

党的十七届五中全会提出坚持把科技进步和创新作为加快转变经济发展方式的重要支撑,并对全面推进国家创新体系、加快建设创新型国家作出重要部署。区域创新体系是国家创新体系的重要组成部分,加强区域创新体系建设对于加快区域科技进步和创新的步伐,推动区域经济和社会的全面、协调、可持续发展具有重要意义。继2010年第6期刊发的一组相关文章后,本期将从区域协调可持续发展、区域科技资源集聚型配置、区域竞争力的测算与评价等方面探讨我国科技创新和区域发展。

——编者

## 区域协调可持续发展潜力的主成分聚类分析

彭玉婷 朱学芳

(南京大学信息管理系,多媒体信息处理研究所,江苏南京 210093)

**摘要:** 根据协调可持续发展相关理论,设计了评价地区协调可持续发展潜力的指标体系,搜集数据并用软件SPSS的主成分聚类方法分析我国31个省区的协调可持续发展潜力现状,对其进行排序和分类的处理研究。将全国省区划分为6个不同类别区域,分析各个区域协调可持续发展潜力的主要特点,提出相关建议,并为我国区域协调可持续发展水平的分层与整体推进提供参考。

**关键词:** 协调可持续发展;指标体系构建;主成分聚类分析

中图分类号: G350

文献标识码: A

DOI: 10.3772/j.issn.1674-1544.2011.01.002

### Principal Component Analysis and Cluster Analysis on Coordinative and Sustainable Regional Development

Peng Yuting, Zhu Xuefang

(Institute of Multimedia Information Processing, Dept. of Information Management, Nanjing University, Nanjing 210093)

**Abstract:** This paper constructs a measurement system of the coordinative and sustainable regional development according to the related theories, analyzes the current status of the coordinative and sustainable development of 31 different regions in China by means of collecting data, ranking and classifying them with SPSS, using Principal Component Analysis (PCA) and clustering analysis. And finally the entire region is classified into six different categories. Besides, the paper studies the main characteristics of coordinative and sustainable development of each category, proposes some positive suggestions, and provides proper references for measuring and promoting the level of coordinative and sustainable development.

**Keywords:** coordinative and sustainable development, construction of measurement system, PCA and clustering analysis

第一作者简介: 彭玉婷(1985-),女,南京大学信息管理系多媒体信息处理研究所硕士研究生,研究方向:信息资源管理。

收稿日期: 2010年3月17日。

## 1 引言

在推进协调可持续发展的进程中，主要面临两个层次的障碍，一是推进各个区域（东中西部）之间的协调可持续发展，一是推进各个区域内（特别是城乡之间）的协调可持续发展。由于一系列历史和现实的原因，前者的矛盾在过去以及当前时期较为突出，后者的矛盾也在逐步显现。许多学者为跨越这两个障碍做了大量的研究。从文献检索结果来看，当前对区域协调可持续发展的定性研究和特定地区的研究居多<sup>[1-2]</sup>，而对我国整体进行协调可持续发展的评价略显不足，如张琳等采用层次分析法对各个区域进行分类，或直接对区域进行聚类<sup>[3]</sup>。虽然这两种分析方法虽然可以对我国各区域的协调可持续发展进行划分，却不能准确地描述造成各个区域间差异的最重要的因素。针对这些考虑，本文建立指标体系，采用主成分分析法对全国31个省区的协调可持续发展潜力进行评估，并用聚类分析进行整体划分。此方法提出各个类别间差异产生的主要原因，并据此对各个省区的协调可持续发展能力进行综合评价，旨在为各个省区协调可持续发展的进一步推进提供重要的参考依据。

## 2 协调可持续发展潜力的评价指标

从广义上讲，区域经济社会的协调可持续发展，是经济、社会和生态的协调可持续发展。这3个方面必然要涵盖人口、资源、环境、经济、科技等方面的因素，因为人口是协调可持续发展的主体，资源是协调可持续发展的物质基础，环境是协调可持续发展的自然生态支撑，经济是协调可持续发展的核心<sup>[4]</sup>，科技是协调可持续发展的驱动力。考虑到以上因素以及数据获取的可行性，选取2007年的截面数据进行分析，构建表1所示的评价指标体系对一个区域的协调可持续发展潜力进行测评<sup>[5-6]</sup>。

本文设计的指标体系共涉及16个基本指标，能够较为全面地衡量一个区域的协调可持续发展潜力。但在实际分析时，由于指标个数较多，且指标间在某种程度上存在信息的重叠，不可避免

地会有一些的相关性。如果整体纳入聚类分析，那么势必会增加分析问题的复杂性。而主成分分析可以把各指标之间相互关联的复杂关系进行简化分析，将原来指标重新组合成一组新的互相无关的几个综合指标来代替原来的指标，同时根据实际需要从中选取几个较少的综合指标，尽可能多地反映原来指标的信息<sup>[7]</sup>。

## 3 协调可持续发展主成分聚类分析

目前，较为普遍的主成分分析法，通常以各个主成分的方差贡献率为权数，对各样本的不同指标进行加权求和，构造主成分的“综合得分”来对样本进行排序。但是这种综合主成分得分所包含的信息量比第一主成分（简称主1，下同）的信息量还要少，故计算“综合得分”且利用“综合得分”的排序进行样本的聚类分析是不合理的<sup>[8-9]</sup>。

使用主成分聚类法可以很好地解决以上的问题。对于多指标系统评估中的排序问题，如果主1的方差贡献率不够大，即其表达的原始数据信息不够充分，仅按照主1得分对样本进行排序评价就不可避免地会有片面性。这时候可以将主成分分析与聚类分析两种统计方法进行结合，即“主成分聚类分析法”。主成分聚类分析法应用于需要分级而不要求具体排名的场合。

本文使用SPSS 13.0对数据进行处理。限于篇幅这里不再详细列出涉及较多的原始数据。以下只列出所需的主要步骤：

(1)将样本数据进行标准化处理，得到16个指标下31个省区样本数据的标准化矩阵。

(2)对原始样本数据进行主成分分析，选定特征值大于1的5个主成分（表2），其累积方差贡献率为80.360%，可以解释原始指标大部分信息。

(3)计算因子载荷矩阵。主1对单位地区生产总值能耗、教育科技支出、万人专利授权量、城乡协调度、耕地资源供给、工业污水排放达标率、工业固体废物处理率以及环境经济协调度有较大的载荷系数，这些指标反映了区域经济与生态的协调性以及区域整体发展的动力因素，故将其定义为区域生态协调与推动潜力因子。主2对

表1 协调可持续发展潜力的评价指标体系

目标	基本层面	基本指标	备注
区域内协调可持续发展潜力	经济发展潜力	地区生产总值年增长率 $X_{11}$	2007与2006年地区生产总值比
		固定资产投资年增长率 $X_{12}$	2007与2006年固定资产投资额比
		居民消费年增长率 $X_{13}$	
		单位地区生产总值能耗率 $X_{14}$	对单位地区生产总值能耗取倒数
	社会发展潜力	人口自然增长率倒数 $X_{21}$	对人口自然增长率取倒数
		教育科技支出比重 $X_{22}$	(教育支出+科学技术支出)/一般公共预算支出
		文化体育支出比重 $X_{23}$	文化体育与传媒支出/一般公共预算支出
		社保和卫生支出比重 $X_{24}$	(社保就业+医疗卫生)/一般公共预算支出
		万人专利授权量 $X_{25}$	专利授权量/年末人口总数(单位:项/万人)
		城乡协调度 $X_{26}$	农村居民人均纯收入/城镇居民人均可支配收入
	生态发展潜力	社会经济协调度 $X_{27}$	地区生产总值年增长率/人口自然增长率
		耕地资源供给 $X_{31}$	人均耕地面积
		水资源供给 $X_{32}$	人均水资源量(立方米/人)
		工业污水排放达标率 $X_{33}$	工业废水排放达标量/工业废水排放量
		工业固体废物处理率 $X_{34}$	工业固体废物综合利用量/(工业固体废物生产量+往年存量)
环境经济协调度 $X_{35}$	环保支出/一般公共预算支出		

注:本表中数据来源于中经网统计数据库(<http://db.cei.gov.cn/>)以及国家统计局2008年统计年鉴。

表2 提取的主成分特征值及方差贡献率

主成分	特征值	方差贡献率 (%)	累积方差贡献率 (%)
1	5.159	32.241	32.241
2	3.125	19.530	51.771
3	1.816	11.350	63.121
4	1.494	9.337	72.459
5	1.264	7.901	80.360

固定资产投资年增长率、人口自然增长率、社保和卫生支出比重以及社会经济协调度有较大的载荷系数,这些指标是社会与经济协调发展的反映,定义为社会经济协调因子。主3对水资源的载荷系数相对较大,定义为能源因子。主4对地区生产总值年增长率和居民消费年增长率的载荷系数相对较大,故将其定义为经济增长因子。主5对文化体育支出比重的载荷系数较大,故将其定义为人文因子。总体来说,5个主成分所表达的含义较为全面地反映了经济、社会和生态方面的内容,因此,本次主成分分析较为合理。

(4)将特征向量矩阵与标准化的样本数据进行矩阵相乘得出5个主成分的得分,再按照主1的得分进行排序(表3)。由于主1的方差贡献率(32.241%)远未达到85%,因此,仅由此得出的排名结果不够准确,还要借助聚类分析进行补充,使排名结果相对来说更为合理、可靠。

(5)按照5个主成分的得分用SPSS对样本系统聚类,得到聚类谱系图(略),自依距离尺度划分为六类较为合理。分别计算各类别样本在主1上的得分均值,确定各类顺序。然后,将各类内部的样本按照主1得分排序,最终可得出科学合理的区域分类和排名。结果见表4。

#### 4 可持续发展潜力区域划分的分析

本文根据主成分聚类分析的结果,将我国31个省区划分为六大类。

第一类,位于我国东部沿海,经济效率居全国前列,社会发展潜力巨大。尤其是上海和北京,拥有最高的政策和智力支持,但这些省区人口密度非常大,生态与经济的协调程度相对较差,如北京与上海的地区生产总值年增长率与人口自然增长率的比值在全国31个省区中处于中等偏下的水平;人均耕地面积、水资源很少,从资源层面上限制了这些地区经济社会的可持续发展。这些地区应该注重经济、社会的发展与生态相协调。

第二类,黑吉辽工业基础雄厚,自然资源丰富,可持续发展的能力很大。此外,东北三省还拥有很高的政策支持。2003年10月,国务院出台了《关于实施东北地区等老工业基地振兴战略的若干意见》。2007年,东北三省的地区生产总

表 3 各省区按照主 1 得分排序结果

排名	省 区	主1得分	排名	省 区	主1得分
1	上 海	4.279702	17	河 北	-0.31787
2	天 津	3.662509	18	重 庆	-0.34536
3	江 苏	3.622406	19	四 川	-0.35674
4	浙 江	3.462157	20	吉 林	-0.38701
5	北 京	3.429947	21	黑 龙 江	-0.52425
6	广 东	2.599836	22	陕 西	-1.23808
7	山 东	1.93479	23	山 西	-1.63169
8	福 建	1.745515	24	云 南	-2.05962
9	辽 宁	0.937614	25	贵 州	-2.07133
10	海 南	0.670287	26	新 疆	-2.30034
11	湖 北	0.513844	27	宁 夏	-2.3619
12	河 南	0.097973	28	内 蒙 古	-2.36207
13	江 西	0.071443	29	甘 肃	-2.98911
14	湖 南	0.018279	30	西 藏	-3.32339
15	广 西	0.009897	31	青 海	-4.58401
16	安 徽	-0.20342			

表 4 区域分类排序表

类别	省 区	主1得分均值
1	上海 天津 江苏 浙江 北京	3.691
2	辽宁 吉林 黑龙江	0.009
3	广东 山东 福建 海南 湖北 河南 湖南 广西 安徽 河北 重庆 四川 陕西	0.395
	江西 山西 云南 贵州 甘肃	-1.736
	内蒙古	-2.362
5	新疆 宁夏 青海	-3.082
6	西藏	-3.323

注：第三类包含省区较多，故又细分为两个子类。

值年增长率和居民消费增长率都处于较高水平，经济得到了迅速发展，而且后劲很足，但是单位地区生产总值能耗很高，人口自然增长率较高，工业固体废弃物处理率较低，这与东北三省以重工业为主的产业结构有很大关系。在某种程度上，这些地区的经济增长是以资源消耗和生态恶化为代价的。所以，黑吉辽三省需注意经济发展与生态发展相协调，控制人口的过量增长。

第三类，绝大多数位于我国中东部。广东、山东、福建、海南、河北、河南、安徽等东部省区对教育科技的投入力度较大，科技推动潜力较大，但是这些地区的人口自然增长率较高，城乡居民收入差距较大，人均资源供给量较低，这些因素可能会成为东部省区协调可持续发展的瓶

颈。这些省区应在大力发展经济的同时，注重社会的协调发展，注重人口质量；在加大科技投入的基础上，注重科技创新，抓住机遇，发挥天时、地利、人和的优势，全面协调可持续发展。

数据表明，湖北、湖南、广西、重庆、四川、陕西地区经济发展潜力很大，人口控制取得显著成效，而且污染治理情况良好。但是单位地区生产总值能耗很高，即其经济的增长在某种程度上是以高能耗为代价的，这与区域经济和生态的协调可持续发展原则是相悖的。此外，这些地区的教育科技、文化体育的投入力度较小，区域科技创新力不足，将成为以后经济持续快速发展的主要限制因素。再者，人均耕地面积以及水资源占有量都低于全国平均水平，尤其是水资源严重匮乏，这不但是资源供给问题，更是生态问题。如何保护生态，保持生态与经济的协调可持续发展将是这些省区急需解决的问题。

江西、山西、云南、贵州、甘肃等 5 省区位于我国中西部，属于经济欠发达地区。从数据中可以看出，这些地区的生产总值增长率、固定资产投资增长率偏低，尤其是居民消费增长率要低于其他大部分地区，而且单位地区生产总值能耗很高，这说明这些地区的经济发展潜力欠缺。从社会发展潜力来看，这些地区在文化体育、社保卫生的投入较多，但由于人口自然增长率较高，



致使经济与社会的协调度相对较低。此外,这些地区的科技创新力明显不足,水资源较为缺乏,固体废弃物的处理能力有待提高。这些地区要想达到协调可持续发展的目标,首先要控制人口数量,提高人口质量,注重科技自主创新,拉动区域内居民消费,增大固定资产投资力度,进而推动整个区域的经济发展,但同时也要协调生态,同步发展。

第四类,从数据可以看出,内蒙古地区的经济发展潜力很大,其地区生产总值年增长率、固定资产投资年增长率和居民消费年增长率都处于很高的水平,单位地区生产总值的能耗很高,经济的增长是以能源的过度消耗为代价的,即倾向于粗放型经济增长模式。此外,内蒙古的社会发展潜力处于我国中等水平,科技创新能力不足,水资源匮乏,污染物处理能力有待进一步提高。近年来,内蒙古地区非常注重生态的保护,而且环保投入力度很大。2010年6月,内蒙古自治区环保局提出要采取8项措施推进松花江水污染防治工作<sup>[10]</sup>,这表明内蒙古生态可持续发展的势头良好。

第五类,从数据看,新疆、宁夏、青海三省区的地区生产总值年增长率、固定资产投资年增长率较低,单位地区生产总值能耗远远高于我国其他省区,这说明新疆、宁夏和青海的经济效率低下,属于粗放型经济增长模式,经济发展潜力不足。此外,由于人口自然增长率很高,导致社会与经济发展不太协调,城乡居民收入差距较大,科技创新力不足,这些都降低了该地区的社会发展潜力。近年来,三地都比较注重生态保护,环保的投入力度很大,但是三地的污染物处理能力有待进一步提高。由数据可知,宁夏的人均水资源占有量仅为全国平均水平的1/40,只有采取相关措施解决水资源供给问题,才能保证区域整体的协调可持续发展。

第六类,西藏地处我国西南边疆,全区土地面积约占全国总面积的12.8%,人口仅占全国总人口的2.149%<sup>[11]</sup>,是我国人口密度最小的地区。在今后的发展中,西藏地区需要加强基础设施建设和生态建设,在大力推动经济社会全面发展的同时提高对生态环境的关注度,提高经济、社会

和生态协调可持续发展的能力。

## 5 结论

本文构建了评价区域协调可持续发展潜力的指标体系,应用主成分聚类分析的方法,通过SPSS软件对我国31个省区的协调可持续发展潜力进行了评估,将其划分为六类,并从经济、社会和生态协调可持续发展潜力3个角度对各类区域进行了分析评价,提出了建议。从整体上看,我国各区域的协调可持续发展潜力呈现从东部到西部的梯度递减趋势,由此可见,经济发展水平在区域协调可持续发展中占有较大比重。希望本文提出的针对性建议可以对各个区域的全面协调可持续发展提供参考。

## 参考文献

- [1] Zuo Guiming. Philosophical Thinking about Total Coordinating Sustainable Development Strategy[J]. Journal of Chuxiong Normal University, 2008, 23(8): 95-99. (in Chinese)  
〔左贵明. 全面协调可持续发展战略的哲学思考[J]. 楚雄师范学院学报, 2008, 23(8): 95-99.〕
- [2] Cai Jie Pursue the Scientific Outlook on Development and Comprehensively Push forward Coordinative and Sustainable Regional Economic Development[J]. Study and Practice, 2007(11): 161-165. (in Chinese)  
〔蔡杰. 坚持以科学发展观为统领, 全面推进区域经济协调可持续发展[J]. 学习与实践, 2007(11): 161-165.〕
- [3] Zhang Lin, Wang Jian. A Comprehensive Assessment on Coordinative and Sustainable Regional Development of China[J]. Estate and Science Tribune, 2008, 7(6): 129-130. (in Chinese)  
〔张琳, 王健. 对我国区域协调可持续发展水平的综合评价[J]. 产业与科技论坛, 2008, 7(6): 129-130.〕
- [4] Zhang Yunfang, Shen Yamin. Factor and Cluster Analysis of the Sustainable Development of Shanxi Regional Society and Economy[J]. Tangdu Journal, 2008, 24(2): 35-38. (in Chinese)  
〔张云芳, 申亚民. 陕西区域社会经济协调可持续发展的因子与聚类分析[J]. 唐都学刊, 2008, 24(2): 35-38.〕
- [5] Hu Zhongjun, Jiang Xiangcheng, Liu Lei. The Etab-

(下转第17页)