ISSN: 2688-9323

# On Production Efficiency and Influencing Factors of Smart Agriculture

## -- Based on DEA Super Efficiency Algorithm and To-bit Model

Zehua He*, Yide Shu, Yujie Zhang, Zehang Tao

School of Statistics and Applied Mathematics, Anhui University of Finance and Economics, Bengbu 233041, China

*hezehua_aufe@163.com

## Abstract

**As China is a largely agricultural country, the long-term balanced and stable development of agriculture will have a direct impact on the national economy and social stability, and now, with the improvement of agricultural production, how to carry out production activities in a highly efficient mode in the process of agricultural development has become a growing concern. With the booming development of information technology in China, the concept of "smart agriculture" has emerged as a result of combining technology and agriculture in agricultural production. Based on the super-efficient DEA algorithm and To-bit model, this thesis investigates the impact of production factors on the production efficiency of smart agriculture in the Yangtze River Delta and Northeast Chiand applies it to the whole country, and then makes suggestions on how to develop smart agriculture in a deep and high level with the research results.**

## Keywords

**Smart Agriculture; Super-efficient DEA Algorithm; To-bit Model; Agricultural Production Efficiency.**

## 1. Introduction

Since ancient times, agriculture has been a basic industry supporting the construction and development of the national economy. In recent years, with the development of social economy and continuous innovation of science and technology, smart agriculture has emerged. In order to ensure the smooth development of smart agriculture, the state attaches great importance to smart agriculture or smart agriculture industrialization and has given a lot of policy support. This shows that the research and promotion of "smart agriculture" is in line with the major needs of China's modern agricultural development and is an important direction of national support. For three consecutive years from 2017 to 2019, the Central Government's No. 1 document has provided guidance on the prospects of smart agriculture. In 2017, the No. 1 document of the Central Government clearly stated the main line of "structural reform on the supply side of agriculture", proposed to accelerate scientific and technological research and development, implement smart agriculture projects, and promote agricultural Internet of Things and agricultural equipment intelligence. 2019 will accelerate the breakthrough of key core technologies in agriculture as the key content, cultivate a number of agricultural strategic Science and technology innovation forces, and promote independent innovation in the fields of biological seed industry, heavy agricultural machinery, smart agriculture, green inputs, etc. In this context, it is important to analyze the development situation of smart agriculture in China, improve the development dilemma of smart agriculture, and study the production efficiency and influencing factors of smart agriculture, which is of great significance to improve the level of smart agriculture in China.

At present, domestic scholars have elaborated the development prospect and development dilemma of smart agriculture from different levels, and conducted in-depth research on smart agriculture. In the study of agricultural production efficiency and influencing factors, the main research methods used by some scholars are super-efficient SBM model [1], super-efficient DEA model and Malmquist index, principal component analysis, TOPSIS entropy weighting method, Delphi method, gray correlation analysis, etc.; in the study of regional agricultural production efficiency and the development situation of agricultural production, scholars have put forward many valuable In the study of regional agricultural production efficiency and agricultural production development, scholars have put forward many valuable problems and solution strategies. In the study of agricultural production efficiency in Hunan Province, Xu Zhengkang proposed to use the total output value of agriculture, forestry, animal husbandry and fishery as the output variable and the total power of agricultural machinery, land input, irrigation input and fertilizer input as the input variables [2]; in the evaluation of agricultural production efficiency in Shanxi Province, Shi Wenwen et al. analyzed the empirical results longitudinally as well as horizontally, and proposed to improve the land transfer rate[3], improve the service system of land management right transfer and other related In a comparative study of agricultural production efficiency in the Yangtze River Delta region, Meng Wei's main measurement method is mathematical and the pure technical efficiency of the agricultural production system in the Yangtze River Delta region is chosen as the object of measurement [4]. Qian Chang [5] makes relevant suggestions for the development of organic agriculture in the Yangtze River Delta region in the new era, for example, strengthening organic ecological cooperation to achieve multi-party synergy; rational use of emerging technologies to enhance economic efficiency, etc. Shao Yaochun analyzed the real-life dilemmas faced at present: low efficiency of traditional agriculture scale, training of traditional agriculture technology out of practice, and insufficient support from the grassroots government [6]. And suggestions are made from three aspects: individual, government, and society. In a measurement analysis of agricultural production efficiency in Northeast China, Zhou Yanhong et al. used the super-efficient DEA model and Malmquist index to conclude that the phenomenon of input redundancy exists in the three Northeastern provinces, and it is necessary to continue to innovate and apply and promote new agricultural technologies, adjust the agricultural structure, and carry out scientific planning and integrated arrangements [7].

Based on the above background, existing studies on smart agriculture and agricultural production efficiency and influencing factors are mainly from multiple perspectives, methods and aspects, which have laid a solid foundation for this paper, but there are not many studies on smart agricultural production efficiency and influencing factors. Based on this, this paper selects the Yangtze River Delta region and the northeast region as the research objects, and firstly speculates the factors that may improve the production efficiency of smart agriculture and conducts regression analysis to screen out the relevant factors and conducts descriptive statistics and principal component analysis method, and then uses Stata to establish the super-efficiency DEA model to further derive the super-efficiency value. Also based on CRITIC weighting method and fuzzy Borda evaluation, specific scores of efficiency are derived, after which the key factors of production efficiency of smart agriculture are modeled based on To-bit model.

ISSN: 2688-9323

## 2. Research Methods, Index Selection, and Data Sources

### 2.1. DEA Super-efficiency Algorithm

Data Envelopment Analysis (DEA) is a non-parametric efficiency evaluation method, which was proposed by American operations researcher in 1978 and the model is a data analysis method that uses linear programming to evaluate the efficiency of multiple input and output indicators. The DEA model does not rely on the subjective setting of input and output index weights in the selection of indicators, nor does it require quantitative processing of data, and the model results are not affected by changes in input units. It can be said that DEA model is widely used and recognized by academics in the study of production efficiency, and many scholars have improved and optimized the model. The super-efficiency DEA model has n decision units, one input unit and one output unit, and the formula for calculating the super-efficiency value of the first decision unit is

$$min\theta$$

$$\text{s.t.}\begin{cases}\sum\limits_{\substack{\sigma=1\\ \sigma\neq\sigma_0}}^{n}\varphi_\sigma X_\sigma+S^-=\theta X_0\\ \sum\limits_{\substack{\sigma=1\\ \sigma\neq\sigma_0}}^{n}\varphi_\sigma X_\sigma-S^-=\theta Y_0\\ \varphi_\sigma\geq 0,\sigma=1,2,3,\ldots n\\ \quad S^+\geq 0,S^-\geq 0\end{cases}$$

In Equation we can see that：θIndicates efficiency, $x_\sigma$ denotes the input quantity of the first decision unit, $y_\sigma$ denotes the output volume of the σ decision unit, $\varphi_\sigma$ indicates the weight factor, $s^+$and $s^-$denote the slack variables and residual variables, respectively.

### 2.2. TOPSIS Entropy Weight Method

The TOPSIS method, called the "Approximating Ideal Solution Ranking Method" and commonly referred to as the distance between superior and inferior solutions method in China, is a common decision making technique for multi-criteria decision analysis of finite solutions. It is a common decision-making technique for multi-criteria decision analysis of finite solutions. This method can reflect the level of the current situation by the distance between the optimal solution and the worst solution, and the model can make full use of the original data information, and the results can accurately reflect the gap between the evaluation solutions. However, the method has a strong subjective selectivity, therefore, to avoid misjudgment caused by subjective factors, the objective assignment method-entropy method will be used. In summary, this study will use the entropy-weighted TOPSIS model to comprehensively evaluate the national agricultural production efficiency and select the two most representative places to study the level of smart agriculture, with the following steps.

**(1) Construction of evaluation matrix**

With m evaluation indicators, n evaluation objects, X represents the original evaluation matrix of agricultural production efficiency：

$$X=\begin{bmatrix}x_{11} & \cdots & x_{1m}\\ \vdots & \ddots & \vdots\\ x_{n1} & \cdots & x_{nm}\end{bmatrix}$$

**(2) Data standardization processing**

When normalizing the data, we first classify the indicators into very large indicators (efficiency indicators) and very small indicators (cost indicators) initially.

ISSN: 2688-9323

indicators (benefit-based indicators) and very small indicators (cost-based indicators), so to facilitate the construction of the model, all indicators are first normalized, and the formula for converting very small indicators to very large indicators is

$$max - x$$

In order to eliminate the influence of different indicator scales, the normalized matrix needs to be normalized to obtain the normalized matrix M after normalizing the original evaluation matrix:

$$M = \begin{bmatrix} x_{11} & \cdots & x_{1m} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nm} \end{bmatrix}$$

Its matrix is normalized and the normalized matrix is denoted as Z. For each element in Z:

$$z_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^{n} x_{ij}^2}}$$

Formula for calculating the score:

$$y = \frac{x - min}{max - min}$$

Standardization matrix:

$$Z = \begin{bmatrix} z_{11} & \cdots & z_{1m} \\ \vdots & \ddots & \vdots \\ z_{n1} & \cdots & z_{nm} \end{bmatrix}$$

Define the maximum value:

$$Z^+ = (Z_1^+, Z_2^+, \dots \dots, Z_m^+)$$
$$= (max\{z_{11}, z_{21}, \dots z_{n1}\}, max\{z_{12}, z_{22}, \dots z_{n2}\}, \dots, max\{z_{1m}, z_{2m}, \dots z_{nm}\})$$

Define the minimum value:

$$Z^- = (Z_1^-, Z_2^-, \dots \dots, Z_m^-)$$
$$= (min\{z_{11}, z_{21}, \dots, z_{n1}\}, min\{z_{12}, z_{22}, \dots, z_{n2}\}, \dots, min\{z_{1m}, z_{2m}, \dots, z_{nm}\})$$

Define the distance of the i-th (i=1,2,3,...,n) evaluation object from the maximum value:

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

Define the distance of the i-th (i=1,2,3,...,n) evaluation object from the maximum value:

$$D_i^- = \sqrt{\sum_{j=1}^{m} (Z_j^- - z_{ij})^2}$$

Then the unnormalized score of the ith (i=1,2,...,n) evaluation object can be derived according to the formula for calculating the score:

$$S_i = \frac{D_i^-}{D_i^+ + D_i^-}$$

According to this formula, Easy to get $0 \le S_i \le 1$, and the larger the $S_i$ the smaller the $D_i^+$.That is, the closer to the maximum value. Of course, in the process of selecting indicators, there are inevitably some special indicators that need our special treatment, for example, when studying the effect of soil pH on plant growth, we need to evaluate the soil quality in various places, then we have to select the PH value as an evaluation indicator, which can be regarded as a floating indicator (floating around a certain value, the closer the better); when studying various nutrients contained in the soil, here It can be considered as an interval indicator (within a certain interval is the best).

ISSN: 2688-9323

Intermediate indicators:

$$M = max\{|x_i - x_{best}|\}, \widetilde{x_i} = 1 - \frac{|x_i - x_{best}|}{M}$$

Interval-type indicators:

$$M = max\{a - min\{x_i\}, max\{x_i\} - b\}$$

$$\widetilde{x_i} = \begin{cases} 1 - \frac{a-x}{M}, x < a \\ 1, \quad a \le x \le b \\ \quad 1 - \frac{x-b}{M} \end{cases}$$

In addition to defining the relevant indicators, we also need to take into account some interference factors, it is the influence of some interference factors, resulting in errors between the actual and theoretical, so we add the "Offset Factor" (Offset Factor) in the model, and use the capital letter O to indicate:

$$O = \frac{I_{ij}\omega_i}{\sum_{i=1}^{m} I_{ij}\omega_i}$$

### 2.3. To-bit Model

To-bit model is a regression model used to deal with the case where the dependent variable is truncated (censored) or restricted (bounded). In smart agriculture research, productivity is often considered as a variable with a lower bound, for example, it is not reasonable for productivity to be below 0. In reality, there are many samples that do not reach this lower bound due to technology, management, and other reasons. Therefore, the To-bit model can be more accurate and robust in dealing with variables with lower bounds such as productivity.

The idea of the To-bit model is to first set the lower bound to a cut-off value, and then use ordinary OLS regression for samples below the cut-off value; for samples above the cut-off value, set the dependent variable to the cut-off value, and at the same time correct the model error by estimating the residual variance of the above regression to obtain more accurate estimation results.

When studying the productivity of smart agriculture, the To-bit model can be modeled in combination with other influencing factors, such as land use, mechanization degree, and farming methods, so as to explore the influence of these factors on productivity. In addition, the To-bit model can also combine other spatial statistical models, such as Moran index and spatial Durbin model, to consider the influence of spatial autocorrelation on production efficiency, so as to obtain more accurate and comprehensive conclusions.

In conclusion, the To-bit model has high practical value in the study of production efficiency of smart agriculture, and can improve the accuracy of the model by solving the problems of the existence of lower bounds, while combining with other statistical models can get more comprehensive and accurate conclusions.

## 3. Analysis and Solution of the Model

### 3.1. DEA Model Solution Results

The specific steps of the analysis are as follows:

(1) Combine the benefits analysis table with the distribution trend chart of the benefits analysis to analyze the combined benefits, technical benefits and scale benefits of each unit separately and explore whether the redundancy and outputs are inadequate.

(2) Further explore the scale benefits of the decision unit by the type of scale payoff analysis and effectiveness analysis.

(3) The quadrant analysis can explore the input-output distribution of each decision unit.

(4) The input and output of non-DEA effective decision units are adjusted according to the difference-in-difference analysis to achieve the optimal relative efficiency.

**Table 1.** Benefit Analysis

| Decision-making unit | Technology Benefits | Scale benefits | Comprehensive benefits | The relaxation variable S- | The relaxation variable S+ | Validity |
|---|---|---|---|---|---|---|
| 1 | 1.000 | 1.000 | 1.000 | 0.000 | 0.000 | DEA strongly effective |
| 2 | 1.000 | 1.000 | 1.000 | 0.000 | 0.000 | DEA strongly effective |
| 3 | 1.000 | 1.000 | 1.000 | 0.000 | 0.000 | DEA strongly effective |
| 4 | 1.000 | 1.000 | 1.000 | 0.000 | 0.000 | DEA strongly effective |
| 5 | 1.000 | 1.000 | 1.000 | 0.000 | 0.000 | DEA strongly effective |
| 6 | 1.000 | 1.000 | 1.000 | 0.000 | 0.000 | DEA strongly effective |
| 7 | 1.000 | 0.995 | 0.995 | 12904.675 | 18699.682 | Non-DEA valid |
| 8 | 1.000 | 0.974 | 0.974 | 76235.075 | 0.000 | Non-DEA valid |
| 9 | 1.000 | 1.000 | 1.000 | 0.000 | 0.000 | DEA strongly valid |
| 10 | 1.000 | 0.997 | 0.997 | 32740.801 | 86833.858 | Non-DEA valid |
| 11 | 1.000 | 0.998 | 0.998 | 58702.686 | 186616.734 | Non-DEA valid |
| 12 | 1.000 | 1.000 | 1.000 | 0.000 | 0.000 | DEA strong valid |
| 13 | 1.000 | 1.000 | 1.000 | 0.000 | 0.000 | DEA strong valid |
| 14 | 1.000 | 0.980 | 0.980 | 46638.372 | 0.000 | Non-DEA valid |
| 15 | 1.000 | 0.939 | 0.939 | 60834.400 | 41989.394 | Non-DEA valid |

Table 1 will only preview the first 15 rows of data (the difference is demonstrated if there is not enough), for the full data, please click the download button to export.

The BCC model (VRS) decomposes the integrated efficiency into technical efficiency and scale efficiency.

The integrated technical efficiency (overall efficiency, OE) reflects the production efficiency of the input factors of the decision unit at a certain (optimal scale), and is a comprehensive measurement and evaluation of the resource allocation capacity, resource use efficiency and other aspects of the decision unit. If the value is equal to 1, it means that the input and output structure of the decision unit is reasonable and the relative efficiency is not optimal, and there may be different degrees of input redundancy and output deficiency.

● Technical efficiency (TE) reflects the production efficiency due to the influence of management and technology, and its value is equal to 1, which means the input factors are fully utilized and the output is maximized with a given input combination.

The scale efficiency (SE) reflects the production efficiency due to the scale factor, and is usually analyzed in combination with the scale payoff table. If the scale payoff is decreasing (not increasing or decreasing or less than 0 or greater than 0), it means the service scale is too large and there is a risk of over-expansion.

● The slack variable S- (difference variable) refers to the amount of input that can be reduced to achieve the target efficiency, i.e., the difference between the actual value and the target value of non-DEA effective units, and the slack variable S+ (excess variable) refers to the amount of output that can be increased to achieve the target efficiency, i.e., the difference between the target value and the actual value of non-DEA effective areas.

● Validity analysis combines the composite efficiency index, S- and S+, a total of three indicators, can determine DEA validity, if the composite efficiency = 1 and both S- and S+ are 0,

ISSN: 2688-9323

then "DEA strong validity", if the composite efficiency is 1 but S- or S+ is greater than 0, then "DEA weak validity ", if the comprehensive benefit < 1 is "non-DEA valid".

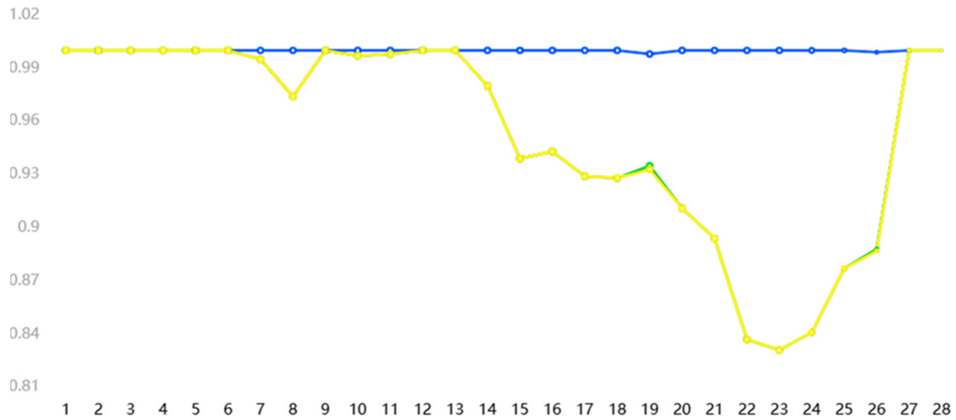


**Figure 1.** Benefit validity analysis

Figure 1 shows the benefit analysis diagram. Where the X-axis represents the decision unit and the Y-axis represents the benefit value.

● The payoff to scale will change with different production scale.

When the lambda weight is <1, the production scale is small and the input-output ratio increases rapidly with the increase of scale, which is called increasing returns to scale (IRS) (small scale can be expanded to increase the efficiency).

When lambda weight = 1, production reaches its peak and output is proportional to input and reaches the optimum production scale, which is called fixed payoff of scale.

● When the scale payoff coefficient (lambda weight) > 1; the scale of production is too large and output slows down, it is called diminishing returns to scale (DRS), that is, when the input increases, the proportion of output increase will be less than the proportion of input increase (too large can reduce the scale increase benefit).

**Table 2.** Pay-for-Scale Analysis

| Items | Payoff of scale factor | Type |
|---|---|---|
| 1 | 1.000 | Compensation for size fixed |
| 2 | 1.000 | Compensation for size fixed |
| 3 | 1.000 | Compensation for size fixed |
| 4 | 1.000 | Compensation for scale fixed |
| 5 | 1.000 | Compensation for scale fixed |
| 6 | 1.000 | Compensation for size fixed |
| 7 | 0.989 | Increasing returns to scale |
| 8 | 0.940 | Increasing returns to scale |
| 9 | 1.000 | Compensation for size fixed |
| 10 | 0.983 | Increasing returns to scale |
| 11 | 0.971 | Increasing returns to scale |
| 12 | 1.000 | Compensation for size fixed |
| 13 | 1.000 | Compensation for size fixed |
| 14 | 0.962 | Increasing returns to scale |
| 15 | 0.917 | Increasing returns to scale |

ISSN: 2688-9323

**Table 3**. Input redundancy analysis

| Decision-making unit | Relaxation variable S-analysis | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Diesel usage | Amount pesticides used | Number poultry | Number animal heads | Rural water consumption | Agricultural power consumption | Amount of agricultural film used | Coal use | Fertilizer use | Rice field area | Natural gas usage |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7 | 0 | 0 | 84.961 | 303.535 | 4.078 | 888.208 | 9118.916 | 2289.107 | 0 | 177.164 | 38.706 |
| 8 | 0 | 0 | 469.239 | 1277.075 | 82.794 | 1725.981 | 57695.407 | 13658.552 | 99.238 | 1017.458 | 209.329 |
| 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 10 | 0 | 0.852 | 419.014 | 612.277 | 0 | 1921.398 | 24057.269 | 5064.123 | 155.882 | 413.342 | 96.645 |
| 11 | 0 | 0 | 877.525 | 0 | 0 | 2572.918 | 43583.864 | 10306.61 | 321.206 | 843.089 | 197.474 |
| 12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 14 | 69.495 | 0 | 1056.06 | 1514.788 | 41.438 | 464.342 | 37241.203 | 5271.969 | 266.065 | 564.809 | 148.204 |
| 15 | 118.358 | 0 | 1830.893 | 2559.411 | 0 | 2840.088 | 46160.346 | 6009.271 | 437.088 | 688.522 | 190.423 |

| Decision-making unit | Input redundancy rate | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Diesel usage | Amount pesticides used | Number of poultry | Number animal heads | Rural water consumption | Agricultural power consumption | Amount of agricultural film used | Coal use | Fertilizer use | Rice field area | Natural gas usage |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7 | 0 | 0 | 0.006 | 0.006 | 0.001 | 0.024 | 0.004 | 0.005 | 0 | 0.006 | 0.007 |
| 8 | 0 | 0 | 0.035 | 0.024 | 0.017 | 0.047 | 0.023 | 0.032 | 0.021 | 0.034 | 0.036 |
| 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 10 | 0 | 0.004 | 0.031 | 0.011 | 0 | 0.049 | 0.01 | 0.012 | 0.033 | 0.014 | 0.017 |
| 11 | 0 | 0 | 0.063 | 0 | 0 | 0.066 | 0.018 | 0.024 | 0.066 | 0.028 | 0.034 |
| 12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 14 | 0.036 | 0 | 0.068 | 0.028 | 0.009 | 0.011 | 0.015 | 0.013 | 0.05 | 0.019 | 0.025 |
| 15 | 0.059 | 0 | 0.112 | 0.046 | 0 | 0.065 | 0.019 | 0.014 | 0.079 | 0.023 | 0.032 |

Table 3 show the preview results:

Input redundancy analysis (variance analysis) is used to analyze how much input reduction is required for each variable to reach the target efficiency.

● The slack variable S- (difference variable) refers to the amount of input reduction required to achieve the target efficiency.

● The input redundancy ratio is the ratio of "excess inputs" to existing inputs, and a larger value means more "excess inputs".

ISSN: 2688-9323

**Table 4.** Output shortfall analysis

| Decision-making unit | Relaxation variable S+ analysis | | | Underproduction rate | |
|---|---|---|---|---|---|
| | Carbon emissions | Agricultural chemical oxygen demand emissions | Aggregate | Carbon emissions | Agricultural chemical oxygen demand emissions |
| 1 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 2 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 3 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 4 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 5 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 6 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 7 | 0.000 | 18699.682 | 18700.000 | 0.000 | 0.007 |
| 8 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 9 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 10 | 0.000 | 86833.858 | 86834.000 | 0.000 | 0.038 |
| 11 | 0.000 | 186616.734 | 186617.000 | 0.000 | 0.086 |
| 12 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 13 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 14 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 15 | 0.000 | 41989.394 | 41989.000 | 0.000 | 0.025 |

Table 4 shows the preview results:

The output shortage analysis (excess variables analysis) is used to analyze how much output needs to be increased for each variable to reach the target efficiency.

S-(excess variables) refers to the amount of output that can be increased to achieve the target efficiency.

● Underproduction rate refers to the ratio of "underproduction" to output, and a larger value means more "underproduction".

## 3.2. TOPSIS Entropy Weight Method Model Solving and Analysis

The specific steps of the analysis are as follows:

(1) Prepare the data and homogenize the trend with the magnitude problem.

(2) Confirm the weight of each indicator, you can use entropy weight method, custom weights (need to handle by yourself, you can use quantitative-AHP).

(3) Find the optimal and inferior matrix vectors (automatically processed by the system).

(4) Calculate the evaluation object and positive ideal solution distance D+ or negative ideal solution distance D-, respectively.

(5) Combine the distance values to calculate the composite degree score C value, and rank them to draw a conclusion.

The table 5 shows that the weight of rice field area is 1.817%, the weight of coal use is 13.303%, the weight of agricultural chemical oxygen demand emission is 6.459%, the weight of natural gas use is 1.598%, the weight of agricultural electricity consumption is 33.388%, the weight of rural water use is 5.166%, the weight of diesel use is 8.604% The weight of agricultural film use is 5.83%, the weight of animal husbandry head is 2.401%, the weight of poultry number is 3.832%, the weight of carbon emission is 6.856%, the weight of fertilizer use is 5.725%, the weight of pesticide use is 5.022%, where the maximum index weight is agricultural electricity consumption (33.388%) and the minimum value is natural gas use (1.598%).

**Table 5.** Calculation of indicator weights

| Entropy method | | | |
|---|---|---|---|
| Items | Information entropy value e | Information utility value d | Weighting(%) |
| Rice field area | 0.977 | 0.023 | 1.817 |
| Coal use | 0.835 | 0.165 | 13.303 |
| Agricultural chemical oxygen demand emissions | 0.92 | 0.08 | 6.459 |
| Natural gas use | 0.98 | 0.02 | 1.598 |
| Agricultural electricity consumption | 0.585 | 0.415 | 33.388 |
| Rural water consumption | 0.936 | 0.064 | 5.166 |
| Diesel use | 0.893 | 0.107 | 8.604 |
| Amount of agricultural film used | 0.928 | 0.072 | 5.83 |
| Number of animal heads in animal husbandry | 0.97 | 0.03 | 2.401 |
| Number of Poultry | 0.952 | 0.048 | 3.832 |
| Carbon emissions | 0.915 | 0.085 | 6.856 |
| Amount of fertilizer use | 0.929 | 0.071 | 5.725 |
| Amount of pesticide use | 0.938 | 0.062 | 5.022 |

**Table 6.** Calculation of indicator weights

| Index valuet | Positive ideal solution distance (D+) | Negative ideal solution distance (D-) | Overall score index | Sort |
|---|---|---|---|---|
| 1995 | 0.7680 | 0.5494 | 0.4170 | 5 |
| 1996 | 0.7641 | 0.5667 | 0.4258 | 4 |
| 1997 | 0.7679 | 0.5203 | 0.4039 | 6 |
| 1998 | 0.7666 | 0.4977 | 0.3936 | 8 |
| 1999 | 0.7773 | 0.4847 | 0.3840 | 9 |
| 2000 | 0.7891 | 0.4524 | 0.3644 | 10 |
| 2001 | 0.7900 | 0.4421 | 0.3588 | 12 |
| 2002 | 0.7870 | 0.4335 | 0.3552 | 13 |
| 2003 | 0.8009 | 0.4141 | 0.3408 | 16 |
| 2004 | 0.8049 | 0.3983 | 0.3310 | 18 |
| 2005 | 0.8084 | 0.3822 | 0.3210 | 21 |
| 2006 | 0.8088 | 0.3671 | 0.3122 | 24 |
| 2007 | 0.8117 | 0.3501 | 0.3013 | 25 |
| 2008 | 0.8074 | 0.3436 | 0.2985 | 26 |
| 2009 | 0.8059 | 0.3410 | 0.2973 | 27 |

The table 6 shows the preview results:

D+ and D- values, these two values represent the distance (Euclidean distance) between the evaluation object and the optimal or inferior solution (i.e., A+ or A-), respectively. The practical meaning of these two values is that the distance between the evaluation object and the optimal or inferior solution, the larger the value indicates the farther the distance, the larger the D+ value of the research object, the farther the distance from the optimal solution; the larger the D- value, the farther the distance from the inferior solution. The most understood research object is the one with the smaller D+ value and the larger D- value.

The comprehensive degree score C value, C = (D-) / (D+ + D-), is calculated by the formula in which the numerator is the D- value and the denominator is the sum of D+ and D-; the larger the D- value is relatively, the further the research object is from the worst solution, then the better the research object is; the larger the C value is, the better the research object is.

**Table 7.** Calculation of indicator weights

| Item | Positive ideal solution | Negative ideal solution |
|---|---|---|
| Rice field area | 0.99999991 | 9e-8 |
| Coal use | 1 | 0 |
| Agricultural chemical oxygen demand emissions | 1 | 0 |
| Natural gas use | 0.99999992 | 8e-8 |
| Agricultural electricity consumption | 1 | 0 |
| Rural water consumption | 0.99999994 | 6e-8 |
| Diesel use | 0.9999998 | 2e-7 |
| Amount of agricultural film used | 1 | 0 |
| Number of animal heads in animal husbandry | 1 | 0 |
| Number of Poultry | 0.99999999 | 1e-8 |
| Carbon emissions | 0.99999997 | 3e-8 |
| Amount of fertilizer use | 0.99999993 | 7e-8 |
| Amount of pesticide use | 0.99999928 | 7.2e-7 |

The table 7 is a preview of the results:

Positive and negative ideal solutions (not distance), these two values represent the maximum value or minimum value of the evaluation index (i.e. optimal solution or inferior solution), these two values are used to calculate the D+ or D- value, the size of these two values does not have much significance

## 3.3. To-bit Model Solution Results

The specific steps of the analysis are as follows:

(1) Perform descriptive statistics on the distribution status of the independent and dependent variables.

(2) Likelihood ratio test of the model to test whether the model is valid and whether the coefficients are all zero.

(3) Obtain the model formula and analyze the significance of each coefficient.

(4) Use the formula for prediction and analysis.

**Table 8.** Likelihood ratio test results

| Likelihood ratio test value | df | P |
|---|---|---|
| 12 | 0.000*** | 142.574 |

The table 8 shows the results of the model likelihood ratio test, and the p-value is analyzed, if the value is less than 0.05, the model is valid; otherwise, the model is not valid.The results of the likelihood ratio chi-square test of the model show a significance p-value of 0.000***, which presents significance at the level and rejects the original hypothesis, thus the model is valid. Table 9 shows the regression results

The formula for the model is: Carbon emissions = -10,446.556 + 0.181 x natural gas use - 0.004 x number of poultry + 1.907 x rural water use + 0.0 x number of animal heads in livestock + 0.0 x agricultural electricity consumption - 1.25 x diesel use + 0.515 x fertilizer use + 0.012 x coal use + 5.698 x pesticide use + 0.004×Paddy field area-0.0×Film use+0.0×Agricultural chemical oxygen demand emission. The field const significance P value is 0.098*, which does not present significance at the level and rejects the original hypothesis, so the coefficient of const term is significant.

The p-value of natural gas use is 0.575, which is not significant at the level and the hypothesis is rejected, so the coefficient of natural gas use is significant.

**Table 9.** Tobit regression results

| Items | Coefficient | Standard Error | t | P | Coefficient 95% confidence interval | |
|---|---|---|---|---|---|---|
| | | | | | Upper limit | Lower limit |
| Constants | -10446.556 | 6318.818 | -1.653 | 0.098* | 1938.1 | -22831.212 |
| Natural gas usage | 0.181 | 0.322 | 0.561 | 0.575 | 0.813 | -0.451 |
| Number of poultry | -0.004 | 0.027 | -0.153 | 0.878 | 0.05 | -0.058 |
| Rural water consumption | 1.907 | 0.37 | 5.16 | 0.000*** | 2.631 | 1.183 |
| Number of animal heads in animal husbandry | 0 | 0 | 0.432 | 0.665 | 0.001 | 0 |
| Agricultural power consumption | 0 | 0 | 0.147 | 0.883 | 0.001 | 0 |
| Diesel usage | -1.25 | 0.393 | -3.181 | 0.001*** | -0.48 | -2.02 |
| Fertilizer use | 0.515 | 0.163 | 3.159 | 0.002*** | 0.834 | 0.195 |
| Coal use | 0.012 | 0.006 | 2 | 0.045** | 0.024 | 0 |
| Amount of pesticides used | 5.698 | 3.871 | 1.472 | 0.141 | 13.285 | -1.889 |
| Rice field area | 0.004 | 0.198 | 0.021 | 0.983 | 0.392 | -0.383 |
| Amount of agricultural film used | 0 | 0 | -0.564 | 0.573 | 0 | -0.001 |
| Agricultural chemical oxygen demand emissions | 0 | 0 | 0.807 | 0.420 | 0.001 | 0 |
| Note: ***, **, * represent 1%, 5%, 10% significance levels, respectively | | | | | | |

## 4. Conclusion and Recommendations

Based on the DEA super-efficiency algorithm and To-bit model, this study explores the production efficiency of smart agriculture and its influencing factors. The results show that technological progress, management level and market development are important influencing factors in smart agriculture production in the Yangtze River Delta region, while land scale, capital input and labor input have less influence on smart agriculture production efficiency. In addition, this study also found that smart agriculture production efficiency as a whole showed a trend of increasing year by year, but there were significant differences in smart agriculture production efficiency in different regions. Finally, based on the findings of this study, some policy recommendations are proposed, such as strengthening technological innovation and management improvement and optimizing land use structure, in order to improve the production efficiency of smart agriculture and promote sustainable agricultural development. In summary, this study conducted an in-depth study on the production efficiency of smart agriculture and its influencing factors, and revealed important factors affecting the production efficiency of smart agriculture by applying the DEA super-efficiency algorithm and To-bit model, and proposed effective policy recommendations. We believe that the findings of this study will make important contributions to the development of smart agriculture and sustainable agricultural development. Future research can further extend the time span and explore the evolution of the production efficiency of smart agriculture and its influencing factors in depth to provide more scientific basis for sustainable agricultural development.

## Acknowledgments

This work is supported by Anhui University of Finance and Economics Undergraduate Research and Innovation Fund Project: A study on production efficiency and influencing factors of smart agriculture based on DEA super efficiency algorithm and To-bit model (No.: XSKY22235).

ISSN: 2688-9323

## References

[1] Jiyu Jiang, Yu Wan, Teng Xu, Lin Xiong, "Based on DEA-Tobit model Comprehensive Analysis of Agricultural Production Efficiency in Anhui Province", Journal of Liaoning University of Technology (Social Science Edition), Vol.22,No.2,2020.

[2] Zhengkang Xu, "Analysis of agricultural production efficiency in Hunan Province", Cooperative Economy and Technology, No3s, 2022.

[3] Wenwen Shi, Jilong Wang, Xia Li, "Evaluation of Agricultural Production Efficiency in Shanxi Province in the Context of Rural Revitalization",Contemporary Rural Finance, No.1,2022.

[4] Wei Meng, "A comparative study of agricultural production efficiency in the Yangtze River Delta region", Industry and Technology Forum, Vol.18,No.12,2019.

[5] Chang Qian, "The current situation of organic agriculture in the Yangtze River Delta in the new era and suggestions for its development", South China Agriculture, Vol.15 No.24, 2021.

[6] Yaochun Shao, "Traditional Agricultural Scale Management in Yangtze River Delta Region The Realistic Dilemma and the Optimization Path", Journal of Nantong University, Vol.36,No.5,2020.

[7] Yanhong Zhou, Yiru Wang, Dan He, Mingji Jin, "Based on super-efficient DEA model and Malmquist index Analysis of Agricultural Production Efficiency Measurements in Three Northeastern Provinces", Northern Horticulture, 2022(03):145-151.