

基于参数自适应算法的环塔里木盆地红枣优生区划分

金玉¹, 徐焕良¹, 梁丰志², 王江波², 王浩云^{1*}

(¹南京农业大学人工智能学院计算机系, 南京 210095; ²塔里木大学植物科学学院, 新疆阿拉尔 843300)

摘要:【目的】进一步优化新疆红枣优生区, 促进新疆红枣产业的提质增效。【方法】现有的优生区划分算法, 由于其参数设值受人为主观影响较大, 使优生区划分呈现较大的不确定性。以环塔里木盆地 19 个地区 33 个采样点的灰枣和骏枣的 12 个果实品质指标为数据集, 提出了参数自适应的红枣优生区划分算法。【结果】首先对枣样本品质数据集进行 300 次有放回的随机抽样, 对每一次抽取的样本进行参数自适应的主成分分析, 并将分析结果进行融合得到红枣的主要品质指标。在此基础上, 对各产区枣样本数据进行 150 次参数自适应的聚类分析, 根据每次聚类结果所对应的红枣主要品质指标, 利用无向加权图进行融合, 得到不同品质指标所对应的优生区划分结果。确定灰枣和骏枣的主成分, 在主要品质指标上将灰枣和骏枣优生区划分为 4 类。【结论】提出新疆环塔里木盆地各产区灰枣和骏枣的主要品质指标为总酸含量、总糖含量、单果质量和制干率, 为新疆红枣区域化发展提供了依据。

关键词: 红枣; 优生区; 果实品质; 参数自适应算法

中图分类号: S665.1

文献标志码: A

文章编号: 1009-9980(2022)01-0095-09

Optimal-adaptive zone division of Chinese jujube in Tarim Basin based on Parameter Adaptive Algorithm

JIN Yu¹, XU Huanliang¹, LIANG Fengzhi², WANG Jiangbo², WANG Haoyun^{1*}

(¹Department of Computer Science, College of Artificial Intelligence, Nanjing Agricultural University, Nanjing 210095, Jiangsu, China; ²Faculty of Plant Sciences, Tarim University, Alar 843300, Xinjiang, China)

Abstract: 【Objective】Xinjiang has a unique advantage in light and heat resources, and the fruit in Xinjiang has very high quality. In response to the national policy of vigorously developing the jujube industry, under the goal of “improving the quality and efficiency” of the jujube industry in Xinjiang, the optimal-adaptive zone division should be taken more into account to meet the demand for the jujube quality. The existing algorithm of optimal-adaptive zoning presents great uncertainty because the parameter values are greatly influenced by human subjectivity. 【Methods】Taking 12 fruit quality indexes of Huizao jujube and Junzao jujube from 33 sampling points in 19 regions around Tarim Basin as data set, a parameter adaptive algorithm was proposed to divide the optimal-adaptive zones of Chinese jujube. The parameter adaptive principal component analysis algorithm adopted the method of constant extraction and replacement to do m random sampling of n sample data sets. Due to the randomness, m data sets were different. This method can play the role in expanding the data set, avoiding the algorithm targeting only to fixed data sets. The experiment was set to select the principal components with a contribution rate of 50%–90% and a step size of 10% for each group of sampled data set. The ranking of the selected principal components’ contribution rate was scored to obtain the principal components of group M and the scores of each group of principal components. Finally, the comprehensive score of the principal components of group M was calculated to obtain the comprehensive selection results of princi-

收稿日期: 2021-07-01 接受日期: 2021-09-11

基金项目: 中央高校基本科研业务费专项(KYLH202006, KYZ201914); 南京农业大学-塔里木大学科研合作联合基金(NNLH202006); 新疆生产建设兵团南疆重点产业支撑计划(2017DB006); 国家自然科学基金(31601545)

作者简介: 金玉, 女, 在读硕士研究生, 研究方向: 作物表型参数测量。Tel: 15366030816, E-mail: 2209188349@qq.com

*通信作者 Author for correspondence. Tel: 13236598613, E-mail: wanghy@njau.edu.cn

pal components. Parameter adaptive clustering analysis algorithm was carried out to select quality index of dimension reduction after the data sampling of p , a sample data set, the K values for each group of sampling data set for 3, 4, 5, clustering analysis and calculation of each group of data similarity index, each group of data clustering results with the ideas of the undirected weighted graph clustering, and clustering adjacency matrix. Each vertex of the adjacency matrix was the sampling area, and the weight of the edge was the comprehensive similarity index of the index of the vertex area after clustering fusion. After that, the final result of the optimal-adaptive zone division was obtained by continuously adjusting the threshold and screening the clustering results. 【Results】 The principal components of jujube were determined as follows: vitamin C (Vc), total sugar, ratio of sugar to acid, total flavonoids, total acid, SSC drying rate and fruit weight. The results of the optimal zone division of jujube were obtained by parameter adaptive cluster analysis algorithm. The fruits of 12, 33, 47 and 48 clusters had high content of Vc, high content of total acid and sour taste. 12 Tuan, 3 Tuan, 41 Tuan, 42 Tuan, 47 Tuan, 48 Tuan, Luopu, Yutian and Ruoqiang had large horizontal and longitudinal diameters, large single fruit and large fruit shape. In 29 Tuan, 46 Tuan, 50 Tuan, Pishan, Ruoqiang, Yutian, Zepu and Luopu, the fruit drying rate was high; the fruits of 46 Tuan, 50 Tuan, Minfeng, Qimo and Shaya were high in Vc and total sugar contents and sweet in taste. The principal components of jujube were determined as follows: Vc, total sugar, ratio of sugar to acid, total flavonoids, total acid, SSC, drying rate and weight per fruit. The results of optimal region division of jujube were obtained by parameter adaptive cluster analysis algorithm. The fruits of 224 Tuan, 29 Tuan, 41 Tuan, 46 Tuan, 47 Tuan, Yutian Tuan and Zepu Tuan had large horizontal and vertical diameters, large single fruit and large fruit shape. The fruits of 12 Tuan, 3 Tuan, Luopu, Minfeng, Pishan and Yutian were high in total sugar content, high in Vc content and sweet in taste. The fruit drying rate of the 224 Tuan, 29 Tuan, 12 Tuan, 33 Tuan, 46 Tuan and 47 Tuan was high. In the 224 Tuan, 41 Tuan, 50 Tuan, Shaya and Zepu, fruits contained high total acid. In addition, the comprehensive fruit quality of Huizao jujube and Junzao jujube in some areas overlaid each other, and the fruit quality advantages of the two types of regions were also presented. The regional advantages of each region were pointed out, and the classification results were comprehensive and had practical guiding significance. 【Conclusion】 A parameter adaptive partitioning algorithm was proposed, which solved the fluctuation of experimental results and experimental errors caused by the traditional principal component analysis algorithm and clustering algorithm due to artificial parameters, small sample size and unstable clustering results. The main quality indexes of jujube were determined as total acid, total sugar, weight per fruit and drying rate, which provided a theoretical basis for the regional development of fine varieties of jujube in the Tarim Basin of Xinjiang.

Key words: Red jujube; Good production area; Fruit quality; Adaptive parameter adjustment algorithm

新疆是我国久负盛名的“瓜果之乡”，具有得天独厚的光热资源，是我国乃至世界优质果品基地^[1-2]。枣树耐旱、耐瘠薄，喜光热^[3]。新疆红枣产区光热资源丰富、干旱少雨，昼夜温差大，且 $10\text{ }^{\circ}\text{C}\geq$ 积温 $3500\text{ }^{\circ}\text{C}$ 以上^[4-5]，地势平坦，适宜机械化作业^[6-7]。骏枣、灰枣是新疆的区域特色主栽品种，品质优、商品性好；与原产地相比，总糖含量高、糖酸比高、果实品质更好^[8-10]。

现有算法中大多基于传统的主成分分析和聚类

算法，如袁辉等^[11]对新疆红枣品质的研究，利用主成分分析法选取前4个主成分并采用聚类算法人为设定欧氏距离阈值为10，将50批次不同产地和品种的红枣聚为3类；蔡海兰等^[12]对浮梁红茶品质研究，利用主成分分析法选取前3个主成分并通过系统聚类将样本分为2类；郭家刚等^[13]对蓝莓品质研究，利用主成分分析法选取前5个主成分并采用聚类算法人为设定欧氏距离阈值为15，将8个品种的蓝莓聚为4类。传统算法只针对固定样本集且受人为设定参数

影响极大,使得试验结果不具有稳定性和普适性。笔者采用对数据集不断抽样的方式,使算法不仅适用于固定样本集,同时能覆盖较大参数范围并对结果进行融合,避免算法结果对参数极大的依赖性,确保试验结果的稳定性和准确性,以更好地指导和优化新疆红枣产区的优生区地位。

1 材料和方法

1.1 灰枣和骏枣试验数据

确定具有代表性的18个地区县(区)为灰枣采样点,具有代表性的15个地区县(区)为骏枣采样点。塔里木大学王江波团队完成对数据集的采集^[14-15]。

18个灰枣种植园主要位于塔里木盆地(36°55'~41°50'N、76°06'~88°09'E)周边,海拔856~1410 m,年日照时数北部约3 000 h,南部不到3 000 h,多风沙和浮尘天气。年均温9~11 °C,南部略高于北部。大陆性由西向东加强,冬季东部比西部冷。盆地地貌呈环状分布,边缘是与山地连接的砾石戈壁,中心是辽阔沙漠,边缘和沙漠间是冲积扇和冲积平原,并有绿洲分布,地质构造上是周围被许多深大断裂所限制的稳定地块,盆地地势西高东低,微向北倾,盆地水分主要来自西风气流,盆地降水稀少,盆地本身无法形成径流^[6]。

18个灰枣采样点分别为环塔盆地东部采样点:新疆生产建设兵团29团(29T)、33团(33T)、且末(QM)和若羌(RQ);环塔盆地西部采样点:新疆生产建设兵团42团(42T)、41团(41T)、48团(48T)、46团(46T)和泽普(ZP);环塔盆地南部采样点:新疆生产建设兵团47团(47T)、皮山农场(PS)、洛浦(LP)、于田(YT)和民丰(MF);环塔盆地北部采样点:新疆生产建设兵团12团(12T)、3团(3T)、50团(50T)和沙雅(SY)。

15个骏枣采样点分别为环塔里木盆地东部采样点:新疆生产建设兵团29团(29T)、33团(33T);环塔里木盆地西部采样点:新疆生产建设兵团41团(41T)、46团(46T)、泽普(ZP);环塔里木盆地南部采样点:新疆生产建设兵团47团(47T)、224团(224T)、皮山农场(PS)、洛浦(LP)、于田(YT)、民丰(MF);环塔里木盆地北部采样点:新疆生产建设兵团12团(12T)、3团(3T)、50团(50T)、沙雅(SY)。

所研究灰枣和骏枣的果实内在和外在品质指标

相同。果实内在品质指标为可溶性固形物含量、总糖含量、总酸含量、糖酸比、总黄酮含量、维生素C(Vc)含量和蛋白质含量;果实外在品质指标为纵径、横径、单果质量、可食率及制干率。

1.2 参数自适应优生区划分算法

采用参数自适应的优生区划分算法,包括参数自适应的主成分分析和聚类分析算法。该算法分为两部分,先是通过参数自适应的主成分分析算法选取主成分,然后对降维的结果通过参数自适应的聚类分析算法对环塔里木盆地红枣产区进行优生区划分。此算法笔者在文中首次提出,输入数据应满足高斯分布,可弥补传统算法受限于样本量、人工确定阈值、对初始的簇中心敏感等缺点。

参数自适应的主成分分析算法流程图如图1所示,算法框架图如图2所示,采取不断抽取放回方式,对 n 个样本数据集做 m 次的随机采样^[17-18],由于随机性, m 个数据集各不相同,此方式可起到扩充数据集的作用,同时避免算法只针对固定数据集。试验设定对每组抽样数据集选择贡献率在50%~90%的主成分且步长为10%,对选取的各主成分贡献率排序情况进行打分,共得到 m 组主成分及各组主成分得分,最后对 m 组主成分计算综合得分得到主成分综合选择结果。

参数自适应的聚类分析算法流程图如图3所示,算法框架图如图4所示,对选取的品质指标降维后的数据进行 p 次抽样,得到 p 个抽样数据集,对每组抽样数据集设定 K 值为3、4、5进行聚类分析,并计算每组数据的相似度指数,将每组数据的聚类结果以无向加权图作聚类融合,得到聚类邻接矩阵,此矩阵的每个顶点为采样地区,边的权值为聚类融合后顶点地区指标具有的综合相似度指数。之后通过不断调整优生区划分阈值,对聚类结果进行筛选,得到最终的优生区划分结果。

2 结果与分析

2.1 参数自适应的主成分分析算法

2.1.1 参数自适应的灰枣主成分选取 通过传统的PCA算法可得前4个主成分累计方法贡献率约为80%,较大地保留了原始信息,故选取了前4个主成分作主成分分析。所选取的主成分分别为横径、纵径、总酸含量、单果质量、总黄酮含量、总糖含量、Vc含量和糖酸比。

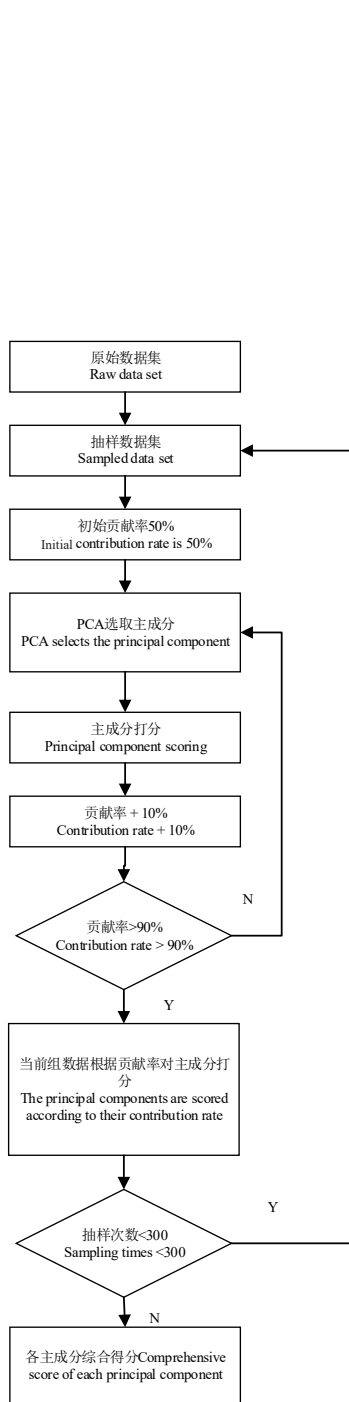


图 1 参数自适应的主成分分析流程图
Fig. 1 Flow chart of parameter adaptive principal component analysis algorithm

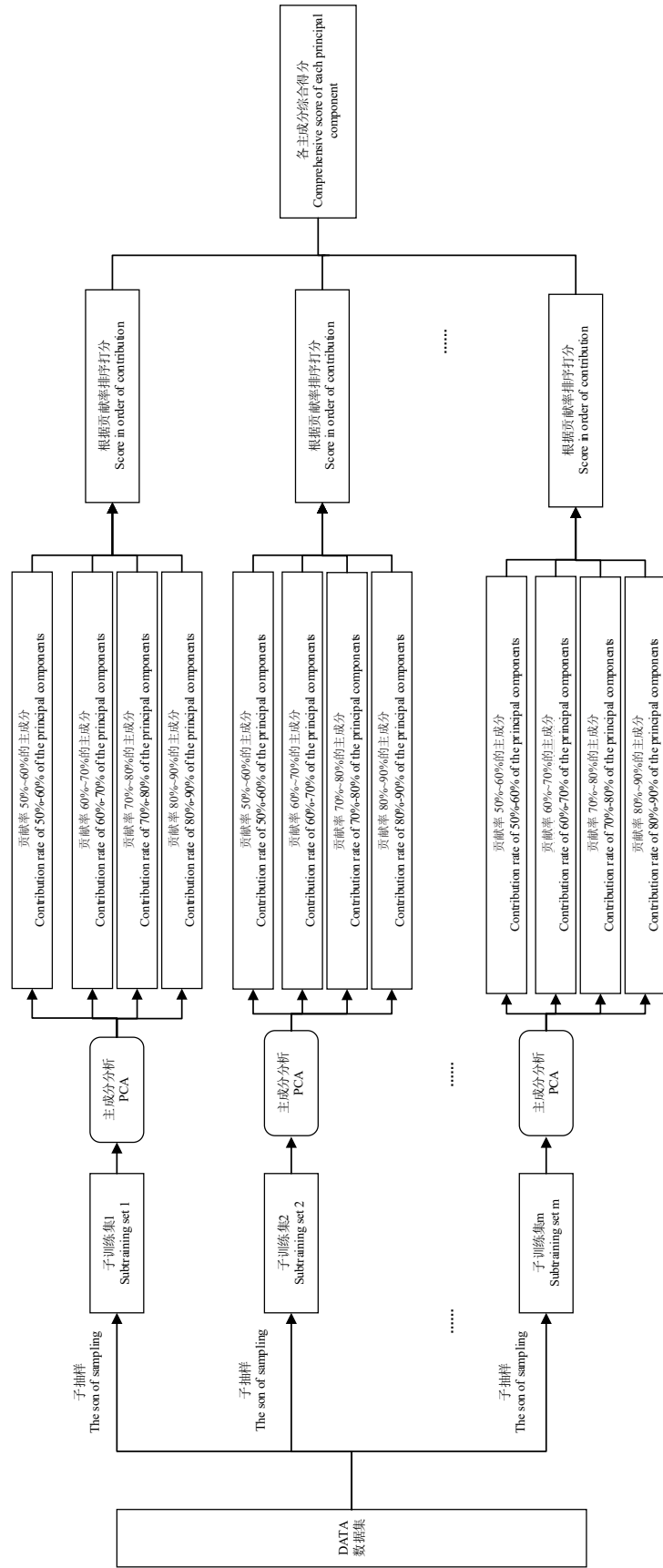


图 2 参数自适应的主成分分析框架图
Fig. 2 Frame diagram of parameter adaptive principal component analysis

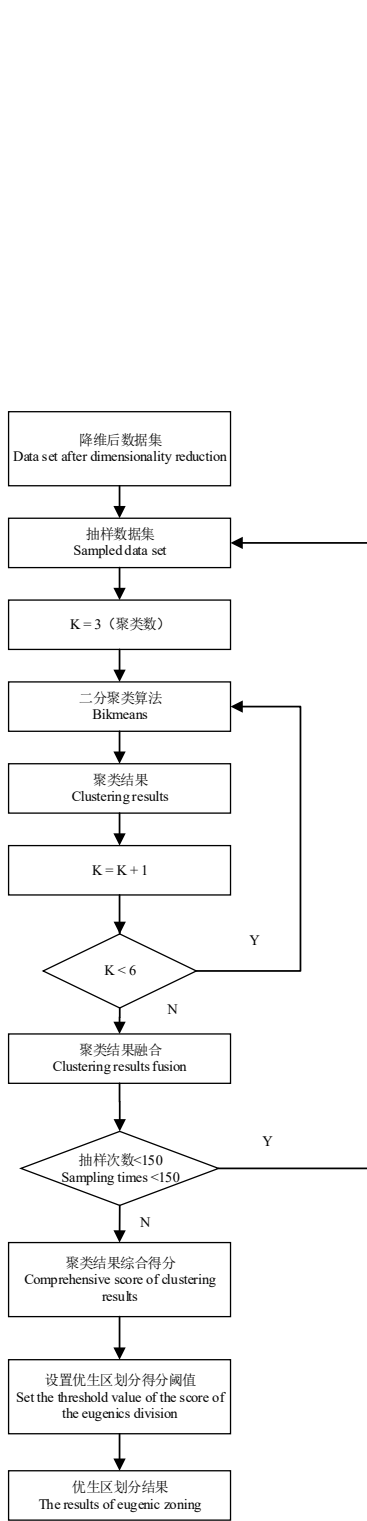


图 3 参数自适应的聚类分析流程图
Fig. 3 Flowchart of parameter adaptive clustering analysis algorithm

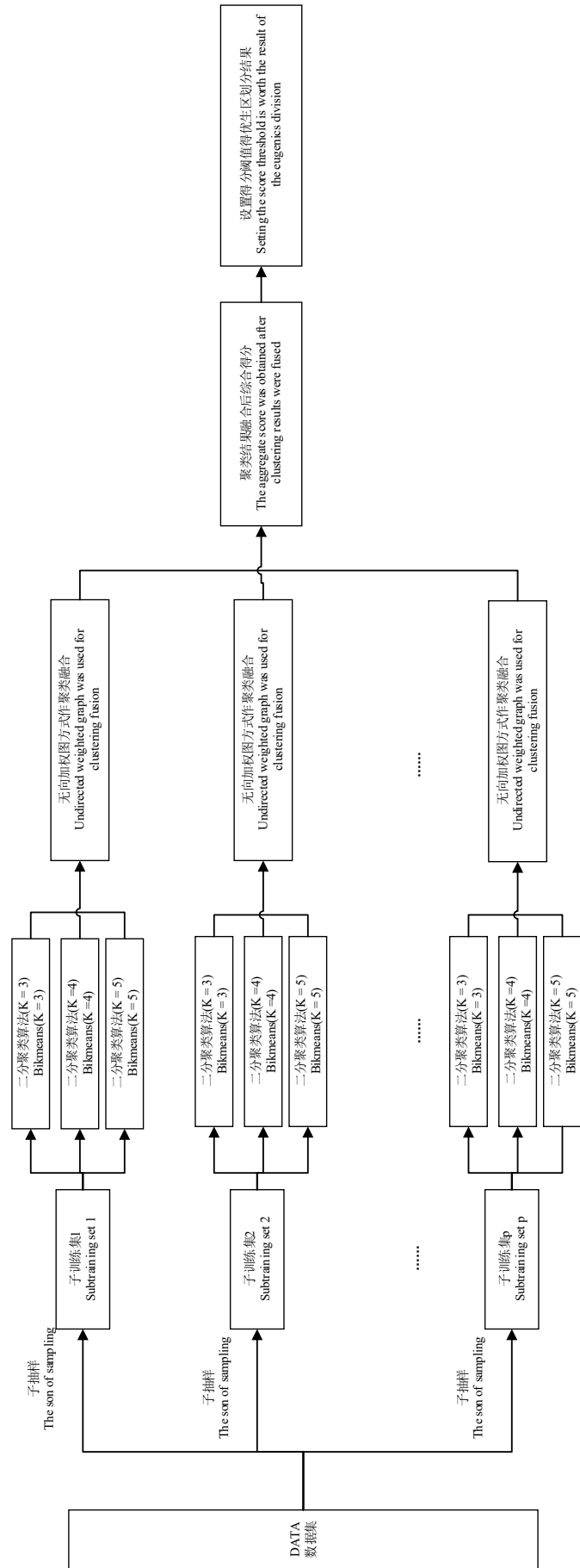


图 4 参数自适应的聚类分析框架图
Fig. 4 Frame diagram of parameter adaptive cluster analysis

从环塔里木盆地18个采样点灰枣54组数据中选取主成分,通过300次有放回的抽样,每次抽取30组数据,选取贡献率在50%~90%的主成分,最后将每次抽样选取的主成分累加,得表1,表中数据为各主成分综合得分。

表1 12个灰枣成分指标选取次数
Table 1 The selection times of the composition indexes of 12 gray jujube

果实指标 Index of fruit	综合得分 Composite scores
维生素C含量 Vitamin C content	293
可溶性固形物含量 Soluble solids content	241
蛋白质含量 Protein content	112
制干率 Dried rate	147
纵径 The longitudinal diameter	87
横径 The transverse diameter	90
单果质量 Single fruit weight	143
可食率 Edible rate	120
总酸含量 Total acid content	246
总糖含量 Total sugar content	278
糖酸比 Sugar acid ratio	262
总黄酮含量 Total flavonoids content	255

由表1可知,试验选取的主成分分别为Vc含量、总糖含量、糖酸比、总黄酮含量、总酸含量、可溶性固形物含量、制干率和单果质量,选取的主成分外在品质和内在品质,能够较为全面地保留原有数据的全部信息。

2.1.2 参数自适应的骏枣主成分选取 同2.1.1,所选取的主成分为横径、纵径、单果质量、总酸含量、糖酸比、总糖含量、可溶性固形物含量和制干率。

从环塔里木盆地15个采样点的45组数据中选取主成分,通过300次有放回的抽样,选取贡献率在50%~90%的主成分,将每次抽样选取的主成分进行叠加,得表2,表中数据为各主成分的综合得分。

由表2可知,试验选取的主成分为单果质量、制干率、总酸含量、可溶性固形物含量、Vc含量、总糖含量、糖酸比和总黄酮含量,同上,选取的主成分能够较为全面地保留原有数据的全部信息。

2.2 参数自适应的聚类分析算法

2.2.1 参数自适应的灰枣产区聚类 采用无向加权图,将18个采样点作为无向图的顶点,进行150次聚类分析,每两个地区相似度指数为边的权值,得表3。通过表3,可将环塔里木盆地灰枣采样点的18个地区聚为4类,聚类结果如表4所示。

由表4可得,第一类地区包括12团、33团、47

表2 12个骏枣成分指标选取次数

Table 2 The selection times of 12 components indexes of jujube

果实指标 Index of fruit	综合得分 Composite scores
维生素C含量 Vitamin C content	220
可溶性固形物含量 Soluble solids content	228
蛋白质含量 Protein content	87
制干率 Dried rate	252
纵径 The longitudinal diameter	97
横径 The transverse diameter	139
单果质量 Single fruit weight	278
可食率 Edible rate	157
总酸含量 Total acid content	240
总糖含量 Total sugar content	219
糖酸比 Sugar acid ratio	193
总黄酮含量 Total flavonoids content	177

团、48团,其果实Vc含量和总酸含量高,口感偏酸;第二类地区包括12团、3团、41团、42团、47团、48团、洛浦、于田、若羌,其果实横纵径和单果质量大,果实较大;第三类地区包括29团、46团、50团、皮山、若羌、于田、泽普、洛浦,其果实制干率高;第四类地区包括46团、50团、民丰、且末、沙雅,其果实Vc和总糖含量高,果实甜。同时,可得部分地区分属两类,12团、47团、48团既属第一类地区,也属第二类地区,即营养价值高且果实大;洛浦、于田、若羌既属第二类地区,也属第三类地区,即单果质量大且制干率高;46团、50团既属第三类地区,也属第四类地区,即Vc、总糖含量和制干率都高。发现第一类地区和第四类地区无重叠,则说明果实口感酸度和甜度具有独立性,符合现实口感规律。

参数自适应的主成分分析和聚类分析算法得到优生区划分结果,根据表4可得,灰枣优生区大部分地区有聚类趋势。

2.2.2 参数自适应的骏枣产区聚类 采用无向加权图,将15个地区作为无向图的顶点进行150次迭代聚类分析,每两个地区相似度指数为边的权值,可得表5。

通过表5可将环塔里木盆地骏枣采样点的15个地区分为4类,聚类结果如表6。

表6可知,第一类地区包括224团、29团、41团、46团、47团、于田、泽普,其果实横纵径大,单果质量大,果实果型较大;第二类地区包括12团、3团、洛浦、民丰、皮山、于田,其果实总糖含量高,Vc含量高,口感偏甜;第三类地区包括224团、29团、12团、

表3 灰枣各地区聚类次数
Table 3 Cluster times of Huizao region

项目 Item	12团 12T	29团 29T	33团 33T	3团 3T	41团 41T	42团 42T	46团 46T	47团 47T	48团 48T	50团 50T	洛浦 LP	民丰 MF	皮山 PS	且末 QM	若羌 RQ	沙雅 SY	于田 YT	泽普 ZP	
12团 12T	0																		
29团 29T	87	0																	
33团 33T	180	15	0																
3团 3T	267	85	39	0															
41团 41T	240	21	36	384	0														
42团 42T	60	219	0	150	216	0													
46团 46T	165	153	303	73	45	54	0												
47团 47T	246	105	180	267	222	54	171	0											
48团 48T	257	33	252	222	213	63	189	230	0										
50团 50T	158	145	318	52	30	69	259	156	188	0									
洛浦 LP	200	54	24	339	405	261	37	188	174	32	0								
民丰 MF	148	5	414	13	12	0	271	148	222	290	8	0							
皮山 PS	123	309	150	79	30	138	211	138	106	202	44	130	0						
且末 QM	180	15	450	39	36	0	303	180	252	318	24	414	150	0					
若羌 RQ	158	355	54	148	84	138	165	160	94	154	89	34	274	54	0				
沙雅 SY	169	162	291	80	48	51	273	175	188	254	38	259	221	291	176	0			
于田 YT	93	438	18	93	27	207	159	114	39	150	54	6	318	18	366	168	0		
泽普 ZP	154	356	27	175	111	153	146	169	87	129	114	9	265	27	326	157	366	0	

表4 灰枣聚类结果
Table 4 Clustering results of Huizao

类别 Class	地区 Region	特点 Characteristic
1	12团、33团、47团、48团 12T,33T,47T,48T	Vc含量高;总酸含量高 High vitamin C content; high total acid content
2	12团、3团、41团、47团、48团、42团、洛浦、于田、若羌 12T,3T,41T,47T,48T,42T,LP,YT,RQ	单果质量大;横纵径较高 Single fruit is significant; high horizontal and vertical diameter
3	29团、皮山、若羌、于田、泽普、46团、50团、洛浦 29T,PS,RQ,YT,ZP,46T,50T,LP	制干率高 High dried rate
4	46团、50团、民丰、且末、沙雅 46T,50T,MF,QM,SY	总糖含量高;Vc含量高 High total sugar content; high vitamin C content

33团、46团、47团,其果实制干率高;第四类地区包括224团、41团、50团、沙雅、泽普,其果实总酸含量高。同时,可得部分地区分属两类,224团、41团、46团、47团既属第一类地区,也属第三类地区,即制干率高,果实大;于田既属第一类地区,也属第二类地区,即单果质量大,总糖和Vc含量高;泽普既属第一类地区,也属第四类地区,即单果质量大,总酸含量

高;沙雅既属第三类地区,也属第四类地区,即总酸含量和制干率都高。也发现第二类和第四类地区无重叠,说明果实口感酸度和甜度具有独立性,符合现实口感规律。

参数自适应的主成分分析和聚类分析得到优生区划分结果,根据表6可得骏枣优生区大部分地区有聚类趋势。

表 5 骏枣各地区聚类次数

Table 5 Clustering times of Junzao region

地区 Region	12团 12T	224团 224T	29团 29T	33团 33T	3团 3T	41团 41T	46团 46T	47团 47T	50团 50T	洛浦 LP	民丰 MF	皮山 PS	沙雅 SY	于田 YT	泽普 ZP
12团 12T	0														
224团 224T	0	0													
29团 29T	0	324	0												
33团 33T	52	12	0	0											
3团 3T	288	0	0	3	0										
41团 41T	96	216	216	63	42	0									
46团 46T	0	300	382	0	0	216	0								
47团 47T	0	306	384	0	0	216	444	0							
50团 50T	68	162	101	147	18	144	58	60	0						
洛浦 LP	308	0	0	21	372	65	0	0	22	0					
民丰 MF	311	0	0	29	369	69	0	0	42	340	0				
皮山 PS	291	18	22	17	351	78	26	24	33	348	320	0			
沙雅 SY	192	96	68	128	84	158	46	48	191	130	138	133	0		
于田 YT	96	201	255	1	150	158	297	297	45	124	123	134	59	0	
泽普 ZP	0	316	288	36	0	202	264	270	198	0	0	18	108	177	0

表 6 骏枣聚类结果

Table 6 Clustering result of Junzao

类别 Class	地区 Region	特点 Characteristic
1	224团、29团、41团、47团、于田、泽普 224T, 29T, 41T, 47T, YT, ZP	单果质量大; 横纵径较高 Single fruit is significant; high horizontal and vertical diameter
2	12团、3团、洛浦、民丰、皮山、于田 12T, 3T, LP, MF, PS, YT	总糖含量高; Vc含量高 High total sugar content; high vitamin C content
3	224团、29团、12团、33团、46团、47团 224T, 29T, 12T, 33T, 46T, 47T	制干率高 High dried rate
4	224团、41团、50团、沙雅、泽普 224T, 41T, 50T, SY, ZP	总酸含量高 High total acid content

3 讨 论

3.1 主成分分析

在前人的研究中,果实品质主成分的选取主要是基于传统的主成分分析方法,传统PCA选取的主成分会随贡献率变化呈现一定的波动性。传统PCA选取的灰枣主成分中同时选取横纵径和单果质量,未选取蛋白质含量、制干率、可溶性固形物含量、可食率,由灰枣品质数据的相关系数矩阵可得,单果质量与横纵径的相关系数大于0.75,与可食率的相关系数大于0.6,相关性较强;制干率与蛋白质含量、可溶性固形物含量的相关系数大于0.5,相关性较强。

骏枣主成分同时选取横纵径和单果质量,未选取Vc含量、总黄酮含量、可食率、蛋白质含量,由骏枣品质数据的相关系数矩阵可得单果质量与横纵径的相关系数大于0.8,与可食率的相关系数大于0.6,相关性较强;Vc、总黄酮含量与其他指标的相关系数小于0.5,相关性较弱;蛋白质含量与制干率的相

关系数大于0.5,相关性较强。

通过上述分析发现,灰枣和骏枣共同选取的主成分为Vc含量、总糖含量、糖酸比、总酸含量、可溶性固形物含量、制干率、单果质量和总黄酮含量,可得灰枣和骏枣选取的主成分一致。使用PCA方法降维处理后,仍可保留原始数据集的大部分变量特征值。本文中灰枣和骏枣未选取的主成分有横纵径、蛋白质含量和可食率,由相关系数矩阵可得单果质量与横纵径、可食率相关性较强,蛋白质含量与制干率相关性较强,试验证明改进后的PCA方法寻找主成分更为谨慎,对特征值描述更加清晰。

3.2 优生区划分

将传统聚类算法的聚类结果与参数自适应的聚类分析算法结果对比,传统聚类算法的聚类结果中各类别地区无明显地区优势,且在综合得分评价上也无聚类趋势。通过设定K值只能将某个地区划分到某一类别,无法得到某地区同时具备多类地区优势的更加全面的结果。

参数自适应的聚类分析算法不需要人为规定K值且聚类结果稳定,笔者将环塔里木盆地灰枣和骏枣采样区划分为4类,分别在总糖含量、总酸含量、制干率和单果质量上表现优异,有针对性地指出各地区的优势。同时,对灰枣和骏枣优生区划分对比发现,灰枣在单果质量上表现优异的地区与骏枣在总糖含量上表现优异的地区有较大重合,灰枣在制干率上表现优异的地区与骏枣在单果质量上表现优异的地区有较大重合,灰枣在总酸含量上表现优异的地区与骏枣在制干率上表现优异的地区有较大重合。可得在不同地区,灰枣和骏枣在某一指标特征上表现的优异性不同。

4 结 论

参数自适应的优生区划分,首先利用参数自适应的主成分分析确定灰枣和骏枣的主成分,然后通过参数自适应的聚类分析得到灰枣和骏枣的主要品质指标为总酸含量、总糖含量、单果质量和制干率。且部分地区灰枣、骏枣存在果实综合品质重叠现象,同时具有两类地区的果实品质特征,有针对性地指出每个地区的优势,划分结果全面且具有现实指导意义。

参考文献 References:

- [1] 刘聪. 新疆红枣的功能性成分及加工工艺的研究[D]. 杭州:浙江大学,2014.
LIU Cong. Studies on functional components and processing technology of Xinjiang jujube dates[D]. Hangzhou: Zhejiang University, 2014.
- [2] 王谦. 新疆农业气候资源的分析与评价[J]. 干旱地区农业研究, 1992, 10(2): 91-97.
WANG Qian. Analysis and evaluation of agricultural climate resources in Xinjiang[J]. Agricultural Research in the Arid Areas, 1992, 10(2): 91-97.
- [3] 漆联全. 新疆红枣产业的现状、要求及其发展趋势[J]. 新疆农业科学, 2010, 47(S2): 8-12.
QI Lianquan. The present situation, requirements and development trends of Xinjiang red date industry[J]. Xinjiang Agricultural Science, 2010, 47(S2): 8-12.
- [4] 普宗朝, 张山清, 李景林, 王命全. 近 50 a 新疆 $\geq 0^{\circ}\text{C}$ 持续日数和积温时空变化[J]. 干旱区研究, 2013, 30(5): 781-788.
PU Zongchao, ZHANG Shanqing, LI Jinglin, WANG Mingquan. Spatiotemporal change of duration and accumulated temperature of temperature $\geq 0^{\circ}\text{C}$ in Xinjiang in recent 50 years[J]. Arid Zone Research, 2013, 30(5): 781-788.
- [5] 管岑澜. 新疆红枣发展现状与对策研究[D]. 武汉:华中农业大学, 2010.
GUAN Qinlan. Study on the potential and strategy to develop Jujube in Xinjiang province[D]. Wuhan: Huazhong Agricultural University, 2010.
- [6] 张山清, 普宗朝, 李景林, 李新建, 吉春容, 傅玮东, 王命全, 李晓菊, 郑新倩. 气候变化对新疆红枣种植气候区划的影响[J]. 中国生态农业学报, 2014, 22(6): 713-721.
ZHANG Shanqing, PU Zongchao, LI Jinglin, LI Xinjian, JI Chunrong, FU Weidong, WANG Mingquan, LI Xiaoju, ZHENG Xinqian. Impact of climate change on the division of jujube planting zones in Xinjiang[J]. Chinese Journal of Eco-Agriculture, 2014, 22(6): 713-721.
- [7] 周丽, 杨伟志, 王长柱, 李新岗. 新疆红枣优生区研究[J]. 果树学报, 2015, 32(3): 453-459.
ZHOU Li, YANG Weizhi, WANG Changzhu, LI Xingang. Superior production region of Chinese jujube in Xinjiang[J]. Journal of Fruit Science, 2015, 32(3): 453-459.
- [8] 鲁周民, 刘坤, 闫忠心, 李新岗. 枣果实营养成分及保健作用研究进展[J]. 园艺学报, 2010, 37(12): 2017-2024.
LU Zhoumin, LIU Kun, YAN Zhongxin, LI Xingang. Research status of nutrient component and health functions of *Ziziphus jujuba* Mill. [J]. Acta Horticulturae Sinica, 2010, 37(12): 2017-2024.
- [9] 付明鑫, 张磊, 曾胜河, 高志建. 新疆兵团农田土壤的钾素养分管理[J]. 新疆农业科学, 2010, 47(2): 340-344.
FU Mingxin, ZHANG Lei, ZENG Shenghe, GAO Zhijian. The management of soil potassium nutrient in Xinjiang production and construction corps[J]. Xinjiang Agricultural Sciences, 2010, 47(2): 340-344.
- [10] 刘杰超, 刘慧, 吕真真, 张春岭, 焦中高. 不同新疆红枣营养成分比较分析[J]. 中国食物与营养, 2018, 24(4): 31-35.
LIU Jiechao, LIU Hui, LÜ Zhenzhen, ZHANG Chunling, JIAO Zhonggao. Comparison of nutritional composition of jujube from Xinjiang province[J]. Food and Nutrition in China, 2018, 24(4): 31-35.
- [11] 袁辉, 王建玲, 远辉. 基于主成分分析和聚类分析对新疆红枣的品质评价[J]. 食品工业, 2020, 41(9): 305-309.
YUAN Hui, WANG Jianling, YUAN Hui. Quality evaluation of Xinjiang jujube based on principal component analysis and cluster analysis[J]. The Food Industry, 2020, 41(9): 305-309.
- [12] 蔡海兰, 杨普香, 朱凤新, 鲍润元, 李琛. 基于主成分和聚类分析的浮梁红茶品质评价[J]. 食品研究与开发, 2019, 40(16): 100-104.
CAI Hailan, YANG Puxiang, ZHU Fengxin, BAO Runyuan, LI Chen. Comprehensive evaluation of Fuliang black tea quality based on principal component and cluster analysis[J]. Food Research and Development, 2019, 40(16): 100-104.
- [13] 郭家刚, 杨松, 伍玉茜, 陈敏. 基于主成分与聚类分析的蓝莓品质综合评价研究[J]. 食品研究与开发, 2020, 41(12): 53-60.
GUO Jiagang, YANG Song, WU Yuhuan, CHEN Min. Comprehensive evaluation of blueberry quality based on principal component and cluster analysis[J]. Food Research and Development, 2020, 41(12): 53-60.
- [14] 梁丰志. 环塔里木盆地灰枣及骏枣果实品质分析及优势产区划分[D]. 塔里木: 塔里木大学, 2021.
LIANG Fengzhi. Fruit quality analysis and dominant zone division of 'Huizao' and 'Junzao' around tarim basin[D]. Tarim: Tarim University, 2021.
- [15] 梁丰志, 童盼盼, 张亚若, 金玉, 王浩云, 吴翠云, 王江波. 新疆南疆不同灰枣产区果实品质分析及优生区划分[J]. 华中农业大学学报, 2021, 40(4): 123-132.
LIANG Fengzhi, TONG Panpan, ZHANG Yaru, JIN Yu, WANG Haoyun, WU Cuiyun, WANG Jiangbo. Analysis of fruit quality and division of eugenic zone in different gray jujube producing regions in southern Xinjiang[J]. Journal of Huazhong Agricultural University, 2021, 40(4): 123-132.
- [16] 李硕, 沈彦俊. 气候变暖对西北干旱区农业热量资源变化的影响[J]. 中国生态农业学报, 2013, 21(2): 227-235.
LI Shuo, SHEN Yanjun. Impact of climate warming on temperature and heat resource in arid northwest China[J]. Chinese Journal of Eco-Agriculture, 2013, 21(2): 227-235.
- [17] 王奕森, 夏树涛. 集成学习之随机森林算法综述[J]. 信息技术, 2018, 12(1): 49-55.
WANG Yisen, XIA Shutao. A survey of random forests algorithms[J]. Information and Communications Technologies, 2018, 12(1): 49-55.
- [18] WANG Y C, CHENG C H. A multiple combined method for rebalancing medical data with class imbalances[J]. Computers in Biology and Medicine, 2021, 134(1): 104527.