

## 基于分类回归树模型的奉节脐橙品质影响因素研究

王佳彬, 张云逸, 李瑜\*, 鲍丽然 (重庆市地质矿产勘查开发局川东南地质大队, 重庆 400038)

**摘要** 为明确土壤特性(pH、有机碳、氮、磷、钾、钙、镁、硫、全铁、锰、硼、钼、铜、锌)对奉节脐橙品质(可溶性固形物、横径)的影响,引入分类回归树模型(CART),将土壤特性作为自变量,脐橙品质作为因变量进行建模分析。结果表明,锌是影响脐橙可溶性固形物最关键的因素,而氮是影响脐橙横径最关键的因素。pH、钙与果实横径达显著正相关,相关系数分别为0.346和0.400。CART模型可以很好地用来研究土壤特性对脐橙品质的影响,就可溶性固形物树模型而言, $R^2$ 、RMSE、MAE分别为0.85、0.46、0.32;就横径树模型而言, $R^2$ 、RMSE、MAE分别为0.89、1.58、1.07。

**关键词** CART;土壤特性;可溶性固形物;横径

中图分类号 S666.4 文献标识码 A

文章编号 0517-6611(2022)06-0135-05

doi: 10.3969/j.issn.0517-6611.2022.06.032



开放科学(资源服务)标识码(OSID):

### Research on Factors Influencing the Quality of Fengjie Navel Orange Based on Classification and Regression Tree Model

WANG Jia-bin, ZHANG Yun-yi, LI Yu et al (Southeast Sichuan Geological Team, Chongqing Bureau of Geology and Minerals Exploration, Chongqing 400038)

**Abstract** In order to clarify the influence of soil characteristics (pH, organic carbon, nitrogen, phosphorus, potassium, calcium, magnesium, sulfur, total iron, manganese, boron, molybdenum, copper, zinc) on the quality (soluble solids, horizontal diameter) of Fengjie navel orange, the classification and regression tree model (CART) was introduced, and the soil characteristics were used as independent variables, and the quality of navel oranges was mostly used as the dependent variable for modeling analysis. The results showed that zinc was the most critical factor affecting the soluble solids of navel oranges, and nitrogen was the most critical factor affecting the horizontal diameter of navel oranges. pH, calcium and fruit transverse diameter reached a significant positive correlation, the correlation coefficients were 0.346 and 0.400, respectively. The CART model could be used to study the influence of soil characteristics on the quality of navel oranges. For the soluble solids tree model,  $R^2$ , RMSE, and MAE were 0.85, 0.46, and 0.32, respectively; for the transverse diameter tree model,  $R^2$ , RMSE and MAE were 0.89, 1.58 and 1.07, respectively.

**Key words** CART; Soil characteristics; Soluble solids; Transverse diameter

奉节拥有悠久的脐橙种植历史,自1953年从江津果树所引进凤园脐橙,经过不断选育新品种,目前已经拥有凤早、凤园、凤晚、班菲尔、纽荷尔、福本、鲍威尔、切斯勒特等多个品种。2006年2月经国家工商总局商标局批准正式使用“奉节脐橙·FJQCH”地理标志证明商标。2019年,脐橙种植面积已达1.5万 $\text{hm}^2$ ,年产量32.47万 $\text{t}^{[1]}$ 。果实品质是外观品质、风味品质、营养品质及加工品质等多种因素的复合体,包括果实形状、大小、果皮色泽、可溶性固形物、有机酸等要素<sup>[2]</sup>。在脐橙众多的品质指标中,可溶性固形物含量和横径是其中最为关键的指标。可溶性固形物不仅是果实成熟度的重要标识,同时也是决定果实口味及营养的关键因素<sup>[3-6]</sup>。横径是《GB 12947—2008 鲜柑橘》中划分果实级别的唯一外在指标。随着人们生活水平的提高,人们对柑橘品质的要求也越来越明确具体,因此有效控制果实品质显得十分重要。果实品质除受降雨、日照、温度等气候因素的影响外<sup>[7-9]</sup>,主要受到土壤特性的影响<sup>[10-16]</sup>。由于脐橙属于多年生植物,其土壤中相关元素的含量会长期影响脐橙的生长,因此研究土壤特性对品质的影响对于果园管理及品质把控具有重要意义。

## 1 材料与方法

### 1.1 研究区概况

**基金项目** 重庆市规划和自然资源局地质调查项目(渝规资[2019]128号)。  
**作者简介** 王佳彬(1992—),男,四川简阳人,工程师,从事土地质量地质调查、地球化学勘查研究。\*通信作者,高级工程师,从事土地质量调查研究。

**收稿日期** 2021-07-07; **修回日期** 2021-07-26

109°1'17"~109°45'58"E, 30°29'19"~31°22'33"N,东邻巫山、南靠恩施、北接巫溪、西连云阳,全域面积4 098.44  $\text{km}^2$ 。奉节属于四川盆地山地地貌,整体呈东南、东北高而中西部稍平缓。境内水系主要有长江(长约41.5 km)、梅溪河、大溪河、石笋河、草堂河、朱衣河等(图1)。气候属于典型的亚热带湿润季风气候,具有无霜期长、雨量充沛、日照时间长、四季分明等特点。奉节因其得天独厚的气候条件,具备脐橙生长的“无台风、无冻害、无检疫性病虫害”三大生态优势,成为我国主要的脐橙生产基地,曾被中国果品流通协会特别授予“中国橙都”称号。

**1.2 样品采集与分析** 在研究区范围内,共采集根系土壤及脐橙样品33件,采样点分布情况见图2。

**根系土:**在树冠滴水线附近采样,避开施肥、滴灌头湿润区,挖取1个采样穴,在5~40 cm土层中均匀采样,每10株的土壤等量混合均匀为一个样品,用四分法将取土量缩分至1 kg左右,并送实验室进行检测。

**脐橙果实:**每个采样点选择3~5棵植株,按树冠上中部不同方向随机采摘果实40~50个,组成一件分析样品,样品重量为5 kg,装入样袋及时送中国农业科学院柑橘研究所进行检测。

土壤主要分析测试了pH、有机碳、氮、磷、钾、钙、镁、硫、全铁、锰、硼、钼、铜、锌14个指标。

脐橙主要分析测试了可溶性固形物含量以及果实横径2个指标。

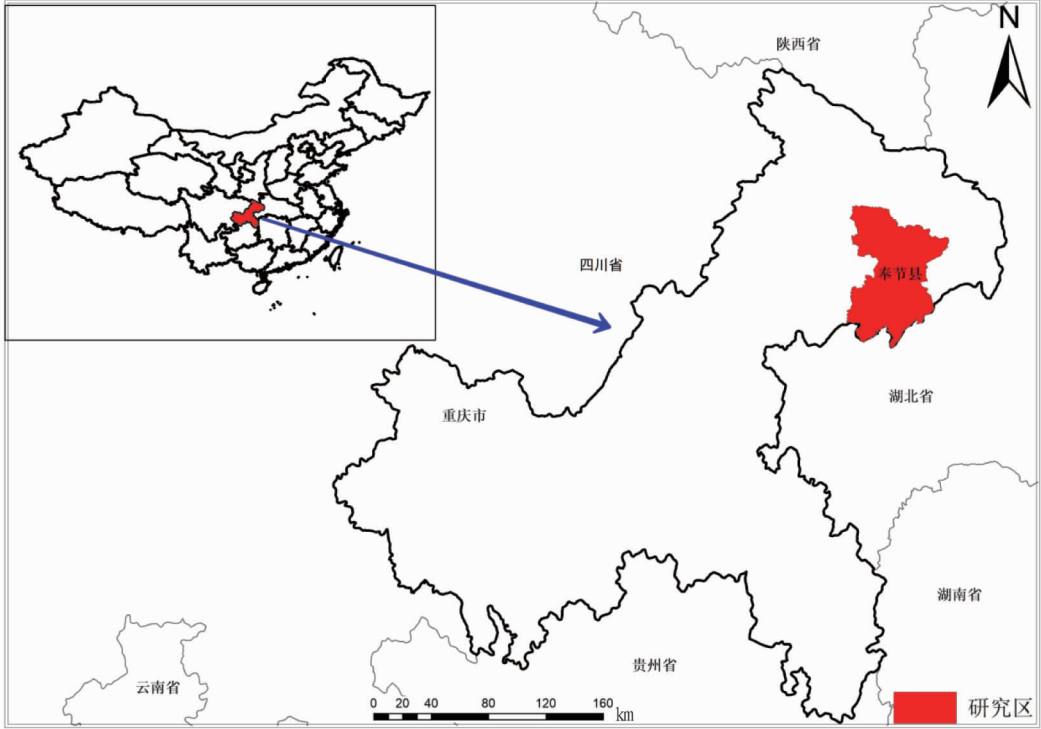


图1 研究区位置图

Fig.1 Location map of the study area



图2 采样点分布

Fig.2 Distribution of sampling points

**1.3 分类回归树模型** 分类与回归树模型(classification and regression tree, CART)是一种典型的监督学习算法,最早由Breiman、Friedman、Olshen和Stone于1984年提出。其基本工作原理是对由测试标量(自变量)和目标变量(因变量)构成的训练集进行循环二分形成二叉树结构。CART树在生长

过程中,主要采用GINI系数作为节点的分裂依据。该算法既可用于分类,也可以用于回归,当自变量是离散型数据时成为分类树,当自变量是连续型变量时成为回归树。CART模型因其具有结构清晰、实现简单、运行速度快、可以处理非线性关系、准确性高等特点被广泛应用于土壤及遥感等领域

域<sup>[17-19]</sup>。分类回归树模型的构建在软件 SPSS v.25.0 中完成。其他相关数据统计在 Excel 2007 中完成。

## 2 结果与分析

**2.1 基本统计** 土壤有机碳含量为 0.25%~2.90%，平均值为 1.26%。土壤 pH 为 4.22~8.37，平均值为 7.47。大量元素氮、磷、钾的含量分别为 568.00~4 281.00 mg/kg、304.00~5 056.00 mg/kg、1.37%~3.27%，平均值分别为 1 567.73 mg/kg、

1 428.45 mg/kg、2.31%。中量元素钙、镁、硫的含量分别为 0.26%~18.87%、0.51%~3.76%、0.01~0.13%，平均值分别为 5.92%、1.15%、0.03%。微量元素铁(全铁)、锰、铜、锌、硼、钼含量分别为 2.39%~4.92%、307.00~1 176.00 mg/kg、11.70~55.20 mg/kg、37.70~138.00 mg/kg、31.00~170.00 mg/kg、0.31~1.47 mg/kg，平均值分别为 3.77%、587.45 mg/kg、29.24 mg/kg、90.12 mg/kg、88.10 mg/kg、0.70 mg/kg(表 1)。

表 1 土壤特性统计  
Table 1 Statistics of soil characteristics

项目 Item	有机碳 Organic carbon//%	pH	氮 N mg/kg	钾 K %	磷 P mg/kg	钙 Ca %	镁 Mg %
最小值 Min	0.25	4.22	568.00	1.37	304.00	0.26	0.51
最大值 Max	2.90	8.37	4 281.00	3.27	5 056.00	18.87	3.76
均值 Mean	1.26	7.47	1 567.73	2.31	1 428.45	5.92	1.15
标准偏差 Standard deviation	0.67	1.08	779.62	0.37	1 084.78	6.25	0.63
偏度 Skewness	0.61	-1.83	1.39	0.01	1.58	0.92	2.53
峰度 Kurtosis	0.01	2.61	3.08	0.97	2.73	-0.53	8.63
项目 Item	硫 S//%	全铁 Fe %	锰 Mn mg/kg	铜 Cu mg/kg	锌 Zn mg/kg	硼 B mg/kg	钼 Mo mg/kg
最小值 Min	0.01	2.39	307.00	11.70	37.70	31.00	0.31
最大值 Max	0.13	4.92	1 176.00	55.20	138.00	170.00	1.47
均值 Mean	0.03	3.77	587.45	29.24	90.12	88.10	0.70
标准偏差 Standard deviation	0.02	0.65	207.85	8.04	18.01	31.62	0.27
偏度 Skewness	2.79	-0.10	0.88	0.58	-0.27	0.35	1.33
峰度 Kurtosis	10.81	-0.93	0.41	2.85	2.14	0.37	1.78

由表 2 可知,可溶性固形物含量为 10.90%~16.10%,平均值为 13.57%,参照 GB 12947—2008 鲜柑橘的相关标准,33 个样品的可溶性固形物含量均达到优等果(可溶性固形物含量 $\geq 10.5\%$ )级别。果实横径为 68.53~85.47 mm,平均值为 76.09 mm,参照 GB 12947—2008 鲜柑橘的相关标准,1 个样品达到 3 L(85.0 mm $\leq$ 横径 $< 95.0$  mm)级别,6 个样品达到 2 L(80.0 mm $\leq$ 横径 $< 85.0$  mm)级别,14 个样品达到 L(75.0 mm $\leq$ 横径 $< 80.0$  mm)级别,7 个样品达到 M(70.0 mm $\leq$ 横径 $< 75.0$  mm)级别,5 个样品达到 S(65.0 mm $\leq$ 横径 $< 70.0$  mm)级别。

表 2 脐橙果实品质  
Table 2 Fruit quality of navel orange

项目 Item	可溶性固 形物含量 Soluble solid content //%	横径 Cross diameter mm
最小值 Min	10.90	68.53
最大值 Max	16.10	85.47
均值 Mean	13.57	76.09
标准偏差 Standard deviation	1.22	4.86
偏度 Skewness	-0.18	0.17
峰度 Kurtosis	-0.23	-0.74

**2.2 相关性分析** 相关性分析结果表明,pH、钙与果实横径达显著正相关,相关系数分别为 0.346 和 0.400。上述土壤特性与可溶性固形物含量均未达显著相关性(表 3)。

**2.3 分类回归树** 将可溶性固形物作为因变量,14 个土壤特性指标作为自变量,设定父节点和子节点分别为 4 和 2,可

溶性固形物生成的 CART 树模型见图 3。将果实横径作为因变量,14 个土壤特性指标作为自变量,设定父节点和子节点分别为 4 和 2,果实横径生成的 CART 树模型见图 4。并将模型的预测结果与实际测量值生成散点图(图 5、6),来检验预测值与真实值的拟合程度。

表 3 相关性分析  
Table 3 Correlation analysis

指标 Index	横径 Cross diameter	可溶性固 形物含量 Soluble solid content
pH	0.346*	-0.054
有机碳(Organic carbon)	0.179	-0.159
氮(N)	0.009	-0.163
磷(P)	-0.158	0.088
钾(K)	-0.244	0.241
钙(Ca)	0.400*	-0.334
镁(Mg)	0.175	0.067
硫(S)	-0.081	-0.076
全铁(Fe)	-0.203	0.062
锰(Mn)	0.029	0.255
硼(B)	0.003	-0.054
钼(Mo)	-0.229	0.173
铜(Cu)	0.134	-0.032
锌(Zn)	0.092	0.167

注: \* 表示线性关系达显著水平( $P < 0.05$ )  
Note: \* indicated significant level( $P < 0.05$ )

分别选取决定系数( $R^2$ )、均方根误差(RMSE)、平均绝对误差(MAE)来检验模型精度。由表 4 可知,可溶性固形物

的  $R^2$ 、RMSE、MAE 分别为 0.85、0.46、0.32;果实横径的  $R^2$ 、RMSE、MAE 分别为 0.89、1.58、1.07。研究表明,当模型的 RMSE、MAE 越小,  $R^2$  越大时,模型精度越好。

上述结果表明,CART 树对可溶性固形物和果实横径的解释率分别达 85%和 89%,说明 CART 树可以很好地用来研究土壤特性对脐橙果实品质的影响。

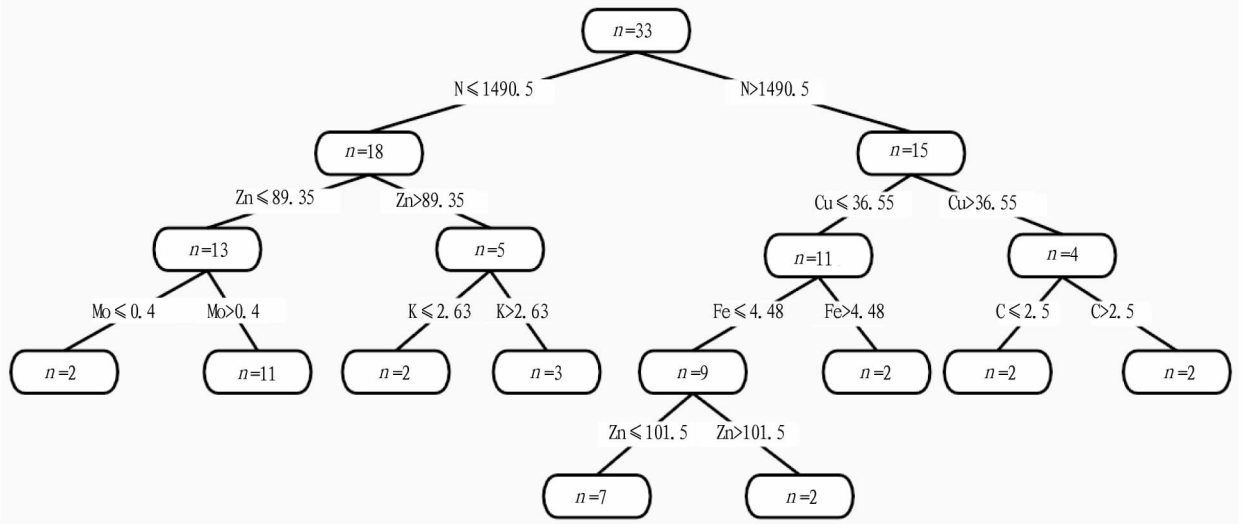


图 3 可溶性固形物含量的 CART 树结构  
Fig.3 CART tree structure of soluble solid content

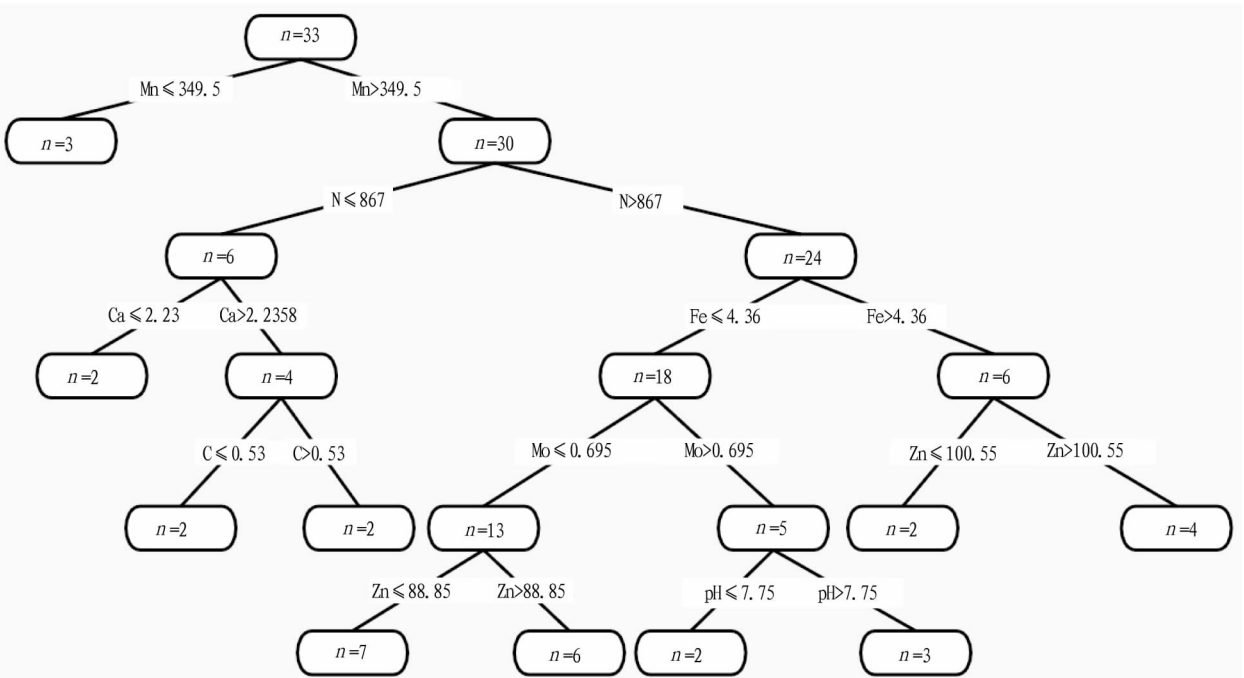


图 4 果实横径的 CART 树结构  
Fig.4 CART tree structure of fruit horizontal diameter

表 4 模型精度  
Table 4 Model accuracy

指标 Index	$R^2$	RMSE	MAE
可溶性固形物含量 Soluble solid content	0.85	0.46	0.32
横径 Cross diameter	0.89	1.58	1.07

2.4 果实品质的主控因素分析 相对重要性图是 CART 模型一个关键的输出结果,表明自变量对模型的重要程度。可溶性固形物和横径的相对重要性见图 7、8。结果表明,锌是

影响奉节脐橙可溶性固形物含量最重要的指标,其他指标的相对重要性表现为有机碳>磷>氮>铜>镁>硫>全铁>钼>锰>钾>硼>钙>pH。氮是影响奉节脐橙横径最重要的指标,其他指标的相对重要性表现为锌>锰>pH>磷>铜>全铁>硫>钼>有机碳>镁>钾>硼>钙。

3 结论

引用 CART 树模型对奉节脐橙果实可溶性固形物、横径与土壤特性之间的关系进行了分析。结果表明,在 14 个土壤特性指标中,锌是影响奉节脐橙可溶性固形物最为关键的

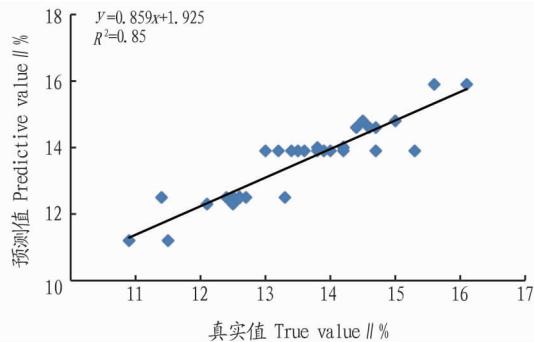


图5 可溶性固形物含量散点图

Fig.5 Scatter plot of soluble solid content

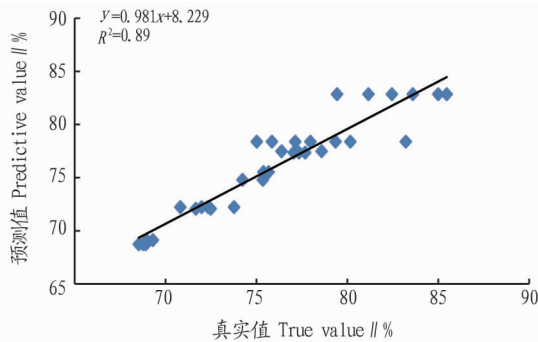


图6 果实横径散点图

Fig.6 Scatter plot of fruit transverse diameter

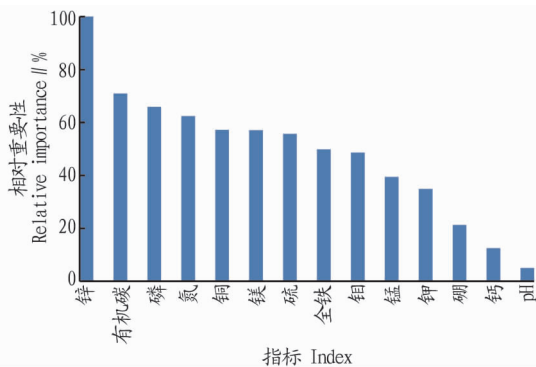


图7 土壤特性对可溶性固形物含量影响的相对重要性

Fig.7 The relative importance of the influence of soil properties on the content of soluble solids

因素,而氮是影响奉节脐橙果实横径最重要的因素。pH、钙与果实横径达显著正相关,相关系数分别为0.346和0.400。由此可见,pH和钙对奉节脐橙横径的作用明显。锌不仅是影响果实可溶性固形物的重要因素,同时也是影响果实横径的次重要因素,说明合理调控土壤锌含量有助于提高奉节脐橙品质。CART模型可以很好地用来研究土壤特性对脐橙品质的影响,就可溶性固形物树模型而言, $R^2$ 、RMSE、MAE分别为0.85、0.46、0.32;就横径树模型而言, $R^2$ 、RMSE、MAE分

别为0.89、1.58、1.07。

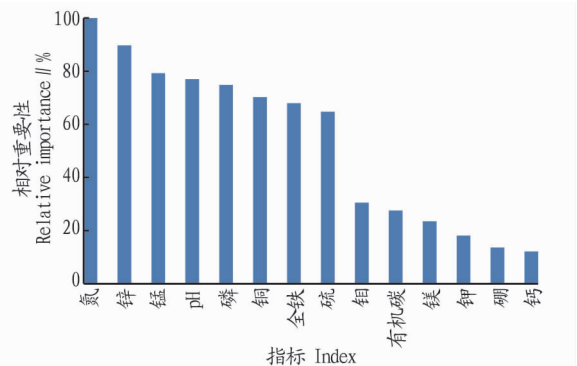


图8 土壤特性对横径影响的相对重要性

Fig.8 The relative importance of soil properties on the influence of transverse diameter

### 参考文献

- [1] 奉节县地方志办公室.奉节年鉴·2020[M].北京:世界图书出版公司,2020.
- [2] 朱丽莎,董超,赵静,等.重庆主栽甜橙品种果实品质比较分析[J].中国南方果树,2018,47(5):128-135.
- [3] 江才伦,彭良志,雷霆,等.柑桔单果间和果实不同部位的可溶性固形物含量差异[J].中国南方果树,2006,35(2):3-5.
- [4] 吴正琴,甘霖.柑桔果实可溶性固形物与品质的关系[J].四川果树科技,1991(1):34-36.
- [5] 余焰文,蔡小琴,蔡哲,等.基于可溶性固形物含量模拟的南丰蜜橘采摘期确定[J].气象与减灾研究,2019,42(4):285-289.
- [6] 刁俊明,曾宪录,朱远平,等.脐橙果实大小对果实感官品质和可溶性固形物含量的影响[J].广东农业科学,2015,42(23):82-85.
- [7] 李泽碧,王正银.柑橘品质的影响因素研究[J].广西农业科学,2006,37(3):307-310.
- [8] 孙丽平,何俊莉,宋春草,等.柑橘果实品质主要影响因素研究进展[J].南方农业,2021,15(1):56-60.
- [9] 芮文婧,王晓敏,高艳明,等.环境对番茄果实品质和风味影响的研究进展[J].安徽农业科学,2017,45(6):34-36.
- [10] 鲍江峰,夏仁学,彭抒昂,等.湖北省纽荷尔脐橙园土壤营养状况及其对果实品质的影响[J].土壤,2006,38(1):75-80.
- [11] 鲍江峰,夏仁学,彭抒昂,等.纽荷尔脐橙园土壤营养与果实品质相关性研究[J].亚热带植物科学,2005,34(1):25-27.
- [12] 鲍江峰,夏仁学,彭抒昂,等.三峡库区纽荷尔脐橙园土壤营养状况及其对果实品质的影响[J].中国土壤与肥料,2006(3):16-20.
- [13] 孔樟良,陈和秀,赵玲玲,等.影响建德市柑橘品质的土壤因素初步研究[J].浙江农业科学,2015,56(2):192-193.
- [14] 罗来辉.施锌对沙田柚果实品质的影响[J].嘉应学院学报,2010,28(5):71-74.
- [15] 沙珉,王道英,吴师金.安远县某地土壤地球化学特征与脐橙品质关系探讨[J].江西地质,2018,19(2):119-124.
- [16] 栾进华,程军,王伟,等.奉节脐橙种植区土壤地球化学特征[J].物探与化探,2011,35(6):829-832.
- [17] 赵萍,傅云飞,郑刘根,等.基于分类回归树分析的遥感影像土地利用/覆被分类研究[J].遥感学报,2005,9(6):708-716.
- [18] MERTENS M, NESTLER I, HUWE B. GIS-based regionalization of soil profiles with Classification and Regression Trees (CART) [J]. Journal of plant nutrition and soil science, 2002, 165(1):39-43.
- [19] WARURU B K, SHEPHERD K D, NDEGWA G M, et al. Estimation of wet aggregation indices using soil properties and diffuse reflectance near infrared spectroscopy: An application of classification and regression tree analysis [J]. Biosystems engineering, 2016, 152:148-164.