_{t-1}^2, \quad (22)$$

(0.015746) (0.026825) (0.019534)

GARCH-RV 模型：

$$h_t = 0.000813 + 0.844486h_{t-1} + 0.113330RV_{t-1}, \quad (23)$$

(0.000249) (0.026943) (0.019885)

传统 MIDAS-GARCH 模型：

$$h_t = 0.051542 + 0.844725h_{t-1} + 0.113136 \sum_{k=1}^8 \phi(k) y_{t-1,k/m}^2, \quad (24)$$

(0.015781) (0.026894) (0.019837)

基于自适应权重的 MIDAS-GARCH 模型：

$$h_t = 0.003650 + 0.863839h_{t-1} + 0.089446 \sum_{k=1}^8 \omega_{t-1,k} y_{t-1,k/m}^2. \quad (25)$$

(0.000699) (0.017559) (0.014975)

模型参数的滞后阶数均依据显著性检验与似然比统计量确定，括号内数值为标准误差。结果显示，在 5% 显著性水平下，所有参评模型（标准 GARCH、GARCH-RV、传统 MIDAS 及自适应 MIDAS）的参数估计均显著。

在参数特征上：GARCH-RV 与自适应 MIDAS 模型的常数项均明显偏低，说明引入已实现波动率或优化权重机制后，模型对截距项的依赖减弱，高频信息的解释力增强；而传统 MIDAS 模型的参数分布则与标准 GARCH 高度趋同，显示其权重机制未带来结构性的参数偏移。

在估计精度（标准误差）上：自适应 MIDAS 模型表现最佳，其各项标准误差显著低于其他模型，证实了动态权重在提升灵活性的同时兼具极高的估计稳定性。相比之下，GARCH-RV 虽在常数项估计上较精准，但波动率项的误差偏大，增加了不确定性；传统 MIDAS 则未表现出超越标准 GARCH 的稳健性。

综上所述，自适应 MIDAS-GARCH 模型凭借最低的常数项依赖和最小的估计误差，展现了最优的参数性质。相较于传统模型对高频信息利用的不足，以及 GARCH-RV 伴随的估计不确定性，自适应模型成功实现了预测准确性与稳健性的双重提升。

2. 结果分析

为客观衡量各模型的预测效能，本研究构建了横向对比分析框架。具体实施中，采用已实现核估计量（RK）作为市场真实波动率的有效代理变量（Proxy），通过考察各模型预测值与该基准值的拟合偏差，来校验其预测精度。具体实证结果如下：

表 4：四种模型预测波动率结果
Tab.4: Volatility Forecasting Results of Four Models

指标	标准 GARCH	GARCH-RV	传统 MIDAS-GARCH	基于自适应权重的 MIDAS-GARCH
MSE	0.3322332	0.3525111	0.3316859	0.2454092
MAE	0.5283462	0.4632915	0.5278649	0.3442092
MSPE	3.7888206	0.5171152	3.7829057	0.2712952
MAPE	1.4490709	0.6988529	1.4477666	0.4731687
R ² LOG	0.8359773	2.1283677	0.8351287	0.7294802

本节选取了四种统计量来量化预测偏差。其中，均方误差（MSE）利用平方机制对极端偏差进行“惩罚”，从而敏锐地捕捉预测结果的离散性；平均绝对误差（MAE）则通过计算偏差绝对值的均值，直观反映了预测值偏离真实水平的整体程度。在相对指标方面，均方百分比误差（MSPE）从相对误差的视角出发，以百分比形式评估偏差大小；而平均绝对百分比误差（MAPE）因具备优良的解释性，常被用于衡量预测误差的平均相对幅度。具体计算公式如下：

$$MSE = \frac{1}{n} \sum_{t=1}^n (\hat{h}_t - h_t)^2, \quad (26)$$

$$MAE = \frac{1}{n} \sum_{t=1}^n \left| \hat{h}_t - h_t \right|, \quad (27)$$

$$MSPE = \frac{1}{n} \sum_{t=1}^n \left(\frac{\hat{h}_t - h_t}{h_t} \right)^2, \quad (28)$$

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left(\left| \frac{\hat{h}_t - h_t}{h_t} \right| \right), \quad (29)$$

$$R^2 LOG = \frac{1}{n} \sum_{t=1}^n \left(\log(h_t) - \log(\hat{h}_t) \right)^2. \quad (30)$$

其中， $\hat{h}_t$ 代表模型得出的波动率预测结果， h_t 为真实波动率预测结果。

表 4 的对比结果显示，自适应权重 MIDAS-GARCH 模型在预测精度与稳健性上均占据绝对优势。

在精度方面（MAE/MAPE）：自适应模型表现最佳，相较于基准 GARCH，其 MAE 和 MAPE 分别大幅下降 34.85% 和 67.35%。GARCH-RV 模型次之（降幅分别为 12.31% 和 51.77%），验证了引入高频信息的有效性。相比之下，传统 MIDAS-GARCH 模型因受限于固定权重，改善幅度微乎其微（不足 1%），仅略优于基准模型。这表明，单纯引入高频数据（传统 MIDAS）或等权聚合（RV）的效果，均不及动态优化的自适应机制。

在稳健性方面（MSE/MSPE/R²LOG）：自适应模型同样表现最稳健，三项指标较基准分别降低 26.13%、92.84% 和 12.74%。值得注意的是，GARCH-RV 模型虽在精度上占优，但在稳健性上表现不稳定，其 MSE 和 R²LOG 反而分别上升了 6.10% 和 155%。

综上所述，自适应权重模型通过灵活的信息融合机制，有效克服了传统固定权重和简单等权聚合的局限，实现了对高频波动特征的精准捕捉，从而在保证预测精度的同时，显著增强了模型的稳健性。

四、结束语

针对混频数据建模中固定权重函数存在的信息错配与僵化问题^[14]，本文构建了引入自适应权重机制的 MIDAS-GARCH 模型，

并基于拟极大似然估计（QMLE）框架推导了参数估计的一致性与渐近正态性。通过上证综合指数高频数据的实证研究及数值模拟发现：

第一，自适应权重函数能够根据日内波动特征动态调整高频信息权重，克服了 Beta、Almon 等传统固定权重的局限性。理论证明与模拟结果均证实，该模型在大样本下具备良好的估计精度与统计稳健性。

第二，在预测效能上，该模型在均方误差（MSE）及平均绝对误差（MAE）等指标上均显著优于标准 GARCH、GARCH-RV

及传统 MIDAS-GARCH 模型。这表明，数据驱动的动态权重机制能更敏锐地捕捉短期微观冲击与长期趋势，显著提升了波动率预测的准确度与稳健性。

尽管本文证实了自适应权重机制在单一股票市场分钟级数据中的有效性，但其在跨资产类别（如外汇、加密货币）、超高频（秒级 /Tick 级）及多因子框架下的适用性仍有待进一步验证。未来的研究可在此框架基础上，融合宏观经济与微观交易行为指标，以构建更具普适性的波动率预测体系，为金融风险管理提供更精准的决策支持。

参考文献

[1]Engle R F. Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of U.K. Inflation[J]. Econometrical, 1981, 50(04): 987-1008.

[2]Bollerslev, Tim. Generalized autoregressive conditional heteroskedasticity[J]. Economics and Econometrics Research Institute (EERI), 1986, (31): 307-327

[3]Taylor S J. Modelling financial time series[J]. John Wiley & Sons, 1986: 296.

[4]Zhang, L., Mykland, P. A., & Ait-Sahalia, Y. A Tale of Two Time Scales: Determining Integrated Volatility With Noisy High-Frequency Data[J]. Journal of the American Statistical Association, 2005(100), 1394 - 1411.

[5]Barndorff - Nielsen O. E, Reinhard H P, Lunde A, et al. Realized kernels in practice: trades and quotes[J]. Econometrics Journal, 2009, 12(03): C1 - C32.

[6]Asgharian H, Hou A J, Javed F. The Importance of the Macroeconomic Variables in Forecasting Stock Return Variance: A GARCH - MIDAS Approach[J]. Journal of Forecasting, 2013, (32): 600-612.

[7] 郑挺国, 尚玉皇. 基于宏观基本面的股市波动度量与预测 [J]. 世界经济, 2014(12): 118-139.

[8]Engle R F, Ghysels E, Sohn B. Stock Market Volatility and Macroeconomic Fundamentals [J]. Review of Economics Statistics, 2013, (95): 776797.

[9]ALMON S. The Distributed Lag Between Capital Appropriations and Expenditures[J]. Econometrica, 1965, 33: 178-196.

[10]Feng, W., Zhang, X., Chen, Y., & Song, Z. Linear regression estimation using intraday high frequency data[J]. AIMS Mathematics, 2023, 8(06): 13123-13133.

[11]Han H, Park J Y. ARCH/GARCH with persistent covariate: Asymptotic theory of MLE[J]. Journal of Econometrics, 2012, 167(01): 95-112.

[12]Buhlmann P. SieveBootstrap for Time Series[J]. Bernoulli, 1997, (03): 123-148.

[13] 李莉丽, 张兴发, 李元, 等. 基于高频数据的日频 GARCH 模型估计 [J]. 广西师范大学学报 (自然科学版), 2021, 39(4) : 68-78.

[14] 许敏. 基于自适应权重函数的 MIDAS-GARCH 模型及其应用研究 [D]. 广州: 广州大学, 2025. DOI: 10.27040/d.cnki.ggzdu.2025.001976.

山西省二手交易市场现状分析

原缓缓, 李雨航, 史华杰, 杨盼

运城学院 数学与信息技术学院, 山西 运城 044000

DOI:10.61369/ASDS.2026040012

摘 要： 习近平同志强调：“要建立绿色低碳循环经济体系，把绿色发展理念贯穿到经济社会发展的各方面和全过程”。山西省正处在能源转型升级的关键时期，绿色循环经济的发展迫在眉睫，其中二手交易市场扮演着不可或缺的角色。年轻群体通过“闲鱼”、“校园圈”流转教材，高净值人群在红布林等平台年交易奢侈品超10亿元，显示出消费分层特征。《十四五循环经济发展规划》等文件推动交易规范化，但实践中商品质检标准缺失、售后维权困难等问题普遍存在。通过调查山西省二手市场买家与卖家的基本情况，分析消费者收入变化、对二手商品信任度、平台服务质量、环保意识等因素对交易频率的影响，并从市场、商家和政策引导结合方面对二手交易市场提出可行性建议。

关 键 词： 低碳绿色发展；循环经济；二手交易市场；Logistic 回归模型

Analysis of the Current Situation of the Second-hand Trading Market in Shanxi Province

Yuan Huanhuan, Li Yuhang, Shi Huajie, Yang Pan

School of Mathematics and Information Technology, Yuncheng University, Yuncheng, Shanxi 044000

Abstract： Comrade Xi Jinping has stressed that "To establish a green and low-carbon circular economy system, the concept of green development should be integrated into all aspects and throughout the entire process of economic and social development." Shanxi Province is currently in a critical period of energy transformation and upgrading. The development of green circular economy is extremely urgent, and the second-hand trading market plays an indispensable role in this process. Young people transfer textbooks through platforms like "Xianyu" and "Campus Circle", while high-net-worth individuals conduct annual transactions of luxury goods worth over 1 billion yuan on platforms such as "Hongbiling", demonstrating the characteristics of consumption stratification. Documents such as the "14th Five-Year Plan for Circular Economy Development" promote the standardization of transactions, but problems such as the lack of commodity quality inspection standards and difficulties in after-sales protection are widespread in practice. By conducting a survey on the basic information of buyers and sellers in the second-hand market in Shanxi Province, analyzing the impact of factors such as changes in consumers' income, trust in second-hand goods, platform service quality, and environmental awareness on transaction frequency, and proposing feasible suggestions for the second-hand trading market from the perspectives of market, merchants, and policy guidance.

Keywords： low-carbon green development; circular economy; second-hand trading market; Logistic regression model

引言

近年来，我国网络零售^[1]快速和二手市场^[2, 3]快速发展，导致闲置商品市场规模突破万亿。作为绿色循环消费的重要实践载体^[4]，二手商品交易不仅承载着资源再生的经济价值，更蕴含着生态文明建设的社会意义。2024年国务院办公厅进一步要求建设集中规范的二手市场，制定网络交易管理办法^[5]，强化信息安全保障，明确电子产品信息清除规范，并出台企业交易评价机制，促进二手交易便利化、标准化。山西省作为能源转型大省和晋商文化的发源地，更应加大对经济的低碳化引导和对晋商文化中务实、诚信等核心价值观的

作者简介：

原缓缓（1998—），女（汉族），山西省运城人。助教，硕士，2023年入职运城学院数学与信息技术学院，研究方向为社会经济统计和应用数学。E-mail: 15536076332@163.com；

李雨航（2004—），男（汉族），山西省吕梁人。运城学院数学与信息技术学院在读本科生，研究方向为应用统计；

史华杰（2004—），男（汉族），山西省阳泉人。运城学院数学与信息技术学院在读本科生，研究方向为应用统计；

杨盼（2003—），男（汉族），山西省运城人。运城学院数学与信息技术学院在读本科生，研究方向为应用统计。

挖掘^[6, 7]。为深入了解二手交易市场的发展现状以及存在的问题，明确未来二手交易市场的发展方向。本文基于线下访谈法、文献调查法、案例分析法和问卷调查法对二手交易市场的买方和卖方购买（或销售）二手商品的类型、影响因素等展开调查，以达到更好地了解目前的二手交易市场并为未来的发展方向提供合理可行化建议。

一、调查方案设计与实施

（一）方案设计

基于文献调查法，综合文献资料，聚焦二手市场的核心议题（如信任机制、定价策略、物流服务），为后续研究提供理论支撑。

基于案例分析法，选择3-5个具有代表性的二手交易平台^[8]（如闲鱼、多抓鱼、线下跳蚤市场）作为分析对象，通过平台公开数据、用户投诉案例及媒体报道，总结交易纠纷高频环节（如虚假描述、售后缺失）。

基于问卷调查法，利用问卷星和问卷网生成链接，通过社交媒体（微信、微博）及二手平台用户群定向投放，覆盖不同地域及年龄层。

（二）调查内容

基于前期线下走访调查，了解到18岁以下的青少年消费群体在二手市场消费群体占比很少，并考虑到未成年人的购买能力有限，所以本次调查针对山西省范围内18岁以上消费群体来对二手市场进行详细分析。

（三）抽样设计

1. 预调查

为了测试问卷的合理性和准确性是否达到预计标准，在问卷中设置甄别题（如“我国的首都是？”）。在正式调查前，采用线上的形式进行预调查，发放问卷100份，回收率为100%，其中有4份甄别题回答错误，属于无效问卷，有效回答率为96%。

2. 样本量的确定

为了解山西省18岁以上人群在二手市场的消费情况，采用分层抽样方法。按照合理的抽样方案^[9]和山西省各市的成年人口数^[10]投放问卷，得到的最终样本量为1112份。其中，太原市171人，运城市150人，临汾市127人，吕梁市106人，长治市103人，大同市97人，晋中市109人，晋城市71人，忻州市85人，朔州市50人，阳泉市43人。

（四）问卷设计

1. 问卷结构

问卷主要分为五部分：消费者基本信息、二手交易角色（买家、卖家）、二手交易行为与习惯、消费观念与社会影响、二手市场评价与未来展望。

2. 数据处理

本问卷实发1324份，利用Python中的pandas库进行数据清洗，将答题时间少于60秒、甄选题回答错误以及重复数据视为无效问卷，共删除113份问卷，最终得1212份问卷，问卷回收率为91.5%，将调查所得数据录入Excel表中。

3. 问卷效度和信度检验

（1）信度检验

信度分析^[11]用来测量研究对象的稳定性和可靠性程度，由表1可得Cronbach's α 系数均大于0.8，问卷设计合理。

表1：信度值

量表	Cronbach's α 系数	项数
二手交易平台的使用频率	0.815	7
购买二手商品重视的因素	0.820	6
对二手市场的评价	0.845	6

（2）效度检验

效度分析^[11]用于检验测量工具能否准确反映研究者所关注的目标特征，即结果与理论预期的契合程度，表2-表4的KMO检验值均处于0.6以上，说明效度良好。

表2：二手交易平台的使用频率

KMO 和巴特利特检验		
KMO 取样适切性量数		0.699
	近似卡方	238.704
巴特利特球形度检验	自由度	21
	显著性	.000

表3：购买二手商品重视的因素

KMO 和巴特利特检验		
KMO 取样适切性量数		0.805
	近似卡方	636.194
巴特利特球形度检验	自由度	15
	显著性	.000

表4：对二手市场的评价

KMO 和巴特利特检验		
KMO 取样适切性量数		0.813
	近似卡方	953.044
巴特利特球形度检验	自由度	15
	显著性	.000

二、描述性分析

（一）买家特征分析

1. 购买的商品类型

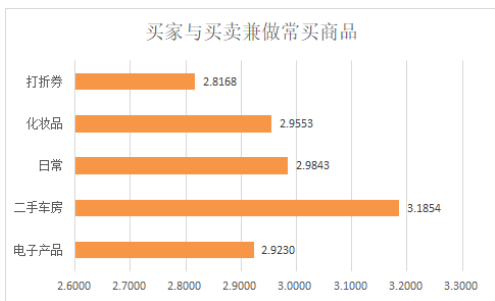


图1：购买商品类型图

如图1所示,打折券平均数为最低,意味着其购买频率最高,化妆品、日常商品以及电子产品平均数居中,意味着购买频率较高,二手车房平均数最多,意味着其并不是很常购买的商品。

2. 商品的重视因素

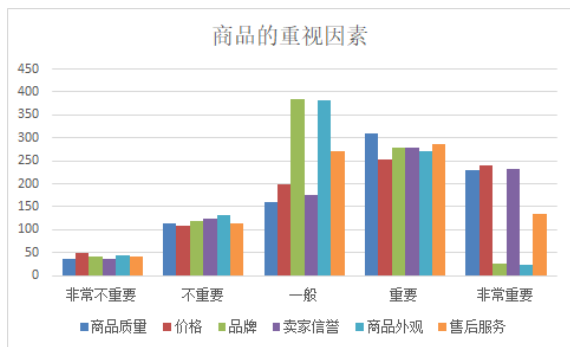


图2：购买商品因素图

从图2可知,被调查者在购买二手商品时的重视因素中认为商品质量、价格与卖家信誉非常重要的人居多,品牌与商品外观相对来说重视程度较为一般。

3. 判断二手商品质量的标准

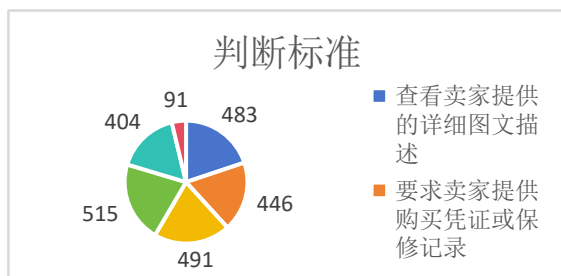


图3：质量判断标准图

如图3分析可得,在判断二手商品的质量时,被调查者有大部分关注买家评价与卖家的信用分,共515人,有较大部分关注的是卖家对商品的详细信息以及平台所提供的验机验货服务,两者均为490人左右,其次是卖家提供的凭证或保修记录,为446人,与卖家视频通话或直播验货的人的关注度也并不低,为404人,其他相对较少。说明买家在进行判别商品质量时最为关注的信息主要来自评论和其他买家的打分、商品的详细描述和平台的验机验货。由此可知平台有必要做好服务和监管,如杜绝刷评论现象出现,杜绝平台信誉功能造假以及对商品的描述要足够详细。

（二）卖家特征分析

1. 商品类型

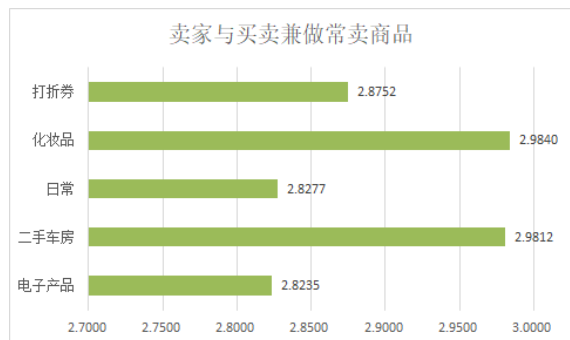


图4：售卖商品类型图

如图4可知,日常生活用品与电子产品的平均数为最低,意味着其售卖频率最高,打折券平均数为第二低,意味着售卖频率较高,化妆品与二手车房平均数较大,意味着其并不是很常购买的商品,可见化妆品与二手车房市场空间还可以有很大提升。

2. 出售二手商品的原因分析

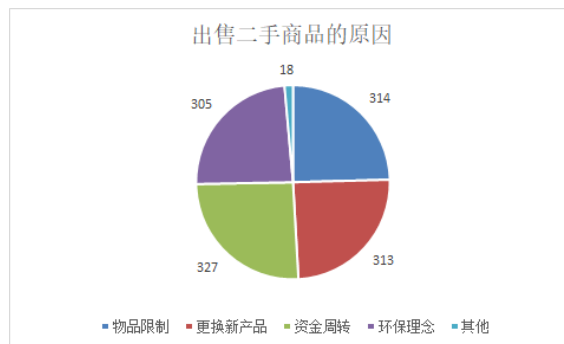


图5：售卖商品原因图

从图5可知卖家在出售二手商品时,主要是考虑到物品限制、更新换产品、资金周转和环保理念这四个因素。

3. 出售二手商品时存在的问题



图6：售卖商品困难图

卖家在出售二手商品时首要问题是难以找到合适的买家,其次为定价困难与买家砍价太狠,担心交易安全与认为物品描述和拍照麻烦次之,见图6。这些足以说明在进行售卖二手商品时定价与商品的需求是最大的困难。

三、影响因素分析

了解消费者二手交易频率的变化及其影响因素,对于平台优化服务、制定市场策略具有重要意义。通过 Logistic 回归模型,探讨消费者收入变化、对二手商品信任度、平台服务质量、环保意识以及与平台和商品无关的因素对二手交易频率的影响,为相关决策提供数据与理论支持。

（一）模型

Logistic 回归模型适用于二分类问题,能够通过自变量预测因变量的概率。公式如下:

$$P(y = 1/x) = \frac{1}{1 + e^{-(w_0 + w_1x_1 + w_2x_2 + w_3x_3 + w_4x_4 + w_5x_5)}}$$

其中, $p(y = 1/x)$ 表示增加二手交易频率的概率, w_0 为截距,

w_1 至 w_3 为自变量系数。

(二) 结果分析

1. 模型系数及显著性

表5：模型系数及显著性

变量	系数	显著性水平
截距项	0.74	<0.001
收入变化	0.46	<0.01
对二手商品的信任度提高 / 降低	0.21	<0.05
二手平台服务改进 / 变差	0.05	>0.10
环保意识增强	-0.21	<0.05
与二手平台、商品无关	0.21	<0.05

由表5可知，收入变化的系数最大为0.46 (>0) 且显著 ($p<0.01$)，表明收入是影响交易频率的最大因素，收入增加会显著提高用户增加二手交易频率的概率。为此可根据用户的收入水平进行分层，针对高收入群体推出一些高品质、高价值的二手商品，如高端电子产品、奢侈品等；对于中低收入群体，则重点推荐性价比高的日常用品。对二手商品的信任度系数为0.21 (>0) 且显著 ($p<0.05$)，意味着消费者的信任是比较重要的因素，对二手商品信任度的提高，能够有效消除用户在交易过程中的顾虑，如担心商品质量、真伪等问题，从而促使他们更愿意参与二手交易，增加交易频率。

环保意识增强的系数显著 ($p<0.05$) 但系数为负，这可能是部分环保意识较强的用户认为二手交易过程中存在一些不环保的环节，如包装、运输等从而影响了他们参与二手交易的积极性。平台服务不显著 ($p>0.10$)，说明目前平台在服务方面的改进措施未能显著影响用户是否增加二手交易频率的决策。这可能是因为平台所做的服务改进没有切中用户的核心需求，或者消费者认为平台对卖家和商品质量的监管太过薄弱。平台可以建立严格的商品质量检测机制，对上架的二手商品进行全面检测和认证，确保商品质量符合一定标准。同时，加强对卖家的管理和监督，对提供虚假商品信息或出售劣质商品的卖家进行严厉处罚，提高用户对二手商品的信任度。

与二手平台、商品无关的因素对用户增加二手交易频率起到了抑制作用，这些因素可能包括个人的时间成本、生活方式的改变等。例如，用户可能因为工作繁忙、生活节奏加快等原因，没有足够的时间和精力去参与二手交易。

2. 模型评估

表6：模型评估指标

准确率	精确率	召回率	F1分数
0.78	0.79	0.76	0.77

从表6中可知，模型整体准确率为78%，说明所得结论具有一定的可信度。在识别“增加交易频率”用户时，精确率和召回率均接近80%，表明模型具备较好的预测能力。

四、结论与建议

(一) 结论

1. 行为逻辑：理性与感性的双重驱动

在卖家出售原因中，“闲置处理”、“资金周转”、“环保理念”

占比均衡，说明其兼具实用与理想诉求，但实际交易中，定价困难与买家砍价是最大痛点，尤其在高单价商品领域，卖家对智能定价工具与第三方仲裁机制的需求迫切。

2. 信任机制：质量与售后仍是核心短板

尽管技术手段（如 AI 验货、区块链溯源）逐步普及，商品质量参差不齐与售后缺位仍是市场痼疾。30% 的负面评价直指假货问题，微博调查显示仅15% 用户对维权结果满意。买家判断商品质量时，最依赖买家评价（42%）、商品详情（40%）与平台验货服务（38%），但对“直播验货”、“卖家信誉分”等新功能信任度有限，反映技术应用尚未完全消解信息不对称。

卖家矛盾同样突出。奢侈品、化妆品等非标品流通效率低下，因缺乏统一分级标准，交易纠纷频发。红布林等垂直平台虽提供专业鉴定，但高达40% 的佣金加剧了卖家负担，导致出现“平台吸血”的抱怨。

3. 循环价值：环保理念与行为存在悖论

二手交易在资源循环的贡献已被广泛认可，闲鱼数据显示二手交易用户年均减碳量相当于种植350万棵树，然而研究发现环保意识与交易频率呈负相关。高环保意识群体可能因反对过度消费而减少交易以及对物流包装等环节的碳排放产生顾虑，环保意识与交易频率呈负相关提示单纯强调环保难以驱动市场增长，需把个人贡献过量化（如碳积分）、优化绿色服务（如共享包装）增强获得感。

(二) 建议

1. 商家层面：提升服务能力，满足细分需求

(1) 场景化运营，激活用户参与

进行校园生态圈打造，如毕业季推出“教材传承计划”，鼓励毕业生捐赠书籍并录制使用攻略视频，捐赠者可上榜“校友公益榜”并获得企业实习机会；发起情感营销活动，比如开展“旧物故事”征集，挖掘二手商品背后的温情回忆（如祖传首饰、童年玩具），通过短视频、播客传播，让“闲置”变“情怀”。

(2) 与品牌联动，推动循环消费

发起商品终身流通计划，购买新品时登记商品编号，未来通过平台转卖可享品牌补贴或折扣券，形成“买新 - 用旧 - 回收”闭环，吸引品牌商家入驻，同时吸引消费者购买。

2. 二级市场层面：优化规则与技术，打造高效平台

(1) 建立标准化信任体系，让交易更放心

对商品进行分级认证。针对衣服、家具等非标品，制定官方质检标准，按新旧划分“优品 / 良品 / 特惠”等级，强制卖家标注。高价商品（如奢侈品）联合中检集团提供“一物一码”鉴定，结果同步至区块链防伪。

进行信用评价改革。卖家信用分挂钩交易行为——多次违规者限制发布高价商品，优质卖家获得优先推荐；买家高信用可享“纠纷先行赔付”，减少维权等待。

纠纷处理透明化。成立由平台、消协、律师组成的独立调解委员会，直播听证会并公示证据链，恶意欺诈者纳入征信黑名单。

(2) 技术赋能，让买卖更智能

AI 工具助力定价。开发“智能估价助手”，综合市场供需、

历史成交价生成建议区间，卖家可浮动调整并说明原因（如“急售优惠”）；手机、家电等标品支持“拍照自动估价”。

精准匹配需求。优化搜索算法，增加等小众标签推荐；建立区域库存数据库，向卖家推送“附近急需商品清单”（如高校区教材、商务区办公设备）。

验货功能升级。要求卖家上传高清视频、360度全景图及细节放大图，开通语音描述功能，方便老年用户操作。

（3）区域协同，盘活资源流通

建立省内物流网络。在太原、运城等活跃城市设二手商品集散中心，提供翻新、仓储服务；吕梁、忻州等欠发达地区设“社区回收站”，通过定期物流专线全省联通。

助力乡村振兴。将二手农机、家具定向输送至农村，政府补贴运输费，既消化城市闲置资源，又支持乡村建设。

3. 社会层面：政策引导与价值升级，推动可持续发展

（1）政策支持，降低交易成本

个人卖家年交易额低于10万元免税，企业参与官方回收计划可抵税；商场、小区开设“二手专区”免租金。同时建立绿色物流，将二手运输纳入城市绿色货运体系，新能源运输企业奖励碳积分；社区、高校布设智能回收柜，支持一键呼叫上门取件。进行新职业扶持，认证“二手质检师”、“闲置规划师”为新职业，提供免费培训，下岗职工、大学生从业可领岗位补贴。

（2）量化环保价值，倡导绿色生活

可视化环保贡献，订单页显示“本次交易减少碳排放量”，年终生成用户“绿色成就报告”供社交分享。同时进行积分激励计划，二手消费累计“环保积分”可兑换公交卡充值、共享单车

周卡，或参与“线上种树”公益活动。

（3）重塑社会认知

通过宣传引导消费者从省钱到责任，使得二手消费从“图便宜”转变为“理性选择+环保担当”的新生活方式。

五、结束语

本文通过对全省1212份有效问卷的数据整理和建模分析，尝试描绘出当下二手消费群体的真实面貌。无论是买家看重的商品质量与卖家信誉，还是卖家面临的定价难与买家砍价压力，背后都指向一个共同的问题：信任。研究发现，虽然收入水平是影响交易频率的主要因素，但消费者对二手商品的信任度同样起着关键作用。与此同时，环保意识虽然被广泛提及，却并未直接推动更多人参与交易，这一矛盾也提醒我们，绿色理念的落地还需要更具体的激励和配套服务。

通过这次调查，我们看到了二手市场的潜力，也看清了它的短板——标准缺失、售后难保、信息不对称仍是阻碍。未来的发展，既需要平台在规则和技术上持续完善，也离不开政策引导和社会认知的转变。

二手交易不只是买卖，更是一种资源的延续和生活态度的表达。希望这篇研究能为山西乃至更广地区的二手市场发展提供一点参考，也让更多人看到——那些被闲置的物品，依然有被重新珍视的可能。

参考文献

- [1] 孙珂. 混合运营模式下网络零售平台二手市场策略研究 [D]. 中南财经政法大学, 2022.
- [2] 郑乾, 张新宇, 白昌山, 等. 浅析共享经济下二手市场的发展现状 [J]. 商场现代化, 2019, (10): 182-183.
- [3] 张洁. RE 公司二手电商的商业模式创新研究 [D]. 北京交通大学, 2023.
- [4] 孙竹寒, 李易周. 循环经济背景下实体二手市场发展路径的可行性实证分析 [J]. 现代商贸工业, 2023, 44(19): 94-96.
- [5] 刘诚. 二手电商平台推动绿色发展 [J]. 东北财经大学学报, 2025, (01): 78-87.
- [6] 张磊, 李熙尉, 樊园杰. 山西省打造能源革命综合试点先行区背景下能源产业转型路径研究 [J]. 煤, 2023, 32(08): 19-21+47.
- [7] 段晋云, 丁晓武. 晋商精神融入晋地中学生思想政治教育的价值探析 [J]. 中北大学学报 (社会科学版), 2022, 38(05): 106-112.
- [8] 侯丽芳, 邝小燕, 吴慧娴, 等. 循环经济中二手电的商发展研究——以“闲鱼”为例 [J]. 再生资源与循环经济, 2024, 17(09): 7-10.
- [9] 郑卫军, 何凡. 现况调查的样本量计算方法 [J]. 预防医学, 2020, 32(06): 647-648.
- [10] 金卫东. 第七次全国人口普查结果分析与解读 [J]. 教学考试, 2022, (27): 21-25.
- [11] 郭涛, 余汶铭, 田井逸, 等. 药学科本科科研兴趣调查问卷的信度和效度分析 [J]. 黑龙江科学, 2021, 12(07): 24-26.

体育赛事网络关注度与举办地旅游客流量的格兰杰因果检验——以2025年江苏省足球比赛（“苏超”）为例

王超，蔡捷*

山东财经大学，山东，济南 250002

DOI:10.61369/ASDS.2026040013

摘 要： 在体育产业与文旅经济深度融合的背景下，厘清网络关注度与旅游客流之间的动态关系，对于优化赛事资源配置和提升区域经济效益具有重要意义。本文以2025年江苏省足球比赛为研究对象，基于2024年1月至2025年12月的月度时间序列数据，构建向量自回归（VAR）模型，探讨网络关注度与举办地旅游客流量之间的因果关系与传导路径。研究表明：第一，网络关注度是旅游客流量的格兰杰原因，而反向因果关系不成立，验证了“注意力到消费力”的单向传导机制；第二，网络关注度对旅游客流量的影响存在显著滞后周期，最优滞后阶数为2期，表明关注度上升后其对旅游客流量的带动效应在1至2个月内逐步释放；第三，脉冲响应分析表明，网络关注度冲击对旅游客流量呈现“先升后降”的动态路径，响应在滞后1期达到峰值，之后逐步衰减。本文为赛事主办方与文旅部门优化宣传时点和构建“体育+文旅”融合发展模式提供量化决策参考。

关 键 词： 体育赛事；网络关注度；旅游客流量；“苏超”联赛；VAR模型

The Granger Causality Test of the Network Attention of Sports Events and the Tourist Flow of the Host City — Taking the 2025 Jiangsu Football Match (‘Suchao’) as an Example

Wang Chao, Cai Jie*

Shandong University of Finance and Economics, Jinan, Shandong 250002

Abstract： Under the background of the deep integration of sports industry and cultural tourism economy, it is of great significance to clarify the dynamic relationship between network attention and tourism passenger flow for optimizing the allocation of event resources and improving regional economic benefits. This paper takes the football match in Jiangsu Province in 2025 as the research object. Based on the monthly time series data from January 2024 to December 2025, a vector autoregressive (VAR) model is constructed to explore the causal relationship and transmission path between network attention and tourist flow in the host city. The results show that : first, network attention is the Granger cause of tourist flow, and the reverse causality is not established, which verifies the one-way transmission mechanism of 'attention to consumption'; second, there is a significant lag period in the impact of network attention on tourist flow, and the optimal lag order is 2 periods, indicating that the driving effect of network attention on tourist flow is gradually released within 1 to 2 months after the increase of network attention. Thirdly, the impulse response analysis shows that the impact of network attention on tourist traffic presents a dynamic path of 'first rise and then fall', and the response reaches the peak in the lag 1 period, and then gradually decays. This paper provides a quantitative decision-making reference for the event organizers and cultural tourism departments to optimize the publicity time point and construct the integrated development model of 'sports + cultural tourism'.

Keywords： sports events; online attention; tourist flows; “Suchao” League; VAR model

引言

目前，体育强国与健康中国战略深度融合，体育产业正成为促进消费升级、提升城市形象和推动区域经济发展的关键引擎^[1]。大型

基金项目：2024年度教育部人文社会科学研究青年基金项目 (24YJC890002)。

作者简介：王超（第一作者），硕士研究生，研究方向：体育统计，邮箱：2540311419@qq.com。

通讯作者：蔡捷，山东财经大学副教授，博士，研究方向：体育统计。

体育赛事具备注意力集聚效应和文化感染力，能够带动举办地的交通、住宿、餐饮、旅游等相关产业的协同发展^[2]。在这一背景下，如何评估体育赛事的经济效益，以及识别赛事流量向经济增量的转化路径，成为学界与业界共同关注的核心议题。

以2025年江苏省足球赛事（以下简称“苏超”联赛）为例，其“十三太保”的地域竞争格局不仅在线上引爆现象级的社交话题与网络关注，更在线下促进文旅消费需求。由于该赛事的热度不断提升，主客场变化、对手实力、赛事结果等因素引发网络关注度的波动，并且江苏省经济基础扎实和旅游市场成熟，使该赛事成为检验网络关注度向旅游经济转化的理想案例，具有时代特征与研究价值。在此背景下，厘清由体育赛事所激发的网络关注度与举办地实际旅游经济收益之间的动态因果关系，具有重要的理论价值和现实意义。

然而，现有研究多数研究聚焦于奥运会、世界杯等顶级赛事的宏观效应评估^[3]，对区域性、高频次赛事的关注相对较少，难以揭示体育赛事热度波动对旅游经济的影响。另一方面研究多运用事件研究法和静态面板回归，验证赛事对经济是否存在影响^[4]，但对二者之间的因果方向和时滞作用缺乏探讨。

因此，本研究通过运用时间序列模型，定量揭示体育赛事注意力转化为旅游经济消费力的传导路径、作用强度及滞后周期。本文以2025年“苏超”联赛为研究对象，构建向量自回归（VAR）模型，检验网络关注度与举办地旅游客流量之间的因果关系。本文旨在回答三个核心问题：网络关注度与旅游客流量是否存在因果关系？反向因果关系是否成立？若存在因果关系，其最优滞后期是多少？通过上述研究，本文试图为理解区域性体育赛事中注意力向经济效益的传导机制提供经验证据，并为赛事主办方优化宣传投放时点、文旅部门实现流量向留量的高效转化提供量化决策依据。

一、文献综述与理论基础

（一）体育赛事旅游效应研究

体育赛事对旅游经济的带动效应，已成为学界的普遍共识。围绕这一问题，学界经历了从宏观定性描述到微观因果识别的演进过程，研究重心逐渐从是否存在影响转向影响如何产生、何时发生、为何存在差异。

1. 赛事效应的存在性检验与因果识别困境

早期研究主要聚焦于证实大型体育赛事对旅游经济的正向作用。Preuss（2007）系统构建大型赛事经济影响的评估框架，并区分直接效应、间接效应与引致效应，为后续研究奠定了方法论基础^[5]。Lee与Taylor（2005）运用投入产出表方法测算2002年韩日世界杯的经济贡献^[6]，Massiani（2018）则采用可计算一般均衡模型评估大型赛事对区域经济的影响^[7]。师傅与任保平（2021）的研究进一步表明，大型体育赛事能够助推主办城市高质量发展，并且体育产业发展与旅游服务规划是释放赛事效应的关键路径^[8]。

然而，这一阶段的研究普遍面临因果识别的局限。无论是投入产出分析还是事件研究法，都难以排除宏观趋势、政策变动、外部冲击等混淆因素对估计结果的干扰。近年来，Fourie与Santana-Gallego（2022）采用面板数据回归，通过控制国家与年份固定效应估计赛事对国际旅游流的因果影响^[9]。最新的研究开始采用堆叠双重差分法与扩展双向固定效应模型^[10]，为识别赛事旅游效应的因果路径提供了更可靠的工具。

2. 网络关注度的预测能力与滞后效应

在赛事效应的宏观评估的基础上，部分学者开始聚焦于网络关注度与旅游经济的关联。Li与Law（2021）探索网络搜索数据在旅游需求预测中的应用，发现其对短期预测的改进效果最为显著^[11]。许艳等（2021）基于百度指数对乌镇景区网络关注度进行研究，表明“东高西低”的空间分布特征与显著的季节性波动^[12]。

部分学者也关注二者之间的时间滞后关系。Bangwayo-Skeete与Skeete（2015）采用谷歌趋势数据预测加勒比地区游客量，发现

网络关注度领先旅游需求1到4周，且领先期长度因目的地类型、客源地距离而异^[13]。刘培学等（2021）基于南京钟山风景名胜区手机大数据与百度指数，采用VAR模型的脉冲响应函数分析发现，PC端搜索的波动冲击会引起游客量在滞后5期（日度数据）达到响应峰值，揭示了不同搜索终端在传导效率上的差异^[14]。

现有文献主要聚焦于奥运会、世界杯等顶级赛事的旅游效应评估，以及网络关注度对常规旅游场景的研究。因此，本文探讨网络关注度对区域体育赛事的经济效应。

（二）理论分析

1. 注意力经济理论

注意力经济理论表明在信息过载的时代，注意力成为稀缺资源，能够有效吸引并维持注意力的主体将获得竞争优势^[15]。体育赛事凭借媒体曝光与文化感染力，能够在短时间内集聚大量社会注意力，而网络关注度正是这种注意力的量化表征，如百度指数、抖音热度等指标反映特定主题的注意力强度。当形成注意力冲击，就为后续消费转化创造条件^[16]。因此，网络关注度先于实际旅游客流量的变化，这正是本文格兰杰因果检验的核心命题。

2. 消费者购买决策过程理论

消费者购买决策过程理论将消费者决策划分为问题认知、信息搜索、方案评估、购买决策和购后行为这五个阶段^[17]。在体育赛事中，网络关注度对应于决策过程中的信息搜索阶段，消费者通过搜索引擎、社交媒体收集赛事和目的地信息；而旅游消费的出行观赛、住宿、餐饮等则对应于购买决策与购后行为阶段。从信息搜索到实际出行，消费者需要经历方案评估、行程安排等多环节，这一过程天然存在时间差^[18]。在“苏超”赛事场景中，本文将通过VAR模型精确估计其长度。

3. 信息扩散理论

根据信息扩散理论，信息在人群中的传播经历启动期、爆发期与衰减期三个阶段，并遵循S型曲线^[19]。在体育赛事周期中，赛前预热期开始进行赛事信息传播作为启动期，网络关注度逐步上升。当赛事进入高潮，网络关注度达到峰值到达爆发期，进入信息爆发期阶段。当体育赛事进入后期，关注度逐渐下降，信息

传播进入衰减期^[20]。因此，不同阶段的信息扩散特征决定了其对旅游经济的传导效率差异，本文通过脉冲响应函数反映其冲击。

基于以上分析，注意力经济理论解释网络关注度为何是旅游客流量的先行指标，消费者决策过程理论解释二者之间存在滞后的原因，信息扩散理论解释不同赛事阶段效应的异质性，这三个理论构成本文的理论分析框架。因此，本文采用格兰杰因果检验和脉冲响应函数实证识别网络关注度与旅游客流量的因果方向与最优滞后期，为理论机制提供实证证据。

二、研究设计

（一）数据来源

本文考虑时间要覆盖完整的赛事周期，并引入2024年同期数据作为对照，控制季节性波动与年度趋势，提高模型稳健性，所以选取2024年1月到2025年12月的月度数据。数据来源于百度指数平台和江苏省文化和旅游厅。

（二）变量说明

1. 网络关注度（Attention）

百度指数以网民在百度的搜索量为数据基础，以关键词为统计对象，科学分析并计算出各个关键词在百度网页搜索中搜索频次的加权。根据搜索来源的不同，百度指数分为PC搜索指数与移动搜索指数，本文采用二者之和作为原始搜索指数。

本文选取关键词主要与“苏超”赛事高度相关，能够反映公众对赛事的关注程度，并且尽可能覆盖赛事本身、地域特色和网络文化等多个角度，以全面捕捉网络关注度。由于，“十三太保”作为江苏省十三个地级市的网络代称，能够反映地域竞争格局引发的网络讨论；“南哥”是南京拟人化的网络昵称，能够体现赛事举办地南京的独特网络文化；而“比赛第一，友谊第十四”是“苏超”赛事期间流行的网络热梗，能够反映赛事的娱乐化传播特征。因此，本文选取关键词有“苏超”、“苏超排名”、“十三太保”、“南哥”和“比赛第一，友谊第十四”。

本文对各关键词的月度搜索指数进行加总以衡量对“苏超”赛事的网络关注度，同时进行对数处理以消除量纲影响并降低数据偏态性，计算公式如下：

$$Attention_t = \ln \left(\sum_{i=1}^5 Index_{it} \right)$$

2. 旅游客流量（Tourism）

由于考虑旅游客流量的测度面临数据可得性的现实约束，本文选取江苏省星级酒店月度入住率作为旅游客流量的代理变量。原因在于星级酒店入住率具有规范的统计口径并且数据连续。其次，该指标能够反映过夜游客规模，与异地观赛人群高度相关。最后，由于江苏省星级酒店分布广泛，能够综合反映各赛事举办城市的旅游接待情况。

3. 控制变量（Holiday）

本文尽可能排除其他因素影响，引入节假日虚拟变量Holiday，主要考虑劳动节和国庆节的影响，所以在5月和10月取值为1。

（三）模型设定

为捕捉网络关注度与旅游客流量之间的动态关系，本文构建

向量自回归模型。模型设定包含网络关注度（ $Attention_t$ ）与旅游客流量（ $Tourism_t$ ）的二维内生变量，同时将节假日虚拟变量（ $Holiday_t$ ）作为外生变量纳入模型。模型的简化形式为：

$$y_t = c + \sum_{i=1}^p \Phi_i y_{t-i} + \Psi \cdot Holiday_t + \varepsilon_t$$

其中， $y_t = (Attention_t, Tourism_t)'$ 为内生变量向量， c 为常数项向量， p 为滞后阶数， ε_t 为白噪声扰动项。

三、实证分析

（一）平稳性检验

由于本文数据是时间序列数据，为避免伪回归，首先进行单位根检验。采用ADF检验判断序列平稳性，检验结果如表1所示。在1%的显著性水平下，旅游客流量（Tourism）与网络关注度（Attention）均拒绝存在单位根的原假设，表明二者均为平稳序列。

表1：变量的 ADF 单位根检验 Table 1: ADF Unit Root Test of Variables				
变量	ADF 值	1% 临界值	P 值	结论
Tourism	-7.670	-3.750	0.000	平稳
Attention	-31.184	-3.750	0.001	平稳

（二）最优滞后阶数选择

根据AIC、HQIC和SBIC准则确定VAR模型最优滞后阶数，结果如表2所示。FPE、AIC与HQIC准则均建议滞后2阶为最优选择，LR检验则建议滞后4阶。由于样本量较小，考虑自由度损失问题，本文选取滞后2阶作为模型设定。

表2：滞后阶数选择结果 Table 2: The Selection Results of Lag Order			
滞后阶数	LogL	LR	FPE
0	-8.077		.009392
1	-1.299	13.556	.007146
2	4.287	11.173	.006202*
3	-7.205	5.8374	.007194
4	12.898	11.385*	.00656
滞后阶数	AIC	HQIC	SBIC
0	1.00771	1.02715	1.10729
1	.72992	.788234	1.02864*
2	.511294*	.668482*	1.06916
3	.679423	.815487	1.37644
4	.570159	.685098	1.40632

注：*表示该准则建议的最优滞后阶数。

（三）格兰杰因果检验

为识别网络关注度与旅游客流量之间的因果关系，采用格兰杰因果检验，结果如表3所示。

表3：格兰杰因果关系检验结果 Table 3: Granger Causality Test Results			
原假设	χ^2 统计量	P 值	结论
Attention 不是 Tourism 的格兰杰原因	5.5148	0.000	拒绝
Tourism 不是 Attention 的格兰杰原因	2.9169	0.000	拒绝

结果可知，在1%显著性水平上，拒绝了网络关注度不是旅游客流量的格兰杰原因的原假设，表明网络关注度是旅游客流量的

格兰杰原因。同时，无法拒绝“旅游客流量不是网络关注度的格兰杰原因”的原假设，表明反向因果关系并不成立。这一结果验证本文的核心研究假设，支持注意力经济理论阐述的注意力集聚是后续消费转化的前提条件，而非相反。

（四）VAR 模型结果及脉冲响应分析

1. 稳定性检验

本文构建旅游客流量和网络关注度的 VAR(2) 模型进行估计，并加入节假日虚拟变量作为外生控制变量。图 1 结果显示所有特征多项式根的倒数均处于单位圆内，说明模型满足稳定性条件，表明脉冲响应函数的标准误估计可靠。

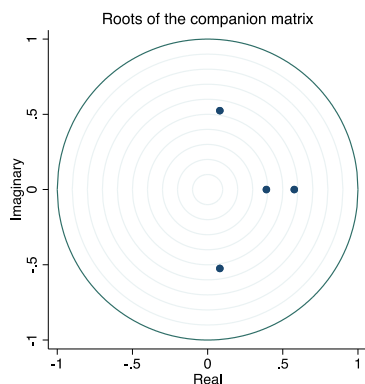


图 1：单位圆检验

Figure 1: Unit Circle Test

2. 脉冲响应分析

本文为进一步探讨网络关注度冲击对旅游客流量的动态影响路径，计算了脉冲响应函数。由于样本量有限，为避免小样本误差，增强脉冲响应的稳定性，本文利用 Bootstrap 方法进行脉冲响应的统计推断，重复抽样 1000 次，并计算 95% 置信水平的置信区间。

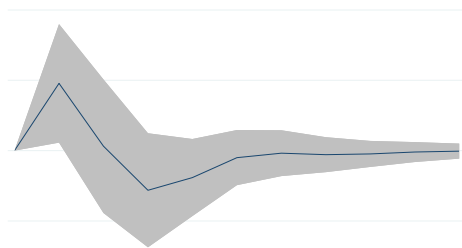


图 2：脉冲响应图

Figure 2: Impulse Response Diagram

由图 2 可知，给网络关注度一个单位标准差的正向冲击，酒店入住率在当期产生正向影响。随后，响应幅度在滞后 1 期达到峰值，并在滞后 2 期保持正向影响，逐步衰减并趋于平稳。这一动态路径呈现先升后降的特征，与赛事影响的实际传导规律相吻合。这一路径变化表明网络关注度上升后，旅游经济效应在 1 到 2 个月内达到峰值。因此，这是文旅部门和政府机关营销宣传的关键时期。

四、研究结论与政策建议

（一）研究结论

第一，网络关注度与旅游客流量之间存在显著的格兰杰因果关系，实证检验了体育赛事注意力到旅游经济消费力的单向传导路径，并且反方向因果并不成立，表明网络关注度能够有效转化为旅游经济效益。

第二，网络关注度对旅游客流量的影响存在显著的滞后周期，根据准则结果滞后阶数为 2 阶。VAR 模型估计进一步证实，网络关注度的滞后 1 期与滞后 2 期系数均显著为正，表明在 1 到 2 个月内，网络关注度对旅游客流量有显著促进作用。

第三，脉冲响应函数显示网络关注度冲击对旅游客流量呈现“先升后降”的影响，并且在当期就有正向响应，在滞后 1 期后达到峰值，滞后 2 期保持显著，之后逐步衰减并趋于平稳。

（二）政策建议

第一，搭建网络关注度监测平台，以实现流量向客流的精准转化。本研究结论表明网络关注度对旅游经济具有显著正向滞后影响。因此，本文建议体育赛事主办方和文旅部门联合建立网络关注度动态监测平台，实时监测百度指数和抖音热度等网络关注度指标的波动，以进行旅游产品推荐、交通住宿规划和景区服务介绍等工作，确保网络流量红利高效转化为旅游经济增量。

第二，把握赛事宣传时点，确定营销最佳时期。脉冲响应结果显示，网络关注度冲击对旅游客流量的带动效应在滞后 1 期达到峰值。因此，本文建议体育赛事主办方将宣传策略重点投放在赛前 1-2 个月的预热期，并在峰值期前完成旅游产品的精准推送，并通过社交媒体运营、话题营销等手段持续放大网络关注度，以最大化营销投入的边际效益。

第三，借赛事热度打造文旅产品，推动赛事流量向城市留量转化。研究结果显示，网络关注度对旅游客流量的带动作用主要集中在赛事前后 1 到 2 个月，之后会逐步减弱，说明赛事带来的热度难以持久。因此，建议文旅部门抓住赛事期间的高关注度，同步推出文创周边、主题线路、特色美食等衍生内容，将赛事热度沉淀为城市文旅资产，实现从赛事流量到城市留量的跨越。

参考文献

- [1] 黄海燕, 康露. 新时代体育产业高质量发展的理论逻辑与实施路径 [J]. 体育科学, 2022, 42(1): 15-34, 58.
- [2] 史瑞应, 宋瑞. 大型体育赛事对城市旅游产业的影响评估——来自北京 2022 年冬奥会赛前阶段的经验证据 [J]. 北京体育大学学报, 2022, 45(4): 52-62.
- [3] 孔庆波, 熊禄全. 大型体育赛事运营之影响力经济形成与挖掘 [J]. 体育文化导刊, 2014(7): 99-101, 125.
- [4] 陈晓红, 李杨扬, 汪阳洁, 等. 网红城市、流量效应与旅游发展 [J]. 管理科学学报, 2022, 25(1): 1-22.

- [5]Preuss H. The conceptualisation and measurement of mega sport event legacies[J]. *Journal of Sport & Tourism*, 2007, 12(3-4): 207-228.
- [6]Lee C K, Taylor T. Critical reflections on the economic impact assessment of a mega-event: The case of 2002 FIFA world cup[J]. *Tourism Management*, 2005, 26(4): 595-603.
- [7]Massiani J. Assessing the economic impact of mega events using computable general equilibrium models: Promises and compromises[J]. *Economic Modelling*, 2018, 75: 1-9.
- [8] 师傅, 任保平. 大型体育赛事助推城市高质量发展的效应研究——基于第14届全运会的分析 [J]. *西安体育学院学报*, 2021, 38(2): 134-139.
- [9]Fourie J, Santana-Gallego M. Mega-sport events and inbound tourism: New data, methods and evidence[J]. *Tourism Management Perspectives*, 2022, 43: 101002.
- [10]Wooldridge J M. Two-way fixed effects, the two-way mundlak regression, and difference-in-differences estimators[J]. *Empirical Economics*, 2025, 69(5): 2545-2587.
- [11]Li X, Law R, Xie G, 等. Review of tourism forecasting research with internet data[J]. *Tourism Management*, 2021, 83: 104245.
- [12] 许艳, 陆林, 赵海溶. 乌镇景区网络关注度动态演变与空间差异分析 [J]. *经济地理*, 2020, 40(7): 200-210.
- [13]Bangwayo-Skeete P F, Skeete R W. Can google data improve the forecasting performance of tourist arrivals? Mixed-data sampling approach[J]. *Tourism Management*, 2015, 46: 454-464.
- [14] 刘培学, 朱知沛, 张捷, 等. 旅游在线搜索与客流波动的动态关联研究——以南京钟山风景名胜区为例 [J]. *旅游学刊*, 2021, 36(11): 95-105.
- [15]Loewenstein G, Wojtowicz Z. The economics of attention[J]. *Journal of Economic Literature*, 2025, 63(3): 1038-1089.
- [16] 孔庆波, 熊祿全. 大型体育赛事运营之影响力经济形成与挖掘 [J]. *体育文化导刊*, 2014(7): 99-101, 125.
- [17] 汪旭晖, 聂可显, 陈荣. “解释行为”还是“解释反应”? 怎样的在线评论更有用——基于解释类型的在线评论对消费者购买决策的影响及边界条件 [J]. *南开管理评论*, 2017, 20(4): 27-37.
- [18] 李莉, 王静. 从“观望者”到“购买者”——中国旅游电子商务消费者购买决策行为探析 [J]. *旅游学刊*, 2008(5): 49-56.
- [19]Spann B, Mead E, Maleki M, 等. Applying diffusion of innovations theory to social networks to understand the stages of adoption in connective action campaigns[J]. *Online Social Networks and Media*, 2022, 28: 100201.
- [20] 阮文奇, 张舒宁, 李勇泉. 大型体育赛事网络信息扩散的时空规律及机制——以国际田联路跑金标赛事为例 [J]. *上海体育学院学报*, 2020, 44(2): 74-86.

