**Canonical Correlation between the Leaf Quality Indicators of "Moderate Aroma" Flue-cured Tobacco**

Lin MENG¹, Yuangang DAI², Chengdong WANG¹, Shusheng WANG¹, Wenjing SONG¹, Yuanhua WU¹, Yimin XI¹*

1. Tobacco Research Institute of Chinese Academy of Agricultural Sciences, Qingdao 266101, China; 2. Technological Center of China Tobacco Hunan Industrial Co., Ltd., Changsha 410007, China

**Abstract** In order to find out the correlation between tobacco quality evaluation indicators in China's traditional "moderate aroma" tobacco-producing areas and simplify the tobacco quality evaluation indicators, we evaluate the appearance quality and smoking quality of 143 flue-cured tobacco leaf samples in China's "moderate aroma" tobacco-producing areas, test the physical traits and chemical component, and analyze the canonical correlation between four quality evaluation indicators. The results show that there is significant or extremely significant canonical correlation between four evaluation indicators (tobacco smoking quality, chemical component, appearance quality and physical trait quality); the cumulative variance contribution rate of evaluation indicators is in the order of chemical component (69.17%) > appearance quality (68.76%) > physical traits (64.13%) > appearance quality is most closely related to physical traits (93.84%). The individual indicators for tobacco quality evaluation make different contribution to the correlation between quality evaluation indicators. The chemical component evaluation indicators mainly include total sugar and ratio of total sugar to betaine; sensory taste indicators mainly include aroma volume, smoke concentration, irritation and softness degree; physical trait evaluation indicators mainly include leaf weight, leaf length and leaf density; appearance quality indicators mainly include leaf organizational structure, color, maturity and identity. Studies have shown that in the large-scale eco-region, using canonical correlation analysis to simplify tobacco quality evaluation indicators is feasible.

**Key words** Flue-cured tobacco, Quality, Evaluation indicators, Canonical correlation

1 Introduction

The tobacco quality evaluation indicator includes appearance quality, physical properties, chemical component, smoking quality and safety, and it is a complex and comprehensive concept. Various quality evaluation indicators consist of different interrelated individual evaluation indicators, and the level of balance and coordination determines the use value of tobacco in the cigarette industry [1-2]. The relationship between tobacco quality evaluation indicators has been great concern among tobacco researchers. Di Huihui et al. [3] and Wang Yujun et al. [4] study the correlation between the chemical component and physical properties of flue-cured tobacco; Wei Chunyang et al. [5] and Huang Qingfen et al. [6] study the correlation between appearance quality and smoking quality of tobacco; Gao Jiahe et al. [7], Bi Shufeng et al. [8] and Chang Aixia et al. [9] study the influence of chemical component of flue-cured tobacco on smoking quality; Yang Yingming et al. [10] and Cai Xianjie et al. [11] study the correlation between tobacco appearance quality and chemical component. Although there are many results of research in this area, most researchers only analyze the relationship between two or three groups, and the research methods are limited to simple correlation [12-18] or regression analysis [14]. When researching the relationship between two groups of variables with different nature, we can employ the canonical correlation analysis which can fully reflect the intrinsic link between the variables. Bao Zichao et al. [15] study the relationship between different tobacco indicators (appearance quality, physical properties, chemical component and smoking quality), but the simple correlation analysis is still used as the method for research. Deng Xiaohua et al. [16] conduct the canonical correlation analysis of four quality evaluation indicators concerning the tobacco samples in main flue-cured tobacco growing areas of Hunan Province, and the results show that it is feasible to use canonical correlation analysis for simplification of quality evaluation indicators. Although there are many results of research in this area, few of people conduct comprehensive and systematic analysis of traditional "moderate aroma" flue-cured tobacco in main tobacco producing areas. In this study, with the 143 tobacco samples from China's typical "moderate aroma" tobacco producing areas as the materials, we evaluate the appearance quality and smoking quality and measure the physical properties and chemical component. Using canonical correlation analysis, we study the main indicators that affect the evaluation of flue-cured tobacco quality, in order to provide a reference for improving flue-cured tobacco quality evaluation system and revealing the relationship between flue-cured tobacco quality evaluation indicators.

2 Materials and methods

2.1 Materials We collected 143 C3F flue-cured tobacco samples in the typical "moderate aroma" flue-cured tobacco producing areas from 26 counties (cities) in Hunan Province, Hubei Prov-
ince, Shandong Province, Liaoning Province, Guizhou Province, Shaanxi Province, Chongqing Province and Jilin Province in 2011. There were 4 to 7 samples in each county, and the weight of each sample was about 5 kg. The main varieties are Yunyan 87, K326.

### 2.2 Experimental methods

#### 2.2.1 Determination of physical indicators

The determination of physical indicators in this experiment is based on the method of Bao Zichao et al.\[^{15}\], including single leaf quality, leaf length, leaf width, leaf thickness, leaf density, equilibrium moisture content and percentage of stem in tobacco leaf.

#### 2.2.2 Determination of appearance quality

The determination of tobacco appearance quality indicators is based on the method of Bao Zichao et al.\[^{15}\], including maturity, oil content, structure, color, identity, and color intensity.

#### 2.2.3 Determination of chemical components

The determination of chlorogenic acid, rutin and scopoletin is based on the method of Zhuang Yadong et al.\[^{17}\]; the determination of total alkaloids is based on YC/T160–2002; the determination of total sugar and reducing sugar is based on YC/T159–2002; the determination of total nitrogen is based on YC/T161–2002; the determination of potassium ion is based on YC/T173–2002; the determination of chloride ion is based on YC/T162–2002; the determination of organic acids (oxalic acid, citric acid, malic acid, linoletic acid, oleic acid and hard fatty acid) is based on the method of Jin Yongming et al.\[^{18}\].

#### 2.2.4 Evaluation of sensory quality

The evaluation standard for the sensory quality of flue-cured tobacco is based on Evaluation Methods for Sensory Quality and Style of Tobacco (Trial Version), developed by Zhengzhou Tobacco Research Institute of China National Tobacco Corporation. The specific scoring standard is based on the method of Deng Xiaohua et al.\[^{19}\]. It is completed by the national judges organized by China Tobacco Hunan Industrial Co., Ltd., and the 0–5 equidistant scale scoring method is used. The main evaluation indicators in this study include aroma status (5 points), smoke concentration (5 points), smoking strength (5 points), aroma quality (5 points), aroma volume (5 points), penetration (5 points), refinement degree (5 points), softness degree (5 points), mellowness (5 points), irritation (5 points), dryness (5 points), aftertaste (5 points) and miscellaneous odors (total scores of various odors).

### 2.3 Data processing

The canonical correlation analysis is performed using SAS 9.2 statistical analysis software, and it is analyzed according to the linear expression of standardized canonical variables.

### 3 Results and analysis

#### 3.1 Canonical correlation analysis of physical traits and smoking quality

The canonical correlation analysis results of tobacco physical traits and smoking quality and the composition of canonical variables can be shown in Table 1. The correlation coefficient of the first two canonical variables of physical characteristics and smoking quality reaches a significant level \((P<0.01)\), and the cumulative variance contribution rate reaches 64.13%, suggesting that there is a highly significant correlation between tobacco smoking quality and physical traits. From the linear expression of canonical variable composition, the amount of load on leaf weight \((P_1)\), leaf length \((P_2)\), aroma quality \((S_4)\) and aroma volume \((S_5)\) in the first set of canonical variables is relatively large, mainly showing that the score of aroma quality and aroma volume is negatively correlated with leaf weight but positively correlated with leaf length; the amount of load on leaf length \((P_3)\), smoke concentration \((S_1)\), penetration \((S_6)\) and softness degree \((S_8)\) in the second set of canonical variables is relatively large, mainly showing that the leaf length is negatively correlated with the score of smoke concentration, penetration and softness degree.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Canonical correlation coefficient</th>
<th>Canonical variable composition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical traits</td>
<td>Smelling quality</td>
<td>0.5839**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(V_1 = -1.0249P_1 + 0.4239P_2 - 0.3877P_3 + 0.3727P_4 + 0.7620P_5 + 0.1343P_6 + 0.1383P_7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(W_1 = 0.0981S_1 - 0.1148S_2 - 0.2812S_3 + 1.3312S_4 + 0.6989S_5 - 0.3832S_6 - 0.4756S_7 + 0.4061S_8 - 0.5279S_9 - 0.2331S_{10} + 0.5965S_{11} - 0.0458S_{12} + 0.33665S_{13})</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.5437**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(V_2 = 0.1806P_1 + 0.0503P_2 - 0.3069P_3 - 0.3952P_4 - 0.5212P_5 - 0.1919P_6 + 0.3266P_7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(W_2 = -0.1315S_1 + 0.5208S_2 - 0.3942S_3 + 0.3031S_4 - 0.4906S_5 + 0.8067S_6 - 0.4105S_7 + 0.8452S_8 + 0.0594S_9 + 0.4122S_{10} + 0.1657S_{11} - 0.4267S_{12} + 0.2110S_{13})</td>
</tr>
</tbody>
</table>

Note: "**" indicates the 5% significance level; "***" indicates the 1% significance level.

#### 3.2 Canonical correlation analysis of appearance quality and smoking quality

The correlation coefficients between appearance quality and the first and second set of canonical variables of smoking quality reach a very significant level \((P<0.01)\) and a significant level \((P<0.05)\) respectively (Table 2), and the cumulative variance contribution rate of the two sets of canonical variables reaches 68.76%, showing that there is a significant canonical correlation between tobacco appearance quality and smoking quality. As can be seen from the linear expression of canonical variables, there is a large load on smoke concentration \((S_2)\), smoking strength \((S_1)\), aroma volume \((S_5)\), color \((A_3)\) in the first set of canonical variables, mainly suggesting that color is negatively correlated with the score of smoke concentration but positively correlated with the score of smoking strength and aroma volume; there
is a large load on maturity \( (A_1) \), color \( (A_2) \), smoking strength \( (S_1) \) and irradiation \( (S_{10}) \) in the second set of canonical variables, suggesting that maturity is positively correlated with the score of irritation but negatively correlated with the score of smoking strength, and color intensity is negatively correlated with the score of irritation but positively correlated with the score of smoking strength.

### 3.3 Canonical correlation analysis of appearance quality and physical traits

As shown in Table 3, there is a significant correlation between physical traits and appearance quality of tobacco. The correlation coefficient of the first two sets of canonical variables reaches a very significant level \( (P < 0.01) \), and the correlation coefficient of the third set of canonical variables reaches a significant level \( (P < 0.05) \). The cumulative variance contribution rate of the first three sets of canonical variables is up to 93.14%. As can be seen from the linear expression of canonical variables, there is a large load on oil content \( (A_1) \) and leaf length \( (P_1) \) in the first set of canonical variables, reflecting the positive correlation between oil content and leaf length; there is a large load on identity \( (A_1) \), structure \( (A_2) \), color intensity \( (A_3) \), leaf weight \( (P_2) \) and leaf length \( (P_3) \) in the second set of canonical variables, mainly suggesting that leaf weight is positively correlated with identity and color intensity but negatively correlated with organizational structure, and leaf length is negatively correlated with identity and color intensity but positively correlated with organizational structure; there is a large load on maturity \( (A_1) \), oil content \( (A_2) \), color intensity \( (A_3) \), leaf density \( (P_2) \) and leaf width \( (P_3) \) in the third set of canonical variables, mainly showing that the leaf density is positively correlated with maturity and oil content but negatively correlated with color intensity, and leaf width is positively correlated with maturity and oil content but negatively correlated with color intensity.

### 3.4 Canonical correlation analysis of chemical component and smoking quality

The correlation coefficient of chemical component and the first two sets of canonical variables of smoking quality reaches a very significant level \( (P < 0.01) \), and the correlation coefficient of chemical component and the third set of canonical variables reaches a significant level \( (P < 0.05) \). The cumulative variance contribution rate of the first three sets of canonical variables is 69.17%, indicating that there is a significant correlation between smoking quality and chemical component of tobacco. From the linear expression of canonical variables (Table 4), it can be found that there is a large load on refinement degree \( (S_1) \), irradiation \( (S_{10}) \), dryness \( (S_{11}) \), total sugar \( (C_1) \), total alkali \( (C_4) \), ratio of total sugar to betaine \( (C_7) \) and stearic acid \( (C_{17}) \) in the first set of canonical variables, mainly suggesting that the score of smoke refinement degree and irritation is positively correlated with stearic acid and total sugar content but negatively correlated with total alkali content and ratio of total sugar to betaine, the score of dryness is negatively correlated with total sugar content and stearic acid content but positively correlated with total alkali content and ratio of total sugar to betaine; there is a large load on softness degree \( (S_7) \), reducing sugar \( (C_6) \) and potassium \( (C_6) \) in the second set of canonical variables, show that the score of softness degree is negatively correlated with potassium content but positively correlated with reducing sugar content; there is a large load on irritation \( (S_{10}) \) and ratio of total sugar to betaine \( (C_7) \) in the third set of canonical variables, mainly showing the
negative correlation between irritation score and ratio of total sugar to betaine.

### 3.5 Canonical correlation analysis of physical traits and chemical component

The correlation coefficient of physical traits and the first set of canonical variables of chemical component reaches a very significant level \((P < 0.01)\), and the correlation coefficient of physical traits and the fourth set of canonical variables reaches a significant level \((P < 0.05)\). The cumulative variance contribution rate of the first four sets of canonical variables reaches 88.72%, indicating that there is a significant canonical correlation between physical traits and chemical component of tobacco. From the linear expression of canonical variable composition, it can be found that there is a large load on percentage of stem in tobacco leaf \((P_1)\) and oxalic acid \((C_{11})\) in the first set of canonical variables, mainly reflecting the positive correlation between percentage of stem in tobacco leaf and oxalic acid content; there is a large load on leaf density \((P_2)\), palmitic acid \((C_{14})\) and oleic acid \((C_{16})\) in the second set of canonical variables, indicating that leaf density is positively correlated with palmitic acid but negatively correlated with oleic acid; there is a large load on leaf density \((P_2)\), leaf length \((P_3)\), total sugar \((C_9)\), reducing sugar \((C_6)\) and palmitic acid \((C_{14})\) in the third set of canonical variables, suggesting that leaf density and leaf length are positively correlated with total sugar content but negatively correlated with reducing sugar and palmitic acid content; there is a large load on leaf weight \((P_3)\), leaf length \((P_3)\), total nitrogen \((C_7)\), ratio of total sugar to betaine \((C_7)\), scopoletin \((C_{10})\) and palmitic acid \((C_{16})\) in the fourth set of canonical variables, showing that leaf weight is positively correlated with total nitrogen and ratio of total sugar to betaine but negatively correlated with scopoletin and palmitic acid content, and leaf length is negatively correlated with total nitrogen and ratio of total sugar to betaine but positively correlated with scopoletin and palmitic acid content.

### Table 4 Canonical correlation between chemical component and smoking quality of tobacco

<table>
<thead>
<tr>
<th>The first set of variables</th>
<th>The second set of variables</th>
<th>Canonical correlation coefficient</th>
<th>Canonical variable composition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chemical component</td>
<td>Smoking quality</td>
<td>0.7899 * *</td>
<td></td>
</tr>
</tbody>
</table>

### Table 5 Canonical correlation of physical traits and chemical component of tobacco

<table>
<thead>
<tr>
<th>The first set of variables</th>
<th>The second set of variables</th>
<th>Canonical correlation coefficient</th>
<th>Canonical variable composition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical component</td>
<td>Chemical component</td>
<td>0.7757 * *</td>
<td></td>
</tr>
</tbody>
</table>

Note: "*" indicates the 5% significance level; "**" indicates the 1% significance level.
3.6 Canonical correlation analysis of appearance quality and chemical component

The correlation coefficient of appearance quality and the first two canonical variables of chemical component reaches a very significant level \((P < 0.01)\), and the cumulative variance contribution rate of the two is 63.66\%, indicating that there is a significant canonical correlation between appearance quality and chemical component of tobacco. From the linear expression of canonical variable composition (Table 6), it can be found that there is a large load on identity \((A_1)\), structure \((A_2)\), total sugar \((C_1)\), potassium \((C_3)\) and ratio of total sugar to betaine \((C_5)\) in the first set of canonical variables, mainly suggesting that the score of identity is negatively correlated with total sugar and potassium content but positively correlated with ratio of total sugar to betaine, and leaf organizational structure is positively correlated with total sugar and potassium content but negatively correlated with ratio of total sugar to betaine; there is a large load on structure \((A_3)\), color \((A_6)\), total sugar \((C_1)\), total alkali \((C_4)\), palmitic acid \((C_{14})\), linoleic acid \((C_{15})\), oleic acid \((C_{16})\) and stearic acid \((C_{17})\) in the second set of canonical variables, mainly showing that leaf organizational structure is negatively correlated with total sugar, palmitic acid, total alkali and linoleic acid but positively correlated with oleic acid and stearic acid.

### Table 6 Canonical correlation analysis of appearance quality and chemical component of tobacco

<table>
<thead>
<tr>
<th>The first set of variables</th>
<th>The second set of variables</th>
<th>Canonical correlation coefficient</th>
<th>Canonical variable composition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appearance quality</td>
<td>Chemical component</td>
<td>0.6682 **</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(V_1 = -0.0985A_1 + 0.2090A_1 + 0.4029A_1 - 0.6089A_1 + 0.6884A_1 - 0.1108A_1 )</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(W_1 = 1.0690C_1 - 0.2599C_1 - 0.2290C_1 - 0.2190C_1 - 0.0192C_1 + 0.6470C_1 - 0.8258C_1 - 0.1418C_1 + 0.0971C_1 )</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(C_4 + 0.4099C_{10} - 0.3442C_{10} + 0.0568C_{10} + 0.3072C_{10} + 0.0053C_{10} - 0.1976C_{10} - 0.1854C_{10} + 0.2740C_{17} )</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.5799 **</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(V_2 = -0.0344A_1 + 0.0199A_1 + 0.3136A_1 - 0.1225A_1 - 0.6385A_1 + 0.9392A_1 )</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(W_2 = 0.7401C_1 - 0.2054C_1 + 0.3797C_1 + 0.7033C_1 - 0.1197C_1 + 0.3845C_1 - 0.2187C_1 - 0.0355C_1 + 0.4838C_1 )</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(C_4 + 0.3350C_{10} + 0.3633C_{10} - 0.1696C_{10} + 0.0447C_{10} + 0.7965C_{14} + 0.7073C_{14} - 0.8252C_{16} - 0.6841C_{17} )</td>
<td></td>
</tr>
</tbody>
</table>

Note: "**" indicates the 1% significance level.

4 Conclusions and discussions

In this paper, we analyze 43 individual indicators of four groups of tobacco quality evaluation indicators, and the results show that there are varying degrees of linkages between the quality indicators which affect the quality of tobacco. According to the cumulative contribution rate of significant canonical variables, we can judge the degree of correlation between them. From the cumulative variance contribution rate of significant canonical variables of smoking quality and the other three groups of quality evaluation indicators, chemical component \((69.17\%) > \) appearance quality \((68.76\%) > \) physical traits \((64.13\%)\), showing that chemical component is most closely related to smoking quality. Results indicate that physical traits are also most closely related to appearance quality \((93.84\%)\), and this result is consistent with the findings of Deng Xiaohua et al.\(^{[19]}\) in the producing areas of Hunan. Thus, in the same production area or different production areas, it is feasible to use the physical traits of flue-cured tobacco leaf to reflect the appearance quality and use chemical component to reflect the intrinsic quality of tobacco. In addition, the individual indicators in various quality evaluation indicators also directly or indirectly affect the quality of tobacco to varying degrees. Leaf length and single leaf quality in tobacco physical traits are closely related to sensory taste, mainly affecting smoke concentration, penetration, softness degree, aroma quality and aroma volume in sensory taste indicators. Maturity and color in tobacco appearance quality indicators are strongly correlated with sensory taste, mainly affecting aroma volume, smoking strength, irritation and smoke concentration in smoking quality indicators. Tang Ruoyun et al.\(^{[20]}\) study the middle leaf of tobacco in tobacco producing areas of South Hunan, and also find the correlation between color and maturity, and aroma volume, consistent with the findings in this paper. Total sugar, total alkali, ratio of total sugar to betaine, stearic acid, reducing sugar, potassium, ratio of total sugar to betaine in tobacco chemical component evaluation indicators have a great impact on sensory taste, mainly affecting irritation, softness degree, dryness and refinement degree in smoking quality indicators. Tissue structure, maturity, oil content, identity and color in tobacco appearance quality indicators are closely related to leaf length, leaf weight and leaf density in physical trait indicators. Percentage of stem in tobacco leaf, leaf density, leaf length and leaf weight in tobacco physical trait indicators are closely related to total sugar, reducing sugar, total nitrogen, ratio of total sugar to betaine, palmitic acid, oxalic acid, oleic acid and scopoletin in tobacco chemical component indicators. Structure and identity in tobacco appearance quality indicators are closely related to potassium, total sugar, ratio of total sugar to betaine, palmitic acid, linoleic acid, oleic acid and stearic acid in tobacco chemical component. In the evaluation of the quality of tobacco, the comprehensive analysis of all individual indicators concerning tobacco quality is time-consuming and not conducive to grasping the essence of the problem. Therefore, when testing the quality of tobacco, it is necessary to select some indicators that can reflect the quality of tobacco from a number of tobacco quality evaluation indicators, and the canonical correlation analysis in this paper just solves this problem.

References


(To page 49)
ucation, convey market awareness and competition awareness, and promote liberation of ideas of farmers and herdsmen and improvement of overall quality. With the aid of high pertinence and practical feature of vocational education, it is recommended to help farmers and herdsmen to grasp a vocational skill in the situation of low overall educational level. Besides, local government should increase investment in skill training of farmers and herdsmen, and actively organize multi-level and multi-skill training for farmers and herdsmen on a regular interval or from time to time in the manner of pre-assigned, order and entrusted training. Apart from some common skill training, it is recommended to strengthen training of farmers and herdsmen in safeguarding rights and professional quality, enhance guidance of rural labor transfer, and establish integrated training and employment service mechanism.

Agricultural, animal husbandry, and agricultural development departments and colleges and universities should enhance effort of converting technologies, and use advanced technologies to promote incremental increase of farmers and herdsmen.

References


(From page 45)