

第六届湖南省研究生数学建模竞赛

题目 基于 TOPSIS 和指数平滑的黄桃质量评价与贮藏模型

摘要：

黄桃果肉金黄、果汁丰富，商用价值较高，但比较容易腐烂。为了更好的探究影响黄桃品质的因素，寻找影响黄桃运输保存的有利条件，对题设问题进行分析。

针对问题一，通过主成分分析模型得到机械指标的主要因素为果实硬度和 TSS 含量，人工评分主要因素为质地和味道，通过灰色关联度模型得到主要因素之间的灰色关联系数以及权重，客观的表示机械指标与人工评分之间的关系。果实硬度和 TSS 含量与质地和味道意见的强相关性也表明了机械指标与人工评分具有强一致性，最后通过 TOPSIS 向商家推荐的黄桃 2018 年 6 月 11 日批次第 10 天（6 月 20 日）为品质最优的黄桃，TOPSIS 模型综合了机械指标与人工评分，而专家经验和普遍认知则是双评分机制，各评分之间相互独立，在结论上的差异表现为 TOPSIS 模型考虑更为全面，而专家经验和普遍认知相对片面。

针对问题二，通过多元线性回归模型建立人工指标与机械指标之间的函数关系时，是的可以依据机械指标得到回归分析的人工指标，从而判断得出样本 A、B、F、G 为优质，样本 C、D、E、H、I 为非优质。

针对问题三，通过指数平滑模型和灰色预测模型分别对黄桃品质在不同的贮藏条件下的品质变化，得到结果与普遍认为相一致，推荐给商家的较好的贮存方法为先对黄桃预冷处理，在 8°C 冷柜中存放 3-5 天，再转至 1°C 冷库中冷藏。在该模型下，果农需要在预冷阶段进行 2 次有损测试，在冷藏阶段进行 7 天 1 次的有损测试。

关键词：主成分分析 灰色关联度 TOPSIS 指数平滑预测 灰色预测

一、问题重述

关于黄桃，有一则有趣的神话传说。相传齐天大圣孙悟空大闹天宫时将瑶池蟠桃打落坠入人间，而文殊菩萨的坐骑狮子仙为求长生不老苦苦寻找后享用的蟠桃核，便种在了这湖南桂东的小村庄里[1]。现如今，狮子仙留下的仙桃成为了湖南桂东、炎陵两地老百姓脱贫致富的黄金桃，为当地经济的发展带来了动力。黄桃分布在我国长江流域及北部地区，营养十分丰富，含有多种人体所需必须氨基酸及营养保健成分，具有降低血糖血脂、提高免疫力和延缓衰老等功效，因此食用价值和保健价值都很高，深受国内外消费者喜爱，具有广阔的市场需求[2, 3, 4]。

然而，黄桃成熟于6月中下旬，此时正处于炎热的夏季，在高温下，果实容易褐变腐烂，品质下降。温度影响了黄桃采摘后贮藏寿命的长短，贮藏温度稍高，黄桃呼吸作用强，代谢正常，口味较好，但贮藏保鲜期较短；贮藏温度低，黄桃呼吸作用降低，贮藏时间延长，但口味容易下降[5]。由于黄桃对温度敏感、耐贮性较差的性质，其保鲜问题成为一个难题。同时为了满足消费者对高品质水果的需求，常需对黄桃的品质进行分析。对于黄桃的品质分析一般可以根据机械指标或者是人工评分。机械指标大多是用仪器对果实进行有损测试得到的指标，包含：可溶性固形物（TSS）、硬度、色差等等。人工评分是由人品尝后对果实的味道、色泽、质地、芳香等性质评价后所得的分数。专家经验认为“优质”黄桃的机械指标是果实硬度 $>1.5\text{kg/cm}^2$ ，TSS含量下降率 $<20\%$ ，而人工评分认为“优质”黄桃的标准是色泽、质地和味道这三个指标均分 ≥ 80 分。普遍认知是人工评分比机械指标得到的结果更为苛刻。

因此，如何利用机械指标和人工评分综合评判果实的综合品质，同时科学地刻画黄桃品质与贮藏运输的关系，以满足市民对高品质果实的需求和帮助农户与商户更好的采摘、运存与售卖是本文研究需要探讨的问题。S市某黄桃种植基地提供了2018年至2020年的黄桃测评数据，依据实验数据，本文需建立模型解决以下问题：

问题1：根据实验的描述和数据，分析机械指标、人工评分之间的关系与两类方法的评判一致性，并为商家推荐优质黄桃构建黄桃品质的评价模型。根据构建的评价模型结论，与“专家经验”和“普遍认知”进行对比，分析差异。

问题2：结合问题1中的模型和结论，根据2020年度的机械指标数据建立模型判断以下9个样本的人工评分等级（“优质”和“非优质”）并给出结果。

样本	品质	样本	品质	样本	品质
样本 A		样本 D		样本 G	
样本 B		样本 E		样本 H	
样本 C		样本 F		样本 I	

问题3：根据实验数据，综合考虑贮藏环境和贮藏时长，构建黄桃品质随贮藏环境和贮藏时间变化的预测模型，并给出合适的贮藏方式和时间。分析数据结果与上述“普遍认为”是否一致，并按照模型给出下一年度用该模型预测黄桃采摘贮藏期需要的有损测试次数。

二、问题分析

2.1 问题一

对于问题 1，将问题 1 分解为三个小问。

Q1:分析机械指标、人工评分的关系与两类方法的评判一致性

Q2:构建黄桃品质分析的定量模型，并为商家推荐优质黄桃

Q3:比较本文的模型结论与“专家经验”“普遍认知”的差异

根据专家经验，机械指标中主要考虑果实硬度和可溶固形物总量（TSS）含量下降率，而人工评分主要考虑色泽、质地和味道这三个指标。在评价黄桃品质时，专家经验使用 TSS 下降率，但是在分析机械指标与人工评分关系和一致性时，TSS 下降率并非是必要考虑因素，在反应变化趋势上 TSS 含量具有明显的直观优势，故分析机械指标、人工评分的关系与一致性时用 TSS 含量代替 TSS 下降率，直观上，很难直接的看出人工评分（色泽、质地、味道）和机械指标（果实硬度、TSS 含量、色差）之间存在一对一的映射关系。但整体上，两组数据都具有较为平稳的趋势和一直的浮动性，说明它们实际上可能存在一定的关联，且在普遍认知中，在优质黄桃的评判上机械指标较容易满足，而人工评分要求更高，所以不由此可推测出机械指标与人工评分之间存在多对一的映射关系。本文通过主成分分析法先提取出机械指标中贡献率最大的主成分因素，再通过灰色关联度模型得到主成分因素与人工评分各项指标的关系，评判两类方法的一致性，最后利用机械指标的逐个对人工评分指标求关联度系数，客观准确的描述机械指标与人工评分的关系。

因为专家经验中考虑了果实硬度、TSS 下降率（机械指标）和色泽、质地、味道（人工评分），故将这 5 个指标组合在一起分析，建立优劣解距离（TOPSIS，下文统称 TOPSIS）模型。TOPSIS 法是一种常用的组内综合评价方法，能充分利用原始数据的信息，其结果能精确地反映各评价方案之间的差距。其模型思想是基于归一化后的原始数据矩阵，采用余弦法找出有限方案中的最优方案和最劣方案，然后分别计算各评价对象与最优方案和最劣方案间的距离，获得各评价对象与最优方案的相对接近程度，以此作为评价优劣的依据。因为 TOPSIS 根据最优方案的距离来进行排序，故其可以作为评价方法为商家推荐最优质的黄桃。

2.2 问题二

对于问题 2，由于 2020 年度的评测中取消了人工评分，使得机械指标与人工评分的双评分机制不再适用，且基于灰色关联度的 TOPSIS 模型也缺少了色泽、质地和味道等人工评分数据，无法对 2020 年的样本数据进行准确评估。在问题一中，由主成分分析法提取了影响机械指标和人工评分的主要因素，灰色关联度模型给出了机械指标与人工评分指标之间的关联系数和权重，但是仍然没有依据观测值给出二者之间的关系式。所以在问题二中，需要根据具有现实意义的观测值，通过对机械指标进行数据处理、分析得到对应的人工评分，且由于单个人工评分数据受多个机械指标因素影响，故本文选用多元线性回归模型求解机械指标与人工评分数据之间的函数关系式。多元线性回归常常处理一种现象与多个因素相联系的问题，通过多个自变量的最优组合共同估计因变量，处理结果的实际意义较大。考虑到灰色关联度模型中的数据选择和 2020 年的机械数据，多元线性回归模型处理时，依然选择果实硬度、TSS 含量、色差、色泽、质地和味道共 6 个指标。

2.3 问题三

对于问题 3，根据问题 1 中的 TOPSIS 模型以及专家经验可得，黄桃品质受温度和时间两个维度的影响，考虑温度只有冷柜的 8°C 和冷库的 1°C 两种情况，所以对于温度的使用分类讨论，定性分析。考虑时间维度上黄桃品质是连续变化的，故选用时间序列类的模型进行对黄桃品质进行预测。构建黄桃随贮藏时间变化的预测模型中，本文选用指数平滑模型进行黄桃短期品质预测，选用灰色预测模型进行黄桃品质长期趋势预测，为商家推荐合适的贮藏方式和时间。根据预测模型的预测结果，可以判断出是否与“普遍认为”的结论一致，并结合环境的不同，探讨黄桃品质的变化规律，给出果农需要进行的有损测试周期，确定测试次数。

三、问题假设与符号说明

3.1 问题假设

1. 假设果实硬度由果皮硬度与果肉硬度共同影响。
2. 假设测试员主观影响因素忽略不计。
3. 假设更换贮藏环境期间受到的影响忽略不计。
4. 假设果实品质仅由机械指标与人工评分影响。

3.2 符号说明

1. TSS_0 表示黄桃样本 TSS 初始值， TSS_i 表示该样本第 i 次测量时的 TSS 含量。
2. L_0 表示黄桃测量时黄桃的明暗度指标， a_0 表示黄桃测量时黄桃的红绿色指标和 b_0 为黄桃测量时黄桃的黄蓝色指标。
3. ΔE 表示测量的色差值。
4. a_1 表示果实硬度数据， a_2 表示 TSS 含量数据， a_3 表示色差数据。
5. M 表示指标 x 可能取的最大值， m 表示指标 x 可能取的最小值。
6. X 表示归一化矩阵。
7. β 表示最小二乘法通过矩阵运算求解的系数矩阵。
8. B 表示累加序列。
9. C 表示紧邻均值序列。
10. Q 表示相对残差。

四、问题一的模型建立与求解

4.1 数据预处理

在对 2018 年、2019 年黄桃的机械指标和人工评分时发现，由于实验记录问题，2018 年的机械指标带皮硬度、去皮硬度、TSS 含量与色差并不是一一对应的，且因为测量实验为有损实验，所以对于两年数据中的机械指标与人工评分都无法一一对应某一个特定样本。此外，2019 年数据中存在果实色差、果皮色差和冷库 6 月 17 日批次的人工评分（色泽、质地、味道）数据缺失，6 月 14 日批次的果皮色差 a 第 12 天数据异常，最后，对于色差指标而言，包含了不同条件下存储

的果实和果皮色差数据，且包含其相应的明暗度（黑白）、红绿色、黄蓝色、彩度（色彩饱和的程度或纯粹度）和色调角等，色差的指标因素过于冗余，不利于数据分析。

针对以上问题，进行如下数据处理：

1. 数据不对应问题：虽然单个样本的数据无法一一对应，但是对于每一批样本而言，无论是机械指标还是人工评分，都是对于这一批样本进行测量和评估的，所以当研究粒度扩大到每一批次的时候，机械指标和人工评分是具有分析意义的。而代表每一批次的机械指标和人工评分，本文选择对每一批次的样本进行均值化处理，使用平均值代表，均值化公式如公式（1）所示：

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad \text{—— (1)}$$

以带皮硬度为例， $\bar{x}$ 表示特定批次黄桃带皮硬度的平均值， n 表示该批次的样本数量， x_i 表示第 i 个样本。

2. 数据缺失问题：对于缺失的机械指标和人工评分，本文选择使用三次样条插值法进行数据补缺，三次样条插值法如公式（1）所示：

$$y = a_i + b_i x + c_i x^2 + d_i x^3 \quad \text{—— (2)}$$

三次样条是指每个小区间的曲线都是如公式（2）所示的一个三次方程，三次样条方程需要满足自变量与因变量一一对应且曲线光滑，即在任何插值点处的函数、一阶导数和二阶导数都是连续的。

题设机械指标含有 TSS 下降率，但并未给出 TSS 下降率，只给出了 TSS 含量，故本文根据 TSS 的下降幅度与初始值的比值作为 TSS 的下降率，具体如公式（3）所示：

$$TSS_i = \frac{TSS_i - TSS_0}{TSS_0} \times 100\% \quad \text{—— (3)}$$

TSS_0 为 TSS 含量初始值， TSS_i 为该样本第 i 次测量时的 TSS 含量。

3. 数据异常问题：对于明显异常的机械指标，本文选择按照数据缺失值的方式，先清除数据，在进行插值补缺。

4. 数据冗余问题：对于果实硬度问题，本文选择使用带皮硬度与去皮硬度的平均值。对于色差明显冗余的机械指标，本文选择使用 CIE Lab 容差公式（Lab 色差公式）对其进行计算，给出可代表对应批次黄桃的色差值，CIE Lab 容差公式如公式（4）所示：

$$\Delta E = \sqrt{(L^* - L_0)^2 + (a^* - a_0)^2 + (b^* - b_0)^2} \quad \text{—— (4)}$$

其中， L_0 、 a_0 和 b_0 分别表示黄桃测量时黄桃的明暗度指标、红绿色指标和黄蓝色指标初始值， L^* 、 a^* 和 b^* 分别表示本次黄桃测量时黄桃的明暗度指标、红绿色指标和黄蓝色指标， ΔE 表示本次测量的色差值，2019 年数据中包含果皮色差和果实色差，本文选择取均值代表对应黄桃的色差。

4.2 机械指标与人工评分的关系问题

4.2.1 问题的分析

根据专家经验，机械指标中主要考虑果实硬度和可溶固形物总量（TSS）含量下降率，而人工评分主要考虑色泽、质地和味道这三个指标。但是由于果实硬度需要综合考虑带皮硬度和去皮硬度，为了更好的反应果实硬度，本文选择他们的均值作为果实硬度的代表，且可溶固形物总量（TSS）含量下降率在 $[0, 1]$ 的区间内，使得该指标与其他指标在统计分析时变化不明显，在评价黄桃品质时，根据专家经验则必须使用 TSS 下降率，但是在分析机械指标与人工评分关系和一致性时，TSS 下降率并非是必要考虑因素，在反应变化趋势上 TSS 含量具有明显的直观优势，故进行数据的关联性分析时用 TSS 含量代替 TSS 下降率，2018 年与 2019 年机械指标与人工评分的折线图分别如图 1、图 2 所示。

直观上，很难直接的看出人工评分（色泽、质地、味道）和机械指标（果实硬度、TSS 含量、色差）之间存在一对一的映射关系。但整体上，两组数据都具有较为平稳的趋势和一直的浮动性，说明它们实际上可能存在一定的关联，且在普遍认知中，在优质黄桃的评判上机械指标较容易满足，而人工评分要求更高，所以不难推测出机械指标与人工评分之间存在多对一的映射关系。

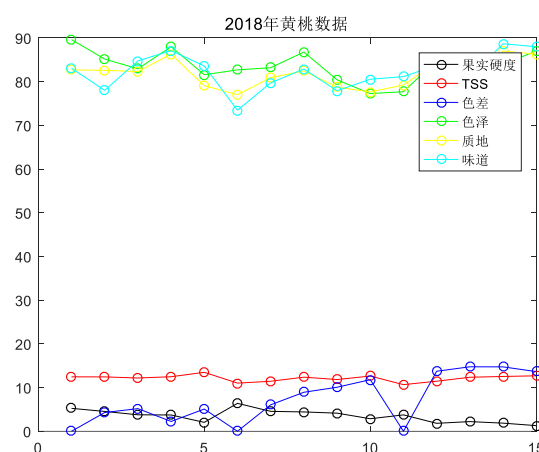


图 1 2018 年机械指标与人工评分折线图

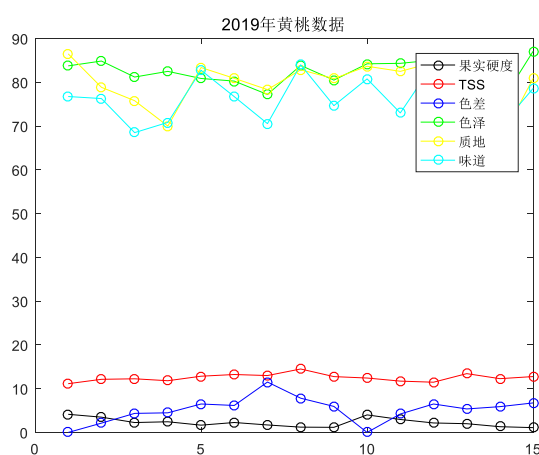


图 2 2019 年机械指标与人工评分折线图

经统计，2018 年中，机械指标对 3 批冷库（1℃）存储黄桃共计 15 次的测量结果均为优质，黄桃优质率为 100%，人工评分对 3 批黄桃共计 15 次的测量结

果为 8 次优质, 7 次非优质, 黄桃优质率为 53.33%; 机械指标对 3 批冷柜 (8℃) 存储黄桃共计 15 次的测量结果为 14 次优质, 1 次非优质, 黄桃优质率为 93.33%, 人工评分对 3 批黄桃共计 15 次的测量结果为 10 次优质, 5 次非优质, 黄桃优质率为 66.67%。2019 年中, 机械指标对 3 批冷库 (1℃) 存储黄桃共计 29 次的测量结果均为优质, 黄桃优质率为 100%, 人工评分对 3 批黄桃共计 29 次的测量结果为 19 次优质, 10 次非优质, 黄桃优质率为 65.52%; 机械指标对 3 批冷柜 (8℃) 存储黄桃共计 29 次的测量结果为 15 次优质, 14 次非优质, 黄桃优质率为 51.72%, 人工评分对 3 批黄桃共计 29 次的测量结果为 13 次优质, 16 次非优质, 黄桃优质率为 44.83%。

统计数据中, 如普遍认知所述一致, 机械指标判断的黄桃优质率均大于人工评分判断的黄桃优质率, 机械指标比人工评分更容易满足。这就意味着人工评分比机械指标具有更高的要求、更严格的条件, 例如在某些指标里, 机械指标判断是需要测量到其值, 就可以给出结果。但是人工评分是, 测试员对于色泽、质地和味道的判断是综合性的, 可能由多个机械指标共同决定, 为了更准确客观的判断机械指标与人工评分的关系和一致性, 本文通过主成分分析法求得机械指标与人工评分中首要影响因素, 并利用灰色关联度模型对首要影响因素进行关联度分析, 求证机械指标与人工评分之间的关系与一致性。

4.2.2 问题模型建立

主成分分析模型:

主成分分析模型作为分析模型, 主要是为了确定机械指标与人工评分的首要影响因素。附件 1 至附件 7 中, 机械指标包含了带皮硬度 (左、右)、去皮硬度 (左、右)、TSS 含量 (左、右)、TSS 下降率、果实色差 (L、a、b)、果皮色差 (L、a、b)、单果重、纵径、横径、侧径, 人工评分包括色泽、质地、芳香、味道。

根据题设要求, 本文在众多指标中确定了果实硬度、TSS 含量、色差、色泽、质地和味道 6 个满足如下条件的指标:

1. 机械指标的评价标准主要以果实硬度和 TSS 下降率为主, 人工评分主要以色泽、质地和味道为主。
2. 数据处理中对于冗余数据的进行了处理, 对于同一指标选取其一种最合理的代表方式, 例如色差项。
3. 选择较为全面的数据, 即 2018 年与 2019 年均包含的数据, 避免单次出现存在的偶然性。

确定的 6 个指标中, 果实硬度、TSS 含量和色差属于机械指标, 色泽、质地和味道属于人工评分, 要准确客观的分析机械指标和人工评分之间的关系与一致性, 即需要探讨果实硬度、TSS 含量、色差对果实色泽、质地、味道的关系。

因为果实硬度、TSS 含量和色差这 3 个机械指标对人工评分产生了一定的影响, 所以令 X_1 、 X_2 、 X_3 依次代表 3 个因素的不同批次的历史数据, 构成初始矩阵 $X = (X_1, X_2, X_3)^T$, 通过正交变换使各数据之间线性不相关, 由此列出 3 个成分表达式, 如公式 (5) 所示:

$$\begin{cases} F_1 = a_{11}X_1 + a_{12}X_2 + a_{13}X_3 \\ F_2 = a_{21}X_1 + a_{22}X_2 + a_{23}X_3 \\ F_3 = a_{31}X_1 + a_{32}X_2 + a_{33}X_3 \end{cases} \quad \text{—— (5)}$$

因为各成分之间通过线性变换不相关,所以任意两个表达式的协方差矩阵均为零,其方差为相关系数,即 $\text{Cov}(F_i, F_j)=0$, 且 $\text{Var}(F_i, F_j)=a_i$;其中 $i, j=1, 2, 3$ 。进一步计算 3 个因素历史数据的均值, 如公式 (6) 所示:

$$\bar{x} = \frac{1}{p} \sum_{j=1}^p x_{jk} \quad \text{—— (6)}$$

由此, 可以计算出各个因素的样本均值, 再根据公式 (7) 计算出协方差矩阵 S , 公式 (7) 如下所示:

$$s_{ij} = \text{cov}(x_i, x_j) = \frac{1}{n-1} \sum_{l=1}^n [(x_i - \bar{x}_i)(x_j - \bar{x}_j)] \quad \text{—— (7)}$$

由公式 (7) 易知, 样本协方差矩阵 S 为:

$$S = \begin{pmatrix} s_{11} & s_{12} & s_{13} \\ s_{21} & s_{22} & s_{23} \\ s_{31} & s_{32} & s_{33} \end{pmatrix}$$

设 3 个因素的特征根为 $\lambda_1, \lambda_2, \lambda_3$, 根据公式 (8) 计算出各个因素特征根的解, 公式 (8) 如下所示:

$$|S - \lambda I| = 0 \quad \text{—— (8)}$$

最后根据公式 (9) 得出 3 个因素对色泽、质地和味道 (人工评分) 的贡献率, 贡献率 r 越大, 影响越大, 贡献率 r 最大的因素成为影响色泽、质地和味道 (人工评分) 的首要因素, 贡献率第二大的因素成为影响色泽、质地和味道 (人工评分) 的次要因素, 公式 (9) 如下所示:

$$r_i = \frac{\lambda_i}{\sum_{j=1}^p \lambda_j} \quad \text{—— (9)}$$

r_i 为第 i 个因素的贡献率, λ_i 为第 i 个因素的特征值。

灰色关联度模型:

灰色关联度模型作为求证模型, 它能刻画出机械指标中果实硬度、TSS 含量、色差与人工评分中色泽、质地、味道之间的关系, 从而进一步求证主成分分析的结果是否客观准确。

令列向量 a_1, a_2, a_3 依次代表果实硬度、TSS 含量和色差三个因素的数据, 构成初始矩阵 $y = (a_1, a_2, a_3)^T$, 进行灰色关联度分析。

第一步, 求初值项: 对 y 的每一项指标求均值得到 y_1 , 则初值项 y_2 可由公式 (10) 求得, 公式 (10) 如下所示:

$$y_2(i, j) = \frac{y(i, j)}{y_1(i, j)} \quad \text{—— (10)}$$

第二步，求差序列：差序列 y_3 可由公式 (11) 求得，公式 (11) 如下所示：

$$y_3(i-1, j) = |y_2(i, j) - y_2((i-1), j)| \quad \text{—— (11)}$$

第三步，求两极差：令 $a=1$ 求最小值, $b=0$ 求最大值，带入循环求解最大值和最小值，其中最大值 $b = \max_i \max_j y_3(i, j)$ ，最小值 $a = \min_i \min_j y_3(i, j)$ 。从而得到两级差 b 和 a 。

第四步，求关联系数 y_4 ：由两级差和差序列根据公式 (12) 推算出关联系数，公式 (12) 如下所示：

$$y_4(i, j) = \frac{a + 0.5b}{y_3(i, j) + 0.5b} \quad \text{—— (12)}$$

第五步，求灰色关联系数：得到关联系数后通过公式 (13) 对关联系数求均值，公式 (13) 如下所示，其中 $i=1, 2, 3$ 。

$$y_5(i) = \frac{\sum_{j=1}^{n-1} y_4(j)}{n-1} \quad \text{—— (13)}$$

$y_5(1)$ ， $y_5(2)$ ， $y_5(3)$ 分别为果实硬度、TSS 含量、色差与色泽、质地、味道之间的灰色关联系数，通过分析比较果实硬度、TSS 含量、色差的灰色关联系数，可以得出它们分别与色泽、质地和味道之间的关联度大小。

4.2.3 问题模型求解

将 2018 年和 2019 年的数据合并在一起，构建冷柜（8℃）和冷库（1℃）下机械指标和人工指标的主成分分析模型，共计 4 个分析模型。而构建色泽与机械指标的主成分分析模型、味道与机械指标的主成分分析模型和质地与机械指标的灰色关联度模型，则共计 6 个分析模型。

果实硬度、TSS 含量和色差在不同的冷藏环境下，对机械指标评价的贡献率如图 3 所示：

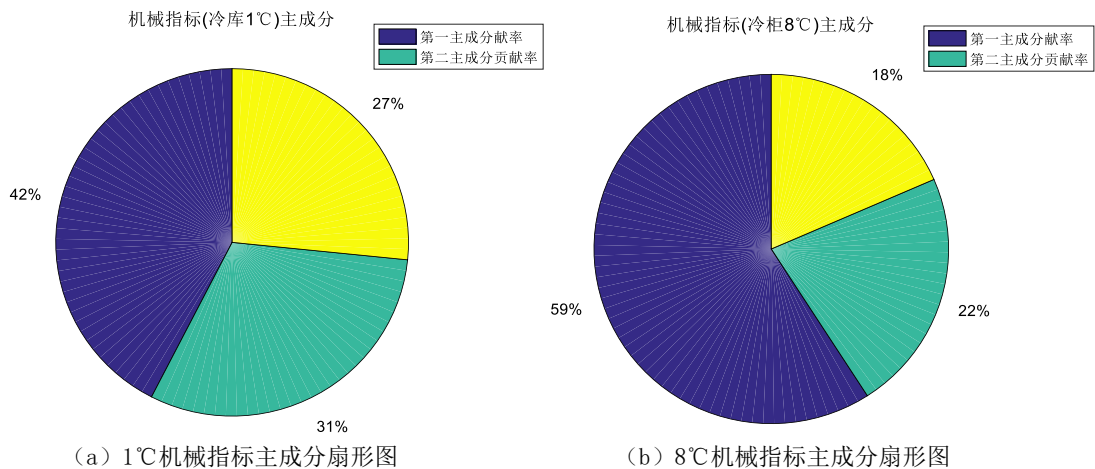


图 3 机械指标主成分贡献率扇形图

机械指标的主成分分析结果如表 2 所示，其包含了果实硬度、TSS 含量和色差在冷库（1℃）和冷柜（8℃）条件下，对机械指标评判影响的特征值、贡献率、累积贡献率和关系系数矩阵。

表 2 机械指标主成分分析结果

机械指标	特征值	贡献率	累积贡献率	主成分系数 1	主成分系数 2	主成分系数 3
果实硬度 1℃	1.2734	42.4%	42.4%	-0.6278	0.3343	0.7029
TSS 含量 1℃	0.9286	31.0%	73.4%	0.4552	0.8902	-0.0169
色差 1℃	0.7980	26.6%	100%	0.6314	-0.3094	0.7111
果实硬度 8℃	1.7736	59.1%	59.1%	0.6035	0.1163	0.7888
TSS 含量 8℃	0.6721	22.4%	81.5%	-0.5550	0.7716	0.3108
色差 8℃	0.5543	18.5%	100%	-0.5725	-0.6254	0.5302

在主成分分析模型中，单个影响因素的贡献率要超过 $1/n$ （ n 为影响因素总数）视为单个影响因素对指标有影响。当 m 个（ $m < n$ ）影响因素的累积贡献率达到 80% 时，该 m 个影响因素视为主要影响因素。

在冷库（1℃）中，只有果实硬度的贡献率（42.4%）超过了 33.33%，但任何 2 个指标的累积贡献率均未超过 80%，故在 1℃ 的情况下，仅有果实硬度对机械指标有微弱的突出影响，各指标影响差别不大。

在冷柜（8℃）中，只有果实硬度的贡献率（59.1%）超过了 33.33%，而果实硬度和 TSS 含量的累积贡献率（81.5%）超过了 80%，故在 8℃ 的情况下，果实硬度和 TSS 含量对机械指标有主要影响，色差的影响不大。

根据关系系数矩阵，我们可以得到以下成分关系式：

在冷柜（8℃）中：

第一主成分 $y_1 = 0.6035x_1 - 0.5550x_2 - 0.5725x_3$ ；

第二主成分 $y_2 = 0.1163x_1 + 0.7716x_2 - 0.6254x_3$ ；

x_1, x_2, x_3 表示果实硬度、TSS 含量和色差，从第一主成分来看，它在果实硬度上是正载荷，而在 TSS 含量和色差上是相近负载荷，说明在机械指标的评价中，一个黄桃的硬度越高，TSS 含量与色差变化越小，则这个黄桃越好吃，所以可认为第一主成分为果实硬度。从第二主成分来看，它在 TSS 含量上有较大的正载荷，在色差上有较大负载荷，说明在机械指标的评价中，一个黄桃的 TSS 含量越高，色差变化越小，则这个黄桃越好吃，所以可认为第二主成分为 TSS 含量。

因为冷库（1℃）的数据贡献率不明显，这可能是由于在温度更低的冷库中，各因素的变化被衰减，使得各因素对机械指标评价黄桃品质的影响都差别不大，故无法列举主成分关系式。

接下来看人工评分部分。果实硬度、TSS 含量和色差在不同的冷藏环境下，对人工评分的贡献率如图 4 所示：

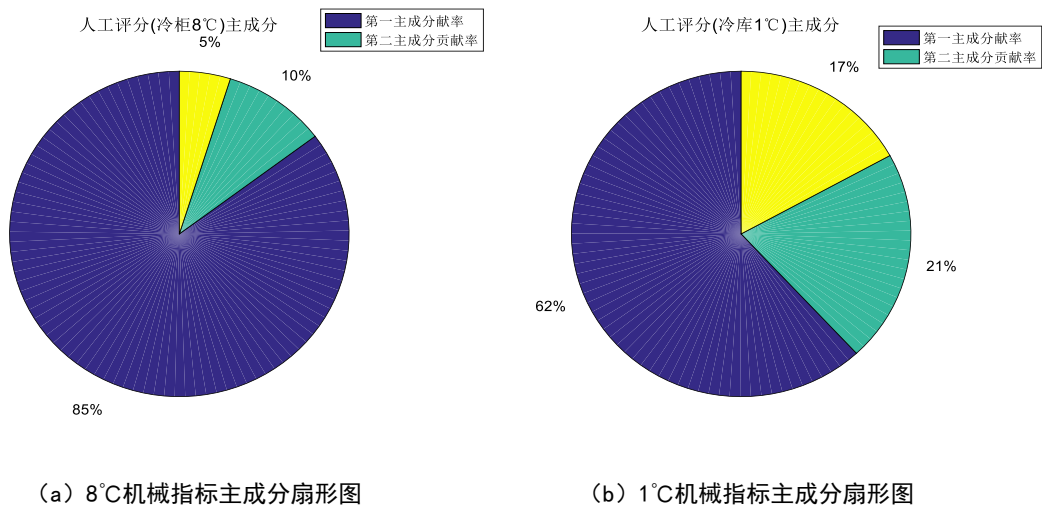


图 4 机械指标主成分贡献率扇形图

人工评分的主成分分析结果如表 3 所示。

表 3 人工评分主成分分析结果

人工评分	特征值	贡献率	累积贡献率	主成分系数 1	主成分系数 2	主成分系数 3
色泽 1℃	0. 5100	17. 00%	17. 00%	0. 5662	-0. 6872	-0. 4551
味道 1℃	0. 6332	21. 11%	38. 11%	0. 5618	0. 7258	-0. 3969
质地 1℃	1. 8568	61. 89%	100%	0. 6031	-0. 0309	0. 7971
色泽 8℃	0. 1458	4. 86%	4. 86%	0. 5727	-0. 6476	-0. 5027
味道 8℃	0. 2967	9. 89%	14. 75%	0. 5643	0. 7562	-0. 3314
质地 8℃	2. 5575	85. 25%	100%	0. 5947	-0. 0939	0. 7985

在冷库（1℃）中，只有质地的贡献率（61.89%）超过了 33.33%，质地和味道的累积贡献率（83%）超过了 80%，故在 1℃的情况下，质地和味道对人工评分有主要影响，色泽影响不大。

在冷柜（8℃）中，只有质地的贡献率（85.25%）超过了 33.33%，且质地的贡献率超过了 80%，说明在 8℃的情况下，质地对人工评分的影响显著，色泽和味道的影响不大。

根据关系系数矩阵，我们可以得到以下成分关系式：

在冷库（1℃）中：

第一主成分 $y_1=0.5662x_1+0.5618x_2+0.6031x_3$ ；

第二主成分 $y_2=-0.6873x_1+0.7258x_2-0.0309x_3$ ；

x_1, x_2, x_3 表示色泽、味道和质地，从第一主成分来看，它在色泽、味道和质地上是相近正载荷，说明在人工评分的评价中，一个黄桃的色泽分数越高，味道分数越高，质地分数越高，则这个黄桃越好吃。说明第一主成分上各主成分影响

相差不多，质地的影响偏大。所以可认为第一主成分为质地。从第二主成分上来看，他在味道上有较大的正载荷，说明在人工评分的过程中，一个黄桃的味道分数越高，则这个黄桃越好吃。所以可认为第二主成分为味道。

在冷柜（8℃）中：

第一主成分 $y_1=0.5727x_1+0.5643x_2+0.5947x_3$ ；

第二主成分 $y_2=-0.6476x_1+0.7562x_2-0.0939x_3$ ；

因为冷柜（8℃）与冷库（1℃）的人工评分数据的成分关系式基本一致，各因素的影响一致，所以可认为在冷柜的条件下质地为第一主成分，味道为第二主成分。

对于机械指标而言，果实硬度与 TSS 含量是关键影响因素；对于人工评分而言，质地与味道是关键的影响因素；在机械指标和人工评分的跨度上，各因素之间的关联性和一致性问题，接下来通过灰色关联度进行分析。

分别构造色泽、味道和质地与机械指标（果实硬度、TSS 含量和色差）的初始矩阵，通过灰色关联度模型，得到在不同冷藏条件（1℃和 8℃）下，果实硬度、TSS 含量、色差分别与色泽、味道和质地之间的权重关系，分别如图 5、图 6、图 7 所示。

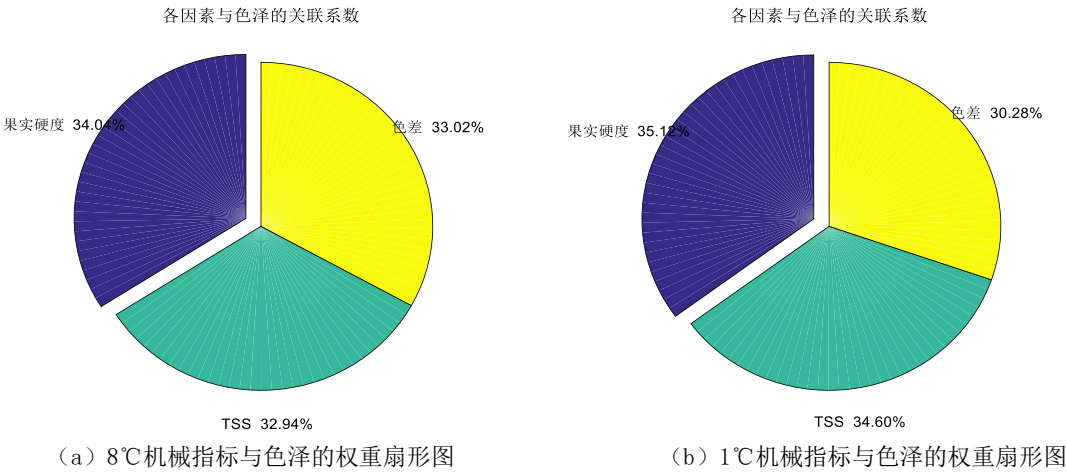


图 5 机械指标与色泽的权重扇形图

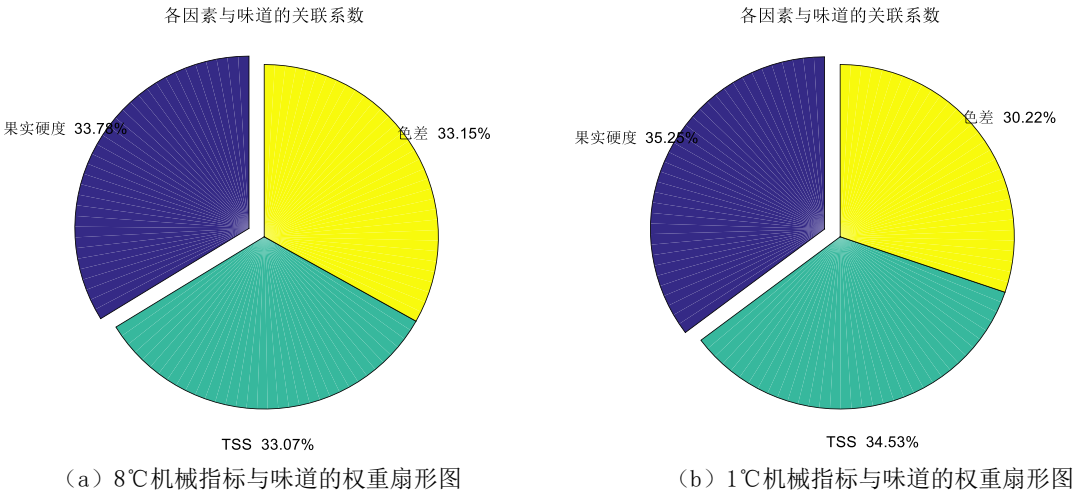


图 6 机械指标与味道的权重扇形图

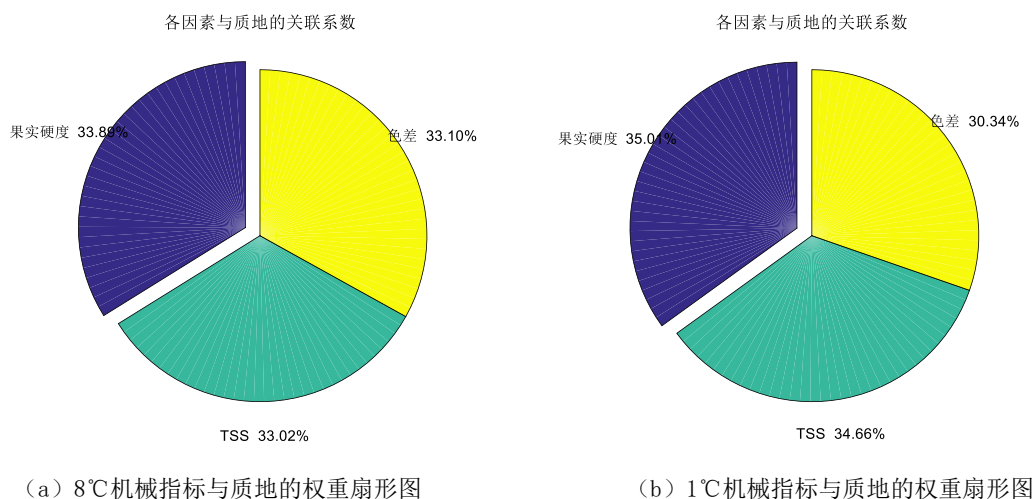


图 7 机械指标与质地的权重扇形图

通过灰色关联度分析发现，在冷柜（8℃）和冷库（1℃）的条件下，果实硬度、TSS 含量和色差对色泽、味道和质地三个人工评分的关联性为相近正相关，均在 30%—36%之间，说明机械指标对人工评分的影响具有较强的一致性，机械指标与人工评分的灰色关联度如表 4 所示。

表 4 人工评分主成分分析结果

灰色关联度	色泽 1℃	色泽 8℃	味道 1℃	味道 8℃	质地 1℃	质地 8℃
果实硬度	0.8452	0.7144	0.8400	0.7064	0.8327	0.7097
TSS 含量	0.8328	0.6915	0.8227	0.6915	0.8245	0.6915
色差	0.7289	0.6932	0.7200	0.6932	0.7217	0.6932

从表中可以直观的看出，机械指标中对于人工评分指标的关联度上：果实硬度>TSS 含量>色差，果实硬度和 TSS 含量作为机械指标的第一主成分和第二主成分因素，同时也是色泽、味道和质地三个人工评分指标中关联度最高的两项，这说明机械指标和人工评分两种方法在判断黄桃品质上具有较强的一致性，他们之间的关系与果实硬度和 TSS 含量具有强相关性，而不同的点在于他们对于指标的要求不同，即对于测量指标的权重分配上具有不同的评价方式。在机械指标中，果实硬度和 TSS 含量作为第一主成分和第二主成分，二者对机械指标评价结果的贡献率在 80%—90%之间，这也是为什么机械指标中只有果实硬度大于 1.5kg/cm²和 TSS 下降率小于 20%这两个要求，即题设中普遍认知里机械指标比较容易满足的原因。而人工评分则综合考虑了果实硬度、TSS 含量和色差的影响，虽然果实硬度和 TSS 含量关联度较高，但色差的相关性也不容忽视，任意两个指标的关联性在 65%左右，这就意味着如果另外一个指标过差，则对于黄桃品质的人工评分就无法大于 80 分，这也是题设中普遍认知里人工评分要求更高的原因。

4.3 构建黄桃品质分析的定量模型问题

4.3.1 问题的分析

本问题的解决，旨在建立一套更加科学的黄桃品质评价指标，综合考虑机械指标与人工评分。构建黄桃品质分析的定量模型，对黄桃品质进行评价，从而选出优质黄桃，并为商家进行推荐。

专家经验在对黄桃品质进行评价时，主要考虑机械指标和人工评分的双评分机制，其中机械指标包括果实硬度与 TSS 含量下降率指标，人工评分包括色泽、质地和味道，本部分将这 5 个指标组合在一起分析，建立优劣解距离（TOPSIS，下文统称 TOPSIS）模型。

4.3.2 问题模型建立

TOPSIS 法根据有限个评价对象与理想化目标的接近程度进行排序，结果能精确的反映各评价方案之间的差距。TOPSIS 法首先通过归一化处理得到原始数据矩阵，然后找出不同批次所有黄桃的品质最优黄桃和品质最差黄桃，再计算剩下的黄桃与品质最优黄桃或品质最差黄桃的贴近程度，最终得到黄桃品质的优劣排序。

黄桃品质分析的定量模型采用基于灰色关联度的 TOPSIS 评价模型。

算法步骤：

Step1: 指标属性正向化

TOPSIS 法使用距离来度量两个样本之间的差距，若使用距离尺度来度量则需要对目标指标的属性进行同向化处理，通常情况下采用成本型指标向效益型指标转化，即数值越大评价越高。下面给出不同类型指标的同向化方法：

（1）极小型指标：期望值越小越好，包含以下两种情况，如公式（14）和公式（15）所示：

$$x' = \frac{1}{x} (x > 0) \quad \text{—— (14)}$$

$$x' = M - x \quad \text{—— (15)}$$

其中 M 为指标 X 可能取的最大值。

（2）中间型指标：期望值取中间值最好，如公式（16）所示：

$$x' = \begin{cases} 2 \frac{x-m}{M-m}, m \leq x \leq \frac{1}{2}(M+m) \\ 2 \frac{M-x}{M-m}, \frac{1}{2}(M+m) \leq x \leq M \end{cases} \quad \text{—— (16)}$$

其中 M 为指标 x 可能取的最大值，m 为指标 x 可能取的最小值。

（3）区间型指标：期望值落在某一特定区间最好，如公式（17）所示：

$$x' = \begin{cases} 1 - \frac{a-x}{a-a^*}, x < a \\ 1, a \leq x \leq b \\ 1 - \frac{x-b}{b^*-b}, x > b \end{cases} \quad \text{—— (17)}$$

其中 $[a, b]$ 为指标 x 的最佳稳定区间, $[a^*, b^*]$ 为最大容忍区间。

在构建黄桃品质评价模型时, 需要考虑将机械指标中“硬度”、“TSS 含量下降率”以及人工评价中“质地”、“味道”、“色泽”等指标进行正向化处理以满足模型的使用条件。由题设可知, 人工评价各指标数值越大越好, 因此均为最大型指标。机械指标中, 黄桃硬度的最佳区间大于 $1.5\text{kg}/\text{cm}^2$, TSS 含量下降率小于 20%, 所以将黄桃硬度及 TSS 含量下降率设为区间型指标进行处理。

Step2: 构造归一化初始矩阵

根据数据信息可知, 黄桃在 2018 年 6 月 11 日、6 月 14 日、6 月 17 日分 3 批采摘, 每批间隔 3 天取样检测, 重复 5 次, 共计 15 批次黄桃样本; 在 2019 年 6 月 14 日、6 月 17 日、6 月 20 日分 3 批采摘, 每批间隔 2 天取样检测, 共计 32 批次黄桃样本。根据本问题需要对该 47 批次黄桃样本作为待评价对象, 评价指标有“硬度”、“TSS 下降率”、“质地”、“味道”、“色泽”五个指标, 则原始数据矩阵构造为:

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{15} \\ x_{21} & x_{22} & \cdots & x_{25} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{n5} \end{bmatrix} (n = 47)$$

接下来构造加权规范矩阵, 将属性进行规范化处理, 即每一列元素都除以当前列向量的范数, 如公式 (18) 所示:

$$z_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^n x_{ij}^2}} \quad \text{—— (18)}$$

由此得到归一化处理后的标准矩阵 Z :

$$Z = \begin{bmatrix} z_{11} & z_{12} & \cdots & z_{15} \\ z_{21} & z_{22} & \cdots & z_{25} \\ \vdots & \vdots & \ddots & \vdots \\ z_{n1} & z_{n2} & \cdots & z_{n5} \end{bmatrix} (n = 47)$$

Step3: 确定最优方案和最劣方案

最优方案 Z^+ 由 Z 中每列元素的最大值构成:

$$Z^+ = (\max \{Z_{11}, Z_{21}, \dots, Z_{n1}\}, \max \{Z_{12}, Z_{22}, \dots, Z_{n2}\}, \dots, \max \{Z_{15}, Z_{25}, \dots, Z_{n5}\})$$

最劣方案 Z^- 由 Z 中每列元素的最小值构成:

$$Z^- = (\min \{Z_{11}, Z_{21}, \dots, Z_{n1}\}, \min \{Z_{12}, Z_{22}, \dots, Z_{n2}\}, \dots, \min \{Z_{15}, Z_{25}, \dots, Z_{n5}\})$$

Step4: 计算各评价对象与最优方案、最劣方案的接近程度

根据得出的最优方案和最劣方案, 计算各评价对象与最优方案和最劣方案的接近程度, 具体如公式 (19) 和 (20) 所示。

$$D_i^+ = \sqrt{\sum_{j=1}^m \omega_j (Z_j^+ - z_{ij})^2} \quad \text{—— (19)}$$

$$D_i^- = \sqrt{\sum_{j=1}^m \omega_j (Z_j^- - z_{ij})^2} \quad \text{—— (20)}$$

其中 ω_j 为第 j 个属性的权重，权重根据灰色关联度模型确定。

Step5: 灰色关联度确定指标权重

根据前文提到的灰色关联度模型中的公式（13），求出各评价对象的灰色关联系数，结合公式（21）求出各个评价对象的权值，公式（21）如下所示：

$$w_i = \frac{y_i}{\sum_{a=1}^n y_a} \quad \text{—— (21)}$$

w_i 表示第 i 个评价对象的权值， y_i 代表第 i 个评价对象的灰色关联系数。

Step6: 计算各指标与最优方案的近似度

得到各个黄桃评价指标的权重后，计算各个指标与最优方案即最优品质的黄桃的近似度，如公式（22）所示：

$$C_i = \frac{D_i^-}{D_i^+ + D_i^-}, 0 \leq C_i \leq 1 \quad \text{—— (22)}$$

其中 C_i 越接近于 1 则表明评价对象品质越优。

Step7: 根据 C_i 大小进行排序

4.3.3 问题模型求解

根据以上分析，可根据各指标与最优方案的近似度 C_i 得到最终黄桃品质的排序。以下是 2018 年冷柜 1℃、2018 年冷柜 8℃、2019 年冷柜 1℃、2019 年冷柜 8℃黄桃样本根据基于灰色关联度 TOPSIS 模型综合考虑机械指标（硬度、TSS 下降率）与人工评分（色泽、质地、味道）后所得的综合得分指数与品质排序。其中硬度、TSS 下降率、色泽、质地、味道项的数值均是原属性数据经过正向化和归一化所得的评分。

表 5 2018 年冷柜 1℃黄桃基于灰色关联度的 TOPSIS 评价模型得分

	日期	硬度	TSS 下降率	色泽	质地	味道	正理想解	负理想解	综合得分 指数	排序
6.11 批次	6.16	0.302	0.291	0.284	0.299	0.306	0.033	0.157	0.825	5
	6.21	0.302	0.369	0.265	0.286	0.276	0.025	0.172	0.872	4
	6.26	0.302	0.338	0.312	0.296	0.293	0.015	0.167	0.918	2
	7.1	0.302	0.337	0.308	0.298	0.283	0.018	0.167	0.904	3
6.14 批次	6.19	0.302	0.290	0.292	0.280	0.280	0.037	0.156	0.809	7
	6.24	0.302	0.188	0.285	0.288	0.299	0.073	0.142	0.661	9
	6.29	0.302	0.372	0.310	0.302	0.285	0.010	0.175	0.946	1

	7.4	0.302	0.334	0.267	0.262	0.271	0.035	0.164	0.824	6
6.16	6.21	0.302	0.177	0.295	0.284	0.302	0.076	0.141	0.650	10
批次	6.26	0.302	0.095	0.292	0.300	0.300	0.107	0.138	0.563	11
	7.1	0.000	0.276	0.273	0.285	0.270	0.143	0.071	0.332	12
	7.6	0.302	0.256	0.274	0.281	0.296	0.049	0.150	0.755	8

表 6 2018 年冷柜 8℃黄桃基于灰色关联度的 TOPSIS 评价模型得分

	日期	硬度	TSS 下降率	色泽	质地	味道	正理想解	负理想解	综合得分 指数	排序
6.11	6.14	0.116	0.461	0.295	0.271	0.289	0.099	0.186	0.654	6
批次	6.17	0.331	0.431	0.287	0.295	0.289	0.017	0.224	0.929	2
	6.20	0.331	0.454	0.305	0.303	0.302	0.004	0.231	0.983	1
	6.23	0.331	0.257	0.283	0.290	0.277	0.081	0.179	0.689	5
6.14	6.17	0.000	0.364	0.288	0.277	0.284	0.155	0.141	0.477	11
批次	6.20	0.331	0.158	0.300	0.288	0.289	0.118	0.162	0.579	7
	6.23	0.331	0.276	0.278	0.271	0.276	0.076	0.183	0.707	4
	6.26	0.331	0.102	0.268	0.280	0.272	0.141	0.154	0.522	8
6.16	6.19	0.331	0.280	0.294	0.291	0.296	0.071	0.185	0.724	3
批次	6.22	0.331	0.075	0.271	0.281	0.280	0.151	0.152	0.502	9
	6.25	0.331	0.034	0.291	0.308	0.305	0.165	0.152	0.479	10
	6.28	0.000	0.000	0.301	0.306	0.302	0.232	0.027	0.103	12

表 7 2019 年冷柜 1℃黄桃基于灰色关联度的 TOPSIS 评价模型得分

	日期	硬度	TSS 下降率	色泽	质地	味道	正理想解	负理想解	综合得分 指数	排序
6.14	6.17	0.200	0.146	0.188	0.165	0.178	0.049	0.106	0.683	16
批次	6.19	0.200	0.125	0.181	0.188	0.193	0.055	0.104	0.653	18
	6.21	0.032	0.131	0.175	0.178	0.183	0.093	0.054	0.366	29
	6.23	0.200	0.126	0.174	0.179	0.167	0.057	0.102	0.641	21
	6.25	0.200	0.168	0.182	0.181	0.180	0.040	0.111	0.736	12
	6.27	0.200	0.222	0.190	0.190	0.176	0.020	0.125	0.862	8
	6.29	0.200	0.102	0.181	0.180	0.175	0.065	0.099	0.603	22
	7.1	0.200	0.000	0.186	0.185	0.180	0.103	0.091	0.468	27
	7.3	0.200	0.137	0.186	0.188	0.174	0.052	0.105	0.671	17
6.17	6.20	0.200	0.121	0.193	0.188	0.191	0.056	0.103	0.645	19
批次	6.22	0.200	0.266	0.185	0.192	0.197	0.003	0.138	0.976	1
	6.24	0.000	0.229	0.191	0.188	0.195	0.091	0.091	0.499	25
	6.26	0.200	0.163	0.185	0.189	0.190	0.040	0.111	0.734	13
	6.28	0.200	0.247	0.181	0.182	0.185	0.012	0.132	0.917	6
	6.30	0.200	0.254	0.191	0.187	0.184	0.009	0.134	0.939	4

	7.2	0.200	0.121	0.181	0.189	0.193	0.057	0.103	0.645	20
	7.4	0.030	0.222	0.192	0.188	0.183	0.079	0.088	0.528	24
	7.6	0.200	0.238	0.183	0.186	0.179	0.015	0.129	0.896	7
	7.8	0.200	0.142	0.192	0.188	0.192	0.048	0.107	0.689	15
	7.10	0.000	0.174	0.189	0.189	0.195	0.097	0.070	0.419	28
6.20 批次	6.23	0.200	0.253	0.187	0.189	0.193	0.006	0.134	0.956	2
	6.25	0.200	0.000	0.187	0.185	0.196	0.103	0.092	0.471	26
	6.27	0.200	0.217	0.185	0.187	0.186	0.020	0.124	0.860	9
	6.29	0.200	0.264	0.184	0.186	0.181	0.009	0.136	0.935	5
	7.1	0.200	0.080	0.180	0.185	0.182	0.073	0.096	0.569	23
	7.3	0.200	0.249	0.188	0.192	0.193	0.007	0.133	0.947	3
	7.5	0.200	0.211	0.189	0.188	0.190	0.022	0.123	0.849	10
	7.7	0.200	0.212	0.188	0.184	0.186	0.022	0.123	0.847	11
	7.9	0.200	0.154	0.188	0.187	0.184	0.044	0.109	0.712	14

表 8 2019 年冷柜 8℃黄桃基于灰色关联度的 TOPSIS 评价模型得分

	日期	硬度	TSS 下降率	色泽	质地	味道	正理想解	负理想解	综合得分 指数	排序
6.14 批次	6.16	0.000	0.171	0.209	0.202	0.200	0.111	0.077	0.409	24
	6.18	0.229	0.156	0.200	0.194	0.180	0.051	0.123	0.707	14
	6.20	0.229	0.211	0.204	0.179	0.186	0.035	0.134	0.793	10
	6.22	0.229	0.082	0.200	0.213	0.217	0.074	0.117	0.612	17
	6.24	0.229	0.033	0.198	0.207	0.201	0.093	0.111	0.543	19
	6.26	0.229	0.063	0.191	0.200	0.185	0.084	0.110	0.569	18
	6.28	0.229	0.000	0.207	0.212	0.221	0.105	0.113	0.518	21
	6.30	0.229	0.093	0.199	0.207	0.196	0.071	0.115	0.620	16
6.17 批次	6.19	0.229	0.199	0.208	0.211	0.192	0.032	0.135	0.808	8
	6.21	0.229	0.171	0.210	0.216	0.221	0.039	0.132	0.771	12
	6.23	0.229	0.183	0.205	0.203	0.221	0.035	0.132	0.788	11
	6.25	0.229	0.262	0.185	0.170	0.183	0.032	0.146	0.819	6
	6.27	0.229	0.258	0.215	0.207	0.206	0.010	0.150	0.935	2
	6.29	0.149	0.260	0.205	0.205	0.203	0.038	0.127	0.770	13
	7.1	0.010	0.183	0.190	0.185	0.193	0.107	0.076	0.415	23
	7.3	0.073	0.239	0.153	0.146	0.171	0.087	0.098	0.530	20
	7.5	0.000	0.177	0.192	0.188	0.189	0.112	0.074	0.398	25
	7.7	0.046	0.242	0.167	0.175	0.171	0.091	0.097	0.516	22
6.20 批次	6.22	0.229	0.224	0.206	0.212	0.213	0.019	0.142	0.880	3
	6.24	0.229	0.265	0.176	0.186	0.179	0.030	0.147	0.829	5
	6.26	0.229	0.209	0.200	0.200	0.191	0.030	0.135	0.817	7
	6.28	0.131	0.241	0.188	0.200	0.182	0.051	0.114	0.690	15
	6.30	0.229	0.216	0.204	0.206	0.214	0.023	0.139	0.858	4

7.2	0.229	0.272	0.197	0.203	0.218	0.010	0.153	0.940	1
7.4	0.229	0.230	0.191	0.184	0.169	0.036	0.138	0.794	9
7.6	0.112	0.000	0.188	0.171	0.179	0.122	0.054	0.306	26

4.3.4 模型结论与“专家经验”和“普遍认知”的差异

“专家经验”给出了评价黄桃是否优质的准确数据，并采用双评分机制，考虑机械指标与人工评分二者对黄桃品质的影响，相较于“专家经验”而言，TOPSIS 模型在权重分配上给出了更加具体的分配方式，并且以给出的指标“果实硬度 $>1.5\text{kg/cm}^2$ ”和“TSS 含量下降率 $<20\%$ ”的最劣方案，向上寻找更优化的解，综合考虑 5 种评价指标对黄桃品质的影响。

“普遍认知”认为机械指标较容易满足，而人工评分要求更高，即权重更偏向于人工评分，对机械指标没有提出更高的满足要求，容易受测试员主观感受的影响，相较于“普遍认知”而言，TOPSIS 提高了机械指标的满足标准，提出了更加客观的评价方法，同时也结合人工评分，给出更加科学的评判果实品质的方法。

五、问题二的模型建立与求解

5.1 问题模型的分析

2020 年度的评测中取消了人工评分，这使得原先的“专家经验”评估方式不再完整，机械指标与人工评分的双评分机制不再适用，且基于灰色关联度的 TOPSIS 模型也缺少了色泽、质地和味道等人工评分数据，也无法对 2020 年的样本数据进行评估。在问题一中，主成分分析法提取了机械指标和人工评分的主要因素，灰色关联度模型给出了机械指标与人工评分指标之间的关联系数和权重，但是仍然没有依据观测值给出两者之间的关系式。所以在问题二中，需要根据具有现实意义的观测值，通过机械指标得到对应的人工评分，且单个人工评分数据受多个机械指标因素影响，故本文选用多元线性回归模型求解机械指标与人工评分数据之间的函数关系式。选用多元线性回归模型是因为在回归分析中，多元线性回归常常处理一种现象与多个因素相联系的问题，通过多个自变量的最优组合共同估计因变量，处理结果的实际意义较大。考虑到灰色关联度模型中的数据选择和 2020 年的机械数据，多元线性回归模型处理时，依然选择果实硬度、TSS 含量、色差、色泽、质地和味道共 6 个指标。

5.2 问题模型的建立

将质地、色泽和味道这三个人工评分指标作为因变量，果实硬度、TSS 含量、色差为自变量，探求自变量与各个因变量之间的关系式，以质地为例，多元线性回归方程设为：

$$y_{\text{质地}} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 \quad \text{—— (23)}$$

其中 x_1 , x_2 , x_3 为自变量果实硬度、TSS 含量和色差。建立模型后要进行最小二乘法进行参数估计，普通最小二乘法通过最小化误差的平方和寻找最佳函数，通过矩阵运算求解系数矩阵，具体如公式（24）所示：

$$\beta' = (X^T X)^{-1} X^T y = (\sum x_i x_i^T)^{-1} (\sum x_i y_i) \quad \text{--- (24)}$$

X 是自变量矩阵，β 为最小二乘法通过矩阵运算求解的系数矩阵。进一步，可以求解标准误差和置信区间，其中标准误差用作回归的优度指标。

$$SE = \sqrt{\frac{\sum (y - y')^2}{n - 3}} \quad \text{--- (25)}$$

置信区间 = $y' \pm t_p SE$ ，其中 t_p 是自由度为 $n - k$ 的 t 统计量数值表中的数值，n 是观察值个数，k 为因变量包括在内的变量个数，这里为 4。

5.3 问题模型的求解

利用多项式线性回归模型出色泽、质地和味道在冷柜和冷库条件下与机械指标的关联性，具体表格因篇幅过长附在附录中，因为 2020 年样本数据均是在冷柜中，所以本文从冷柜中选取了回归效果最好的三条多项式回归方程，具体如表 9、表 10 所示。

表 9 冷柜（8℃）贮藏方式下色泽与机械指标的回归方程系数

年份	贮藏方式	批次	指标	b0	b1	b2	b3
2018	冷柜 8℃	6.11	色泽	110.7351	-0.0454	-1.6291	-1.3521
2018	冷柜 8℃	6.11	质地	173.7016	-2.2151	-6.2848	-0.9687
2018	冷柜 8℃	6.11	味道	234.9486	-6.569	-9.4574	-2.0627

表 10 冷柜（8℃）贮藏方式下色泽与机械指标的回归方程检验值

年份	贮藏方式	批次	指标	r2	F	p	s2
2018	冷柜 8℃	6.11	色泽	0.9535	6.8312	0.2725	2.1028
2018	冷柜 8℃	6.11	质地	0.7309	0.9055	0.6295	7.0354
2018	冷柜 8℃	6.11	味道	0.9689	10.3953	0.2233	1.3908

其中 b0 是常数项，b1, b2, b3 为多项式系数，依次对应果实硬度、TSS 含量和色差，r2 是残差，F 是 F 检验的统计量观测值，值越大回归效果越好。p 为 F 检验的 p 值，小于 0.001 时回归效果显著，s2 为方差项。

由表 9，表 10 可知，色泽、质地和味道的回归方程残差较高，回归效果较好，机械指标和人工评分回归关系式如下：

色泽 = $110.7351 - 0.0454 \times \text{果实硬度} - 1.6291 \times \text{TSS 含量} - 1.3521 \times \text{色差}$

质地 = $173.7016 - 2.2151 \times \text{果实硬度} - 6.2848 \times \text{TSS 含量} - 0.9687 \times \text{色差}$

味道 = $234.9486 - 6.569 \times \text{果实硬度} - 9.4574 \times \text{TSS 含量} - 2.0627 \times \text{色差}$

代入 2020 年样本 A-I 数据，具体如表 11 所示

表 11 多元线性回归方程求得样本 A-I 的人工评分

	硬度	TSS	色差	色泽	质地	味道	综合
样本 A	6.23	13.2	3.8483	83.7449	73.2143	83.7449	82.2347
样本 B	1.345	12.9	9.4381	76.8973	80.5056	76.8973	80.1001

样本 C	1.69	14	9.9891	74.3448	72.2945	74.3448	75.6613
样本 D	2.99	13.8	6.9149	78.7682	73.6498	78.7682	79.0620
样本 E	1.27	12.9	11.4069	74.2387	78.7646	74.2387	77.7474
样本 F	0.51	12.8	10.2354	76.0202	82.2114	76.0202	80.0839
样本 G	2.265	13.5	6.9956	79.1806	77.0629	79.1806	80.4747
样本 H	1.03	13.4	8.8002	76.9596	78.6790	76.9596	79.5328
样本 I	0.815	14.3	6.3196	78.8573	75.9019	78.8573	79.8721

题设要求结果如下表所示：

样本	品质	样本	品质	样本	品质
样本 A	优质	样本 D	非优质	样本 G	优质
样本 B	优质	样本 E	非优质	样本 H	非优质
样本 C	非优质	样本 F	优质	样本 I	非优质

六、问题三的模型建立与求解

6.1 问题模型的分析

在前文的 TOPSIS 模型和专家经验中，对于黄桃品质的评价都受到了温度和时间两个维度的影响，对于温度而言，只有冷柜的 8℃和冷库的 1℃两种情况，所以对于温度的使用分类讨论，定性分析。因为在时间维度上，对于黄桃品质的评价是连续变化的，故可以选用时间序列类的模型进行对黄桃品质进行预测。构建黄桃随贮藏时间变化的预测模型中，本文选用指数平滑模型进行黄桃短期品质预测，选用灰色预测模型进行黄桃品质长期趋势预测，为商家推荐合适的贮藏方式和时间。根据预测模型的预测结果，可以判断出是否与“普遍认为”的结论一致，并结合环境的不同，探讨黄桃品质的变化规律，给出果农需要进行的有损测试周期，确定测试次数。数据的选择上，2018 年的数据测量间隔较大，插值误差较大，故不选用数据样本。因为 2019 年的数据中 6 月 17 日批次的样本集更广泛，进行插值处理后的数据可信度更高，且针对黄桃品质随时间和温度双维度的变化问题需要控制黄桃批次的变量，故样本集选用 2019 年 6 月 17 日在冷柜和冷库中的综合评分数据。

6.2 问题模型的建立

指数平滑模型：

为了所求预测更能符合实际情况，在移动平均法的基础上采用时间序列分析法，在数据预测移动平均的过程中赋予权值，不舍弃过去的的数据，兼容全期平均和移动平均所长，根据过去真实值的趋势情况确定误差最小的平滑系数。针对本题本文所采用的是三次指数平滑模型，模型如公式（26）所示：

$$\begin{aligned}
 s_i &= a(x_i - p_{i-k}) + (1-a)(s_{i-1} + t_{i-1}) \\
 t_i &= \beta(s_i - s_{i-1}) + (1-\beta)t_{i-1} \\
 p_i &= \gamma(x_i - s_i) + (1-\gamma)p_{i-k}
 \end{aligned}
 \quad \text{—— (26)}$$

其中, x_i 表示第 i 期的真实值, s_i 表示一次预测值, t_i 表示二次预测值, p_i 表示三次预测值, a 表示一次最优平滑系数, β 表示二次最优平滑系数, γ 表示三次最优平滑系数, 累加三次指数平滑预测公式如 (27) 所示:

$$x_i + h = s_i + h \times t_i + p_{i-k} + (h \bmod k) \quad \text{--- (27)}$$

h 表示黄桃品质往后预测的天数, 将本文所采集的黄桃品质数据代入 x_i 形成一系列预测值, 对于一次, 二次, 三次预测值的预测结果作为下一次的真实值代入 x_i 实现后续预测。

灰色预测模型:

灰色预测模型中, 利用 2019 年 6 月 17 日样本集作为原始序列, 根据公式 (28) 生成累加序列 B 。

$$B(i) = \left\{ \sum_{j=1}^i A(j) \mid i = 1, 2, \dots, n \right\} \quad \text{--- (28)}$$

累加生成新序列后, 根据公式 (29) 生成新的紧邻均值序列 C 。

$$C(i) = \frac{B(i) + B(i-1)}{2} \quad \text{--- (29)}$$

再根据公式 (30) 求出发展系数 a 和灰作用量 b 。

$$A(i) = \begin{bmatrix} -C(i), 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} \quad \text{--- (30)}$$

通过公式 (31) 求得归为矩阵 c , 其中 $B = [-C, 1]$, $Y = Bc^T$, $c^T = \begin{bmatrix} a \\ b \end{bmatrix}$ 。

$$c = \begin{bmatrix} a^T \\ b^T \end{bmatrix} = (B^T B)^{-1} B^T Y \quad \text{--- (31)}$$

得到方差比后, 根据公式 (32) 得到后续数据 $F(i)$ 。

$$F(i) = (A(1) - \frac{b}{a})e^{a(i-1)} + \frac{b}{a} \quad \text{--- (32)}$$

在 $F(i)$ 的基础上对原先数据进行累减, 还原出新的预测序列 G , 预测序列 G 在真实值 1 到 n 计做 H , 即 $H=G(1:n)$, 而原式序列和预测序列差记做 ϵ , 即 $\epsilon=A-H$, 第 i 项的 ϵ 与真实值的比为 δ , 真实值记为 $A(i)$, $\delta=|A(i)-H(i)/A(i)|$, 最后我们可以根据公式 (33) 求出相对残差 Q 。

$$Q = \frac{\sum_{i=1}^n \delta(i)}{n} \quad \text{--- (33)}$$

最小方差比 s 是通过分别计算 ϵ 和 δ 的标准差 (分别记为 $S1$ 和

S2) 并求两个结果之商的来的, 即 $s=S1/S2$, 通过计算每个残差减去残差的均值之后, 结果小于 0.6745 倍 A 的方差的概率, 我们称之为小误差概率, 即认为所预测的黄桃品质误差在可接受范围之内, 视作正确数据。

6.3 问题模型的求解

对 2019 年 6 月 17 日批次的黄桃进行指数平滑预测, 结果如下图 15 所示:

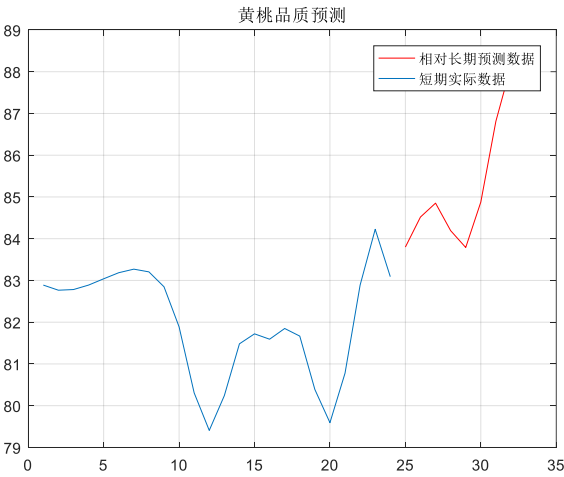


图 8 黄桃品质冷库环境指数平滑预测图

未来一周的黄桃品质预测值为: 84.852, 84.193, 83.79, 84.877, 86.814, 88.137, 87.647。

对冷柜内黄桃进行指数平滑预测, 结果如下图 16 所示:

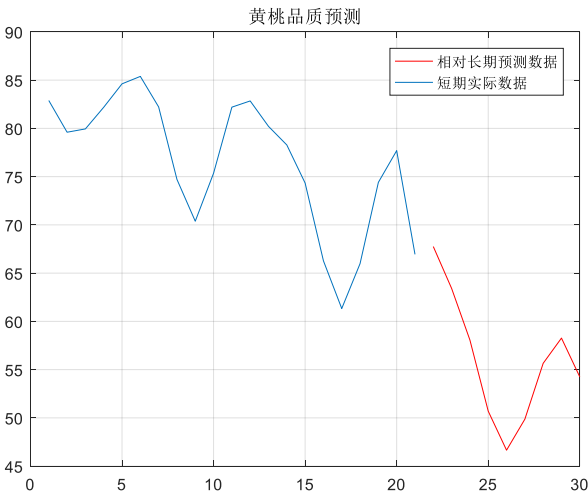


图 9 黄桃品质冷柜环境指数平滑预测图

未来一周的黄桃品质预测值为: 58.064, 50.695, 46.652, 49.866, 55.652, 58.271, 54.249。

通过指数平滑模型的接过来看, 当黄桃存储在冷库 (1℃) 时, 其品质前期较差, 后期保持较好; 当黄桃存储在冷柜 (8℃) 时, 其品质前期较好, 后期保持较差, 品质下降较快。同 “普遍认为” 一致, 在冷酷贮藏的情况下, 低温会影响黄桃代谢, 风味有所下降, 但保质期较长; 在冷柜贮藏的情况下, 黄桃代谢较

好，但保质期较短。

对冷库内黄桃进行灰色预测，结果如下图 17 所示：

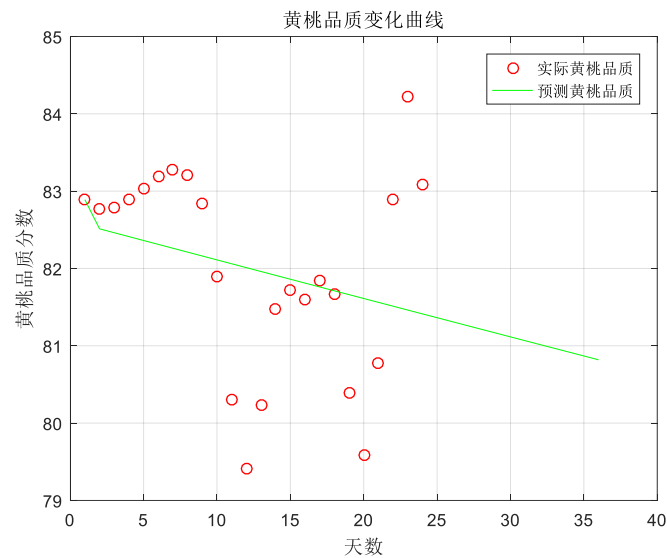


图 10 黄桃品质冷库环境灰色预测图

未来一周的黄桃品质预测值为：81.412，81.363，81.313，81.263，81.214，81.164，81.115。

对冷柜内黄桃进行灰色预测，结果如下图所示：

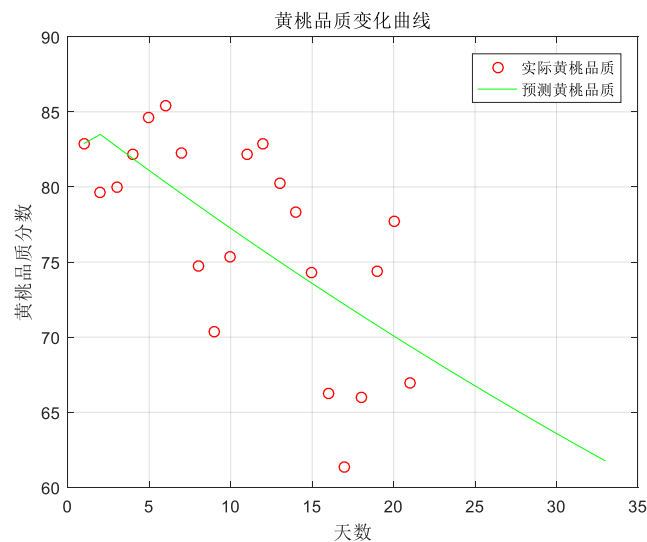


图 11 黄桃品质冷柜环境灰色预测图

未来一周的黄桃品质预测值为：67.412，66.76，66.113，65.473，64.84，64.212，63.59。

在灰色预测中，黄桃在冷库中贮存时，品质会突然下降，然后品质下降速度变缓；在冷柜贮存中，品质会突然上升，然后下将速度要远大于在冷库中贮

存的速度，品质逐渐变差。这说明合理的贮藏黄桃的方式，应该先在冷柜（8℃）中贮存 3-5 天时间进行预冷，然后在贮存在冷库中（1℃）中，冷库的低温降低黄桃代谢，使其保质期得以延长。在灰色预测中，冷库 30 天后的预测值仍有 80.2523 的评分，则认为如果按照在冷柜中预冷，在放置于冷库中存储，则会有 60 天左右的保质期（低于 80）。

对于黄桃品质的周期性变化，2019 年 6 月 11 日批次的黄桃品质表现如图 19 所示：

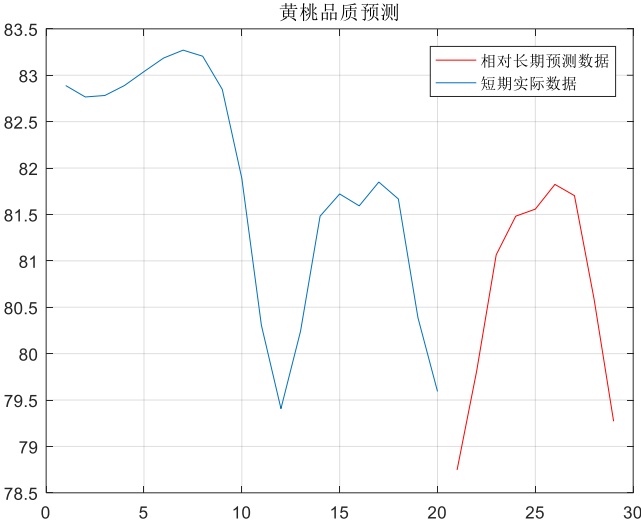


图 12 6.11 黄桃品质冷库环境指数平滑预测图

6 月 11 日数据体现冷库环境下，黄桃品质存在 7 天的周期性，成稳定下降趋势，故在由预冷的冷库情况下，冷库存储时期每 7 天检测一次。而预冷阶段只有 3 到 5 天，故预冷阶段初始时检测一次，转入冷库时检测一次，共 2 次。

七、模型评价

7.1 模型优点

1. 本文对数据进行了严谨的预处理，使得数据更为科学准确。
2. 本文采用主成分分析提取机械指标和人工评分的主要影响因素，是的模型研究复杂度降低，同时避免了认为选择变量带来的主观性。
3. 本文采用了灰色关联度模型和多元线性规划对各个指标的关系和一致性进行分析，相互验证，更有科学性。
4. 本文使用了基于灰色关联度的 TOPSIS 模型，利用后验权重作为评价基础，具有较强的可信度。
5. 本文使用指数平滑预测模型和灰色预测模型对黄桃品质从温度和时间两个维度进行考量，是结果更加全面。

7.2 模型缺点

1. 在主成分分析模型建立中，为了降低维度，减少计算复杂度而提取了主成分，对于果壳半径，部分色差数据没有进行考虑。
2. 对于环境进行定性分析，实际上，温度、湿度等环境因素是连续变量，不

能简单的定性分析。

八、参考文献

- [1] 百度百科. 桂东黄桃. <https://baike.baidu.com/item/%E6%A1%82%E4%B8%9C%E9%BB%84%E6%A1%83/19830379>. 2021. 11. 22.
- [2] 李明丽. 黄桃罐头加工工艺. 农业知识. 33: 59. 2015.
- [3] 杨 权. 降血糖血脂保健黄桃罐头的制备方法: cN106071663A[P]. 2016. 11. 09.
- [4] 张洪礼, 马玉华, 王宇, 王壮, 王立娟, 曾帆, 周俊良. 锦绣黄桃品质的影响因素及贮藏保鲜技术研究进展. 贵州农业科学. 47:91-95. 2019
- [5] 林震山, 王则金, 陈观文, 郭吴珉. 建宁黄桃最适贮藏温度研究. 福建泉州. 福建省科协第十六届学术年会福建省制冷学会分会场——福建省制冷学会 2016 年学术年会. 2016
- [6] 焦艺, 毕金峰, 刘璇, 曾目成, 阮卫红, 陈芹芹. 桃品质评价研究进展. 农产品加工(学刊). 7: 55-59, 62. 2014.
- [7] 张海英, 韩涛, 王有年, 李丽萍. 桃果实品质评价因子的选择. 农业工程学报. 22(8):235-239. 2006
- [8] 牟红梅, 于强, 李庆余, 王义菊, 姜福东, 李元军, 薛敏, 王兆龙. 基于主成分分析的烟台地区西洋梨果实品质综合评价. 果树学报. 36(8): 1084-1092. 2019.
- [9] 郭家刚, 杨松, 伍玉菡, 陈敏. 基于主成分与聚类分析的蓝莓品质综合评价研究. 食品研究与开发, 41(22): 53-60. 2020
- [10] 王丹丹, 师建华, 李燕, 张庆银, 齐连芬, 董灵迪. 基于主成分与聚类分析的辣椒主要农艺性状评价. 中国瓜菜. 34(2): 47-53. 2021
- [11] 王琴, 王建友, 韩宏伟, 李勇, 蒋江照, 巴图巴雅尔, 毛金梅, 刘凤兰. 南疆地区甜樱桃品种果实品质测定与评价. 食品工业科技. 40(8):215-220. 2019
- [12] 张海明, 邵建辉, 曹慧玲, 王淼, 朱怡凡, 朱书生, 何霞红, 马春花. 引种无核葡萄在昆明地区的表现及果实品质综合评价. 西部林业科学. 50(2):99-108. 2021
- [13] 田冰, 黄山, 孙晔, 闫宇杰, 金康康. 基于 SOM 神经网络聚类和灰度 TOPSIS 评价法的岩爆预测. 中国矿业. 30(1):188-192. 2021
- [14] 范洁, 杨世海, 武文广, 曹晓, 陈浩. 基于组合赋权的非介入式终端辨识能力灰色 TOPSIS 评价. 电力工程技术. 39(6):110-116, 131. 2020
- [15] 安雪芳. 基于熵权的 TOPSIS 评价法构建煤炭企业战略风险评价新模型及应用. 中国矿业. 23(12):37-41. 2014

附录一：

表 12 冷柜（8℃）贮藏方式下色泽与机械指标的回归方程系数

年份	贮藏方式	批次	指标	b0	b1	b2	b3
2018	冷柜 8℃	6.11	色泽	110.7351	-0.0454	-1.6291	-1.3521
2018	冷柜 8℃	6.14	色泽	9.7234	5.775	3.4306	0.7816
2018	冷柜 8℃	6.16	色泽	125.3022	-8.5313	-1.463	-0.5927
2018	冷柜 8℃	综合	色泽	55.8312	0.7074	2.0989	-0.0669
2019	冷柜 8℃	6.14	色泽	61.4952	0.4196	1.8758	-0.8232
2019	冷柜 8℃	6.17	色泽	50.5682	1.4623	2.6545	-1.0229
2019	冷柜 8℃	6.20	色泽	89.3448	3.0464	-1.3065	0.2178
2019	冷柜 8℃	综合	色泽	66.8536	2.7937	0.7525	-0.224

表 13 冷柜（8℃）贮藏方式下色泽与机械指标的回归方程检验值

年份	贮藏方式	批次	指标	r2	F	p	s2
2018	冷柜 8℃	6.11	色泽	0.9535	6.8312	0.2725	2.1028
2018	冷柜 8℃	6.14	色泽	0.5436	0.3971	0.7893	22.6059
2018	冷柜 8℃	6.16	色泽	0.9321	4.5733	0.3281	4.6201
2018	冷柜 8℃	综合	色泽	0.166	0.7299	0.5554	14.9946
2019	冷柜 8℃	6.14	色泽	0.7984	6.6022	0.0344	1.7702
2019	冷柜 8℃	6.17	色泽	0.3733	1.3897	0.323	56.66
2019	冷柜 8℃	6.20	色泽	0.4235	1.2242	0.3923	16.7424
2019	冷柜 8℃	综合	色泽	0.3052	3.6605	0.0258	24.1556

表 14 冷库（1℃）贮藏方式下色泽与机械指标的回归方程系数

年份	贮藏方式	批次	指标	b0	b1	b2	b3
2018	冷库 1℃	6.11	色泽	54.8673	7.0808	0.3552	-5.2628
2018	冷库 1℃	6.14	色泽	46.9188	8.2011	-0.4573	-1.6838
2018	冷库 1℃	6.16	色泽	-83.0622	-9.6787	19.538	-1.022
2018	冷库 1℃	综合	色泽	45.9902	0.7114	3.0883	-0.3205
2019	冷库 1℃	6.14	色泽	107.1051	-0.9564	-1.6559	-0.3979
2019	冷库 1℃	6.17	色泽	80.5634	-0.1158	0.1549	0.0197
2019	冷库 1℃	6.20	色泽	83.0138	0.1538	0.0315	-0.2501
2019	冷库 1℃	综合	色泽	85.3141	-0.6673	0.0445	-0.2913

表 15 冷库（1℃）贮藏方式下色泽与机械指标的回归方程检验值

年份	贮藏方式	批次	指标	r2	F	p	s2
2018	冷库 1℃	6.11	色泽	0.5123	0.3502	0.8104	73.7755

2018	冷库 1℃	6.14	色泽	0.9955	73.9054	0.0853	0.3475
2018	冷库 1℃	6.16	色泽	0.9902	33.5291	0.1261	0.4112
2018	冷库 1℃	综合	色泽	0.2219	1.0455	0.4108	22.4951
2019	冷库 1℃	6.14	色泽	0.4544	1.3883	0.3484	7.8674
2019	冷库 1℃	6.17	色泽	0.0063	0.0148	0.9973	5.9761
2019	冷库 1℃	6.20	色泽	0.252	0.5614	0.6634	3.1532
2019	冷库 1℃	综合	色泽	0.1553	1.5321	0.2307	5.1163

表 16 冷柜（8℃）贮藏方式下质地与机械指标的回归方程系数

年份	贮藏方式	批次	指标	b0	b1	b2	b3
2018	冷柜 8℃	6.11	质地	173.7016	-2.2151	-6.2848	-0.9687
2018	冷柜 8℃	6.14	质地	51.6614	2.4641	1.004	0.6418
2018	冷柜 8℃	6.16	质地	96.5245	-5.2451	0.2181	-0.3756
2018	冷柜 8℃	综合	质地	77.8475	-0.8499	0.5832	-0.0235
2019	冷柜 8℃	6.14	质地	29.9207	2.0435	3.7796	-0.4757
2019	冷柜 8℃	6.17	质地	39.9243	3.7556	2.6932	-0.5852
2019	冷柜 8℃	6.20	质地	120.1054	1.8947	-3.9835	0.3449
2019	冷柜 8℃	综合	质地	68.7341	3.4483	0.2275	-0.0615

表 17 冷柜（8℃）贮藏方式下质地与机械指标的回归方程检验值

年份	贮藏方式	批次	指标	r2	F	p	s2
2018	冷柜 8℃	6.11	质地	0.7309	0.9055	0.6295	7.0354
2018	冷柜 8℃	6.14	质地	0.2273	0.0981	0.9504	16.0913
2018	冷柜 8℃	6.16	质地	0.7089	0.8118	0.652	15.0538
2018	冷柜 8℃	综合	质地	0.1866	0.8414	0.4992	10.6611
2019	冷柜 8℃	6.14	质地	0.2243	0.4818	0.7092	28.5993
2019	冷柜 8℃	6.17	质地	0.4344	1.7924	0.2362	60.6599
2019	冷柜 8℃	6.20	质地	0.7617	5.3277	0.0514	10.6685
2019	冷柜 8℃	综合	质地	0.2698	3.0797	0.0457	36.6245

表 18 冷库（1℃）贮藏方式下质地与机械指标的回归方程系数

年份	贮藏方式	批次	指标	b0	b1	b2	b3
2018	冷库 1℃	6.11	质地	110.4749	1.0107	-2.7728	-0.5065
2018	冷库 1℃	6.14	质地	15.0995	9.6217	1.3042	-1.3976
2018	冷库 1℃	6.16	质地	67.1663	3.1599	-0.2349	0.1396
2018	冷库 1℃	综合	质地	50.5773	0.3888	2.5839	-0.1109
2019	冷库 1℃	6.14	质地	109.2649	-1.1732	-1.6296	-0.5415
2019	冷库 1℃	6.17	质地	85.7553	0.1026	-0.2231	0.0207
2019	冷库 1℃	6.20	质地	81.8656	0.8555	-0.1568	0.1277
2019	冷库 1℃	综合	质地	87.3629	-0.6365	-0.0948	-0.1636

表 19 冷库（1℃）贮藏方式下质地与机械指标的回归方程检验值

年份	贮藏方式	批次	指标	r2	F	p	s2
2018	冷库 1℃	6.11	质地	0.398	0.2204	0.877	4.9842
2018	冷库 1℃	6.14	质地	0.9919	41.0225	0.1142	0.5787
2018	冷库 1℃	6.16	质地	0.7889	1.2458	0.5637	4.0464
2018	冷库 1℃	综合	质地	0.1948	0.887	0.4779	9.8698
2019	冷库 1℃	6.14	质地	0.3457	0.8806	0.5107	17.749
2019	冷库 1℃	6.17	质地	0.0239	0.0571	0.9807	1.6498
2019	冷库 1℃	6.20	质地	0.555	2.0788	0.2217	0.8148
2019	冷库 1℃	综合	质地	0.0939	0.8637	0.4728	6.8079

表 20 冷柜（8℃）贮藏方式下味道与机械指标的回归方程系数

年份	贮藏方式	批次	指标	b0	b1	b2	b3
2018	冷柜 8℃	6.11	味道	234.9486	-6.569	-9.4574	-2.0627
2018	冷柜 8℃	6.14	味道	31.7936	0.6426	3.5439	0.2965
2018	冷柜 8℃	6.16	味道	74.1211	-4.7979	2.3392	-0.6189
2018	冷柜 8℃	综合	味道	76.4262	-2.3117	1.2936	-0.2403
2019	冷柜 8℃	6.14	味道	1.8988	0.776	6.1709	-1.0943
2019	冷柜 8℃	6.17	味道	42.0335	2.1738	2.6894	-0.6039
2019	冷柜 8℃	6.20	味道	96.5886	3.4083	-2.69	1.2667
2019	冷柜 8℃	综合	味道	63.768	2.6522	0.4265	0.2269

表 21 冷柜（8℃）贮藏方式下味道与机械指标的回归方程检验值

年份	贮藏方式	批次	指标	r2	F	p	s2
2018	冷柜 8℃	6.11	味道	0.9689	10.3953	0.2233	1.3908
2018	冷柜 8℃	6.14	味道	0.6715	0.6813	0.6876	16.3766
2018	冷柜 8℃	6.16	味道	0.6779	0.7016	0.6817	17.2697
2018	冷柜 8℃	综合	味道	0.5543	4.5608	0.0261	9.5306
2019	冷柜 8℃	6.14	味道	0.5325	1.8987	0.2477	20.9558
2019	冷柜 8℃	6.17	味道	0.3653	1.343	0.3355	42.3291
2019	冷柜 8℃	6.20	味道	0.5117	1.7462	0.2731	41.8006
2019	冷柜 8℃	综合	味道	0.1134	1.0656	0.3814	39.9759

表 22 冷库（1℃）贮藏方式下味道与机械指标的回归方程系数

年份	贮藏方式	批次	指标	b0	b1	b2	b3
2018	冷库 1℃	6.11	味道	115.0467	-4.1369	-1.2505	0.7169
2018	冷库 1℃	6.14	味道	-10.3882	5.796	5.1672	-0.5465
2018	冷库 1℃	6.16	味道	-8.7048	-2.5212	9.7213	-0.7426

2018	冷库 1℃	综合	味道	44.8425	-0.9151	3.5614	-0.0947
2019	冷库 1℃	6.14	味道	54.1338	1.5272	1.631	-1.0114
2019	冷库 1℃	6.17	味道	108.7946	1.7776	-2.926	-0.1574
2019	冷库 1℃	6.20	味道	74.7717	0.5355	0.3051	-0.0118
2019	冷库 1℃	综合	味道	76.811	-0.1205	0.2113	-0.0824

表 23 冷库（1℃）贮藏方式下味道与机械指标的回归方程检验值

年份	贮藏方式	批次	指标	r2	F	p	s2
2018	冷库 1℃	6.11	味道	0.519	0.3596	0.806	22.307
2018	冷库 1℃	6.14	味道	0.9721	11.6179	0.2116	1.3588
2018	冷库 1℃	6.16	味道	0.6623	0.6537	0.6959	17.199
2018	冷库 1℃	综合	味道	0.2206	1.0378	0.4139	12.8544
2019	冷库 1℃	6.14	味道	0.2978	0.7068	0.5879	12.7838
2019	冷库 1℃	6.17	味道	0.6999	5.443	0.0301	4.0066
2019	冷库 1℃	6.20	味道	0.2626	0.5934	0.6459	6.6396
2019	冷库 1℃	综合	味道	0.016	0.1351	0.9382	12.1746

附录二 源码

```
return list(map(normalization, datas))
```

```
def dataDirection_3(datas, x_min, x_max, x_minimum, x_maximum):
```

```
    def normalization(data):
```

```
        if data >= x_min and data <= x_max:
```

```
            return 1
```

```
        elif data <= x_minimum or data >= x_maximum:
```

```
            return 0
```

```
        elif data > x_max and data < x_maximum:
```

```
            return 1 - (data - x_max) / (x_maximum - x_max)
```

```
        elif data < x_min and data > x_minimum:
```

```
            return 1 - (x_min - data) / (x_min - x_minimum)
```

```
    return list(map(normalization, datas))
```

```
def entropyWeight(data):
```

```
    data = np.array(data)
```

```
    # 归一化
```

```
    P = data / data.sum(axis=0)
```

```
    # 计算熵值
```

```
    E = np.nansum(-P * np.log(P) / np.log(len(data)), axis=0)
```

```

# 计算权系数
return (1 - E) / (1 - E).sum()

def topsis(data, weight=None):
    # 归一化
    data = data / np.sqrt((data ** 2).sum())

    # 最优最劣方案
    Z = pd.DataFrame([data.min(), data.max()], index=['负理想解', '正理想解'])

    # 距离
    weight = entropyWeight(data) if weight is None else
np.array(weight)
    Result = data.copy()
    Result['正理想解'] = np.sqrt(((data - Z.loc['正理想解']) ** 2 *
weight).sum(axis=1))
    Result['负理想解'] = np.sqrt(((data - Z.loc['负理想解']) ** 2 *
weight).sum(axis=1))

    # 综合得分指数
    Result['综合得分指数'] = Result['负理想解'] / (Result['负理想解']
+ Result['正理想解'])
    Result['排序'] = Result.rank(ascending=False)['综合得分指数']

    writer = pd.ExcelWriter('C:\\Users\\76446\\Desktop\\TOPSIS 数据结
果.xlsx')

```