

基于主成分分析法的城市化水平演化研究

——以皖江城市带为例

胡孔玉¹, 王竞钊²

¹上海工程技术大学管理学院, 上海

²武汉工程大学土木工程与建筑学院, 湖北 武汉

收稿日期: 2021年9月22日; 录用日期: 2021年10月15日; 发布日期: 2021年10月26日

摘 要

进入21世纪以来, 我国的城市化进程逐步加快, 城市发展也呈现出多样化的态势。通过选取安徽省城市化水平较高的皖江城市带所包含的城市作为项目研究案例开展实证研究, 基于研究区2009~2018年的相关统计数据, 利用人口城市化率、人均GDP以及土地城市化率三个指标评价了区域人口、经济、空间城市化水平, 并进一步利用主成分分析法整合三个指标获得综合指标值, 对皖江城市带城市化水平进行了综合评价分析。结果表明: 皖江城市带人口城市化增长质量不一, 经济城市化稳定上升, 土地城市化在空间上差异较大, 区域综合城市化水平整体呈波动上升趋势。

关键词

主成分分析法, 皖江城市带, 城市化水平, 演化

Research on the Evolution of Urbanization Levels Based on Principal Component Analysis

—Taking the Wanjiang City Belt as an Example

Kongyu Hu¹, Jingzhao Wang²

¹School of Management Studies, Shanghai University of Engineering Science, Shanghai

²School of Civil Engineering and Architecture, Wuhan Institute of Technology, Wuhan Hubei

Received: Sep. 22nd, 2021; accepted: Oct. 15th, 2021; published: Oct. 26th, 2021

Abstract

Since the beginning of the 21st century, the process of urbanization in China has been gradually accelerated, and the urban development has shown a trend of diversification. In this paper, the cities in Wanjiang City Belt with high urbanization level in Anhui province are selected as project cases to carry out empirical research. Based on the relevant data of the research area from 2009 to 2018, three indexes of population urbanization rate, per capita GDP and land urbanization rate are used to evaluate the population, economy and spatial urbanization level of the region. The paper analyzes the level of urbanization in Wanjiang City Belt by integrating the three indexes with the principal component analysis method. The results show that the growth quality of population in Wanjiang City Belt is different, the spatial difference of land urbanization is great, the economic urbanization increases stably, but the overall urbanization level shows fluctuant rise trend.

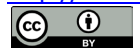
Keywords

Principal Component Analysis, Wanjiang City Belt, The Urbanization Level, Process

Copyright © 2021 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

城市是“以人为主体,以空间和自然环境的合理利用为前提,以聚集经济效益和社会效益为目的,集约人口、经济、科技、文化的空间地域大系统”[1]。人口、经济与空间则是能够显而易见地展现城市发展的重要显性特征,也是衡量一个城市的城市化水平的主要标准。随着国家经济的快速增长,以及城乡一体化发展政策的不断推进,我国城市化过程中累积的问题日益增多,加强对城市化的深入研究尤其现实意义。

纵观已有的城市化相关研究,专家学者们从城市化研究的具体内容,采用的方法,研究的时空尺度,案例地的特点等各个不同角度开展了大量探讨。大部分的国外学者主要以全球或区域等大尺度的城市化发展为研究课题,对其社会、经济发展规律等有着较为系统的探索,而国内学者更倾向于以地区或城市等小尺度的城市化发展为研究课题。城市化的“S”型曲线是诺瑟姆针对城市化发展规律所提出的重要研究结论[2];H·钱纳里也对区域城市化进程与经济发展水平之间的联系进行了探讨[3]。沈建法及其研究团队通过对珠三角发展的研究,对其未来发展态势提出了双轨城市化的理论[4]。刘耀彬通过构建一个耦合关联程度模型,将中国区域城市化与生态进行了关联[5]。从研究方法来看,学者们采用的研究方法也在不断多样化,陈彦光等人通过适当修正 Northam 曲线,改进 Logistic 过程,得到了四阶段结果[6]。武盈盈、张伟基于拔靴滚动检验方法探索了中国城市化与碳排放动态因果关联[7]。从研究角度来看,学者们更偏向于将单一尺度测量所得的城市化水平与生态、产业结构等进行耦合分析,李志翠等人基于面板数据尝试解读产业结构合理化和高级化演变能够正向促进城市化水平的提升[8];黑龙江位于我国东北部,经济发展速度有待提高,徐素波对该省 12 个地级市展开研究发现黑龙江省的城市化发展水平与生态环境耦合度呈现高度协调发展趋势[9]。针对城市化进程的探究,学者们由最初局限的使用定性分析方法逐渐发展为定性与定量结合的思维,专注于专业和局部地区的单项研究也日益增多,并逐渐偏向综合性研究。

皖江城市带承接产业转移示范区包括合肥、芜湖、马鞍山、铜陵、安庆、池州、巢湖、滁州、宣城 9 市全境, 以及六安市的金安区和舒城县, 共 59 个县(市、区) [10] (图 1)。皖江城市带人口约 3222 万余人, 截止到 2018 年国内生产总值达 19,411 亿元。2009 年到 2018 年期间, 该地区人均 GDP 实现了 24,524 元/人到 56,303 元/人的稳步增长, 增长率达 129.58%。由于该地区与长三角在经贸上的往来十分密集, 十年间, 第二产业已经成为皖江城市带经济增长的动力支撑, 第三服务类产业的份额也实现了跨越式提升。

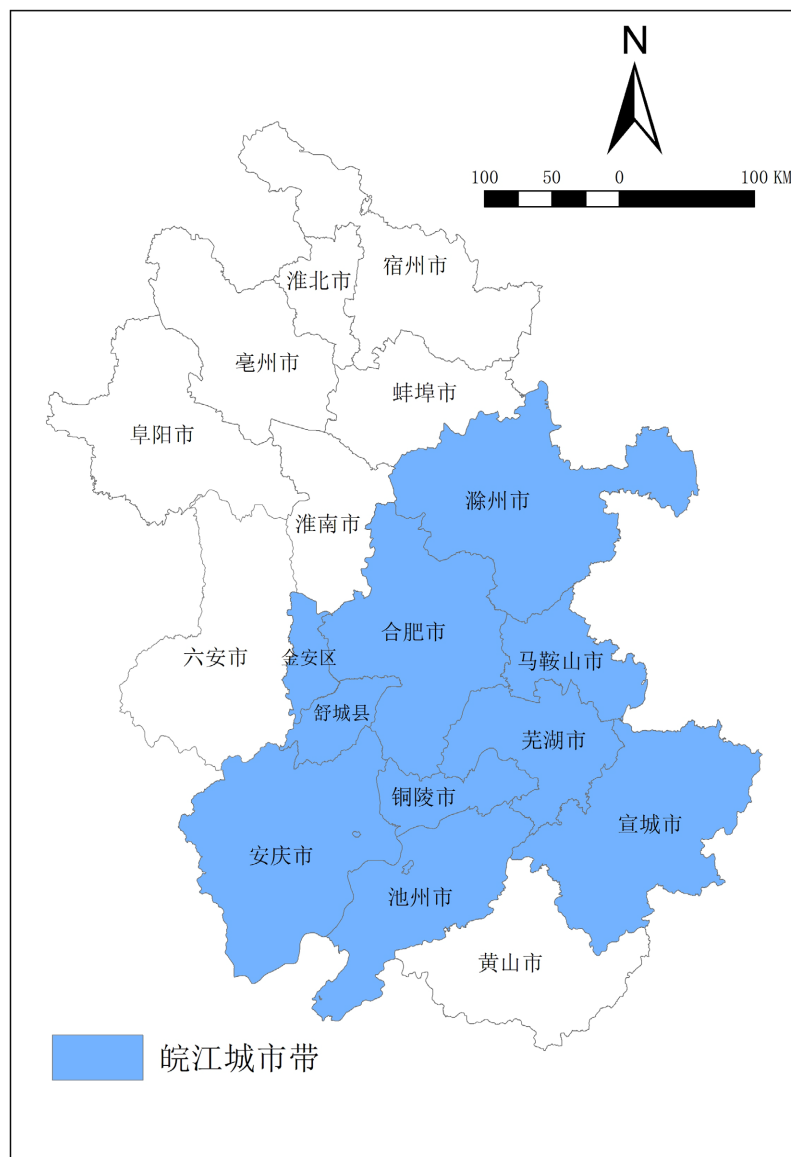


Figure 1. Diagram of the Wanjiang City Belt
图 1. 皖江城市带示意图

皖江城市带是长三角地区不可或缺的组成部分, 在我国实施促进中部崛起战略过程中的发挥着重要作用, 在中西部承接产业转移中扮演着重要角色。近年来运用主成分分析法研究城市化发展水平的研究日益减少, 针对皖江城市带的城市化研究更是屈指可数, 因此本文以皖江城市带范围内的各个城市作为研究对象, 对其城市化演化水平展开分析, 以期为皖江城市带的持续发展提供有益参考。

2. 数据与方法

2.1. 数据来源

本文以皖江城市带 2009 年至 2018 年十年为时间范畴, 在客观公平公正的基础上构建人口城市化、经济城市化、空间城市化共 3 个评价指标。本次研究的数据均来自于 2009 年至 2018 年的《六安统计年鉴》、《安徽省政府信息公开工作年度报告》、《安徽省国民经济和社会发展统计公报》、《中国统计年鉴》、《安徽统计年鉴》。其中 2010~2018 年各市人口城市化率是直接从各年统计年鉴中直接获取, 人口城市化率是每年非农人口与户籍人口的比值, 经济城市化借以城市人均 GDP 为指标, 空间城市化利用土地城市化率, 其数值采用地均 GDP 为指标, 是每年各地区生产总值与每年区域总面积的比值。

2.2. 指标选取

城市化描述的是一个涉及人口、经济、空间三方面的综合复杂的过程, 据此将城市化归纳为人口城市化、经济城市化和空间城市化三个方面。根据所搜集到的数据, 本文采用人口城市化率、人均生产总值、土地城市化率三个指标来综合评价皖江城市带城市化水平(表 1)。

Table 1. Urbanisation evaluation indicator system

表 1. 城市化评价指标体系

系统层	子系统层	指标层(单位)
城市化	人口城市化	人口城市化率(%)
	经济城市化	人均生产总值(元/人)
	空间城市化	土地城市化率(万元/km ²)

2.3. 人口城市化指标计算

人口城市化是指农村人口转换为城市人口的过程, 本文采用人口城市化率来评价人口城市化程度。人口城市化率最直接的标志是城镇人口比重, 其计算公式为:

$$\text{人口城市化率}(\%) = \frac{\text{城镇人口数(非农人口)}}{\text{常住人口数(户籍人口)}} \quad (1)$$

基于数据来源的限制, 2014 年之前(包括 2014 年)巢湖市和六安市的两个地方的人口城市化率依据非农人口和户籍人口计算, 2015 年及以后巢湖市, 六安市人口城市化率依据城镇人口和常住人口计算。

2.4. 经济城市化指标计算

经济城市化代表的是经济结构由农业向工业、服务业转换的过程, 用产业结构描述即为第一产业向第二产业、第三产业转移的过程。本文采用研究区人均生产总值指标来评价经济城市化程度, 计算公式为:

$$\text{人均生产总值(元/人)} = \frac{\text{区域生产总值}}{\text{区域总人口}} \quad (2)$$

2.5. 空间城市化指标计算

空间城市化是人口、经济城市化过程在地域空间的外在表现, 本文选取土地城市化率这个指标来描述研究区空间城市化水平。根据所搜集到的数据, 最终确定通过土地利用效益评价指标——地均生产总值来评价土地城市化率, 计算公式为:

$$\text{土地城市化率(万元/平方千米)} = \frac{\text{地区生产总值}}{\text{区域总面积}} \quad (3)$$

2.6. 城市化水平综合指标计算

为了减少评价指标之间的干扰, 以实现减少选择指标的工作量的最终目的, 本文借用 SPSS 软件, 采用主成分分析法对人口、经济、空间城市化三个指标数据进行处理。主成分分析法是通过恰当的数学变换, 使新变量主成分成为原变量的线性组合, 并选取少数几个在变差总信息量中比例较大的主成分来分析事物的一种方法。运用主成分分析法可消除评价指标之间的相互影响, 同时也可以减少选择指标的工作量[11]。

2.6.1. 原始数据的标准化

设有 n 个样本, p 项指标, 可得数据矩阵 $X = (X_{ij})_{n \times p}$, $i=1,2,\dots,n$ 表示 n 个样本, $j=1,2,\dots,p$ 表示 p 个指标, X_{ij} 表示第 i 个样本的第 j 项指标值。

$$X_{n \times p} = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1p} \\ X_{21} & X_{22} & \cdots & X_{2p} \\ \cdots & \cdots & \cdots & \cdots \\ X_{n1} & X_{n2} & \cdots & X_{np} \end{bmatrix}$$

用 Z-score 法对数据进行标准化变换:

$$Z_{ij} = (x_{ij} - \bar{x}_j) / S_j$$

式中, $\bar{x}_j = (\sum_{i=1}^n x_{ij}) / n$, $S_j^2 = [\sum_{i=1}^n (x_{ij} - \bar{x}_j)^2] / (n-1)$, $i=1,2,\dots,n$, $j=1,2,\dots,p$
标准化矩阵

$$Z_{n \times p} = \begin{bmatrix} Z_{11} & Z_{12} & \cdots & Z_{1p} \\ Z_{21} & Z_{22} & \cdots & Z_{2p} \\ \cdots & \cdots & \cdots & \cdots \\ Z_{n1} & Z_{n2} & \cdots & Z_{np} \end{bmatrix}$$

2.6.2. 求指标数据的相关矩阵

$R = (r_{jk})_{p \times p}$, $j=1,2,\dots,p$, $k=1,2,\dots,p$, r_{jk} 是指标 j 和指标 k 的相关系数。

$$r_{jk} = \frac{\frac{1}{n-1} \sum_{i=1}^n [(x_{ij} - \bar{x}_j)^2]}{S_j} \left[\frac{(X_{ik} - \bar{X}_k)^2}{S_k} \right]$$

即 $r_{jk} = \frac{1}{n-1} \sum_{i=1}^n Z_{ij} Z_{ik}$ 有 $r_{ii} = 1$, $r_{jk} = r_{kj}$, $i=1,2,\dots,n$, $j=1,2,\dots,p$, $k=1,2,\dots,p$

2.6.3. 求相关矩阵 R 的特征根特征向量, 确定主成分

由特征方程式 $|\lambda_p R - R| = 0$, 可求得 p 个特征根 $\lambda_g (g=1,2,\dots,p)$, λ_1 将其按大小顺序排序为 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$, 它是主成分的方差, 它的大小描述了各个主成分在描述被评价对象上所起作用对的大小, 由特征方程式, 每一个特征根对应一个特征向量 Lg 。

将标准化后的指标变量转换为主成分:

$$F_g = l_{g1} Z_1 + l_{g2} Z_2 + \cdots + l_{gp} Z_p \quad (g=1,2,\dots,p)$$

F_1 称为第一主成分, F_2 称为第二主成分, \dots , F_p 称为第 p 主成分。

2.6.4. 求方差贡献率, 确定主成分个数

一般主成分个数等于原始指标个数, 如果原始指标个数较多, 进行综合评价时就比较麻烦。主成分分析法就是选取尽量少的 k 个主成分($k < p$)来进行综合评价, 同时还要使损失的信息量尽可能少。

k 值由方差贡献率 $\sum_{g=1}^k \lambda_g \sum_{g=1}^p \lambda_g \geq 85\%$ 决定。

2.6.5. 对 k 个主成分进行综合评价

先求每一个主成分的线性加权值 $F_g = l_{g1}Z_1 + l_{g2}Z_2 + \dots + l_{gp}Z_p$, $g = 1, 2, \dots, k$, 再对 k 个主成分进行加权求和, 即得到最终评价价值, 权数为每个主成分的方差贡献率 $\lambda_g \sum_{g=1}^p \lambda_g$, 最终评价价值

$$F = \sum_{g=1}^k \left(\lambda_g \sum_{g=1}^p \lambda_g \right) F_g \quad [12] [13] [14] [15]$$

2.7. 计算结果

在使用 SPSS 软件对主成分进行分析的同时, 标准化处理也会同步自动进行。用 KMO 和 Bartlett 检验法得出的结果见表 2。通常 KMO 检验的值比 0.6 大, 表示检验合格, 能进行下一步的主成分分析环节。下表中的结果明显符合这个条件, 说明数据适合因子模型分析。

Table 2. KMO and Bartlett's test
表 2. KMO 和 Bartlett 的检验

取样足够度的 Kaiser-Meyer-Olkin 度量		0.694
	近似卡方	6.859
Bartlett 的球形度检验	df	3
	Sig	0.077

通过计算, 最终评价价值 $F = 0.33F_1 + 0.35F_2 + 0.32F_3$, 其中, F_1 为人口城市化率, F_2 为土地城市化率, F_3 为人均 GDP, 代入标准化变化后的数值, 可得 2009~2018 年三个指标的综合评价价值(F 值)。表 3~6 分别为 2009~2018 年皖江城市带各地区人口城市化率平均值、经济城市化率平均值、土地城市化率平均值和皖江城市带城市化水平综合评价价值。

Table 3. 2009~2018 average urbanization rate of population by region in the Wanjiang City Belt
表 3. 2009~2018 年皖江城市带各地区人口城市化率平均值

城市	合肥市	芜湖市	马鞍山市	铜陵市	安庆市	滁州市
增长率	0.72424	0.63850	0.01412	-0.26715	0.30293	0.34481
城市	巢湖市	池州市	宣城市	金安区	舒城县	
增长率	0.63209	0.28810	0.27593	0.53034	0.84040	

Table 4. 2009~2018 average economic urbanization rate by region in the Wanjiang City Belt
表 4. 2009~2018 年皖江城市带各地区经济城市化率平均值

城市	合肥市	芜湖市	马鞍山市	铜陵市	安庆市	滁州市
平均值	1.34387	1.28478	0.59252	0.62505	1.88499	2.44726
城市	巢湖市	池州市	宣城市	金安区	舒城县	
平均值	2.14234	1.70974	1.98468	2.65786	1.47165	

Table 5. 2009~2018 average land urbanization rate by region in the Wanjiang City Belt
表 5. 2009~2018 年皖江城市带各地区土地城市化率平均值

城市	合肥市	芜湖市	马鞍山市	铜陵市	安庆市	滁州市
平均值	45964.16	16638.60	22256.62	15713.52	17002.66	8202.48
城市	巢湖市	池州市	宣城市	金安区	舒城县	
平均值	1627.03	1930.99	3314.58	864.62	673.86	

Table 6. 2009~2018 comprehensive assessment values of urbanization levels in the Wanjiang City Belt
表 6. 2009~2018 年皖江城市带城市化水平综合评价值

年份	2009	2010	2011	2012	2013
F 值	-1.51103	-1.08028	0.10072	0.20342	-0.23358
年份	2014	2015	2016	2017	2018
F 值	0.76467	-0.04621	-0.17912	1.36057	0.62339

3. 讨论与结论

对人口, 经济, 空间城市化三个指标的研究可知, 三个因素对一个区域的城市化水平的影响是几乎均等的。在区域城市化演化进程中, 皖江城市带人口城市化水平增长一直滞后于经济城市化和土地城市化增长进程, 城市人口增长与土地扩张速率不匹配, 由此导致区域整体城市化水平发展速率不高。

人口城市化增长质量不一。随着安徽省户籍制度改革的推进, 农村人口向城镇地区流动相较之前更加自由, 加快推进了城镇人口增长, 促进了人口城市化率的增长。研究区范围内, 除铜陵市以外, 其他各地区人口城市化水平呈稳定增长趋势, 但增长率差异较大。

经济城市化水平稳定上升。皖江城市带整体在研究期间经济城市化水平稳定上升, 增长幅度维持在 9.88% 左右。其中, 增长速度最快的城市为滁州市, 增长率为 244.73%, 最低的城市是马鞍山市, 增长率为 59.24%。城市化率的提高会带动城市的经济发展, 但经济发展水平高的地区城市化水平不一定高, 二者并不存在明显的正相关关系。

空间城市化差异较大。研究区土地城市化率在空间上差异明显。皖江城市带中部地区土地城市化率普遍较高, 均超过了 15,000 万元/km², 南北地区的城市土地城市化率与之差异较大, 土地城市化率最高的滁州也仅为 8208.48 万元/km²。但总体而言, 皖江城市带的空间城市化演变是一个城市扩张, 侧重于土地资源开发的城市建设过程。

综合城市化水平整体波动上升。2009~2018 年皖江城市带城市化水平整体都处在增长的态势, 但在 2012 年之后该指标出现了短暂性的下降, 后在 2014 年再次上升, 2015 年之后又一次连续两年下降, 但 2016 年下降幅度较小, 在 2017 年大幅度上升, 到 2018 年短暂下降, 综上数据表明皖江城市带的整体城市化水平一直处于波动上升阶段。

2009~2018 年期间, 波动上升成为皖江城市带整体城市化水平的主流态势, 但在城市带内部各个城市之间的城市化水平的发展速度也存在一定的分化, 不平衡性与层次性是最主要的特征之一, 东部地区的城市化水平相较于西部地区具有明显的优势。由于城市化水平与地区经济水平并不存在绝对的正相关关系, 因此城市在增加自身综合实力的基础上要注重区域协调发展, 经济发展不平衡的区域之间可以进行积极的经济交流合作。今后皖江城市带区域城市应做到深入贯彻落实节约集约的发展理念, 尽量减少城镇规模扩张过程中所造成的土地资源的消耗, 提高资源的产出率, 协调人口、经济、空间三者之间的

关系, 打造城市化水平更高的城市。在推进经济发展的同时要注重人口素质发展, 对于像安庆等人口基数大而经济水平相对落后地区, 必须要加大对教育、文化和科技等方面的投资, 优化人口受教育程度结构, 推行鼓励积极就业的政策, 提高劳动力水平, 更要提高生产力水平, 从而促进人口规模与经济水平协调发展。

基金项目

武汉工程大学研究生教育创新基金(CX2020090)。

参考文献

- [1] 盛广耀. 城市化模式研究综述[J]. 城市发展研究, 2011, 18(7): 13-19.
- [2] 陈明星, 叶超, 周义. 城市化速度曲线及其政策启示——对诺瑟姆曲线的讨论与发展[J]. 地理研究, 2011, 30(8): 1499-1507.
- [3] 李晶, 谭少华. 国内外城市化研究进展综述[J]. 山西建筑, 2007, 33(31): 21-22.
- [4] 沈建法, 冯志强, 黄钧尧. 珠江三角洲的双轨城市化[J]. 城市规划, 2006(3): 39-44.
- [5] 刘耀彬, 李仁东, 宋学锋. 中国区域城市化与生态环境耦合的关联分析[J]. 地理学报, 2005, 60(2): 237-247.
- [6] 陈彦光, 周一星. 城市化 Logistic 过程的阶段划分及其空间解释——对 Northam 曲线的修正与发展[J]. 经济地理, 2005, 25(6): 817-822.
- [7] 武盈盈, 张伟. 中国城市化与碳排放动态因果关系研究: 基于拔靴滚动检验方法[J]. 东岳论丛, 2019, 40(10): 84-93.
- [8] 李志翠, 朱琳, 张学东. 产业结构升级对中国城市化进程的影响——基于 1978-2010 年数据的检验[J]. 城市发展研究, 2013, 20(10): 35-40.
- [9] 徐素波. 黑龙江省城市化与生态环境耦合协调发展研究[J]. 林业经济, 2019, 41(7): 94-100.
- [10] 搜狗百科. 皖江城市带[DB/OL]. <https://baike.sogou.com/v3224193.htm>, 2021-03-25.
- [11] 吴学泽, 王心源, 王官勇. 皖江城市带经济差异与其自然地理成因分析[J]. 国土与自然资源研究, 2008(4): 5-6.
- [12] 李艳双, 曾珍香, 张闯, 于树江. 主成分分析法在多指标综合评价方法中的应用[J]. 河北工业大学学报, 1999, 28(1): 94-97.
- [13] 王龙惠. 安徽省新型城市化水平测度——基于 2018 年安徽省统计年鉴的因子分析模型[J]. 湖北科技学院学报, 2019, 39(5): 27-32.
- [14] 韩会然, 焦华富, 李俊峰, 王荣荣. 皖江城市带空间经济联系变化特征的网络分析及机理研究[J]. 经济地理, 2011, 31(3): 384-389.
- [15] 张璐. 第三方支付与商业银行经营绩效的研究[D]: [硕士学位论文]. 绵阳: 西南科技大学, 2021: 49-55.