

中国科技论文外流态势监测指数设计 and 应用*

陈国娇¹ 田瑞强^{1,2}

(1. 中国科学技术信息研究所, 北京 100038; 2. 医学期刊知识挖掘与服务重点实验室, 北京 100038)

摘要: 科研成果是国家走向科技自立自强的重要战略资源, 科技论文是科研成果的重要载体。科技论文的外流为我国科技发展和国家安全带来诸多负面影响。为了全面了解我国科技论文外流情况, 通过采集1987—2022年发表在非中国SCI科技期刊上的中国论文数据, 借助ARMA模型建立中国科技论文外流数量预测模型, 并设计中国科技论文外流态势监测指数进行应用分析。结果表明, 中国科技论文并非一直处于外流的状态, 中国科技论文外流态势监测指数在1988—1995年快速下降, 在1996—2006年快速上升, 在2006年以后逐渐趋于稳定。从外流情况来看, 中国科技论文在1990—2003年未表现出外流特征, 在2004—2021年出现持续外流的特征, 在2022年表现出回流的特征。要进一步扭转我国科技论文外流趋势, 需要从几个方面重点发力: 加快进行科研评价体系改革; 在关键领域鼓励创办优秀英文科技期刊; 扩大国际合作范围, 加深国际合作。

关键词: 科技论文; 论文; 科技期刊; SCI论文; 时间序列分析

中图分类号: G350 DOI: 10.3772/j.issn.1673-2286.2024.01.004

引文格式: 陈国娇, 田瑞强. 中国科技论文外流态势监测指数设计 and 应用[J]. 数字图书馆论坛, 2024, 20(1): 33-38.

新时代发展背景下, 我国科学技术快速发展, 科学家数量日渐增多, 科技论文数量快速增长。2021年SCI收录中国科技论文数量位列世界第一^[1]。我国科技创新水平不断提高, 但中国科技期刊影响力仍然不足, 部分科研人员过度追求高影响因子, 更倾向于将优秀论文发表在国外科技期刊上。2002年7月, 中国科学院院士陈家镛^[2]指出, 我国不恰当地强调在具有高影响因子的SCI期刊上发表论文, 造成我国大量高水平论文在其他国家高影响因子期刊上发表。2005年, 《中国教育报》《光明日报》等多家媒体披露科研论文外流现象^[3]。近几年的统计数据也揭示了这一问题。EI收录的在中国科技期刊上发表的中国科技论文比例从1980年的35.1%下降到2019年的5.0%^[4]。《中国科技期刊发展蓝皮书(2020)》显示, 2010—2019年中国作者向SCI数据库贡献了18.06%的论文, 但发表在中国SCI期刊上的论文

仅占1.72%^[5]。科技论文外流成为中国科技和出版领域重点关注的课题。

1 研究背景

我国科技论文外流问题的形成有其客观原因。一方面, 自21世纪以来, 我国科技快速发展, 尤其是加入世贸组织以后, 我国加快了对外开放的进程。要促进科技进步, 就必须参与全球科技发展。我国科研人员数量快速增长, 科研评价体系也更加注重量化评价指标, 因此我国SCI论文数量快速上升。另一方面, 国外科技期刊发展起步早于我国, 国际出版商通过全球性的推广营销方式, 建立了广为人知的科技期刊品牌, 而我国的SCI期刊数量远远少于主要英语科技期刊出版国家。2020年中国作者在SCI期刊上发表论文55万余篇, 而我

收稿日期: 2023-10-23

*本研究得到中国科学技术信息研究所创新研究基金—青年项目“‘破五唯’背景下学术评价的突破点研究”(编号: QN2023-10)资助。

国SCI期刊数量仅为225种,占SCI期刊总量的1.1%^[1],因此我国SCI期刊客观上无法容纳大量的高水平科技论文。此外,国外科技期刊发展成熟,也更有竞争力。在数字时代,国际出版商能够更快地利用最新的技术手段,其编审系统和流程能够更好地满足作者需求。相比国内同水平科技期刊,国外科技期刊录用率更高、发表周期更短^[6],这也导致了我国的科研人员倾向于向国外科技期刊投稿。

在世界新竞争格局背景下,大量科技论文的外流为我国科技发展带来了诸多负面影响。①知识产权流失和科研经费外流。大量科研人员在国外科技期刊上发表论文,国外出版商或科技期刊通常要求作者签署不合理的版权转让协议,导致了我国科技论文知识产权的流失。同时,科技论文外流也伴随着大量科研经费外流,包括高昂的论文处理费、论文发表服务费、科学基金经费、国外商业全文数据库订阅费等^[7]。②削弱国家自主创新能力。科技论文的大量外流影响最新科研成果在国内的传播和科研成果的及时转化与利用,从长远来看,会削弱国家自主创新能力。③带来科技信息安全隐患。重要领域科技论文大量外流,会导致我国高科技情报外泄。部分国际科技期刊要求作者提供论文的科学研究数据,导致相关科学数据外流,不利于国家信息安全保护。④影响中国科技期刊水平和国际声誉。大量科技论文外流导致我国科技期刊丧失优质稿源,进而极大地阻碍我国科技期刊的发展^[8]。因此,研究中国科技论文外流问题具有重要的现实意义。

不少学者对我国科技论文外流现象展开研究。宋鸿^[9]认为,修订我国的科技评价指标、强化中国科技期刊编委责任和提高中国科技期刊学术质量是减少优秀论文外流的重要途径。曾建勋等^[7]结合国家工程技术图书馆统计数据,分析了我国科技论文外流、科研经费外流和科学数据状况,发现中外科技论文交流逆差巨大。刘丽英等^[10-12]通过对农业工程领域论文外流及其造成的经济损失进行研究和分析,提出端正科研导向、创办具有国际知名度的英文期刊是抑制优秀稿源外流的主要措施。王密平^[13]分析了医学学科在2009—2014年的论文外流情况,发现论文外流的主要原因是我国科技期刊吸引力不足。沈菲飞^[14]统计分析了15所高校与学术论文评价和奖励相关的94份制度文件,深入分析了我国高校学术论文外流现象背后的政策原因,发现评价导向是科技论文外流的重要原因;高校过度依赖国外评价体系评判论文水平,导致马太效应,增大了优秀论文

外流的概率。

现有研究对我国科技论文外流的分析主要建立在中国作者在非中国SCI期刊上发表的论文数量的绝对值和增速的基础上,存在一定的局限性。为了对中国科技论文外流问题进行深入研究,本文对1987—2022年中国科研人员在SCI期刊上发表论文数量的变化趋势进行分析,同时借助自回归滑动平均(Auto-Regressive and Moving Average, ARMA)模型研究发表在非中国SCI期刊上的论文数量的时序变化趋势,根据预测值和实际值的差异设计能够反映中国科技论文外流态势的监测指数,以预测未来中国研究人员在非中国SCI期刊上发表的论文数量。

2 研究方法与数据来源

在我国加快建设国际一流科技期刊,中国高质量科技论文数量持续增长的背景下,需要持续动态监测我国科技论文外流态势,以有效评估我国采取的相关政策是否取得了成效。为此,借助时间序列分析模型,对在非中国SCI期刊上发表的论文数量作长期趋势分析,设计我国科技论文外流态势监测指数。

ARMA模型是时间序列研究的重要方法,以自回归模型与滑动平均模型为基础。其基本思想是:某时间序列是一组时间变量,单个序列值虽然具有不确定性,但整个序列的变化有一定的规律,可以用相应的数学模型近似描述。ARMA模型是非常经典的预测模型,适用范围广泛,可以应用于各种时间序列数据,可以捕捉到时间序列数据中的自相关性和滞后误差,并提供准确的预测结果。其参数具有直观性,表示了过去观测值和过去滞后误差对当前观测值的影响程度,可以帮助理解时间序列数据的动态特征。ARMA模型可以用来预测未来的时间序列值,也可以用来分析序列中的趋势、季节性等特征。利用ARMA模型并使用已知的历史数据,可以生成未来观测值,并估计不确定性。中国科技论文外流数据符合时间序列特征,因此使用ARMA模型对其进行预测。模型检验效果良好,能够有效反映中国科技论文外流数据的时序特征。

研究数据来源于SCI数据库,检索时间为2023年9月。由于我国科学界自1987年开始采用国际通用的科学计量指标评估高校和科研院所的研究成果,检索1987—2022年中国作者在非中国SCI期刊上发表的论文。文献类型选择Article和Review。

3 指数设计

首先, 对1987—2022年中国作者在非中国SCI期刊上发表论文的数量 n_{foreign} 数据做真实值曲线图(见图1), 观察其增长趋势, 可以直观发现中国作者在非中国SCI期刊上发表的论文数量具有明显的指数增长趋势。序列不平稳, 故对原始数据取对数。

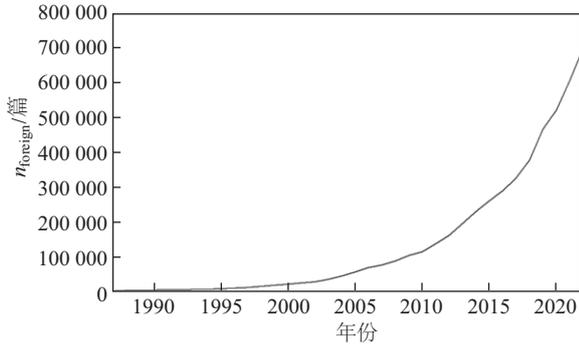


图1 1987—2022年中国作者在非中国SCI期刊上发表的论文数量曲线

取对数后, $\ln n_{\text{foreign}}$ 具有比较明显的线性增长趋势(见图2)。序列不平稳, 需要对 $\ln n_{\text{foreign}}$ 进行一阶差分, 并进行单位根检验。单位根检验结果见表1。 $\ln n_{\text{foreign}}$ 一阶差分后, 单位根检验的 $P < 0.05$, 拒绝原假设, 则认为一阶差分后的序列平稳。使用ARMA模型对 $\ln n_{\text{foreign}}$ 进行长期追踪。对序列 $\ln n_{\text{foreign}}$ 进行自相关检验, 结果(图3)表明自相关性和偏相关性均在2倍误差范围内(见图3虚线范围内), 因此选择AR(1)模型进行时间序列分析, 输出结果见表2。

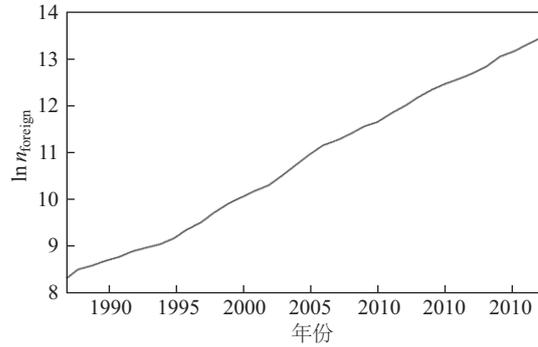


图2 取对数后的中国作者在非中国SCI期刊上发表的论文数量曲线

输出结果通过 F 检验, 表明模型有效, 因此, 整理后得出 $\ln n_{\text{foreign}}$ 模型表达式, 如式(1)所示。

表1 一阶差分后的单位根检验结果

Augmented Dickey-Fuller检验		临界值		
t 统计量	P	1%水平	5%水平	10%水平
-4.393 46	0.001 4	-3.639 41	-2.951 13	-2.614 30

自相关性	偏相关性	自相关系数	偏相关系数	Q 统计量	P
		1 0.296	0.296	3.331	0.068
		2 0.144	0.062	4.147	0.126
		3 -0.099	-0.174	4.546	0.208
		4 -0.158	-0.107	5.595	0.232
		5 -0.054	0.061	5.723	0.334
		6 -0.030	-0.011	5.762	0.450
		7 0.040	0.015	5.837	0.559
		8 0.219	0.216	8.127	0.421
		9 -0.129	-0.303	8.953	0.442
		10 -0.024	0.045	8.983	0.534
		11 -0.271	-0.199	12.941	0.297
		12 -0.349	-0.282	19.815	0.071
		13 -0.074	0.174	20.138	0.092
		14 -0.149	-0.199	21.503	0.089
		15 0.053	0.000	21.688	0.116
		16 0.094	0.072	22.289	0.134

图3 序列自相关检验结果

表2 序列的AR(1)模型时间序列分析结果

变量		数值	变量		数值
C	系数	-34.783 520	AR(1)	系数	1.003 250
	标准误差	71.505 850		标准误差	0.005 103
	t 统计量	-0.486 443		t 统计量	196.615 800
	P	0.629 9		P	0.000 0
R^2		0.999 147	因变量均值		10.881 170
调整 R^2		0.000 121	因变量标准差		1.579 645
回归标准误		0.046 827	Akaike信息准则		-3.229 273
残差平方和		0.072 361	Schwarz准则		-3.140 396
对数似然		58.512 270	Hannan-Quinn准则		-3.198 592
F 统计量		38 657.77	Durbin-Watson统计量		1.341 549
AR特征根		1.00			

$$\ln n_{\text{foreign}, t} = 0.113\ 044\ 283\ 597 + 1.003\ 249\ 938\ 01 \cdot n_{\text{foreign}, t-1} \quad (1)$$

式中： t 表示年份。经转化，得出 n_{foreign} 的估计值模型，如式(2)所示。

$$E(n_{\text{foreign}, t}) = \exp(0.113\ 044\ 283\ 597) \cdot n_{\text{foreign}, t}^{1.003\ 249\ 938\ 01} \quad (2)$$

据此，设计中国科技论文外流态势监测指数，如式(3)所示。

$$i_t = \frac{A(n_{\text{foreign}, t})}{E(n_{\text{foreign}, t})} \quad (3)$$

式中： $E(n_{\text{foreign}, t})$ 表示中国作者在非中国SCI期刊上发表的论文数量在 t 年的估计值，由此可以预测出2023年中国作者在非中国SCI期刊上发表的论文数量为817 464篇； $A(n_{\text{foreign}, t})$ 表示中国作者在非中国SCI期刊上发表的论文数量在 t 年的真实值。

当 $i_t=1$ 时，中国作者在非中国SCI期刊上发表的论文数量的真实值等于估计值，说明中国作者在非中国

SCI期刊上发表的论文数量趋于稳定；当 $i_t>1$ 时，中国作者在非中国SCI期刊上发表的论文数量的真实值大于估计值，说明中国作者论文表现出外流的趋势；当 $i_t<1$ 时，中国作者在非中国SCI期刊上发表的论文数量的真实值小于估计值，说明中国作者论文表现出回流的趋势。

4 结果与分析

根据指数测算公式，得出1988—2022年中国科技论文外流态势监测指数(见表3)。从表3数据来看，1988年和1989年监测指数大于1，但没有表现出持续性特征，可以认为这一时期尚未出现外流趋势；1990—2003年监测指数持续小于1，这一时期中国作者较少在非中国SCI期刊上发表论文；2004—2021年监测指数持续大于1，这一时期论文表现出外流趋势；2022年监测指数小于1，论文有回流迹象，但还需持续观测。

表3 1988—2022年中国科技论文外流态势监测指数

年份	监测指数	状态	年份	监测指数	状态	年份	监测指数	状态
1988	1.010 520		2000	0.991 356		2012	1.006 703	外流
1989	1.003 243		2001	0.990 861		2013	1.009 700	外流
1990	0.998 985		2002	0.988 165		2014	1.010 324	外流
1991	0.992 182		2003	0.994 296		2015	1.008 029	外流
1992	0.990 074		2004	1.001 811	外流	2016	1.004 118	外流
1993	0.983 106		2005	1.008 509	外流	2017	1.001 100	外流
1994	0.975 705		2006	1.012 901	外流	2018	1.000 411	外流
1995	0.973 892		2007	1.008 611	外流	2019	1.004 760	外流
1996	0.979 373		2008	1.007 484	外流	2020	1.001 329	外流
1997	0.980 114		2009	1.008 435	外流	2021	1.000 962	外流
1998	0.987 600		2010	1.003 634	外流	2022	0.999 902	回流
1999	0.991 578		2011	1.006 058	外流			

从中国科技论文外流态势监测指数趋势图(见图4)来看，指数自1988年持续下降，在1995年出现极小值0.973 892，此后快速上升，保持指数增长，至2006年达到峰值1.012 901，此后趋于稳定，并在2022年下降至1以下。这一走势与现实情况相互映衬。20世纪80年代末，南京大学首先将SCI引入科研评价体系，开启了中国科研评价的定量时代^[15]。在当时的现实情境下，SCI因其相对客观的评价标准及其相对超脱的国际身份，迅速被国内科学界视为可尽量避免人为因素影响的科研评价标准，并逐步树立了一定程度上的权威地位。

2002年7月，陈家镛院士指出，我国不恰当地强调在具有高影响因子的SCI期刊上发表论文，造成我国大量高水平论文在其他国家期刊上发表。此后，中国科技论文外流现象引起中国学术界、期刊界和政府部门的强烈关注，相关政府部门和各大科研机构都为扭转我国高水平科技论文外流趋势采取行动。一方面，转变评价导向，如提出“破四唯”“破五唯”；另一方面，加快建设世界一流科技期刊，如实施“中国科技期刊卓越行动计划”。相关举措的实施逐步缓解了我国科技论文外流问题。

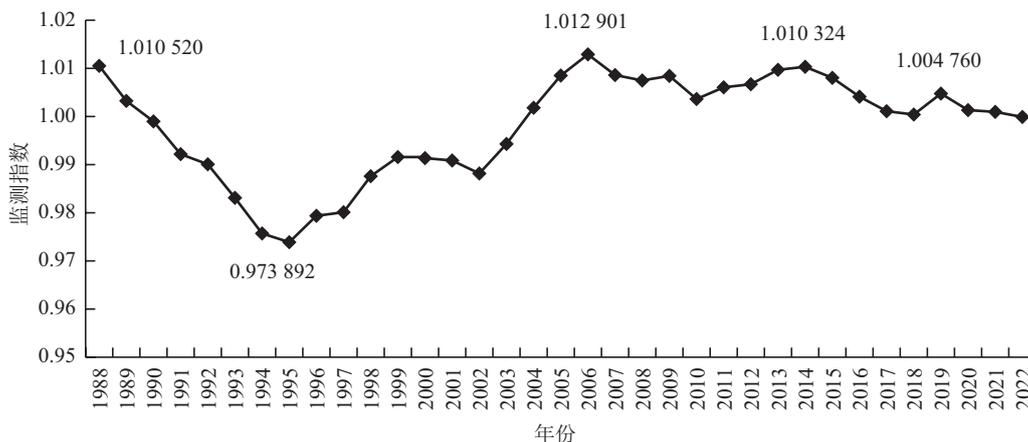


图4 1988—2022年中国科技论文外流态势监测指数趋势

5 结论与建议

通过采集中国1987—2022年发表在非中国SCI期刊上的论文数据,借助ARMA模型建立了中国科技论文外流预测模型,以此设计了中国科技论文外流态势监测指数并进行了应用分析。结果表明,中国科技论文并非一直处于外流的状态,中国科技论文外流态势监测指数在1988—1995年快速下降,1996—2006年快速上升,在2006年以后逐渐趋于稳定。从外流情况来看,中国科技论文在1990—2003年未表现出外流趋势,2004—2021年出现持续外流的特征,并在2022年表现出回流的特征。从现实情况来看,全球每年发表的科技论文总数和主要发达国家每年发表的科技论文总数始终保持一定的增长速度。作为全球科技发展的重要组成部分,中国的科技论文数量也会保持长期增长趋势,但是这并不意味着一定产生论文外流现象。因此,不能仅依据论文绝对数量和增速来判断论文是否处于外流状态,要客观地对未来发展趋势进行测算,进行持续动态监测,正确评判我国国际科技论文数量增长的合理性。

要进一步扭转我国科技论文外流趋势,需要从以下几个方面重点发力。

(1) 加快进行科研评价体系改革。目前国家层面已经明确“破五唯”,但破后立新标是关键,如何推动人才评价体系重建和避免“以刊评文”是目前讨论的焦点。下一步,相关管理部门、科研人员等应该推动同行评议制度建设,加强科研诚信监管,深入探讨如何对科研成果重要性进行判别。

(2) 在关键领域鼓励创办优秀英文科技期刊。

随着我国国际开放的程度不断加深,深度参与全球科技交流不可避免,在关键领域创办更多优秀英文科技期刊势在必行。尽管我国相关部门已经做出许多努力,尤其是“中国科技期刊卓越行动计划”实施以来取得了显著成效,但是我国期刊仍然无法满足我国科研人员的发文需求。从现实情况来看,全球SCI论文总数始终处于增长状态,我国SCI论文数量在未来也会保持这一趋势。从产业规模和发展速度来看,据《乌利希国际期刊指南》统计结果,近10年世界科技期刊总量增幅为20.85%,美国科技期刊总量增幅为25.60%,而中国科技期刊增幅只有7.05%。从科技期刊数量来看,美国有20 573种科技期刊(占世界总数的20.71%),远多于我国科技期刊数量^[5]。因此,我国迫切需要创办优秀英文科技期刊。同时,积极引导未获得CN号的我国优秀英文科技期刊回归中国。在我国战略需求领域创办优秀英文科技期刊,也能在一定程度上防止版权外流。

(3) 扩大国际合作范围,加深国际合作。完善科技论文和科学技术信息交流传播机制是促进科学技术和传播的重要手段。通过更加深入的国际交流与合作可以快速提升我国的科学研究水平,取人之长补己之短。在国家政策的大力支持下,高等院校、科研院所、医院、学/协会、企业等各类创新主体已经成为全球科研生产团队的重要组成部分,要更加积极参与国际学术交流,推动世界科学技术进步。我国科研人员应该积极发起和参与国际大科学计划和大科学工程,提升我国学术话语权。

参考文献

- [1] 中国科学技术信息研究所. 中国科技论文统计报告[M]. 北京: 中国科学技术信息研究所, 2022.
- [2] 陈家辅. 应鼓励科学家在国内期刊发表高水平论文[J]. 分子植物育种, 2003, 1(4): 583-584.
- [3] 刘彩娥. 把论文写在祖国大地上: 国内科研论文外流现象分析[J]. 北京工业大学学报(社会科学版), 2018, 18(2): 64-72.
- [4] 中国科技论文统计与分析课题组. 2019年中国科技论文统计与分析简报[J]. 中国科技期刊研究, 2021, 32(1): 99-109.
- [5] 中国科学技术协会. 中国科技期刊发展蓝皮书(2020)[M]. 北京: 科学出版社, 2021.
- [6] 王芙蓉, 刘明寿. SCI指挥棒对中文科技期刊影响的实证研究: 基于全国16所省属农业高校的数据分析[J]. 编辑学报, 2022, 34(6): 618-623.
- [7] 曾建勋, 杨代庆. 关于扭转我国科技论文外流局面的政策性思考[J]. 编辑学报, 2020, 32(6): 600-604.
- [8] 刘敬义, 王紫霞, 向政, 等. 科技论文稿件外流及其应对策略[J]. 编辑学报, 2008, 20(1): 47-49.
- [9] 宋鸿. 高水平学术论文外流问题应当引起重视[J]. 中国科技期刊研究, 2005, 16(5): 736-739.
- [10] 刘丽英, 魏秀菊, 朱明, 等. 我国科技论文外流的文献经济损失构成及原因分析[J]. 编辑学报, 2015, 27(5): 426-428.
- [11] 刘丽英, 魏秀菊, 王柳, 等. 2003—2012年中国优秀科技论文外流状况定量分析: 以农业工程领域论文为例[J]. 中国科技期刊研究, 2014, 25(4): 478-484.
- [12] 刘丽英, 魏秀菊, 王柳, 等. 科技论文外流文献经济损失估算模型建立及其应用(转载)[J]. 农业工程学报, 2015, 31(18): 311-314.
- [13] 王密平. 我国高水平论文外流状况分析: 基于医学学科[J]. 图书与情报, 2015(1): 49-52.
- [14] 沈菲飞. 政策导向对高校稿件外流的影响: 基于15所高校的制度文本分析[J]. 中国科技期刊研究, 2018, 29(12): 1192-1200.
- [15] 陈晨. 30年科学评价: SCI功与过[EB/OL]. [2023-10-01]. http://www.bic.cas.cn/tpxwgikjxw/200905/t20090508_53221.html.

作者简介

陈国娇, 女, 博士, 助理研究员, 研究方向: 科技期刊评价, E-mail: chengj@istic.ac.cn。
田瑞强, 男, 博士, 副研究员, 研究方向: 科技期刊评价。

Design and Application of the Monitoring Index for the Outflow of Chinese Scientific Papers

CHEN GuoJiao¹ TIAN RuiQiang^{1,2}

(1. Institute of Scientific and Technical Information of China, Beijing 100038, P. R. China;

2. Key Laboratory of Knowledge Mining and Service for Medical Journals, Beijing 100038, P. R. China)

Abstract: Scientific research achievements are important strategic resources for a country to strive for technological self-reliance. Scientific papers serve as an important form of carrying out these research achievements. The outflow of scientific papers has brought many negative impacts on China's scientific and technological development and national security. In order to comprehensively understand the outflow of scientific papers in China, this study collects data on papers published in non-Chinese SCI scientific journals from 1987 to 2022 in China. The ARMA model is applied to establish a predictive model for the outflow of Chinese scientific papers and we design a monitoring index for the outflow trend for application analysis. It is found that Chinese scientific papers have not always been in a state of outflow. The monitoring index for the outflow trend of Chinese scientific papers declines from 1988 to 1995, increases rapidly from 1996 to 2006, and almost remains the same after 2006. From the perspective of outflow, Chinese scientific papers do not show outflow characteristics from 1990 to 2003, but show sustained outflow characteristics from 2004 to 2021, and show reflux characteristics in 2022. To further reverse the outflow trend of Chinese scientific papers, it is necessary to focus on three aspects, including accelerating the reform of the research evaluation system, encouraging the establishment of excellent English scientific journals in key areas, and expanding the scope of international cooperation and promoting in-depth international cooperation.

Keywords: Outflow of Scientific Papers; Scientific Journal; SCI Paper; Time-Series Analysis

(责任编辑: 王玮)