

地理国情分类区划模型构建及实证研究 ——以河南省为例

谢明霞¹,王家耀²,陈科²

(1. 武汉大学资源与环境科学学院, 武汉 430079;

2. 中国人民解放军信息工程大学地理空间信息学院, 郑州 450052)

摘要:地理国情分类区划是认知和分析地理国情的有效途径,是研究认识地理国情类型特征、组合及其演变趋势的区域差异和地理国情空间分异的基础。本文通过对地理国情普查的基本统计数据进行分析,制定了地理国情分类区划的指标体系;基于多维尺度分析和聚类分析,构建了地理国情分类区划模型。选择河南省地理国情普查数据,对其进行分类区划实验。通过对地理国情数据进行多维尺度分析,根据分析的可视化结果,可直观地确定地理国情的分类区划数,并将此设置为初始的聚类数;在此基础上进行 k 均值聚类,可实现河南省地理国情的分类区划,通过对各聚类簇中指标数据特征进行分析,并根据各聚类簇中心对区划各类型进行定义和解释,最后对聚类结果进行空间化表达。结果表明:河南省地理国情可划分为指标数据极值区、指标数据低值区、指标数据中值区和指标数据高值区4类,且根据空间集聚性分析,可判断分类区划的空间集聚效应显著,各区划类型间相互包容和渗透,集聚呈4条西南↔东北的条状带。该项工作可为统筹地理国情与区域可持续发展,制定切实可行的区域发展政策提供科学依据。

关键词:地理国情;分类区划;多维尺度分析;聚类分析;河南省

1 引言

区划就是区域的划分,它既是一种数据分析的手段,也是一种分析结果的表现形式,区划必须适应并能区域的可持续发展服务。地理国情主要是指地表自然和人文地理要素的空间分布、特征及其相互关系(陈俊勇, 2012; 徐德明, 2012)。狭义来看,地理国情是指与地理空间紧密相连的自然环境、自然资源基本情况和特点的总和;广义来看,地理国情是指通过地理空间属性将包括自然环境和自然资源、经济发展状况、政治状况、社会状况、文化传统、科技教育状况、国际环境和国际关系等在内的各类国情进行关联与分析,从而得出能深入揭示经济社会发展的时空演变和内在关系的综合国情(陈俊勇, 2014)(图1)。从世界范围来看,许多国

家和组织都开展了地理国情相关的研究项目或工程,以便更好地服务于本国或地区的资源、环境、能源、社会等领域(Ceccato et al, 2005; Young et al, 2006; Mithal et al, 2011)。地理国情监测与分析已被确定为未来20年中国测绘工作的重点之一(陈俊勇, 2012)。

统计分析是地理国情普查与监测数据向地理国情信息转化的必要手段(张继贤, 2013)。其中,分类、分级是统计分析最为典型和常用的方法。目前,中国地理国情普查工作已结束,获取的数据内容丰富、种类繁多,也非常复杂,但在地理国情综合分析研究方面尚处于探索阶段(武琛, 2012; 李红等, 2014; 刘耀林等, 2014; 王永峰等, 2014)。科学认识地理国情、促进资源的合理利用开发、提高区域社会经济发展和事业水平、促进区域的可持续

收稿日期:2016-05;修订日期:2016-09。

基金项目:中国工程院重大咨询项目(2014-XZ-17);国家自然科学基金项目(41401462) [Foundation: Major Consulting Project of Chinese Academy of Engineering, No.2014-XZ-17; National Natural Science Foundation of China, No.41401462]。

作者简介:谢明霞(1985-),女,湖北武汉人,博士研究生,主要从事地理国情统计分析研究, E-mail: xmx0424@whu.edu.cn。

引用格式:谢明霞,王家耀,陈科. 2016. 地理国情分类区划模型构建及实证研究:以河南省为例[J]. 地理科学进展, 35(11): 1360-1368. [Xie M X, Wang J Y, Chen K. 2016. The model construction and empirical research on classification-based regionalization of geographical national conditions: Take Henan Province as an example[J]. Progress in Geography, 35(11): 1360- 1368.]. DOI: 10.18306/dlkxjz.2016.11.006

续发展,都需要对地理国情进行分类和评价,这是地理国情统计分析的基础(王家耀等, 2016)。

地理国情作为人与自然的复杂空间综合体,它是一个集资源、生态、环境、经济、社会等众多要素于一体的复杂系统(王家耀等, 2016),其区划理论及方法与自然地理区划、经济区划等有着广泛而深刻的联系,它们为地理国情复杂系统区划研究奠定了方法论基础。地理国情区划是地理国情综合统计分析成果的一种表达形式,目的在于实现对抽象、高度综合的地理国情的客观分类描述和量化评价,并对其进行直观表达,从而能科学地揭示资源、生态、环境、经济、社会等要素在地理空间上相互作用、相互影响的内在关系,挖掘和发现社会、经济发展与自然资源、环境和生态等的内在联系和演变规律,实现地理国情普查与监测数据向地理国情信息的转化,为科学管理决策提供可靠依据。地理国情区划可分为分类区划和分级区划 2 种,是在对地理国情普查数据进行基本统计、汇总和分析的基础上,结合相关领域的专题数据,从资源、环境、生态、经济和社会 5 个方面,分别对地理国情进行分类分析和协调性评价分级而确立的分区单元(图 1)。

目前,典型的分类区划包括行政区划、农业区划(陈百明, 1986)、景观生态区划(肖笃宁等, 1998)、地貌区划(沈玉昌等, 1982; 程维明等, 2014)、气候区划(张新时, 1993; 毛飞等, 2011)、经济区划(杨树珍, 1983)、地理区划(郑度等, 2008)等。无论是行政区划、气候区划、地理区划、农业区划、地貌区划、经济区划,还是景观生态区划,都是以区域自然、经济、农业或生态等某一要素为标准来认识地域分异特征,在区划过程中所考虑的综合性,仅仅是自然、经

济、农业或生态某一方面的综合性,但相对于整个区域发展而言,这种综合是不完全的。随着可持续发展理念的普及,尤其是科学发展观的确立,人们认识到区域发展并不仅仅是经济、生态、或社会等某一方面的持续发展,而是由资源、环境、生态、经济、社会等众多要素构成的区域系统的整体演进。

地理国情分类区划是认知和分析地理国情的有效途径,是研究认识地理国情类型特征、组合及其演变趋势的区域差异和地理国情空间分异的基础。根据地理国情空间分异特征,对各区域地理国情要素发展进行调整,可为区域协调发展提供决策依据(谢明霞等, 2015)。本文主要对地理国情的类型区划进行介绍,参考已有的区划理论和方法,根据地理国情普查内容与指标,结合统计分析中分类方法,设计了地理国情分类区划模型,并以河南省为例进行了相关实验分析。

2 关键技术与模型构建

2.1 关键技术

(1) 聚类分析

聚类分析(Cluster Analysis, CS)是指将物理或抽象对象的集合,根据对象间的相似度(或相异度)分组为由类似的对象组成的多个类别的分析过程(图 2),其核心是针对收集的数据特征,选择和设计合适的技术方法,衡量数据间的相似度(或相异度),据此将数据源分类到不同的簇中。聚类源于众多领域,在不同的应用领域中,聚类技术得到了快速发展。现阶段,研究人员根据不同的应用提出了大量的聚类算法,大体可以归纳为以下 5 类:划分方法、层次方法、基于密度的方法、基于网格的方法和基于模型的方法(Han et al, 2007; 谢明霞, 2011)。

(2) 多维尺度分析

多维尺度分析(Multidimensional Scaling, MDS)

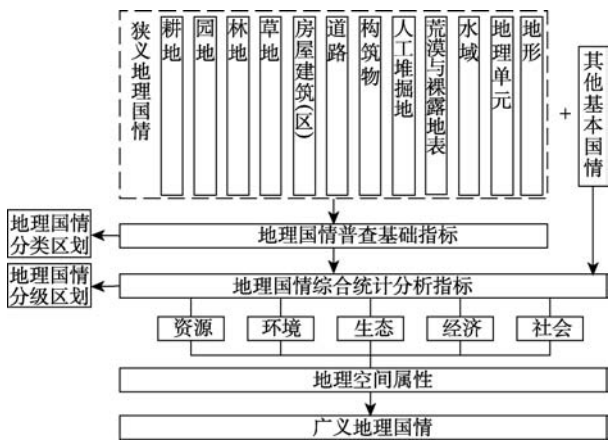


图 1 地理国情概念分析及区划类型

Fig.1 The concept of geographical national conditions and regionalization types

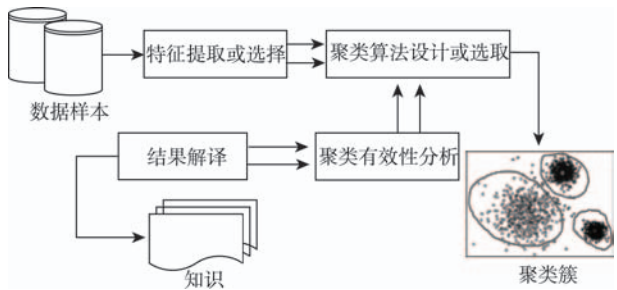


图 2 聚类分析基本过程

Fig.2 Steps of cluster analysis

(Cox et al, 2014)将一组个体间的相异数据经过MDS转换成空间构图,且保留原始数据的相对关系,用于反映多个个体间相似(不相似)程度,通过适当的降维方法,将这种相似(不相似)程度在低维空间中用点与点之间的距离表示,并帮助识别那些影响个体间相似性的潜在因素。因子分析是研究从变量群中提取共性因子的统计技术,其能够在许多变量中找出隐藏的具有代表性的因子,并将相同本质的变量归入一个因子,从而减少变量的数目,同时还可检验变量间关系的假设。多维尺度分析和因子分析、聚类分析之间既存在相同之处,又侧重点各异,相关关系如表1所示。其中,聚类分析可作为多维尺度分析的补充,以进一步获取经过MDS转换的相对关系保持的降维数据分组信息。

2.2 地理国情分类区划模型构建

不同区域的地理国情既有相似性又存在差异性,地理国情分类区划是根据地理国情的总体特征和统计研究目的,在系统深入研究地理国情类型及其组合特征、成因及其异同的基础上,根据相应的指标选取原则,选取地理国情分类指标,将国情总体划分为若干个各具特色的不同区域,以便进一步运用数据挖掘方法探索地理国情复杂系统的内在联系和规律。它是研究地理国情空间分异的基础,对地理单元的地理国情研究具有十分重要的作用。

本文研究地理国情分类区划目的是为了使人更加深刻地认识地理国情,理解地理国情,分析地理国情,以便准确掌握、科学分析资源和环境等的承载能力和发展潜力,从而满足经济和社会发展的需要,以及生态文明建设需要。主要研究内容包括:

(1) 地理国情分类坐标系的构建

在对地理国情指标集进行MDS转换的基础上,利用降维后的维度构建地理国情分类坐标系,为地理国情分类结果可视化提供支持,通过可视化使人们对地理国情分类结果有定性的理解,并有助于对地理国情普查与监测数据的挖掘和探索。

表1 因子分析、聚类分析和多维尺度分析的异同

Tab.1 Similarities and differences of factor analysis, cluster analysis, and multidimensional scaling

方法	分析对象	分析目的
因子分析	相似性	维度缩减,分析维度和因素的含义
聚类分析	相似性(或相异性)	数据分组
多维尺度分析	相异性	维度缩减,分析降维后数据点的位置

(2) 地理国情类型表达方式的设计

对地理国情类型进行表达是地理国情区划分析的基础。在地理国情类型表达方式设计过程中,可参考现有的分类区划方法中对类型的表达方式。本文根据聚类中心对各区划类型进行命名,即首先获取各类型区划中距离聚类中心最近的区域,然后通过对该区域的地理国情特征进行分析和总结,从而实现对地理国情区划类型的定义。

地理国情分类区划方法流程如图3所示。首先根据地理国情普查与监测统计成果和地理国情指标集,设计好地理国情分类原则后对地理国情指标进行MDS降维,根据降维结果建立地理国情分类坐标系,并确定聚类数;然后对MDS降维后数据进行 k 均值聚类,并对各聚类簇进行解释和分类区划分析;最后在地理国情分类的基础上,分析各类型的组合特征、成因等,生成地理国情区划专题图。

3 地理国情分类区划实验分析

3.1 研究区域与数据来源

本文研究区域为位于中国中部、黄河中下游的河南省,该省承东启西,古称天地之中,为中国第一

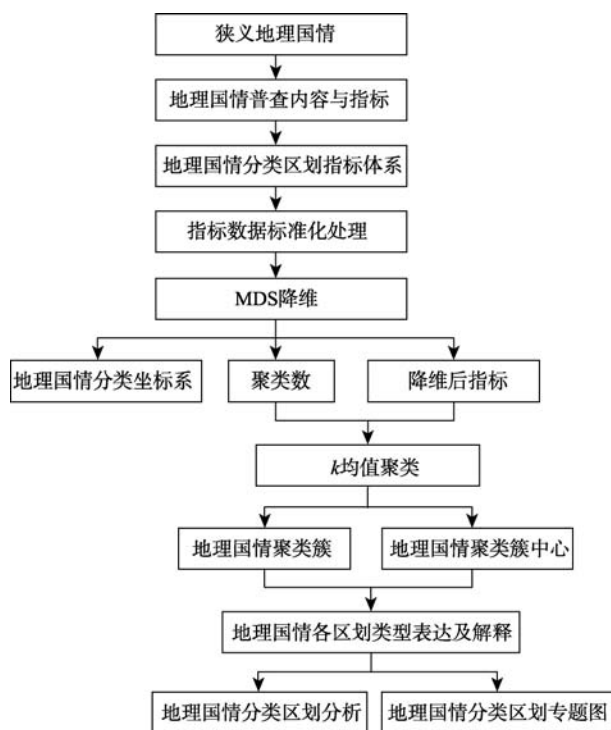


图3 地理国情分类区划模型

Fig.3 Classification-based regionalization model of geographical national conditions

人口大省、第一农业大省、新兴工业大省和劳动力输出大省。东接安徽、山东,北接河北、山西,西接陕西,南邻湖北,横跨海河、黄河、淮河、长江四大水系,土地面积 16.7 万 km²,其中山地丘陵面积 7.4 万 km²,占全省总面积的 44.3%;平原和盆地面积 9.3 万 km²,占总面积的 55.7%。研究单元为地市行政单元,共计 18 个。研究数据来源于 2015 年河南省地理国情普查数据。

3.2 指标选取及信度分析

地理国情分类区划的关键在于正确选择分类指标和划分各类界限,各指标应仅对地理国情作纯粹的描述,而不加任何主观的认识。根据国务院第一次全国地理国情普查领导小组办公室制定的《地理国情普查内容与指标》,地理国情普查内容分为 12 个一级类(图 1 中狭义地理国情所涉及内容)、58 个二级类、135 个三级类,本文以地市为研究单元,从面积、长度和个数等方面,针对地理国情普查内容的 12 个一级类选取相关的地理国情分类区划指标(图 4),由于获取的数据原因,暂未对普查内容中的地形要素进行考虑。根据获取的研究区域地理国情分类区划所需的相关数据,计算分类区划各指标值,利用 SPSS 软件,采用信度 α 系数法对河南省地理国情分类区划指标进行信度分析。通常 α 系数值在 0~1 之间。若 α 系数不超过 0.6,一般认为内部一致信度不足;达到 0.7~0.8 时,表示量表具有相当的信度;达 0.8~0.9 时,说明量表信度非常好。本文的计算结果为:未标准化指标项的 α 值为 0.479,指标的可信度较差,但是经标准化后的指标项的 α 值为 0.834,指标信度非常好。指标数据的标准化对指标的信度分析结果影响较大,说明标准化处理是客观描述和分析对象中必不可少的环节之一。本文选取的地理国情分类区划指标,经过数据标准化

处理后指标的可信度非常好,故能够据此客观的认知和分析地理国情类型。

3.3 河南省地理国情分类区划

根据地理国情分类区划指标体系以及构建的区划模型,首先对指标数据进行标准化处理,然后利用软件 SPSS 22,选择分析→度量→多维刻度,对实验数据进行多维尺度分析。分析结果显示,其 *Stress* 值(压力系数,也称作应力,多维尺度分析的信度估计值)为 0.16018, *RSQ* 值(相关系数平方,即拟合优度,多维尺度分析的效度估计值)为 0.89643,降维后各维度坐标结果如表 2 所示。其中, *Stress* 值越小越好,如果大于 0.2 则是不可接受的。 *RSQ* 的值在 0~1 之间,其值越大越理想;越接近 1,说明拟合程度越好;越接近 0,说明拟合程度越差。一般 *RSQ* 值在 0.60 以上是可接受的,由此可看出文中河南省地理国情多维尺度分析的结果是比较理想的。

利用多维尺度分析结果(表 2 中的维度 1 和维度 2)构建分类坐标系,如图 5 所示。在维度 1 上,南阳在最右侧,信阳、洛阳和郑州在中间,其余的地市则集中在左侧;在维度 2 上,洛阳和郑州在最上部;驻马店、商丘、新乡、周口和信阳则在下部,其余地市则集中在中间。

在多维尺度分析结果可视化(图 5)的基础上,设置 *k* 均值聚类数为 4,按照 4 类分区方案进行聚类,结果如图 6 和表 3 所示。其中,聚类簇 1 包含的地市有:驻马店、商丘、新乡、周口和信阳;聚类簇 2 包含的地市有:洛阳和郑州;聚类簇 3 包含的地市有:三门峡、安阳、济源、鹤壁、许昌、平顶山、濮阳、焦作、开封和漯河;聚类簇 4 只包含南阳。

根据河南省地理国情数据 MDS 降维结果及 *k* 均值聚类结果,对各聚类簇中包含的地市的指标数

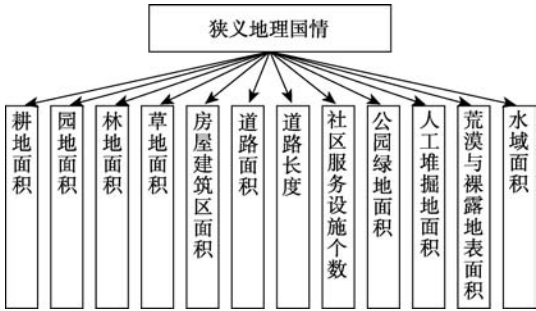


图 4 地理国情分类区划指标

Fig.4 Classification-based regionalization indices of geographical national conditions

表 2 河南省地理国情多维尺度分析

Tab.2 Multidimensional scaling result of geographical national conditions in Henan Province

行政区	维度 1 坐标值	维度 2 坐标值	行政区	维度 1 坐标值	维度 2 坐标值
安阳	-0.4225	0.7576	濮阳	-1.1777	-0.1420
鹤壁	-1.0321	0.2182	三门峡	-0.6020	1.1954
济源	-1.4897	0.2407	商丘	-0.2797	-1.1644
焦作	-0.1989	0.0637	新乡	0.6378	-1.0900
开封	-0.3322	-0.2362	信阳	1.0043	-1.2966
洛阳	0.9954	1.7661	许昌	-0.5465	0.0765
漯河	-1.1636	-0.2370	郑州	1.6840	1.6548
南阳	3.0637	-0.0791	周口	-0.2780	-1.0988
平顶山	0.0623	0.1584	驻马店	0.0753	-0.7874

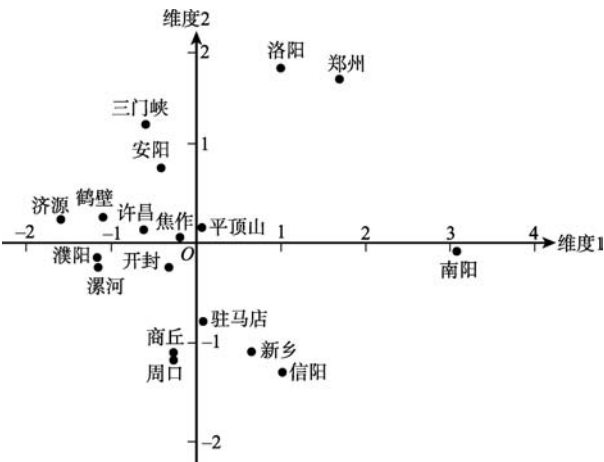


图5 河南省地理国情分类坐标系(欧氏距离模型)

Fig.5 Classification coordinate system of geographical national conditions in Henan Province (distance model: Euclidean)

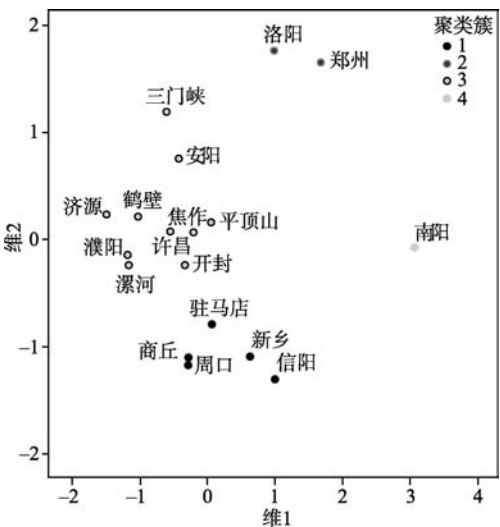


图6 河南省地理国情k均值聚类结果可视化

Fig.6 TVisualization of K-means clustering result for geographical national conditions in Henan Province

表3 河南省地理国情k均值聚类结果

Tab.3 K-means clustering result for geographical national conditions in Henan Province

聚类簇	簇中心坐标		距离簇中心最近的地市	各聚类簇包含的地市
	维度1	维度2		
1	0.2319	-1.0874	驻马店	驻马店、商丘、新乡、周口、信阳
2	1.3397	1.7105	洛阳(郑州)	洛阳、郑州
3	-0.6903	0.2095	许昌	三门峡、安阳、济源、鹤壁、许昌、平顶山、濮阳、焦作、开封、漯河
4	3.0637	-0.0791	南阳	南阳

据进行分析,其中各聚类簇指标数据分析图中,虚线为以聚类簇中心最近的地市构建的线性趋势线(图7)。从图7可以发现:聚类簇1中包含的各地市中地理国情分类区划各项指标值有高有低,发展不均衡,本文将此类区域命名为指标数据极值区;聚类簇2中大部分指标值处于中间值,仅有较少部分指标值处于较高水平,将此类区域定义为指标数据中值区;聚类簇3中大部分指标值处于较低值,将此类区域定义为指标数据低值区;聚类簇4中各项分类区划指标值均较高,将此类区域定义为指标数据高值区(表4)。

根据聚类结果及各聚类的类型表达,构建河南省地理国情分类区划专题图(图8)。从图8可以看出,河南省各地市地理国情分类区划显著而直观地呈现出一定程度的空间集聚现象,且各区域类型间相互包容和渗透,各区域块集聚呈4条西南↔东北的条状带,即“三门峡→济源→焦作→新乡→鹤壁→安阳→濮阳”“洛阳→郑州”“南阳→平顶山→漯河→许昌→开封”和“信阳→驻马店→周口→商

丘”。地理国情极值区中包含的各地市主要集中在东南周边,而地理国情低值区中包含的各地市主要集中在西北周边,中部各类型地理国情区划则呈相互交织和渗透的格局。

4 地理国情分类区划空间集聚性分析

为了对地理国情分类区划结果作进一步分析,从而挖掘其潜在的信息和知识,以下对分类区划结果的空间集聚性进行分析。

定义1:地理国情分类区划类型相关矩阵表示2个区域之间分类区划类型异同的矩阵,即将分类区划类型相同的两地市相关值设为1,不同的设为0。

定义2:地理国情邻接矩阵表示2个区域之间相邻关系以及区划类型异同的矩阵,即将地理位置相邻,且区划类型相同的2个区域间的邻接关系设置为1,反之则为0。假设空间邻接矩阵为 A ,地理国情分类区划类型相关矩阵为 B ,则地理国情邻接矩阵(C)为:

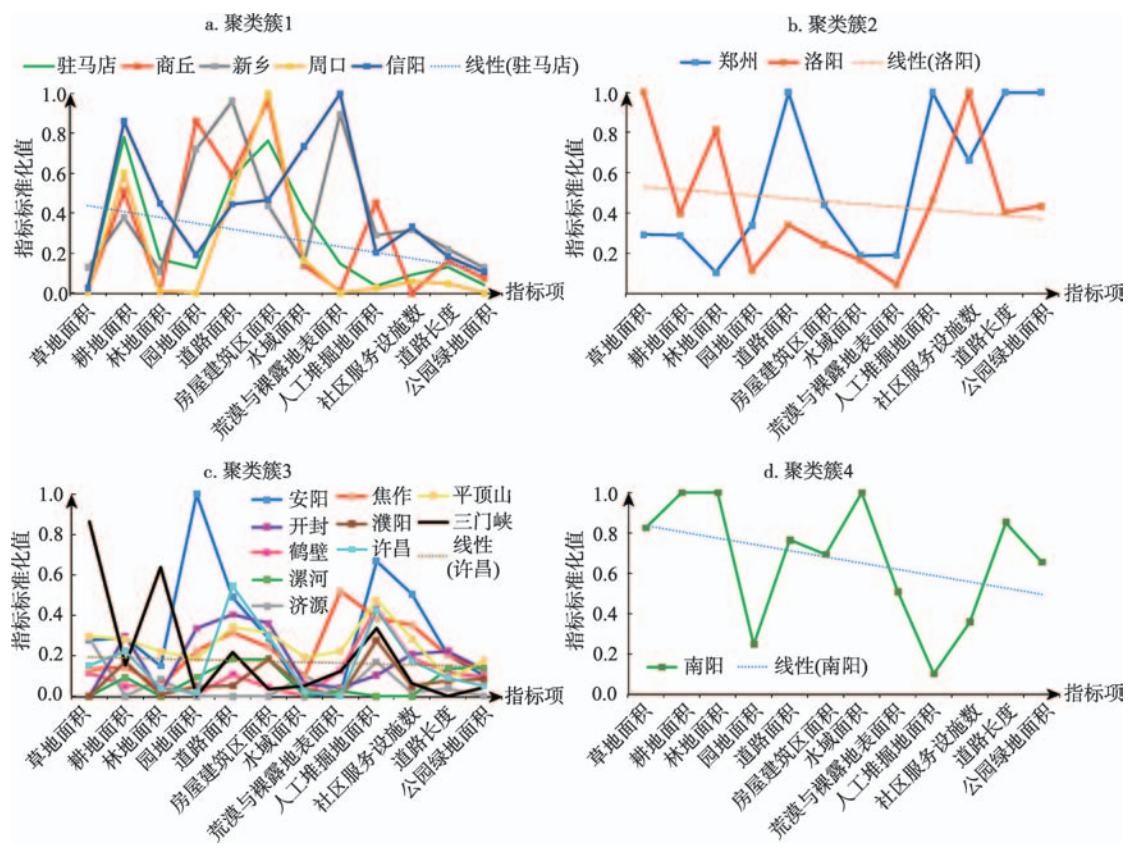


图 7 各聚类簇指标数据分析
Fig.7 Index value analysis for the clusters

表 4 河南省地理国情分类区划类型表达
Tab.4 Types of classification-based regionalization for geographical national conditions in Henan Province

聚类簇	区划类型	区划特征
1	指标数据极值区	各项指标值有高有低,发展不均衡
2	指标数据中值区	大部分指标值处于中间值,仅有较少部分指标值处于较高水平
3	指标数据低值区	大部分指标值处于较低值
4	指标数据高值区	各项分类区划指标值均较高

$$C=A\times B \tag{1}$$

根据邻接区域的分类属性异同性进一步对河南省地理国情分类区划的空间集聚性进行分析。首先获取河南省各地市的空间邻接矩阵 A (表 5);然后获取各地市区划类型相关矩阵 B (表 6),即区划类型相同的 2 个地市的相关值为 1,不同的设为 0;最后,计算区划类型相同的邻接地市占有所有邻接地市的比例,通过比例值判断河南省地理国情分类区划的空间集聚性。

将矩阵 A 和矩阵 B 中的各对应项进行相乘得到地理国情邻接矩阵 C 。计算矩阵 A 和矩阵 C 中元素值为 1 的个数分别为: $\text{sum}(\text{sum}(A,1))=90$ 和

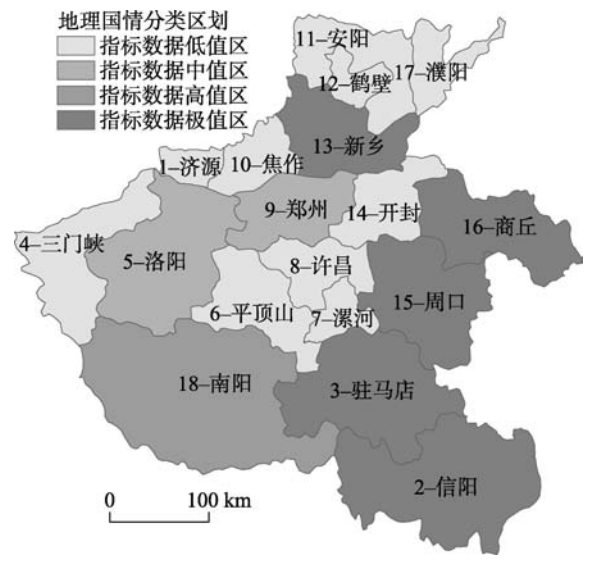


图 8 河南省地理国情分类区划图
Fig.8 Thematic map of classification-based regionalization for geographical national conditions in Henan Province

$\text{sum}(\text{sum}(C,1))=40$,两者之比为 0.4444,即河南省相邻地市中区划类型一致的地市占相邻地市总数的 44.44%,接近一半的相邻地市的区划类型相同。

表5 河南省各地市空间邻接矩阵
Tab.5 Spatial adjacency matrix of all cities in Henan Province

ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	1	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0
2	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
3	0	1	1	0	0	1	1	0	0	0	0	0	0	0	1	0	0	1
4	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1
5	1	0	0	1	1	1	0	0	1	1	0	0	0	0	0	0	0	1
6	0	0	1	0	1	1	1	1	1	0	0	0	0	0	0	0	0	1
7	0	0	1	0	0	1	1	1	0	0	0	0	0	0	1	0	0	0
8	0	0	0	0	0	1	1	1	1	0	0	0	0	1	1	0	0	0
9	0	0	0	0	1	1	0	1	1	1	0	0	1	1	0	0	0	0
10	1	0	0	0	1	0	0	0	1	1	0	0	1	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	1	0
12	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0
13	0	0	0	0	0	0	0	0	1	1	1	1	1	1	0	0	1	0
14	0	0	0	0	0	0	0	1	1	0	0	0	1	1	1	1	0	0
15	0	0	1	0	0	0	1	1	0	0	0	0	0	1	1	1	0	0
16	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0
17	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	1	0
18	0	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	1

注:表5和表6中ID号与图8中各地市注记的编号对应。

表6 河南省各地市分类区划类型相关矩阵
Tab.6 Type correlation matrix of classification-based regionalization of all cities in Henan Province

ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	1	0	0	1	0	1	1	1	0	1	1	1	0	1	0	0	1	0
2	0	1	1	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0
3	0	1	1	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0
4	1	0	0	1	0	1	1	1	0	1	1	1	0	1	0	0	1	0
5	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0
6	1	0	0	1	0	1	1	1	0	1	1	1	0	1	0	0	1	0
7	1	0	0	1	0	1	1	1	0	1	1	1	0	1	0	0	1	0
8	1	0	0	1	0	1	1	1	0	1	1	1	0	1	0	0	1	0
9	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0
10	1	0	0	1	0	1	1	1	0	1	1	1	0	1	0	0	1	0
11	1	0	0	1	0	1	1	1	0	1	1	1	0	1	0	0	1	0
12	1	0	0	1	0	1	1	1	0	1	1	1	0	1	0	0	1	0
13	0	1	1	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0
14	1	0	0	1	0	1	1	1	0	1	1	1	0	1	0	0	1	0
15	0	1	1	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0
16	0	1	1	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0
17	1	0	0	1	0	1	1	1	0	1	1	1	0	1	0	0	1	0
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

因此,可以判断河南省各地市地理国情分类区划的空间集聚效应显著。

5 结论与讨论

5.1 结论

分类是人类认识事物行之有效且简单易行的方法。研究地理国情分类区划,目的在于对抽象、复杂的地理国情有直观的认识,并在此基础上,理

解地理国情,从而分析地理国情。主要结论如下:

(1) 针对分类区划的实质和特点,根据《地理国情普查内容与指标》选取了地理国情分类区划指标(所选取的指标都是对地理国情作基本的描述而不含任何评价指标),利用多维尺度分析和聚类分析方法,构建了地理国情分类区划模型。

(2) 根据河南省地理国情数据多维尺度分析的Stress值和RSQ值,可以判断多维尺度分析的结果较为理想,从可视化的结果可以看出河南省各地市

的地理国情呈现分类聚集形式。

(3) 根据多维尺度分析结果,将河南省地理国情分为4类,并进行 k 均值聚类以实现地理国情分类区划。根据聚类中心,定义河南省地理国情区划类型分别为:指标数据极值区——商丘、新乡、信阳、周口和驻马店,指标数据中值区——洛阳和郑州,指标数据低值区——安阳、鹤壁、济源、焦作、开封、漯河、平顶山、濮阳、三门峡和许昌,指标数据高值区——南阳。

(4) 河南省各地市地理国情分类区划的空间集聚效应显著,且根据邻接区域的分类属性异同性分析对这一结论进行了佐证。河南省地理国情各区域类型间相互包容和渗透,各区域块集聚呈西南↔东北向的条状带。

5.2 讨论

基于以上结论,根据地理国情类型的划分及各区划类型的特点,结合其呈现出的空间集聚性特征,可对相同特征的区域制定相同或相似的区域协调发展策略,因地制宜。地理国情分类区划为科学认识地理国情,促进资源的合理开发利用,提升区域的社会和经济发展,促进地理国情区域协调与可持续发展提供了基础支撑。结合后续开展的地理国情分级区划研究,可进一步明确各类型区划地理国情协调发展的整体状态和具体的改进方向,为统筹地理国情区域发展,制定切实可行的协调发展策略提供决策支持。

参考文献(References)

- 陈百明. 1986. 土地分类体系与土地评价问题探讨[J]. 自然资源, (2): 91-96. [Chen B M. 1986. Tudi fenlei tixi yu tudi pingjia wenti tantao[J]. Natural Resources, (2): 91-96.]
- 陈俊勇. 2012. 地理国情监测的学习札记[J]. 测绘学报, 41(5): 633-635. [Chen J Y. 2012. Study notes on geographic national condition monitoring[J]. Acta Geodaetica et Cartographica Sinica, 41(5): 633-635.]
- 陈俊勇. 2014. 关于地理国情普查的思考[J]. 地理空间信息, 12(2): 1-3. [Chen J Y. 2014. Guanyu Dili guoqing pucha de sikao[J]. Geospatial Information, 12(2): 1-3.]
- 程维明, 周成虎. 2014. 多尺度数字地貌等级分类方法[J]. 地理科学进展, 33(1): 23-33. [Cheng W M, Zhou C H. 2014. Methodology on hierarchical classification of multi-scale digital geomorphology[J]. Progress in Geography, 33(1): 23-33.]
- 李红, 宋尚萍, 王广福. 2014. 基于地理国情普查数据的综合统计分析研究: 以试点区域北安市农业(耕地)变化统计分析为例[J]. 测绘与空间地理信息, 37(6): 137-139. [Li H, Song S P, Wang G F. 2014. Comprehensive statistical

analysis study based on national geographic condition survey data: The case of Bei'an agricultural farmland as the pilot area[J]. Geomatics & Spatial Information Technology, 37(6): 137-139.]

- 刘耀林, 何建华. 2014. 地理国情多层次统计分析指标体系设计[J]. 地理空间信息, 12(3): 1-4. [Liu Y L, He J H. 2014. Dili guoqing duocengci tongji fenxi zhibiao tixi sheji[J]. Geospatial Information, 12(3): 1-4.]
- 毛飞, 孙涵, 杨红龙, 等. 2011. 干湿气候区划研究进展[J]. 地理科学进展, 30(1): 17-26. [Mao F, Sun H, Yang H L, et al. 2011. Research progress in dry/wet climate zoning[J]. Progress in Geography, 30(1): 17-26.]
- 沈玉昌, 苏时雨, 尹泽生. 1982. 中国地貌分类、区划与制图研究工作的回顾与展望[J]. 地理科学, 2(2): 97-105. [Shen Y C, Su S Y, Yin Z S. 1982. Retrospect and prospect of the research work on the classification, regionalization and mapping of the geomorphology of China[J]. Scientia Geographica Sinica, 2(2): 97-105.]
- 王家耀, 谢明霞. 2016. 地理国情与复杂系统[J]. 测绘学报, 45(1): 1-8. [Wang J Y, Xie M X. 2016. Geographical national condition and complex system[J]. Acta Geodaetica et Cartographica Sinica, 45(1): 1-8.]
- 王永峰, 葛亮, 孙忠芳. 2014. 热点分析在国情监测数据分析中的应用初探[J]. 测绘与空间地理信息, 37(6): 84-85, 88. [Wang Y F, Ge L, Sun Z F. 2014. The application of hotspot analysis in national geographic condition monitoring data analysis[J]. Geomatics & Spatial Information Technology, 37(6): 84-85, 88.]
- 武琛. 2012. 地理国情监测内容分类与指标体系构建方法研究[D]. 泰安: 山东农业大学. [Wu C. 2012. Research on content classification and construction of index system of geographical national condition monitoring[D]. Taian, China: Shandong Agricultural University.]
- 肖笃宁, 钟林生. 1998. 景观分类与评价的生态原则[J]. 应用生态学报, 9(2): 217-221. [Xiao D N, Zhong L S. 1998. Ecological principles of landscape classification and assessment[J]. Chinese Journal of Applied Ecology, 9(2): 217-221.]
- 谢明霞. 2011. 高维数据聚类若干关键问题研究[D]. 郑州: 解放军信息工程大学. [Xie M X. 2011. The research on some key issues in high dimensional data clustering[D]. Zhengzhou, China: The PLA Information Engineering University.]
- 谢明霞, 王家耀. 2015. 地理国情分类区划及分级评价模型[J]. 测绘科学技术学报, 32(2): 192-196. [Xie M X, Wang J Y. 2015. The classified regionalization and evaluation model of geographic national condition[J]. Journal of Geomatics Science and Technology, 32(2): 192-196.]
- 徐德明. 2012. 监测地理国情服务科学发展[J]. 中国测绘, (4): 4-5. [Xu D M. 2012. Jiancex dili guoqing fuwu kexue fazhan[J]. China Surveying and Mapping, (4): 4-5.]
- 杨树珍. 1983. 国土整治与经济区划[J]. 地理学报, 50(2):

- 105-112. [Yang S Z. 1983. National land management and economic regionalization[J]. *Acta Geographica Sinica*, 50 (2): 105-112.]
- 张继贤. 2013. 地理国情普查基本统计分析技术规定[R]. 北京: 中国测绘科学研究院. [Zhang J X. 2013. Dili guoqing pucha jiben tongji fenxi jishu guiding[R]. Beijing, China: Chinese Academy of Surveying and Mapping.]
- 张新时. 1993. 研究全球变化的植被—气候分类系统[J]. 第四纪研究, 13(2): 157-169. [Zhang X S. 1993. A vegetation-climate classification system for global change studies in China[J]. *Quaternary Sciences*, 13(2): 157-169.]
- 郑度, 欧阳, 周成虎. 2008. 对自然地理区划方法的认识与思考[J]. 地理学报, 63(6): 563-573. [Zheng D, Ou Y, Zhou C H. 2008. Understanding of and thinking over geographical regionalization methodology[J]. *Acta Geographica Sinica*, 63(6): 563-573.]
- Han J W, Kamber M. 2007. 数据挖掘: 概念与技术[M]. 范明, 孟小峰, 译. 2版. 北京: 机械工业出版社. [Han J W, Kamber M. 2007. Data mining: Concepts and techniques[M]. Fan M, Meng X F, Trans.. 2nd ed. Beijing, China: China Machine Press.]
- Ceccato P, Connor S J, Jeanne I, et al. 2005. Application of geographical information systems and remote sensing technologies for assessing and monitoring malaria risk[J]. *Parassitologia*, 47(1): 81-96.
- Cox T F, Cox M A A. 2000. Multidimensional scaling[M]. 2nd ed. London, UK: Chapman & Hall/CRC.
- Mithal V, Garg A, Boriah S, et al. 2011. Monitoring global forest cover using data mining[J]. *ACM Transactions on Intelligent Systems and Technology*, 2(4): 36.
- Young O R, Lambin E F, Alcock F, et al. 2006. A portfolio approach to analyzing complex human-environment interactions: Institutions and land change[J]. *Ecology & Society*, 11(2): 31.

The model construction and empirical research on classification-based regionalization of geographical national conditions: Take Henan Province as an example

XIE Mingxia¹, WANG Jiayao², CHEN Ke²

(1. School of Resource and Environmental Sciences, Wuhan University, Wuhan 430079, China;

2. Institute of Geospatial Information, PLA Information Engineering University, Zhengzhou 450052, China)

Abstract: Classification based regionalization is an effective way to analyze geographical national conditions and a basis for studying the regional differences in the classification characteristics, combinations, change, and spatial differentiation of geographical national conditions. The index system of the classification-based regionalization of geographical national conditions is determined through the analysis of the basic statistical data of the geographical national conditions survey. Using multidimensional scaling and cluster analysis, the classification-based regionalization model of geographical national conditions was established. The classification-based regionalization experiment was carried out using the geographical national conditions survey data of Henan Province. The number of regions of geographical national conditions was initially intuitively determined through multidimensional scaling on geographic national conditions data of Henan Province. On this basis, the classification-based regionalization was achieved by *K*-means clustering. The type of regions is defined and explained according to the cluster centers and the characteristics of index data in each cluster. Finally, the clustering result is spatialized. The experimental results show that the geographical national conditions in Henan province can be divided into four regions: extreme value region, low value region, median value region and high value region. The spatial agglomeration effect of regions of geographical national conditions in Henan Province is significant, and It is mutual tolerance and penetration between each classification-based regionalization and constructed of four southwest to northeast strip. In the overall planning of regional development and realizing regional coordinated and sustainable development, we can make feasible regional balanced development policies according to the geographical national conditions of each region.

Key words: geographical national condition; classification-based regionalization; multidimensional scaling; cluster analysis; Henan Province