

基于数据分析方法的有居民海岛分类评价研究

张宏晔¹,王娜¹,曹英志¹,张志卫²,赵锦霞²

(1. 国家海洋信息中心 天津 300171;2. 自然资源部第一海洋研究所 青岛 266061)

摘要:海岛是我国国土重要的组成部分,是具有特殊性质的独立地理单元。海岛进行分类评价对海岛的开发和保护意义重大。文章选择 30 个海岛,从环境保护、经济发展和社会发展 3 个方面建立 16 个指标,通过聚类分析、主坐标分析和主成分分析,对海岛进行分类,同时研究海岛不同指标之间的关系。研究发现,海岛可以分为经济发展型、环境发展型和传统发展型 3 类。海岛的经济发展与社会发展表现出了较强的相关性,环境发展与社会发展表现出了较弱的相关性,经济发展与环境发展相关性很弱。通过对海岛进行分类评价,一方面可以深层次地了解海岛现状,同时也可以更好地把握海岛开发利用与环境保护之间的关系。

关键词:海岛分类;海岛环境;海岛发展;数据分析;主成分分析

中图分类号:P748 文献标志码:A 文章编号:1005-9857(2022)07-0090-10

Research of Islands Classification and Evaluation Based on Data Analysis Methods

ZHANG Hongye¹,WANG Na¹,CAO Yingzhi¹,ZHANG Zhiwei²,ZHAO Jinxia²

(1. National Marine Data and Information Service,Tianjin 300171,China;
2. The First Institute of Oceanography,MNR,Qingdao 266061,China)

Abstract: Islands, as an important land part of China, are independent geographical unit with special property. It is of great significance for exploitation and protection of islands that give them a correct classification evaluation. In this study, 30 islands were chosen with 16 different factors include environment protection, economic and social development to realize the classification of them. Meanwhile the relationship between different factors that can give a direction of developing islands was explored. The study showed that the islands could be classified by economic developing islands, environment developing islands and traditional islands. The economic development had a strong correlation to the social development and the environment had a weak correlation to the social development. The economic development had very few correlation to the environment. Through the islands classification and evaluation, on the one hand the current situation of

收稿日期:2021-11-06;修订日期:2022-06-07
基金项目:国家海洋信息中心 2018 年度青年基金项目。
作者简介:张宏晔,工程师,硕士,研究方向为海岛数据管理、海岛资源与环境评价等
通信作者:曹英志,研究员,博士,研究方向为海域海岛监管、海域海岛数据库建设、海域海岛政策等

islands could be deeply discovered, on the other hand the relationship between the exploitation and protection of islands could be grasped in a better field.

Keywords: Island classification, Island environment, Island development, Data analysis, Principal component analysis

0 引言

海岛是我国国土重要的组成部分,也是重要的生产、生活和生态空间的载体^[1]。海岛具有特殊的地理特性,它四面环水,天然独立,每个海岛都构成一个相对独立的地理环境^[2]。同时,海岛本身面积狭小、地域结构简单、生态系统脆弱。我国海岛分布在从南到北的各个海区,海岛数量众多^[3],每个海岛大小、岸线长度、地理位置、生态类型均不同,同时由于海岛的地理位置、人口以及历史发展等原因,每个海岛的自然条件以及开发类型也各不相同。如不因地制宜,合理规划,盲目开发,不仅不利于我国海岛的经济发展,而且会给海岛及周边海域带来非常严重的生态环境问题。建立相对完善、科学、量化的海岛评价方法对海岛进行评价,对海岛资源的合理利用以及海岛的生态环境保护都具有重要的意义。

由于海岛独立的地理特征,以及自然和社会的双重属性,因此可以从多个角度对海岛进行评价,包括海岛的地理特性、地质组成、生态环境,到海岛的经济发展、开发利用类型以及海岛的历史文化等,涉及多学科、多专业。目前国内对于海岛的评价研究则侧重于海岛单一角度,如对海岛的价值体系研究^[1]、生态环境评价研究^[2]、海岛的经济或生态脆弱性研究^[4-5],也有海岛保护及分类规划研究^[6]、无居民开发适宜性研究^[7]和PSR模型对海岛的生态评价研究^[8]等。在国际上,也有一些针对海岛特性进行的定量评价研究,如生态系统安全研究^[9]、基于神经网络的海岛评价研究^[10]和海洋环境影响因子研究^[11]等。这些方面每一项都可以建立相应的指标体系对海岛特定的属性进行评价,但是这些体系大多只是从一个或几个方面对海岛进行研究评价,而如果需要相对全面、客观地评价海岛生态环境和发展状况,则需要对评价指标进行关联,使用综合评价模型对其进行评价。同时,多数评价只是

针对单一海岛,并没有建立不同海岛之间的联系,尤其是不同海岛之间社会发展和自然环境之间的潜在联系。

数据挖掘与分析是从大量数据中发现具有描述性、可理解、能预测的模型的过程^[12]。主要包括探索性因子分析^[13]、频繁模式挖掘^[14]、分类和聚类^[15]。目前,国内外各类数据分析在综合评价体系已经相对成熟,对资源环境的综合评价方法有几十种甚至上百种之多,其中利用数据进行定量分析的评价方法就有几十种,如层次分析法(AHP)^[16]、灰色关联分析法(GIA)^[17]和熵权法(EA)^[18]等,而随着数据挖掘和分析研究的深入,基于统计分析的综合评价方法也有很多种,如主成分分析(PCA)^[19]、因子分析(FA)^[20]和聚类分析^[21]。这些评价方法也应用在对不同类型的资源评价过程中,如城市宜居度分析^[22]、决策方法选择^[23]和能源效用评价^[24]等。数据分析也是土地利用以及区域发展的研究的主要手段。在城市功能区划分类^[25]、乡村旅游资源分类评价^[26]和区域综合性评价^[27]等研究中,均分析了不同类型的地类的功能划分以及空间特征。

本研究基于数据挖掘和数据分析的方法,结合海岛特殊的自然和社会发展性质,构建了基于自然环境、经济发展和社会民生的海岛分类评价体系。基于海岛的自然和社会属性建立了多重评价指标,对我国30个海岛进行分类评价。通过聚类、分类以及主成分分析、主坐标分析等数据分析方法对海岛进行定量分析,以期建立较完善的海岛分类评价方法。对海岛的分类评价,既可以发现海岛的发展现状,探究海岛发展的个性和共性,也可以揭示海岛经济发展和环境保护之间的关系。可以对海岛规划、海岛资源的开发利用及保护提供科学、客观的数据支持及参考借鉴。

1 研究区域情况与数据来源

1.1 实例海岛及选取依据

与人为的城市或者行政区划不同,海岛四面

环水,海洋天然地将海岛划分成为独立的地理单元。但是由于不同海岛的自然环境、人口规模和区位因素等往往存在较大差异,为了客观、全面、具体地评价海岛,一方面选择的海岛需要尽可能多地体现海岛本身属性的差异,如所在省份及海区,近陆距离、主导产业等;另一方面,为避免因统计数据差距过大导致难以标准化,选择的海岛面积人口规模应相对接近,防止选择海岛面积人口规模过大(如,海南岛)或过小(如,永兴岛)。

根据以上原则,本研究在全国范围内选择了 30 个典型的有居民海岛进行评价(表 1)。这 30 个海岛分布在全国各个沿海省(自治区、直辖市),每一个海岛都具有独特的地理环境、开发利用特征、经济发展以及社会环境,这样可以使分类评价的结果更具有客观性,同时也能更清楚地体现各评价指标对海岛分类的影响。

表 1 研究的目标海岛基本情况

标准名称	所属省(自治区、 直辖市)	所属市	所属县
东海岛	广东省	湛江市	麻章区
海陵岛	广东省	阳江市	江城区
六横岛	浙江省	舟山市	普陀区
长兴岛	上海市		崇明县
南田岛	浙江省	宁波市	象山县
琅岐岛	福建省	福州市	马尾区
南日岛	福建省	莆田市	秀屿区
桃花岛	浙江省	舟山市	普陀区
梅山岛	浙江省	宁波市	北仑区
广鹿岛	辽宁省	大连市	长海县
涠洲岛	广西壮族自治区	北海市	海城区
大嵛山	福建省	宁德市	福鼎市
花岙岛	浙江省	宁波市	象山县
湄洲岛	福建省	莆田市	秀屿区
施公寮岛	广东省	汕尾市	城区
獐子岛	辽宁省	大连市	长海县
鹿西岛	浙江省	温州市	洞头县
大万山岛	广东省	珠海市	香洲区
北长山岛	山东省	烟台市	长岛县
灵山岛	山东省	青岛市	黄岛区
连岛	江苏省	连云港市	连云区
枸杞岛	浙江省	舟山市	嵊泗县
新埠岛	海南省	海口市	
东庠岛	福建省	福州市	平潭县
桂山岛	广东省	珠海市	香洲区
下大陈岛	浙江省	台州市	椒江区
花鸟山岛	浙江省	舟山市	嵊泗县
刘公岛	山东省	威海市	环翠区
白沙山岛	浙江省	舟山市	普陀区
龙门岛	广西壮族自治区	钦州市	钦南区

1.2 海岛评价指标体系的构建

海岛作为特殊的地理单元,是一类自然—人工的复合体。这类单元的资源环境不仅仅受地理位置和生态环境等自然因素的影响,也受到其所属的社会环境、经济发展、文化以及民生等社会因素的影响。基于此,在参照城市分类或其他地理分类评价研究的基础上^[25],结合海岛的实际情况以及指标的客观性,本研究分别从经济发展、生态环境、社会民生和文化及规划管理等方面选择了 16 项评价指标(表 2),这些指标对海岛可以进行相对完善的分类评价。为了进一步的数据分析,本研究对各个指标需要通过数字进行描述,即用不同的方法对这些指标进行量化。最终利用通过得到的指标数据化结果对海岛进行分类评价。

1.3 原始数据的收集

为使海岛数据相对标准统一,本研究以 2016 年为时间节点获取原始数据。其中遥感影像是基于 2016 年度拍摄的高分 1 号以及资源 3 号卫星(分辨率 2m)的影像,获取影像后基于 ArcGIS10.5 平台进行解译,计算海岛的植被、开发利用面积和岸线长度等,同时对部分设施(如,防灾减灾和文化体育类)予以标注,从而获得相应的评价指标数据。对于经济和社会类指标,一方面咨询当地相关行政部门获取环境、医疗和社会保障等数据以及经济统计调查数据;另一方面在 2016—2017 年进入各海岛现场进行补充调查。最后,搜集海岛所在县、乡的统计公报、年鉴及政府工作报告以及涉及海岛的其他公开资料,同时参考 2016 年国家海洋局发布的《海岛统计调查公报》^[28]以及《2016 年中国海洋环境状况公报》^[29],对部分评价指标数据进行补充计算。

对获取的数据,根据数据的性质、分类和数据意义等属性信息,对每一项数据进行针对性的数值量化,得到的最终数据为数值型数据。得到数据以后,利用 R 语言平台对数据进行多个类型的数据分析。将海岛进行分类,并探究影响海岛分类指标之间的联系。

表 2 海岛的评价指标

序号	指标名称	指标类别	指标含义	数据来源及量化方法
1	单位面积财政收入	经济发展	—	根据海岛统计调查数据直接量化
2	海岛人均可支配收入		—	根据海岛统计调查数据以及县级政府公报量化
3	植被覆盖率	生态环境状况	海岛植被面积/海岛总面积	遥感解译以及基于 GIS 计算
4	岛陆建设比例		海岛开发利用面积/海岛总面积	遥感解译以及基于 GIS 计算
5	自然岸线保有率		海岛自然岸线长度/海岛岸线总长度	遥感解译以及基于 GIS 计算
6	海岛周边水质指数		海岛周边水质情况,分值越高水质越好	基于公报数据及 GIS 计算
7	海岛污水处理率	社会民生及基础 设施建设	—	海岛所在地相关行政部门提供
8	海岛自然特征指数		反映珍稀濒危物种及栖息地、古树名木、自然和历史人文遗迹等保护情况,以上情况越多分数越高	基于海岛统计调查以及现场调查结果赋值
9	防灾减灾设施		海岛防灾减灾设施完备情况,设施越完备,分值越高	遥感解译以及现场调查结果量化
10	海岛利用特色指数		反映海岛所获荣誉称号,如“十大美丽海岛”,5A 级旅游景区等,荣誉称号越多分数越高	基于公开资料以及现场调查进行赋值
11	规划管理		海岛本身是否有明确的发展规划	基于现场调查结果赋值
12	对外交通条件		反映海岛对外交通的便利程度,越便利分值越高	基于现场调查结果赋值
13	医疗卫生指数		每千名常住人口公共医疗卫生人员数	海岛所在地相关行政部门提供
14	社保状况指数		岛上居民养老保险以及医疗保险覆盖率均值	海岛所在地相关行政部门提供
15	文化体育指数		反映岛上人均拥有公共文化体育设施面积	遥感解译以及现场调查结果量化
16	社会治安指数		岛上警务机构数量以及岛上治安案件发生情况,警务数量越多、治安案件越少,分数越高	基于乡镇政府报告以及现场调查结果赋值

1.4 评价方法的评价体系的建立及数据分析

1.4.1 评价指标的标准化

由于每个评价指标不同,描述评价指标的量纲也就不同,所以,需要对各评价指标数据进行标准化,以便消除不同量纲对数据结构的影响。对数据进行标准化以后,各项指标在相对统一的量纲内,可以相对准确计算每个目标海岛之间的距离。从而进行进一步的统计分析。标准化的方法如下:

$$X_{is} = \frac{X_i - \bar{X}}{X_{rms}}$$

$$X_{rms} = \sqrt{\frac{\sum_{i=1}^n X_i^2}{n-1}}$$

式中: X_{is} 是为指标统一量纲得到的数值; n 为样本数; $\bar{x}$ 为样本均值; X_i 是样本中的具体指标数值; X_{rms} 是样本均方根。

1.4.2 聚类分析

聚类分析是将对象的集合进行分组为由类似的对象组成多个类的分析过程^[30],通过将海岛进行聚类分析,可以研究不同海岛之间的指标特征可能的相似性,从而构建不同海岛的聚类树。本研究采用层次聚类法^[31],将海岛进行聚类,相似程度高的海岛聚类距离较近,而相似程度低的海岛则距离较远。这样,通过聚类分析即可判断海岛数据间的相似程度高低,并且最终得到一个海岛的聚类树。通过聚类树可以分析海岛在评价指标之间的相似性。

1.4.3 主坐标分析

主坐标分析(PCoA)是一种降维排序方法,通过一系列的特征值和特征向量排序从多维数据中提取最主要的元素和结构^[32]。基于距离矩阵来寻找主坐标。即通过计算各个类簇间的距离,可视化表征各个海岛的分类情况,本研究利用基于 Bray-Curtis 距离来进行 PCoA 分析,并选取贡献率最大的主坐标

组合进行二维坐标图展示。如果不同海岛间指标特征越接近,则它们在 PCoA 图中的距离越接近。

1.4.4 主成分分析

主成分分析(PCA)是一种应用方差分解,对多维数据进行降维,从而提取出数据中最主要的元素和结构的方法^[33]。利用主成分分析,可以提取出最大限度反映海岛间差异的 2 个或者 3 个特征,并将特征通过二维或者三维坐标图进行展示,如果海岛的各个指标数据越相似,则它们在 PCA 图中的距离越接近。每个特征与海岛的评价指标均有关联性,提取关联较大的评价指标,可以分析对海岛分类影响较大的指标,从而发现不同海岛指标之间的关联性。

2 研究结果与分析

2.1 海岛最佳类簇数的确定

研究利用 CH 指数可以确定特定聚类效果最佳的类簇的数量。具体计算方法如下:

$$CH_k = \frac{\frac{B_k}{k-1}}{\frac{W_k}{n-k}}$$

式中: n 为样本数; k 为类别数; B_k 为类别间的协方差矩阵; W_k 为类别内的协方差矩阵。该指数越大,说明类别间的协方差与类别内部的协方差之间的比值越大,也就是类别间的差异大于类别内部的差异。即分类效果最佳。最终得到的分类结果如图 1 所示。

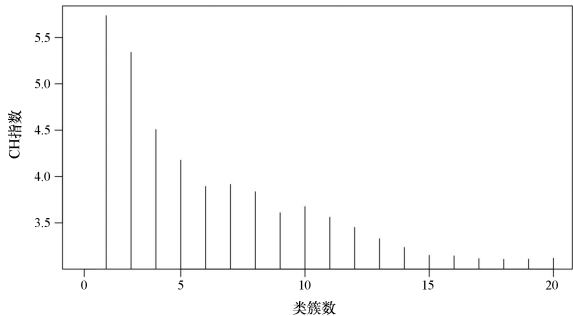


图 1 海岛聚类数量的 CH 分析

由图 1 可见,类别为 2 或者 3 时聚类效果最佳。但是当类簇为 2 时,聚类结果显示一个类簇仅剩余梅山岛一个样本,即梅山岛为一类,其余海岛为另一类。为了更明确地展示所有海岛之间的差异,本

研究选择 3 为聚类最佳类簇数对海岛进行聚类。

2.2 对目标海岛的聚类分析

利用层次聚类法,由海岛的各项指标的表征计算各海岛之间在表征的距离,基于距离得到海岛聚类的效果,并验证聚类分成 3 类的聚类特征。经过多种距离计算方式演算,本研究采用最能直观地表现聚类的效果的方法,即利用离差平方和法(也称为 ward 度量)计算不同类之间的距离^[34]。聚类结果如图 2 所示,在类簇数为 3 的情况下,30 个海岛被分成了 3 个主要类别。

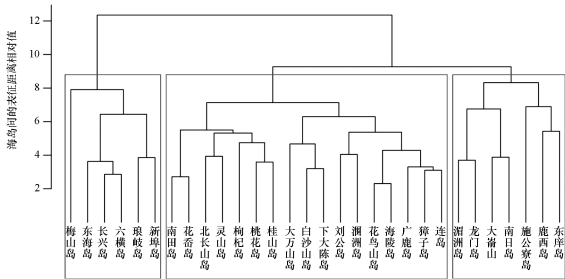


图 2 海岛聚类分析

对海岛和评价指标进行二维聚类,利用聚类热图对海岛和评价指标进行双向聚类(图 3)对层次聚类结果进行验证。从结果可以发现,一方面,双向聚类 and 层次聚类法的聚类结果相同,层次聚类结果合理;另一方面,同一指标类别的指标距离接近,说明同一类别的指标对海岛的描述相对统一,从数据上反映的指标分类较为合理。

同时,通过聚类热图可以进一步表征这些海岛以及指标的双重聚类的情况,并更好地解释每一类海岛的特征。具体海岛分类结果如下。

(1)经济发展型海岛,包括:梅山岛、东海岛、长兴岛、六横岛、琅岐岛和新埤岛。海岛有经济发展类的指标高、社会民生基础指标较高,但是环境相关指标较低等特征,本研究将其称之为经济发展类海岛。

(2)环境发展型海岛,包括:南田岛、花鸟岛、北长山岛、灵长山岛、枸杞岛、桃花岛、桂山岛、大万山岛、白沙山岛、下大陈岛、刘公岛、涠洲岛、花鸟山岛、海陵岛、广鹿岛、獐子岛和连岛。该类海岛的环境相关指标高,社会民生基础指标较 1 类海岛低,但较 3 类海岛高,经济发展类指标较第 1 类低,但较第

3 类高,我们将其称为环境发展类海岛。

(3)传统发展型海岛,包括:湄洲岛、龙门岛、大嵛山、南日岛、施公寮岛、鹿西岛和东庠岛。3 类海

岛环境相关指标介于前两者之间,社会民生基础及经济发展情况较前两者更低的特征,本研究将其称为传统发展类海岛。

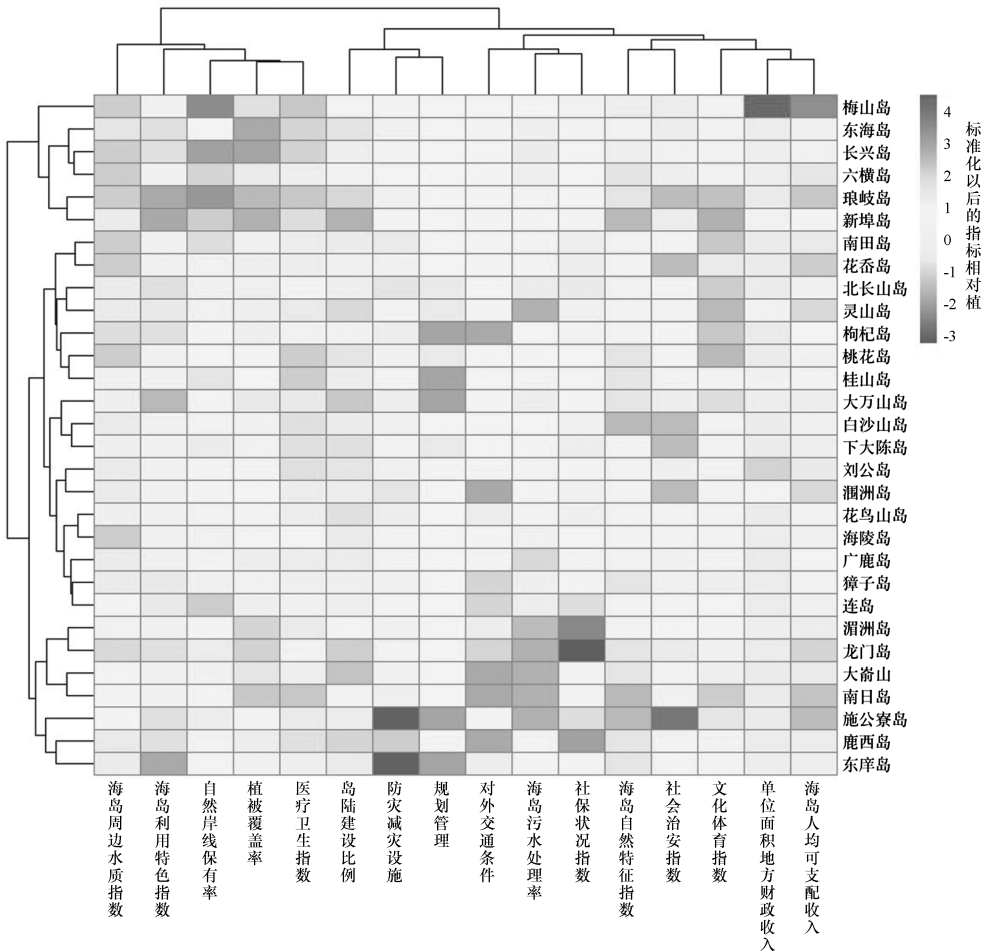


图 3 海岛及评价指标双向聚类热图

2.3 主坐标分析对海岛分类结果的验证

30 个样本进行主坐标分析的结果如图 4 所示, 3 个类别的海岛基本分布在坐标轴不同的区域,验证了聚类结果的正确性,3 个类别的海岛无明显交叉,分类结果良好,环境发展型海岛与传统发展型海岛的距离相对于经济发展型海岛与该两类海岛的距离近。

2.4 海岛评价指标之间的关联性

首先对分析结果进行碎石检验。结果表明,将评价指标分为 2 个或者 3 个主成分即可解释大部分的信息。然后,本研究提取了 3 个主成分进行主成分分析,并制作了主成分分析表(表 3)。主成分分析表主要表现了对海岛评价指标主成分分析的如下特征。

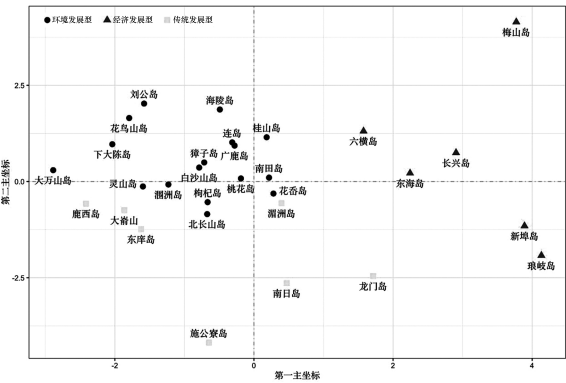


图 4 主坐标分析

可以发现,第一主成分与环境因素(如,岛陆建设比例、植被覆盖率和自然岸线保有率)有关;第二

主成分与经济发展(如,岛上人均可支配收入和单位面积财政收入)有关;第三主成分与海岛基础设施建设(如,规划管理、社保和防减灾措施)有关。3 个主成分对数据的解释度共为 49%,解释度并不高,但是主成分的指标出现了比较高的相似性。为了进一步验证主成分分析结果,将评价指标向量在两个

解释度最高的主成分组成的二维平面中进行投影,制作主成分方向图(图 5),从而研究各个评价指标的相关性。从图中可以发现,社会民生相关基础指标与经济发展相关指标有较强的正相关,但与环境的相关性较弱,经济发展相关指标与环境指标相关度很低。

表 3 主成分分析

海岛描述指标	相关系数			公因子方差	成分唯一性
	第一主成分(PC1)	第二主成分(PC2)	第三主成分(PC3)		
单位面积财政收入	-0.29	0.57	-0.14	0.43	0.57
海岛人均可支配收入	-0.27	0.73	-0.15	0.63	0.37
植被覆盖率	0.89	0.16	-0.18	0.85	0.15
岛陆建设比例	-0.83	-0.27	0.18	0.80	0.20
自然岸线保有率	0.80	-0.12	0.15	0.67	0.33
海岛周边水质指数	0.40	-0.18	-0.02	0.19	0.81
海岛污水处理率	-0.02	0.47	-0.44	-0.41	0.59
海岛自然特征指数	0.18	0.47	0.29	0.33	0.67
防灾减灾设施	-0.33	0.49	0.45	0.56	0.44
海岛利用特色指数	0.48	0.43	0.19	0.45	0.55
规划管理	-0.37	0.25	0.65	0.62	0.38
对外交通条件	-0.29	0.29	-0.48	0.39	0.61
医疗卫生指数	0.67	0.19	0.13	0.50	0.50
社保状况指数	-0.08	0.36	-0.62	0.52	0.48
文化体育指数	0.13	0.51	0.28	0.36	0.64
社会治安指数	0.20	0.31	0.26	0.20	0.80
主成分对数据的解释程度	0.22	0.16	0.11		

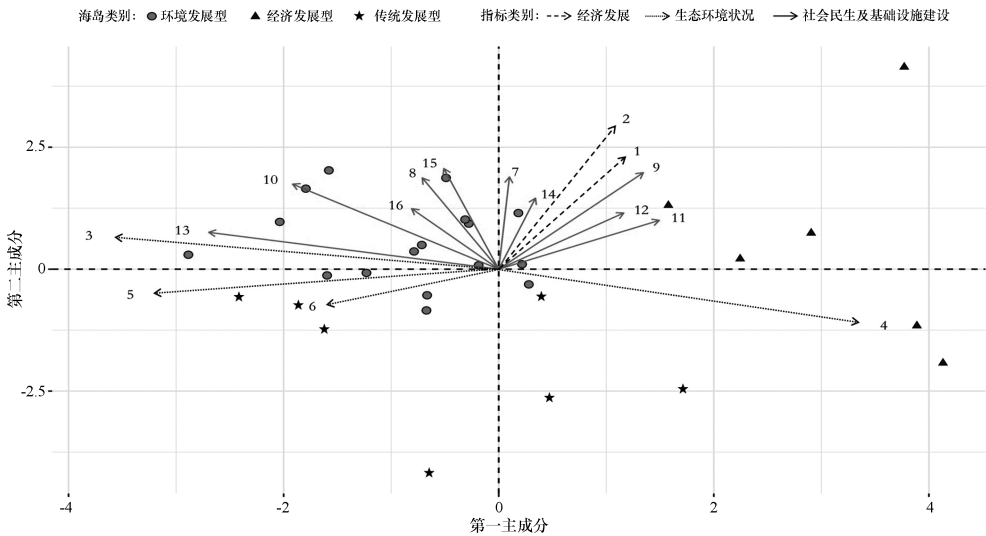


图 5 主成分分析矢量图

注:箭头的数字为评价指标,对应为表 2 的指标序号。

3 讨论与结论

3.1 讨论

海岛作为四面环水且面积较小的地理单元,生态环境复杂度低、人口相对较少、每个海岛的居民生活和产业状况相对简单。但是不同海岛区位、气候、地理、自然和生态类型都不尽相同,同时海岛也是一类自然-人文复合的地理单元^[35],因此对海岛的评价一方面需要针对海岛本身的特点,对能够描述海岛的多种指标综合进行统计分析从而对海岛进行分类。另一方面,针对影响海岛的自然因素和社会因素,若发现不同因素间可能存在的关系,则可以客观全面地对海岛进行评价。从评价结果看,利用聚类方法可以做到将海岛进行分类,且每一类海岛具有一些相关的指标特性,具体如下。

(1)海岛可以分成 3 类,即:经济发展型海岛、环境发展型海岛和传统发展型海岛。

经济发展型海岛:该类型海岛在经济指标上明显高于其他海岛,这可能与海岛发展类型有关,有研究表明第二产业和第三产业的发展对一个地区的经济规模产生较大影响^[36-37],多数的经济发展型海岛岛内有较强的第二、第三产业支柱,也有因区位优势顺应周边区域经济的发展。而其他类型的海岛除旅游业外,岛内基本无第二产业或其他第三产业,或海岛区位相对偏远,较难获得周边优质的经济发展资源。

环境发展型海岛:该类海岛的明显特征是相比于经济发展类海岛,环境指标突出。由于海岛兼具海陆的生态特征,环境优美,旅游产业一直是海岛的优势产业^[38-39],而本研究中分类得到的环境发展型海岛大多是以旅游业为主要发展产业,由于产业的需要优先发展了海岛的环境。部分海岛虽然不以旅游业为主要产业,但是由于岛内开发利用活动较少,环境较好,同时经济相关的指标较传统发展型海岛高,在聚类中更接近环境发展型,因此被归类到环境发展型海岛。

传统发展型海岛:该类海岛的显著特征是经济发展相关指标较低,该类型大多数海岛均以传统的以渔业、养殖业为主的第一产业作为海岛主要产业。作为渔业,海岛第一产业对海岛开发利用规模

较小,对海岛环境的影响有限。因此本类海岛环境指数相对较高。但是其环境指标值整体低于环境发展型海岛,这可能是由于海岛本身经济发展和基础设施发展较差,且对该类海岛环境保护中投入有限,所以环境指标较环境发展型海岛低。

(2)在 3 类海岛的评价指标中,经济水平相关指标与海岛的社会发展评价指标呈现较强的正相关,海岛的环境相关评价指标与社会发展评价指标呈现较弱的相关性,海岛的经济相关指标和环境相关指标无明显相关性。

有些研究表明,经济的快速发展可以导致资源的巨大消耗和对环境的影响^[40-41]。由于海岛的生态结构简单、环境脆弱、经济结构单一,海岛的经济发展很可能会对海岛环境具有较大影响,但是,从本研究的结果来看,海岛的经济发展情况并没有与环境指标显示出较明显的相关性,也就是说,海岛的经济发展和海岛的环境状况相对独立,这可能有以下几个原因。①海岛的开发方式不同,部分经济发展型海岛如梅山岛,其海岛产业以物流、贸易为主,该类产业对环境的影响较低,因此并未对海岛的环境造成较大的影响。②有些经济发展较强的海岛,在建设基础设施的同时,进行了生态修复和环境建设,如长兴岛,在现场对比遥感解译的情况时,发现许多植被覆盖的区域是生态修复工程区,这对本研究的环境指标有正向影响。环境指标较高的经济发展型海岛,其社会民生及基础设施指标中偏环境的指标如海岛自然特征指数和海岛利用特征指数等也较高,侧面反映了这些海岛在经济发展的同时,也在注重并改善海岛本身的生态环境。与之相对的,在经济发展指标反映较差的传统发展型海岛,并没有显示出较高的与生态环境相关指标值。也反映了经济发展与生态环境状况的独立性。生态修复作为人工的改善环境的方法,在其他类型地理单元可以明显改善环境^[42-43],但是海岛作为特殊的环境地理单元,其生态修复的效应可能与其他地区不同。后续研究可以探究海岛生态修复与海岛环境改善之间的关系。

海岛是地理特点和区位因素都具有独特性的地理单元。对其进行科学分类,明确其发展的现

状,研究其发展的成因,对指导海岛地区发展,保护海岛及其周边海域生态环境都具有重要的意义。对于海岛的发展,既要做到兼顾海岛的经济和社会情况,又要保护海岛的生态环境。对于不同类型的海岛,要结合实际情况,有针对性地制订海岛发展的规划和政策,对于经济发展型海岛,需要重点保护和改善海岛的环境,加强海岛环境监测,加强生态修复。对于环境发展型海岛,需要划定生态红线,防止人类活动对海岛环境产生的潜在的压力。对于传统发展型海岛,则需制订明确的海岛发展规划,引导海岛产业发展或转型,使海岛经济向多元化、集约化发展,加强海岛交通及基础设施建设。

3.2 结论

本研究针对海岛的特点建立了一套基于数据挖掘和数据分析的描述海岛发展和环境的分类体系。通过 30 个海岛的数据对体系进行了验证并分析。结果表明,分类体系相对可行,通过分类体系也可以发现不同海岛之间发展因素的内在联系。由于海岛的描述指标众多,且海岛的大小和人口规模都各不相同,海岛分类体系仍需完善。一方面,由于数据量越大,其可挖掘的数据越多^[44],后续会将更多海岛纳入分类体系,可以使海岛分类更精确、结果更明显;另一方面,增加一些更详细的描述指标,如海岛的生态系统相关指标,或对海岛的某个学科领域进行细致分类,可以更深层次地发现海岛发展的内生因素,有利于海岛未来发展规划的制定和管理政策的实施。

参考文献

[1] 刘容子,齐连明.我国无居民海岛价值体系研究[M].北京:海洋出版社,2006.

[2] 张志卫,丰爱平.基于生态系统的海岛保护与利用规划理论与实践[M].北京:海洋出版社.2017.

[3] 自然资源部.2017 年海岛统计调查公报[Z].2017.

[4] 秦伟山,孙剑锋,张义丰,等.中国县级海岛经济体脆弱性综合评价及空间分异研究[J].资源科学,2017,39(9):1692—1701.

[5] 池源,石洪华,郭振,等.海岛生态脆弱性的内涵、特征及成因探析[J].海洋学报,2015,37(12):93—105.

[6] RUTHERFORD J,KOBYRN H,NEWSOME D.A case study in the evaluation of geotourism potential through geographic information systems: application in a geology-rich island

tourism hotspot[J].Current Issues in Tourism,2015,18(3): 267—285.

[7] 陈秋明.基于生态—经济的无居民海岛开发适宜性研究[D].厦门:厦门大学,2009.

[8] 黄备,魏娜,孟伟杰,等.基于压力—状态—响应模型的辽宁省长海海域海洋生物多样性评价[J].生物多样性,2016,24(1): 48—54.

[9] CHEN H S,LIU W Y,HSIEH C M.Integrating ecosystem services and eco-security to assess sustainable development in Liuqiu Island[J].Sustainability,2017,9(6):1002.

[10] DOU J,YAMAGISHI H,POURGHASEMI H R,et al.An integrated artificial neural network model for the landslide susceptibility assessment of Osado Island, Japan [J]. Natural Hazards,2015,78(3):1749—1776.

[11] VANSTOCKEM J,BASTIAENS A,HELSEN K,et al.Is no green roof an island? Impact of the local and regional environment on diversity[C]//Society for Ecological Restoration Europe 2018,Date:2018/09/09—2018/09/13,Location:Reykjavik,Iceland.2018.

[12] WOLD S,ESBENSEN K,GELADI P.Principal component analysis[J].Chemometrics and Intelligent Laboratory Systems, 1987,2(1—3):37—52.

[13] BANDALOS D L,FINNEY S J.Factor analysis:Exploratory and confirmatory [M].New York:The reviewer's guide to quantitative methods in the social sciences.Routledge,2018: 98—122.

[14] HOMAYOUN S,DEGHANTANHA A,AHMADZADEH M,et al.Know abnormal,find evil:frequent pattern mining for ransomware threat hunting and intelligence[J].IEEE transactions on emerging topics in computing,2017,8(2):341—351.

[15] NGUYEN H L,WOON Y K,NG W K.A survey on data stream clustering and classification[J].Knowledge and information systems,2015,45(3):535—569.

[16] SATTY T L,TRAN L T.On the invalidity of fuzzifying numerical judgements in the Analytic Hierarchy Process [J]. Mathematical and Computer Modelling,2007,46(7—8):962—975.

[17] 刘思峰,蔡华,杨英杰,等.灰色关联分析模型研究进展[J].系统工程理论与实践,2013,33(8):2041—2046.

[18] LI A,ISMAIL A B,THU K,et al.Performance evaluation of a zeolite-water adsorption chiller with entropy analysis of thermodynamic insight [J]. Applied Energy,2014,130:702—711.

[19] KAO Y H,ROY B VDirected Principal Component Analysis [J].Operations Research,2014,62(4):957—972.

[20] HSU C C.Evaluation criteria for blog design and analysis of

- casual relationships using factor analysis and DEMATEL[J]. Expert Systems with Applications, 2012, 39(1): 187—193.
- [21] XIE J, GIRSHICK R, FARHADI A. Unsupervised deep embedding for clustering analysis[C]//International conference on machine learning. PLMR, 2016, 48: 478—487.
- [22] 张睿涵. 基于层次分析法的淮海经济区城市宜居度研究[J]. 低碳世界, 2019, 9(7): 339—340.
- [23] 孙晓东, 焦玥, 胡劲松. 基于灰色关联度和理想解法的决策方法研究[J]. 中国管理科学, 2005, 13(4): 63—68.
- [24] MIAN H, YAN W W, JIANG W X, et al. Multi-energy management with hierarchical distributed multi-scale strategy for pelagic islanded microgrid clusters[J]. Energy, 2019, 185.
- [25] 陈小俊. 智慧城市规划中的大数据应用与技术分析[J]. 智能城市, 2019, 5(5): 69—70.
- [26] 陈宇. 湘西少数民族地区乡村旅游资源分类及评价[J]. 中国农业资源与区划, 2019, 40(2): 205—210.
- [27] 卢新海, 沈纬辰, 杨喜, 等. 中国区域实体经济、科技创新、现代金融与人力资源协同发展评价[J]. 统计与决策, 2019(15): 129—132.
- [28] 国家海洋局. 2016 年海岛统计调查公报[Z]. 2016.
- [29] 国家海洋局. 2016 年中国海洋环境状况公报[Z]. 2017.
- [30] KAUFMAN L, ROUSSEEUW P J. Finding groups in data: an introduction to cluster analysis[M]. New Jersey: John Wiley & Sons, 2009.
- [31] GOMEZ L, KARATZAS D. A fast hierarchical method for multi-script and arbitrary oriented scene text extraction[J]. International Journal on Document Analysis and Recognition (IJDAR), 2016, 19(4): 335—349.
- [32] GOWER J C. Principal coordinates analysis[J]. Encyclopedia of biostatistics, 2005(6).
- [33] LEVER J, KRZYWINSKI M, ALTMAN N. Points of significance: principal component analysis[J]. Nature Methods, 2017, 14: 641—642.
- [34] MURTAGH F, LEGENDRE P. Ward's hierarchical agglomerative clustering method: which algorithms implement Ward's criterion? [J]. Journal of classification, 2014, 31(3): 274—295.
- [35] GONZALEZ J, MONTES C, RODRIGUEZ J, et al. Rethinking the Galapagos Islands as a complex social-ecological system: implications for conservation and management[J]. Ecology and Society, 2008, 13(2).
- [36] 李连刚, 张平宇, 关皓明, 等. 基于 Shift-Share 的辽宁老工业基地区域经济弹性特征分析[J]. 地理研究, 2019, 38(7): 1807—1819.
- [37] 陈肯界. 金融促进产业结构升级的区域效应比较研究: 以东莞市和株洲市为例[J]. 东莞理工学院学报, 2019, 26(4): 1—8.
- [38] 胡洁, 段晓峰, 李先杰, 等. 佛得角圣文森特岛旅游发展模式探讨[J]. 海洋经济, 2020, 10(3): 57—64.
- [39] 孙旖繁. 基于区块链技术的海岛文化旅游影响特性分析[J]. 现代制造技术与装备, 2018(5): 180—182.
- [40] 周正柱, 王俊龙. 长江经济带经济增长对工业环境污染的动态影响: 基于状态空间模型研究[J]. 南京工业大学学报(社会科学版), 2019, 18(4): 87—93+112.
- [41] CHENG X, SHUAI C, LIU J, et al. Topic modelling of ecology, environment and poverty nexus: an integrated framework [J]. Agriculture, ecosystems & environment, 2018, 267: 1—14.
- [42] CROUZEILLES R, FERREIRA M S, CHSZDON R L, et al. Ecological restoration success is higher for natural regeneration than for active restoration in tropical forests [J]. Science Advances, 2017, 3(11): e1701345.
- [43] PALMER M A, ZEDLER J B, Falk D A. Ecological theory and restoration ecology[M]. Washington, DC: Foundations of restoration ecology. Island Press, 2016: 3—26.
- [44] LOW Y, BICKSON D, GONZALEZ J, et al. Distributed GraphLab: a framework for machine learning and data mining in the cloud[J]. Proceedings of the VLDB Endowment, 2012, 5(8): 716—727.