Scholar Publishing Group
International Journal of Multimedia Computing
https://doi.org/10.38007/IJMC.2020.010301
ISSN 2789-7168 Vol. 1, Issue 3: 1-10

Tobacco Resource Allocation Based on Factor Analysis of Mathematical Modelling——From the Perspective of Tobacco Farmers

Jing Huang[1], Liang Guo[2], Lingjie Xu[2], Yunfei Ma[3], Dongsheng Mu[2] and Bin Wei[2]

[1]*Guizhou Tobacco Company, Guiyang 550002, China*

[2]*China National Tobacco Corporation Guizhou, Guiyang 550001, China*

[3]*Guizhou Academy of Tobacco Sciences, Guiyang 550081, China*

Keywords: Factor Analysis, Resource Allocation, Tobacco Farmers and Tobacco Leaves

Abstract. Using the factor analysis method of mathematical modeling, the agricultural development level and farmers' income level, tobacco resource allocation mode and factor input mode, tobacco production and management subject and production organization mode, tobacco planting benefit and farmers' income level in typical areas of Guizhou Province were investigated. This paper analyzes the factors and contributions affecting the development status of tobacco industry, and finds out the problems existing in resource allocation mode, factor input mode, tobacco planting scale level and tobacco production organization mode. Finally, collect and sort out the resource allocation data of tobacco planting subjects in 33 districts and counties of flue-cured tobacco counties (districts), and analyze the characteristic factors of tobacco resources.

1. Introduction

The transformation and development of traditional tobacco leaf is a comprehensive subject, which involves the in-depth integration of system, technology, organization, management and so on. The improvement of resource allocation efficiency of tobacco industry is based on the development of traditional agriculture. This is a comprehensive set of reform and innovation, involving the in-depth integration of production relations and productivity. Compared with the simple transformation and development of the traditional tobacco industry, it is more advanced, far-reaching and modern, which not only provides new opportunities for the transformation and upgrading of the traditional tobacco industry, but also creates new opportunities for the investment of the tobacco industry.

At present, the Food and Agriculture Organization of the United Nations has formulated agricultural product development indicators. The Ministry of Agriculture and some local governments (such as Yunnan Province and Fujian Province) have also issued coffee industry development indicators and tea industry development indicators. A large number of enterprises, experts and scholars in the agricultural field at home and abroad are referring to these indicators to evaluate the development trend of the industry, and use them to adjust the scale and quality of the industry and agricultural products, so as to make the industry reach the optimal state (the maximum value of the input-output ratio).

In this paper, the importance and significance of studying the efficiency of large-scale tobacco planting by mathematical methods are expounded through literature review, and the concept and model of factor analysis method are introduced. By using this model, the factors affecting the development of tobacco industry and their contributions were analyzed, and the existing problems in resource allocation mode, factor input mode, tobacco planting scale level and tobacco production organization mode were found out. Finally, the data of tobacco resource allocation in 33 districts and counties of flue-cured tobacco county (city) were collected, and the factor analysis of tobacco resource characteristics was carried out.

2. Literature Review

Under the guidance of the high-quality development of tobacco industry, this paper reviews the related literature of domestic tobacco industry resource allocation. Natural resource endowment of a place (such as climate, temperature, humidity, etc.). Decided its industrial development direction [1]. Dazhou has natural conditions suitable for planting burley tobacco, obvious tobacco leaf style characteristics, preliminary development scale and overall good market prospect. However, compared with other advanced producing areas, Dazhou flue-cured tobacco production has no advantages in economic conditions, social environment, planting scale and production quality. On the contrary, the overall development level is at least five years behind. The average net income per mu and average household income of flue-cured tobacco growers will be much higher than other crops in the short term. However, from the perspective of long-term strategy, the comprehensive index and index of production scale, quality and development of flue-cured tobacco in Dazhou cannot catch up with other producing areas. It is difficult to gain a competitive advantage in the market. In the end, all kinds of resources such as manpower, material resources and financial resources will be wasted and huge economic losses will be caused, and the embarrassing situation of "cutting more and rushing less" cannot be avoided.

Pingdingshan City encountered the problems and contradictions of market economy in the tobacco production process, and considered that large-scale planting (50 mu per household is the best) was the only way out for tobacco production under the current policy system [3]. For the quality stability requirements and outstanding problems of tobacco raw materials, the cigarette industry put forward suggestions of continuous planting, moderate concentration, large-scale production and unified management [4]. It is believed that it is of great significance to popularize and apply new technologies, standardize production, reduce fluctuation range of tobacco leaf quality and realize agricultural mechanization. A survey of five different flue-cured tobacco planting modes in six towns of Guiyang showed that the large-scale planting mode was superior to the decentralized planting mode in cultivation management level, tobacco quality and yield performance, and the best management mode was tobacco companies+employees+science parks [5]. It is suggested to control large-scale tobacco planting and develop moderate scale tobacco planting.

On the basis of constructing the general model of agricultural production function at the farmer level, according to the principle that the factor marginal product is equal to the market value, this paper explores the optimal input of agricultural factors [6]. In order to truly describe the production decision-making environment faced by farmers, the capital constraint condition of farmers' family income is introduced in the analysis of agricultural input [7], the wage gap model under the urban-rural dual structure is introduced in the analysis of labor input, and the land under the household contract management system is divided into initial conditions in the analysis of land factor input [8]. Based on the extended analysis of the general model, this paper probes into the reasons of insufficient incentive or unsmooth circulation of capital, labor and land in agricultural production, and puts forward some policy suggestions on improving factor incentive and resource allocation.

It can be seen from the literature review that more and more scholars use various methods to study the efficiency of tobacco leaf scale planting, but most of them evaluate the efficiency of tobacco leaf scale planting in different regions from a macro perspective. From the microscopic point of view, the efficiency of tobacco planting in large scale is very low. No scholars found that, based on the survey data of a large sample of tobacco farmers, the input method and output method were used to make an empirical study on the scale planting efficiency of tobacco leaves from different angles of tobacco farmers and tobacco companies. Most of the researches on the influencing factors of tobacco leaf scale planting use qualitative methods, and few use quantitative methods to analyze them.

3. Factor Analysis Based on Mathematical Modeling

3.1. The Concept of Factor Analysis

Factor analysis refers to the statistical technique of extracting common factors from variable groups. This method was first proposed by British psychologist C.E. Spearman. He found that students' scores in various subjects have certain relevance. Students with good results in one subject often have better scores in other subjects, thus deducing whether there are some potential common factors, or some general intelligence conditions, that affect students' academic performance. Factor analysis can find hidden representative factors in many variables, classify variables with the same essence into one factor, reduce the number of variables, and can also test the relationship between variables

Factor analysis is a multivariable statistical analysis method that starts from studying the internal correlation of variables and reduces some variables with complicated relationships to a few comprehensive factors. Its basic idea is to classify the observed variables and classify them into the same category with higher correlation, i.e. closely related variables, while the correlation between different categories of variables is lower, so each category of variables actually represents a basic structure, i.e. common factors. The problem under study is to try to describe each component of the original observation by using the sum of the least number of unmeasurable so-called common factor linear functions and special factors.

The factor analysis model is described as follows:

(1) X=(x1, x2, ..., xp)₵ is an observable random vector, the mean vector E(X)=0, the covariance matrix Cov(X)=∑, and the covariance matrix ∑ is equal to the correlation matrix r (as long as the variables are standardized).

(2) F=(F1, F2, ..., Fm)₵ (m<p) is an unmeasurable vector, whose mean vector E(F)=0 and covariance matrix Cov(F) =I, i.e. the components of the vector are independent of each other.

(3) e=(e1, e2, ..., ep)₵ and F are independent of each other, and E(e)=0, the covariance matrix

$\sum$ of E is diagonal matrix, that is, the components E are independent of each other, then the model:

x1 = a11F1+ a12F2 +…+a1mFm + e1

x2 = a21F1+a22F2 +…+a2mFm + e2

………

xp = ap1F1+ ap2F2 +…+apmFm + ep

It is called factor analysis model, because the model is aimed at variables and the factors are orthogonal, it is also called R-type orthogonal factor model.

Its matrix form is: x = af+e.

Among them:

x=, A=, F=, e=

Here,

(1) m £ p;

(2) COV (f, e)=0, i.e. f and e are irrelevant;

(3) d (f) = im, i.e. F1, F2, ..., Fm are uncorrelated and the variance is 1;

D(e)=, i.e e1, e2, ..., ep are not correlated and have different variances.

We call f the common factor or latent factor of x, matrix a the factor load matrix, and e the special factor of x.

A = (aij), aij is factor load. Mathematically, it can be proved that the factor load aij is the correlation coefficient between the I variable and the J factor, which reflects the importance of the I variable on the J factor.

$$X = \left(X_{1,}X_{2,} \dots, X_P\right)^T$$

Factor analysis is a statistical method to simplify and analyze high-dimensional data. Assuming that the P-dimensional random vector satisfies $X = \mu + A\bar{f} = \bar{e}$, $\vec{f} = \left(f_{1,}f_{2,} \dots, f_q\right)^T$ is a q-dimensional random variable, $q \le p$, satisfy $E\vec{f} = 0, E\vec{f}\vec{f}^T = \vec{I}_q$, its weight f_i called common factor, it works for every component of x. $\vec{e} = \left(e_{1,}e_{2,} \dots, e_p\right)^T$ is a P-dimensional unobservable random vector that satisfies.

$$E\vec{e} = 0, E\vec{e}\vec{e}^T = \mathrm{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_P^2) \stackrel{def}{\Rightarrow} \sum$$

And $E\vec{f}\vec{e}^T = 0$, the component of e $\vec{e}_i$ called special factor, it is only for the component of x $\vec{X}_i$ it works.

μ and A are parameter matrices. If x satisfies the above formula, then the random vector x is said to have a factor structure. At this time, easy to calculate:

$$\mathrm{Var}(X) = \overrightarrow{A}\overrightarrow{A}^T + \sum$$

Matrix a is called factor load, its elements a_{ij} is the i-th component $\vec{X}_i$ at factor j f_i load on. Remember $h_i^2 = \sum_{j=1}^q a_{ij}^2$, there are $\mathrm{Val}(X_i) = h_i^2 + \sigma_i^2$, thus it can be seen, h_i^2 reflects the common factor pair. The impact of $\vec{X}_i$, known as common factor pair $\vec{X}_i$ of "Contribution". $h_i^2 \ge \sigma_i^2$, indicates a common factor pair $\vec{X}_i$ of the influence is greater than the special factor. The impact of $\vec{e}_i$, as can also be seen h_i^2 reflects the weight $\vec{X}_i$ public factor f_i degree of dependence.

On the other hand, for a given common factor f_i, remember $g_i^2 = \sum_{i=1}^p a_{ij}^2$, called common factor f_i contribution to x. The greater the value of g_i^2, reflects the common factor f_i greater the impact on x, so g_i^2 is a measure of the importance of public factors.

The task of factor analysis is to sum the correlation matrices of x $\sigma_1^2, \sigma_2^2, \ldots, \sigma_P^2$ set out, through the orthogonal rotation with the largest variance, each column of matrix A is found, and the corresponding "contributions" are ordered: $g_1^2 \geq g_2^2 \ldots \geq g_q^2$.

F1, F2, ..., Fm in the model are called principal factors or common factors, which are common factors in the expressions of all the original observed variables and are independent and unobservable theoretical variables. The meaning of public factors must be determined by combining the practical significance of specific problems. E1, e2, ..., ep are called special factors, which are unique to the component Xi (I = 1,2, ..., p) of vector x, and are independent of each other among the special factors and between the special factors and all common factors. The elements (aij) in the load matrix A in the model are factor loads. Factor load aij is the covariance of xi and Fj, and is also the correlation coefficient of xi and Fj, which indicates the degree to which xi depends on Fj. Aij can be regarded as the weight of the I variable on the common factor J. The larger the absolute value of aij (|aij| £ 1), the greater the dependency between xi and Fj, or the greater the load of common factor Fj on xi. In order to obtain the economic explanation of the result of factor analysis, two statistics in factor load matrix A are very important, i.e. the common degree of variables and the variance contribution of common factors.

The square of the elements in row I of the factor load matrix A is recorded as hi2, which is called the degree of commonality of variables xi. It is the contribution of all common factors to the variance of xi and reflects the influence of all common factors on variable xi. Hi2 indicates that the I-th component xi of x is highly dependent on each component F1, F2, ..., Fm of f.

The square of each element in column j (j =1, 2, ..., m) of factor load matrix a is recorded as gj2, which is called the variance contribution of common factor Fj to x. Gj2 represents the sum of the variances provided by the j-th common factor Fj for each component Xi (I = 1,2, ..., p) of x, which is an index to measure the relative importance of the common factor. The larger gj2 is, the greater the contribution of the common factor Fj to X, or the greater the influence and effect on X. If all gj2 (j =1 1,2, ..., m) of factor load matrix a are calculated and sorted according to size, the most influential common factor can be extracted.

After the establishment of the factor analysis model, another important function is to apply the factor analysis model to evaluate the position of each sample in the whole model, that is, to conduct comprehensive evaluation. For example, after the factor analysis model of regional economic development is established, we want to know the economic development of each region and classify the regional economy into which regions are developing faster, which are moderately developed and which are slower. At this time, it is necessary to express the common factors by linear combination of variables, that is, to estimate its factor score by various index values of the regional economy.

Let the linear combination of common factor f represented by variable x be:

$$Fj = uj1\ xj1 + uj2\ xj2 + \ldots + ujpxjp \quad j=1, 2, \ldots, m$$

This formula is called the factor score function and is used to calculate the common factor score for each sample. If m=2 is taken, the factor scores F1 and F2 of each sample can be calculated by substituting the P variables of each sample into the above formula, and the factor score scatter plot can be made on the plane, so that the samples can be classified or the original data can be further studied.

However, the number m of equations in the factor score function is smaller than the number p of variables, so the factor score cannot be accurately calculated and can only be estimated. There are many methods to estimate factor scores, including regression estimation, Bartlett estimation and Thomson estimation. Here, we compare it with the leading planting industry, with the emerging industry, and apply regression estimation method to cooperative service satisfaction.

3.2. Analysis of Tobacco Resource Allocation

Firstly, we get the average of all the indicators of planting subjects in the primary indicators of tobacco production resource allocation in these 33 districts and counties of Guizhou Province, which are secondary indicators such as age structure, educational background structure, years of tobacco planting, average labor force per household, planting scale, average net income per mu, average net income per household, per capita net income, comparison with leading planting industries and comparison with emerging industries. These 33 counties are Qixingguan District, Dafang County, qianxi county, Jinsha County, zhijin county, Nayong County, Weining County, Hezhang County, Bozhou District, Huichuan District, Tongzi County, Shuyang County, Zheng 'an County, Daozhen County, Wuchuan County, Fenggang County, Meitan County, yuqing county City, Xingyi City, Pu 'an County, Anlong County and Xing Secondly, the average number of planting subjects in the primary indicators of tobacco production resource allocation in 33 counties of Guizhou Province was summarized, and finally, the planting subjects in the primary indicators of tobacco production resource allocation in 33 counties of Guizhou Province were analyzed.

The results of investigation and analysis show that the age structure of tobacco production resources allocation in these 33 districts and counties of Guizhou Province is the largest in Qixingguan District, and its average age structure is 44.1; Among the educational structure of tobacco production resources allocation in these 33 districts and counties of Guizhou Province, Bozhou District is the largest, and its educational structure is 1.99; In the allocation of tobacco production resources in these 33 districts and counties of Guizhou Province, Xingren City ranks first with 1.87 per household labor force; Among the 33 districts and counties in Guizhou Province, Hezhang County ranks first with the planting scale of 52.02. Among the average net income per mu of tobacco production resources allocation in these 33 districts and counties in Guizhou Province, zhenyuan county ranked first with an average net income of 4147.23; Among the average net income of tobacco production resources allocation in these 33 districts and counties of Guizhou Province, Hezhang County ranked first with the average net income of 219,652.42; Among the per capita net income of tobacco production resources allocation in these 33 districts and counties of Guizhou Province, Pu 'an County ranked first with 113,739.14 per capita net income; In the comparison of tobacco production resources allocation with the leading planting industry in these 33 districts and counties of Guizhou Province, Kaiyang County occupies the first place in comparison with the leading planting industry with 30,407.91; In the comparison between the allocation of tobacco production resources and emerging industries in these 33 districts and counties of Guizhou Province, Kaiyang County ranked first in comparison with emerging industries with 38,154.47.

Secondly, the average number of each index of planting subjects in the primary index of tobacco production resource allocation in these 33 districts and counties of Guizhou Province was standardized. Then, the tobacco production data collected from 33 counties (or districts) in the whole province were substituted into SPSS software, and the weights of the primary and secondary indicators (planting subjects) of tobacco production resources allocation in Guizhou were calculated by factor analysis method, which was shown in the following table.

Table 1: Primary and secondary indicators and their weights of tobacco production resources allocation in Guizhou (main part of planting)

Primary indicator	Secondary index	Indicator unit	Indicator weight (%)
Planting main body (19.38%)	age structure	Years old	0.54
	Educational background structure	points	0.91
	Years of planting tobacco	Year	0.28
	Average household labor force	A	2.53
	Planting scale	Mu	3.16
	Average net income per mu	Element	2.52
	Average net income per household	Element	0.92
	Per capita net income	Element	0.91
	Compared with the leading planting industry	points	2.28
	Compared with emerging industries	points	2.59
	Satisfaction with Cooperative Services	points	2.74

Finally, the average number of each index of planting subjects in the first-class index of tobacco production resource allocation in 33 districts and counties of Guizhou Province is standardized and multiplied by the weight, and the resource allocation score is calculated.

4. Results and findings

Through the factor analysis method of mathematical model, the resource allocation of planting subjects in the first-class index of tobacco production resource allocation in 33 counties of Guizhou Province was analyzed, and the scores of tobacco production resource allocation in 33 counties of Guizhou Province were obtained.

Table 2: Score of planting subject index of tobacco resource allocation in 33 counties of Guizhou Province

Area	Score of resource allocation
zhenyuen kounty	0.157674793862706
hezhang county	0.157603510033291
Pu 'an County	0.152344586660932
bozhou district	0.140487907142109
dafang county	0.137434191810517
Xingren city	0.132801529961329
Weining county	0.132115511124642
qi xing guan district	0.129220271275851
shuicheng county	0.126838530388711
suiyang county	0.126600338745521
zhenfeng county	0.124381031078090
Yuqing county	0.121531120874496
zhijin kounty	0.121377980118592
Xingren city	0.121376930542311
xixiu district	0.121223575968544
Meitan County	0.120455142622982
zitongdong railway station	0.119218037467791
zhengan county	0.118829733221934
anlong county	0.117409049386691

qianxi kounty	0.116717150410966
Yanhe county	0.116688212060296
kaiyang county	0.116522688953369
huichuan district	0.116513269688239
pingba qu	0.116512779363369
Shibin	0.116511984563269
na yong	0.115952127788832
Panzhou city	0.115409993689371
dejiang country	0.115348871972316
fenggang county	0.113517295602891
Daozhen County	0.112650907385507
Wuchuan county	0.11233825116395
Weng 'an County	0.111154995558135
jinsha county	0.106582866381137

In the resource allocation of planting subjects in the first-level index of tobacco production resource allocation, the resource allocation of planting subjects in the first-level index of tobacco production resource allocation in 33 districts and counties of Guizhou Province was analyzed.

Table 3: Proportion of the most effective indicators for resource allocation of the main planting bodies in 33 counties of Guizhou Province

Area	The most efficient indicator of resource allocation efficiency	Proportion
kaiyang county	Average household labor force	22%
dafang county	Planting scale	23%
qianxi kounty	Average household labor force	23%
zhijin kounty	Planting scale	27%
qi xing guan district	Average household labor force	23%
hezhang county	Planting scale	24%
Weining county	Average household labor force	22%
jinsha county	Average household labor force	26%
Wuchuan county	Average household labor force	25%
bozhou district	Average household labor force	21%
zitongdong railway station	Average household labor force	24%
suiyang county	Average household labor force	22%
zhengan county	Average household labor force	24%
fenggang county	Average household labor force	25%
Meitan County	Average household labor force	22%
Yuqing county	Average household labor force	23%
Daozhen County	Average household labor force	23%
Panzhou city	Average household labor force	24%
shuicheng county	Average household labor force	24%
Xingren city	Planting scale	24%
anlong county	Planting scale	27%
zhenfeng county	Average household labor force	25%
Pu 'an County	Planting scale	24%
Weng 'an County	Average household labor force	24%
pingba qu	Average household labor force	26%
na yong	Average household labor force	24%

dejiang country	Average household labor force	24%
Yanhe county	Average household labor force	24%
huichuan district	Average household labor force	24%
Shibin	Average household labor force	25%
Xingyi city	Planting scale	23%
zhenyuen kounty	Planting scale	24%
xixiu district	Labor force per household	24%

5. Conclusions

Factor analysis combined with the development trend of agricultural industrialization and the development index of agricultural industry can provide basis for the government and industry authorities to understand and master the development trend of agricultural industry, understand the operation status of agricultural industry market, and formulate relevant industrial policies and development plans. Visualize, intelligently diagnose and predict agricultural development indicators, take agricultural development indicators as a simple and feasible evaluation standard, establish an evaluation system of industrial development level, find out shortcomings of agricultural industry development, and provide scientific basis for optimizing agricultural resource allocation, improving system and mechanism, and strengthening policy regulation. Provide business information for producers and operators of agricultural products, plant marketable agricultural products, choose business varieties and acquisition opportunities, maintain reasonable inventory, and improve the utilization rate of funds; To provide scientific basis for market managers to further improve their management level and service quality.

The significance of this research lies in studying the evaluation indexes of tobacco resources through factor analysis, establishing evaluation standards, innovating evaluation methods, quantitatively analyzing and scientifically evaluating the development status and level of tobacco industry, which is conducive to objectively and accurately grasping the current situation of the industry and enhancing the pertinence of tobacco work. Secondly, aiming at the problems existing in tobacco production resource allocation, planting scale and organization mode, according to the methods formed by this project, the strategies to solve the problems are put forward on the basis of analysis and evaluation, and the paths, methods, objectives and schemes to solve the problems are formed on the basis of pilot verification, so as to continuously improve the efficiency of tobacco production resource allocation, optimize planting scale and organization mode and improve the benefits and efficiency of industrial development.

Acknowledgements

In 2019, China National Tobacco Corporation Guizhou Province Company Science and Technology Project "201930 Research on Improving the Allocation Efficiency of Guizhou Tobacco Production Resources Based on High Quality Development"

References

[1] Xie J 2017. On the development prospects of tobacco science and technology in China. Acta Tabacaria Sinica 23 (3): 1-7

[2] Yang L Y, Wang L T, Ma J H 2017 Effects of light quality on growth and development, photosynthetic characteristics and content of carbohydrates in tobacco (Nicotiana tabacum, L.)

plants. Photosynthetica 55 (3): 467-477. https://doi.org/10.1007/s11099-016-0668-x
[3] Zeng S, Li Z J 2018 Evaluation of self-development ability of tobacco farmers' cooperative in Hunan Province. Acta Tabacaria Sinica 2018 24 (1): 100-105
[4] Berk J B, Stanton R, Zechner J 2010 Human capital, bankruptcy, and capital structure. The Journal of Finance 65 (3): 891-926. https://doi.org/10.1111/j.1540-6261.2010.01556.x
[5] Mehran H 2005 Technical talents compensation structure, ownership, and firm performance. Jouranal of Financial Economies (38): 163-184. https://doi.org/10.1016/0304-405X(94)00809-F
[6] Grund C 2018 The dispersion of employees' wage increases and firm performance. ILR Review 61 (4): 485-501. https://doi.org/10.1177/001979390806100403
[7] Bickers T C 2010 obacco growers wrestle with acreage decisions. Southeast Farm Press
[8] Roy R 2009 Tobacco production is family tradition for Carolina grower. Southeast Farm Press