

# 基于主成分分析和随机森林回归分析的非遗传播趋势预测

赵 芳, 朱娇蕾, 刘一一

西南交通大学希望学院, 四川 成都

收稿日期: 2025年12月29日; 录用日期: 2026年2月3日; 发布日期: 2026年2月14日

## 摘 要

在“十四五”文化数字化战略与乡村振兴协同推进背景下, 非物质文化遗产的保护与传播面临新机遇与挑战。针对现有研究在非遗传播趋势预测与影响机制量化方面的不足, 研究引入主成分分析与随机森林回归模型, 系统评估数字化旅游对布依族非遗传播效能的影响, 并预测其未来发展趋势。研究首先通过主成分分析对影响非遗传播的多维旅游要素进行降维, 提取关键主成分; 随后利用随机森林回归分析各主成分对非遗传播效果的非线性影响及特征重要性。结果表明, 自然景观、文化服务质量与价格合理性是影响非遗传播的核心因素, 旅游体验提升对传播趋势具有显著正向作用。研究为非遗保护与传播的量化评估提供了方法支撑, 对推动文旅融合与非遗数字化传播具有理论与实践意义。

## 关键词

非物质文化遗产, 文旅融合, 主成分分析, 随机森林

# Predicting the Dissemination Trends of Intangible Cultural Heritage Based on Principal Component Analysis and Random Forest Regression

Fang Zhao, Jiaolei Zhu, Yiyi Liu

Hope College of Southwest Jiaotong University, Chengdu Sichuan

Received: December 29, 2025; accepted: February 3, 2026; published: February 14, 2026

## Abstract

Under the synergistic advancement of the “14th Five-Year Plan” cultural digitalization strategy and

文章引用: 赵芳, 朱娇蕾, 刘一一. 基于主成分分析和随机森林回归分析的非遗传播趋势预测[J]. 社会科学前沿, 2026, 15(2): 470-479. DOI: 10.12677/ass.2026.152153

rural revitalization, the protection and dissemination of intangible cultural heritage (ICH) encounter both new opportunities and challenges. To address the gaps in existing research regarding ICH dissemination trend prediction and quantification of influencing mechanisms, this study introduces Principal Component Analysis (PCA) and Random Forest Regression to systematically evaluate the impact of digital tourism on the dissemination effectiveness of Bouyei ethnic ICH and forecast its future development trends. The study first employs PCA to reduce the dimensionality of multidimensional tourism factors affecting ICH dissemination and extract key principal components. Subsequently, Random Forest Regression is utilized to analyze the nonlinear effects of each principal component on ICH dissemination outcomes and assess feature importance. The results indicate that natural landscapes, cultural service quality, and price rationality are the core factors influencing ICH dissemination, while enhanced tourism experiences exert a significant positive effect on dissemination trends. This study provides methodological support for the quantitative assessment of ICH protection and dissemination, offering both theoretical and practical implications for promoting culture-tourism integration and digital ICH communication.

## Keywords

Intangible Cultural Heritage, Culture-Tourism Integration, Principal Component Analysis, Random Forest

Copyright © 2026 by author(s) and Hans Publishers Inc.

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

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



Open Access

## 1. 引言

宁南县西瑶镇拉落村是四川省布依族文化的重要传承地，2020 年被评为国家级 3A 旅游景区。该村以“非遗三月三”节庆为核心，融合布依族传统手工艺(如蜡染、刺绣) [1]、音乐舞蹈(如八音坐唱) [2]及自然生态景观(如梯田、古寨) [3]，形成“非遗 + 旅游”特色发展模式，配套建设标准化停车场、布依特色民宿和非遗工坊等设施，年接待游客量突破 10 万人次。作为省级非遗旅游示范村，其旅游收入占当地 GDP 比重达 18%，成为非遗活态传承与文旅融合的典型范例。

非物质文化遗产是中华民族的重要文化财富，近年来在国家政策引领下，保护与传承力度不断加强 [4]。以四川宁南县拉落村为代表的“非遗 + 旅游”模式 [5] [6]，推动了非遗的活态传承和文旅融合，取得了积极成效。然而，随着全球化和现代化进程加快，非遗文化面临传承人断层、传播方式单一等挑战，同时也迎来了旅游业发展和数字技术赋能的新机遇。新媒体、短视频平台和大数据、人工智能等技术的应用，为非遗文化的广泛传播和个性化推荐提供了新路径 [7]-[9]。如何科学评估非遗传播的影响因素，预测其发展趋势，成为当前亟需解决的问题。

## 2. 文献综述

近年来，非物质文化遗产的保护与传承逐渐成为学界关注的热点，但在实际发展过程中仍面临诸多挑战。首先，非遗文化的传承存在流失和断层现象。肖远平、王伟杰(2016) [10]通过对我国国家级少数民族非遗名录和传承人的统计，指出非遗项目数量参差不齐、传承人分布不均，且名录评审体系尚需完善。孙九霞、李怡飞(2018) [11]则从社会资本和人力资本角度，分析了非遗传承人地位获得的影响。胡艳丽、张雪梅(2017) [12]关注苗侗地区“非遗传承人”的生产性传承培育，卢颖(2024) [13]从非遗传承与乡村振兴视角下关注文化保护和产业发展问题，温必锋、尹海鹰(2024) [14]争对非遗目前发展挑战问题提出了非

遗传承与创新发展提供了相应策略,朱华明,石巍(2025) [15]对非遗传承中的信任体系缺失与技术赋能不足问题,焦明瑞(2025) [16]提供的策略不仅关注技艺的传承,更旨在推动非遗深度融入当代生活,从而实现其活态存续与价值再生,构建了相应的解决对策这些研究共同反映出非遗传承人队伍建设薄弱、后继乏人,传承链条存在断裂风险。

3. 数据来源与样本结构

3.1. 问卷数据收集

本次问卷主体包含受访者个人信息,受访者对布依族非遗文化认知与参与度:对布依族的非遗文化的了解程度、在受访者所在社区的传承和发展现状、对参与布依族非遗文化活动满意度、宁南县布依三寨的旅游满意度,受访者对布依族非遗数字化发展的认知情况:对数智化宣传方式的接受度、对当前数字化宣传内容的评价、在互联网平台传播的期望,布依族非遗文化数智化传播的改进建议与期望。采用分层抽样的方式,在各省份特别是在宁南县进行抽样调查,共收集有效样本 1871 份。

3.2. 量表信度与效度检验

(1)信度检验

本研究采用 Cronbach’s  $\alpha$  系数检验内部一致性。问卷包含 49 个项目,计算得出 Cronbach’s  $\alpha$  系数为 0.973 (见表 1),表明问卷具有极高的内部一致性,各项目间相关性高,测量的构念稳定一致,为后续研究提供了可靠的数据基础。

Table 1. Cronbach’s  $\alpha$  coefficient test results  
表 1. Cronbach’s  $\alpha$  系数检验结果

Cronbach’s $\alpha$ 系数	项数	样本数
0.973	49	1871

(2) 效度检验

采用 KMO (Kaiser-Meyer-Olkin)检验和巴特利特(Bartlett)球形检验进行效度分析(见表 2)。KMO 值为 0.807 (>0.7),巴特利特球形检验的 p 值为 0.000 (<0.001),表明问卷具有良好的结构效度。

Table 2. KMO values and Bartlett’s test results  
表 2. KMO 值和巴特利特球形检验结果

检验指标	数值
KMO 值	0.807
巴特利特球形值	46783.128
df	1225.000
p 值	0.000

3.3. 受访者人口统计学特征

调查数据表明,样本性别构成呈现相对均衡分布(见表 3),其中女性受访者占比 52.43%,男性占比 47.57%。在年龄结构方面,20~29 岁青年群体占比最高 31.21%,60 岁及以上老年群体占比最低 8.07%,

其余各年龄段分布较为均匀。这一人口统计学特征表明，乡村旅游中的布依族文化体验者以青年群体为主体。

Table 3. Distribution by gender and age  
表 3. 性别及年龄分布表

特征	类别	人数(人)	比例(%)
性别	女性	981	52.43
	男性	890	47.57
年龄	19 岁及以下	272	14.54
	20~29 岁	584	31.21
	30~39 岁	302	16.14
	40~49 岁	310	16.57
	50~59 岁	252	13.47
	60 岁及以上	151	8.07
	合计	1871	100

在民族结构方面，受访者群体呈现显著的民族多样性特征(如下图 1 所示)。汉族受访者占主体地位(64.30%)，其次为基诺族(3.10%)、蒙古族(2.19%)和满族(2.14%)等少数民族群体。该分布特征表明调查样本具有较好的民族代表性，覆盖了较广的民族谱系，此外还表明布依族文化旅游的吸引力呈现跨民族特性。

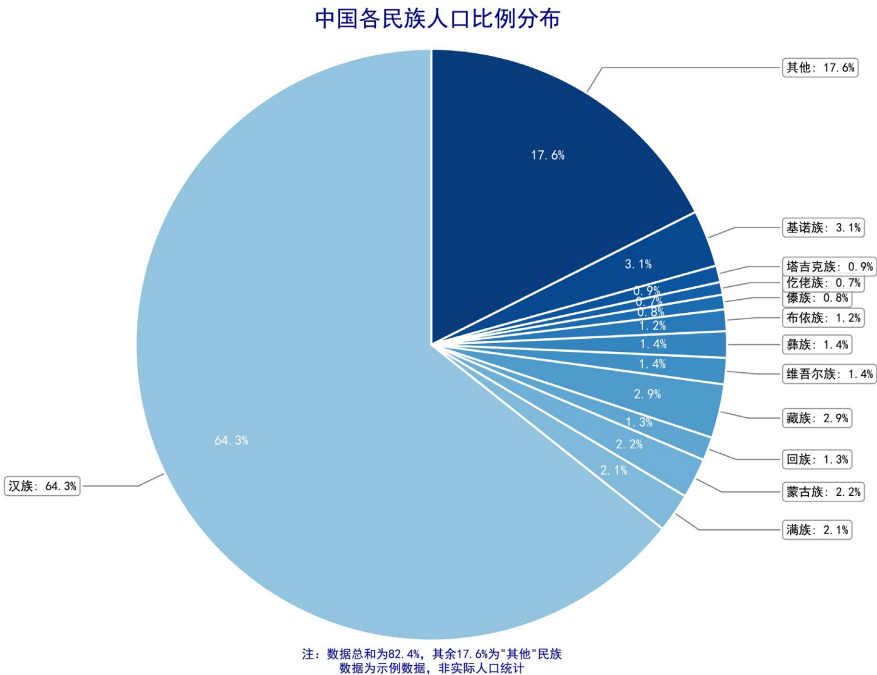


Figure 1. Distribution of ethnic structure  
图 1. 民族结构分布图

4. 基于动态主成分分析的旅游驱动下非遗文化传播机制研究

4.1. 特征选取

研究采用斯皮尔曼相关系数对 20 个指标进行相关性分析。热力图分析结果表明(如下图 2 所示), 各指标间相关性存在显著差异。为避免多重共线性会导致回归系数估计偏误、标准误增大以及模型解释能力下降, 设定对自变量间的相关系数绝对值大于 0.8 的变量进行剔除处理, 满足线性回归模型关于解释变量间独立性的基本假定, 经筛选, 最终选取自然风光、文化体验、交通便利、住宿条件、餐饮服务、当地特色活动、价格合理性、安全性以及旅游设施作为特征变量。

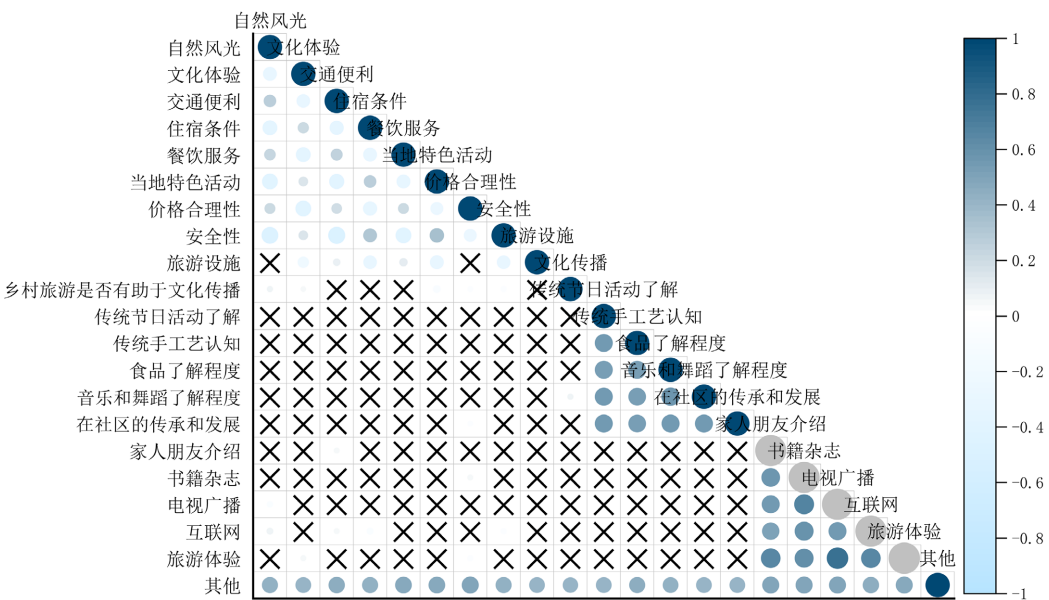


Figure 2. Spearman correlation analysis plot  
图 2. 斯皮尔曼相关性分析图

4.2. 旅游主成分的提取与验证

(1) 计算特征值与特征向量

主成分分析是一种经典的降维技术, 通过正交变换将原始相关变量转换为一组线性无关的主成分。该方法在保留原始数据主要信息的前提下, 能够有效降低数据维度, 消除变量间的多重共线性, 从而提升模型的稳健性和可解释性。在进行主成分分析前, 首先需要验证数据的适用性。如下表 4 所示, KMO 检验值为 0.669, 巴特利特球形检验结果显著( $\chi^2 = 271.218$ ,  $df = 66$ ,  $p < 0.001$ ), 表明数据适合进行主成分分析。

Table 4. Results of KMO and Bartlett's test of sphericity for PCA  
表 4. PCA 的 KMO 和 Bartlett 球形度检验结果表

KMO 值		0.669
巴特利特球形度检验	近似卡方	271.218
	df	66
	p	0.000

通过计算得到各指标的贡献度如下表 5 所示:

**Table 5.** Weights of variables on principal components  
**表 5.** 各个变量在主成分上的权重

主成分	自然 风光	文化 体验	交通 便利	住宿 条件	餐饮 服务	特色 活动	价格 合理性	安全性	旅游 设施
$PC_1$	-0.310	0.352	-0.222	0.447	-0.156	0.427	-0.192	0.458	-0.274
$PC_2$	-0.315	-0.085	0.385	0.151	0.543	0.125	0.497	-0.069	-0.39
$PC_3$	-0.678	-0.226	0.000	-0.034	0.044	0.112	0.022	0.0779	0.682
$PC_4$	-0.202	0.518	0.557	0.121	-0.011	-0.194	-0.444	-0.354	0.064
$PC_5$	0.014	0.314	-0.587	-0.023	0.647	0.045	-0.151	-0.315	0.108
$PC_6$	-0.196	0.604	-0.127	-0.422	-0.217	-0.241	0.528	0.1287	0.006
$PC_7$	0.0523	0.063	0.144	-0.530	-0.101	0.783	-0.064	-0.245	-0.045

## (2) 主成分选择

根据主成分分析的 Kaiser 准则和累积方差贡献率标准, 计算可得到各主成分方差的贡献率, 如下表 6 所示。考虑到旅游相关研究通常要求较高的信息保留度, 本研究采用累积方差贡献率达到 85% 以上的标准进行主成分选择。结果显示, 前 5 个主成分的累积方差贡献率为 88.9%, 能够较好地反映原始变量的信息特征, 因此选择前 5 个主成分进行后续分析。

**Table 6.** Contribution rates of principal components  
**表 6.** 各个主成分贡献率

$PC_1$	$PC_2$	$PC_3$	$PC_4$	$PC_5$	$PC_6$	$PC_7$
21.5%	13.4%	25.4%	11.6%	17.0%	14.9%	14.3%

根据各主成分与旅游要素的载荷关系, 对主成分  $PC_1$  命名为旅游体验满意度因子,  $PC_2$  命名为文化与服务质量因子,  $PC_3$  命名为自然风光与安全性因子,  $PC_4$  命名为便利性与设施因子,  $PC_5$  命名为价格与住宿质量因子, 根据主成分的贡献率可以看出, 自然风光与安全性、旅游体验满意度以及价格与住宿质量是旅游者最关注的三个因素。

## 5. 基于随机森林回归分析的非遗传播研究

### 5.1. 特征选取

随机森林作为一种集成学习方法, 能够有效处理变量间的非线性关系, 对变量分布无严格要求, 更适合本研究的数据特征。因此, 在主成分分析的基础上, 提取前 5 个主成分因子得分作为自变量, 以旅游对非遗文化传播的影响程度(1~5 级有序变量)为因变量, 构建随机森林回归模型, 变量定义如下表 7 所示。

### 5.2. 随机森林回归模型建立

随机森林回归(Random Forest Regression)是一种基于决策树集成的非参数机器学习方法。设训练集为  $D = \{(x_i, y_i)\}_{i=1}^N$ , 其中  $x_i$  为输入特征向量,  $y_i$  为目标变量。随机森林通过有放回地从原始样本中抽取若干



子集，分别训练  $M$  棵回归树  $\{h_m(x)\}_{m=1}^M$ 。对于新样本  $x$ ，其预测值为所有树预测结果的平均值：

$$\hat{y} = \frac{1}{M} \sum_{m=1}^M h_m(x)$$

其中， $h_m(x)$  表示第  $m$  棵回归树对样本  $x$  的预测结果， $M$  为森林中树的总数。

在本研究中，首先采用对特征变量进行标准化处理，获得标准化特征矩阵。随后，初始化随机森林回归模型，设置 `n_estimators = 100`，`random_state = 42`。模型以标准化之后的特征和目标变量进行训练，并利用训练好的模型对数据进行预测分析。

**Table 7.** Features selected for the random forest regression model  
**表 7.** 选取随机森林回归模型的特征

变量类别	符号	变量名称
自变量	$X_1$	旅游体验满意度
	$X_2$	文化与服务质量
	$X_3$	自然风光与安全性
	$X_4$	便利性与设施
	$X_5$	价格与住宿质量
因变量	$Y_0$	旅游是否助于非遗传播

5.3. 随机森林模型检验

为全面评估模型性能，本研究采用均方误差(MSE)、均方根误差(RMSE)和平均绝对误差(MAE)等指标，量化预测值与实际观测值之间的偏差程度。结果显示(如下表 8 所示)，MSE、RMSE 和 MAE 均维持在较低水平，表明模型预测精度较高，拟合效果良好。此外，各特征变量的方差膨胀因子(VIF)均为 1，进一步证实模型不存在多重共线性问题，结果具有较高的可靠性。

**Table 8.** Evaluation of the random forest regression model  
**表 8.** 随机森林回归模型检验

MSE	RMSE	MAE
0.375	0.612	0.511

概率密度分布图(如下图 3 所示)显示，样本预测值主要集中于 0.6~0.75 区间，表明模型评估结果具有显著倾向性：乡村旅游对布依族非物质文化遗产的传播与保护呈现积极促进作用。该结论与预测概率的集中趋势一致。

5.4. 布依族非遗传播因素结果分析

通过主成分贡献率与特征重要性分析(如下图 4 所示)，发现主成分  $PC_3$ (自然风光与安全性)的权重和解释方差均为最高，表明该维度对旅游体验及非遗文化传播具有决定性影响。优越的自然环境与安全保障能够有效提升游客的文化体验深度和非遗传播意愿。其次，主成分  $PC_5$ (价格与住宿质量)和主成分  $PC_2$ (文化与服务质量)也表现出较高的重要性，说明合理的价格、优质的住宿及服务质量对非遗传播同样具有

积极作用。主成分  $PC_4$  (便利性与设施)和主成分  $PC_1$  (旅游体验满意度)的影响相对较弱,但仍具统计学意义,尤其在价格敏感型市场中,  $PC_5$  的显著波动性暗示其可能通过调节游客决策间接影响文化传播效果。

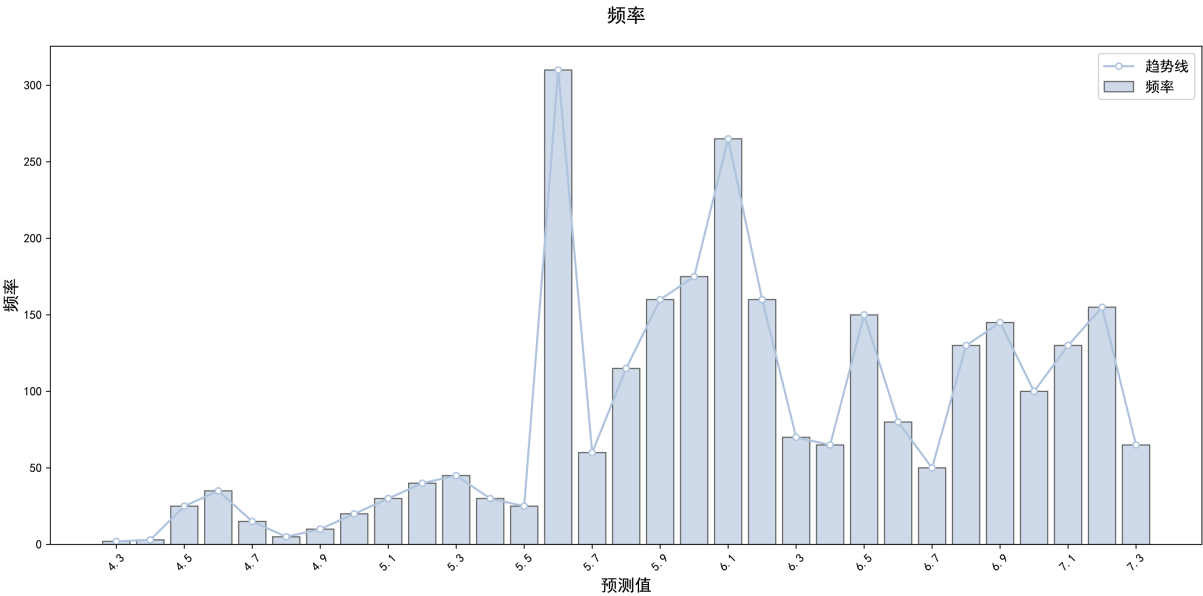


Figure 3. Distribution of predicted probabilities  
图 3. 预测概率分布图

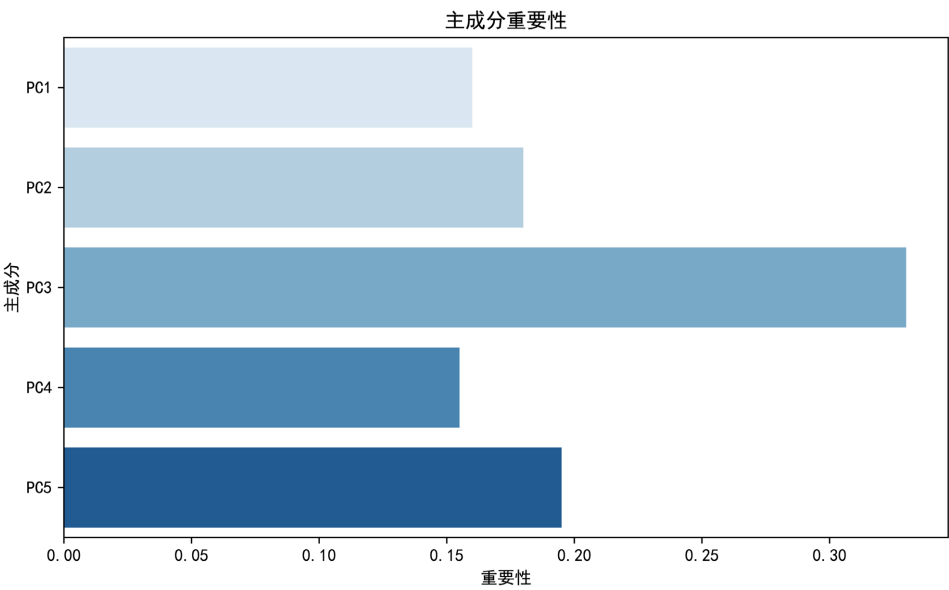


Figure 4. Importance of influencing factors  
图 4. 各因素重要性

## 6. 结论与建议

### 6.1. 结论

本研究采用主成分分析与随机森林回归,评估数字化旅游对布依族非遗传播的影响并预测其趋势。通过主成分降维提取 5 个主成分,识别出影响非遗传播的关键维度:自然风光与安全性是决定性因素;



文化与服务质量重要性较高；旅游体验满意度和便利性与设施影响相对较弱但仍具意义。结合非遗传播的活态性、体验性与文化真实性特征，各因素的重要性可从非遗传承本质进行系统性解释。

首先，自然风光与安全性作为决定性因素，其核心价值根植于非遗的活态传承特性——非遗作为“活态文化基因”，必须依附于特定的自然人文生态空间存在。布依族的稻作文化、蜡染技艺等非遗形式，其技艺展示、仪式展演均与喀斯特山区的梯田景观、河谷环境深度绑定，自然风光不仅是物理空间载体，更是文化意义的符号化表达，直接影响非遗体验的真实性与完整性。而安全性则构成活态传承的前提保障，在非遗传承常需深入偏远村寨的现实背景下，安全保障体系直接决定游客的参与深度，例如对传统村落消防设施、应急通道的改造，能够显著提升游客对“夜祭仪式”“古法造纸”等夜间非遗活动的参与意愿，进而强化活态传承效果。

其次，文化与服务质量的重要性源于非遗传播的体验性本质。非遗传播区别于普通文化传播的核心在于身体实践，游客通过参与非遗工坊的蓝靛染色、铜鼓铸造等生产性体验，实现从旁观者到传承参与者的身份转换。这种体验性特征要求文化展示必须保持原真性，服务供给需具备专业性，二者共同构成体验深度的基础——当游客在“三月三”节庆中通过传承人指导完成祭山仪式时，其获得的文化认知强度远高于被动观赏，这种深度体验正是非遗实现有效传播的关键路径。

最后，旅游体验满意度与便利设施作为辅助因素，其意义体现在非遗传播的可持续性层面。非遗的活态传承需要形成“体验-认同-传播”的闭环，良好的体验满意度能够促使游客转化为非遗文化的自觉传播者，例如通过社交媒体分享蜡染制作过程、苗族飞歌学习视频等内容，形成更快的传播效应。而便利设施则为这种持续体验提供物质支撑，如修建非遗工坊的无障碍通道可保障多元群体参与，完善物流体系能让游客将亲手制作的非遗产品便捷带回家，这种“体验-拥有”的完整链条，进一步强化了文化认同与传播黏性。

## 6.2. 建议

### (1) 加强核心因素

自然风光与安全性是影响布依族非遗传播的决定性因素，需优先通过生态保护(如划定核心景观保护区、实施生态修复工程)与安全设施升级(如完善景区监控系统、配备应急救援设备)双轨并行策略，系统性增强游客体验安全感与自然景观吸引力，从基础层面夯实非遗传播的场景支撑。

### (2) 优化关键维度

文化与服务质量是影响布依族非遗传播的重要因素，需重点通过文化体验深化与服务体系标准化双路径提升。文化层面，结合布依族蜡染、刺绣、八音坐唱等非遗特色，开发“非遗工坊沉浸式体验”“节庆文化深度参与”等项目，强化游客对非遗文化的感知与认同；服务层面，建立涵盖接待流程、讲解服务、应急响应的标准化体系，培训兼具非遗文化素养与服务技能的复合型人才，系统性提升服务质量与游客体验满意度，为非遗传播提供核心内容支撑。

### (3) 平衡辅助要素

旅游体验满意度和便利性与设施是影响布依族非遗传播的辅助性因素，需通过价格体系优化、住宿质量分层提升及便利性条件完善三维度协同推进。价格与住宿层面，结合不同客群需求推出差异化文旅套餐，在保障基础服务质量的前提下实现性价比与体验感的动态平衡；便利性层面，完善景区内外交通接驳系统，升级游客服务中心、导览标识等设施，降低游客参与非遗体验的时间与体力成本，为非遗传播提供无障碍的参与环境。

### (4) 创新传播模式

数字化传播与文旅融合是提升布依族非遗传播效能的关键创新方向，需通过“非遗+旅游+数字

技术”三位一体模式系统推进。数字技术赋能体验升级：利用 VR/AR 技术开发非遗场景沉浸式体验项目(如虚拟古寨漫游、蜡染工艺数字交互)，通过短视频平台(抖音、快手)打造“非遗 + 乡村生活”内容矩阵，扩大传播触达范围；文旅融合深化传播载体：依托拉落村“非遗三月三”节庆 IP，联动周边景区开发非遗主题旅游线路，推出“非遗工坊研学 + 自然景观游览”组合产品，增强传播场景的体验性与互动性；长效机制构建：建立非遗数字资源库(含工艺视频、传承人访谈、节庆影像)，与高校、文旅企业合作开展数字传播人才培养，推动非遗传播从单次体验向持续化、体系化转变，为非遗活态传承与文旅深度融合提供创新动力。

## 参考文献

- [1] 王科敏, 唐婉婷, 周煜, 等. 乡村振兴背景下布依族蜡染在现代手工艺中的创新应用研究[J]. 纺织报告, 2025, 44(10): 146-148.
- [2] 张袁子奇, 杨怡. 布依八音奏新声乡土传承育新人[N]. 曲靖日报, 2025-10-28(003).
- [3] 代诗敏. 黔西南布依族传统聚落景观意象的建构[D]: [硕士学位论文]. 天津: 天津美术学院, 2019.
- [4] 王晨阳. 以习近平文化思想引领中国特色非物质文化遗产保护制度建设[J]. 中国非物质文化遗产, 2026(1): 5-18.
- [5] 朱碧茵. 乡村振兴战略下地方非遗产业化创新路径探索[J]. 农业产业化, 2025(12): 30-32+39.
- [6] 张青, 朱亚成. 乡村全面振兴背景下藏族赛马节与生态旅游融合发展研究[J]. 安徽农业科学, 2025, 53(24): 111-114.
- [7] 郑洪凤. 数字技术在四川省凉山州文化遗产保护中的应用研究[D]: [硕士学位论文]. 成都: 四川省社会科学院, 2025.
- [8] 李文霞, 张婷婷. 数字技术在文化遗产保护中的运用[J]. 文物鉴定与鉴赏, 2023, 263(20): 56-59.
- [9] 朱米娜, 陈旺. 数字技术在非物质文化遗产传承保护中的应用研究——以徽雕艺术为例[J]. 艺术品鉴, 2017(1): 336.
- [10] 肖远平, 王伟杰. 中国少数民族非遗名录及传承人统计分析[J]. 西南民族大学学报(人文社会科学版), 2016, 37(1): 40-45.
- [11] 孙九霞, 李怡飞. 社会资本视角下非遗传承人的地位获得[J]. 湖北民族大学学报(哲学社会科学版), 2018, 36(6): 5-11.
- [12] 胡艳丽, 张雪梅. 苗侗“非遗传承人”生产性传承培育研究[J]. 原生态民族文化学刊, 2017, 9(3): 142-146.
- [13] 卢颖. 非遗传承与乡村振兴中的文化保护和产业发展[J]. 民间故事选刊, 2024(14): 7-9.
- [14] 温必锋, 尹海鹰, 沈素惠. 非遗“活态”传承发展与创新策略的研究[J]. 汽车维修与修理, 2025(24): 4-5.
- [15] 朱华明, 石巍. 非遗传承信任体系建设的对策思考[J]. 质量与认证, 2025(12): 25-26.
- [16] 焦明瑞. 衡水非遗: 从“守护传承”到“融入生活”[N]. 衡水日报, 2025-12-04(004).