International Journal for Housing Science and Its Applications
Publish August 6, 2025. Volume 46, Issue 3 Pages 8000-8018

https://doi.org/10.70517/ijhsa463684

Strategies for Optimizing Resource Allocation and Efficiency Improvement through Data Analytics Algorithms in the Digital Economy

Han Liu[1,*]
[1] College of Economics and Management, China Agricultural University, Beijing, 100083, China
Corresponding authors: (e-mail: lh13263357727@126.com).

Abstract In the era of digital economy, the optimization of resource allocation efficiency has become a key factor for high-quality economic development. Traditional resource allocation methods are difficult to meet the development needs of emerging industries, and data analysis algorithms provide new ideas for optimizing allocation. In this study, you use the data envelopment analysis method to measure the allocation efficiency of science and technology resources in China, and explore the development path of the integration of the digital economy and the real economy based on the fuzzy set qualitative comparative analysis. The study constructed an index system for measuring the efficiency of S&T resource allocation containing two primary indicators, four secondary indicators and six tertiary indicators, and used the DEA-BCC model and DEA-Malmquist model for empirical analysis. The results show that China's S&T resource allocation efficiency as a whole shows an upward trend during 2016-2022, and the average value of the total factor productivity index is 1.074; among the eight economic zones, the Great Northwest Comprehensive Economic Zone has the highest S&T resource allocation efficiency, and the total factor productivity index reaches 1.14; among the 30 provinces, except for Yunnan, Tianjin, and Hainan, S&T resource allocation efficiency of the other 27 provinces is on an upward trend. Based on fsQCA method analysis, it is found that data factor marketization, digital human capital and financial support are necessary conditions for high level of digital-real integration development, and two high level digital-real integration development paths, digital environment-industry-driven under the domination of digital resources and two-way domination of digital manpower-market demand, are identified. The conclusions of the study have important reference value for promoting the deep integration of China's digital economy and real economy, optimizing the efficiency of S&T resource allocation, and providing scientific basis for the formulation of relevant policies.

Index Terms digital economy, science and technology resources, allocation efficiency, DEA-Malmquist model, fuzzy set qualitative comparative analysis, digital-real integration

I. Introduction

With the development of information technology and the accumulation and storage of large amounts of data, data analysis has become an important part of information processing today [1]. Data analysis refers to the use of computer tools for numerical analysis and research to obtain useful information and ultimately use that information to solve problems or improve the decision-making process [2]. Data analysis algorithm is a key step in the process of data analysis, which can effectively solve the problem of large amount of data and complex structure, and provide support for improving the quality and efficiency of decision-making [3], [4]. Data analysis algorithms can be divided into three categories: classification, clustering and prediction, which play an important role in optimizing resource allocation and efficiency improvement [5], [6].

The significance of optimizing resource allocation is to improve the efficiency and effectiveness of resource utilization [7]. Effective resource allocation can avoid the waste and inefficient use of resources, so that resources are fully utilized and produce greater value [8], [9]. At the same time, optimizing resource allocation can also promote sustainable economic development and bring practical benefits to society [10]. Data analysis is the key to optimize resource allocation, through the collection, collation and analysis of relevant data, we can understand the use of resources and potential problems, and provide a reference basis for optimizing resource allocation [11], [12]. Data analysis can start from multiple dimensions, such as the output of resources, the relationship between inputs and outputs, and the efficiency of resource use [13].

As an important driving force of the new round of scientific and technological revolution and industrial change, the digital economy is profoundly changing the global economic pattern and the way of resource allocation. At present, the deep integration of digital technology and all walks of life has promoted the optimization and upgrading of economic structure, given rise to new business forms and new models, and become an important engine for promoting high-quality economic development. The rapid development of the digital economy has brought a new way of thinking for resource allocation, and through the deep mining and analysis of data elements, it can effectively break through the time and space limitations of traditional resource allocation, and realize the efficient flow of resources and accurate matching. In this context, how to give full play to the optimization of the digital economy on resource allocation and enhance the efficiency of resource utilization has become a hot issue of theoretical research and practical exploration. Resource allocation efficiency is the core issue of economic development, especially scientific and technological resources as the key elements of innovation-driven development, and its allocation efficiency directly affects the national innovation capacity and industrial competitiveness. However, in the process of digital economy development, there are obvious differences in the resource endowment of various regions, the problem of uneven allocation of scientific and technological resources still exists, and the degree of integration and development of the digital economy and the real economy varies, which restricts the improvement of the overall economic efficiency. Therefore, in-depth study of the measurement method and enhancement strategy of resource allocation efficiency under the digital economy, and exploration of the effective path of the integration of the digital economy and the real economy are of great practical significance to promote the coordinated development of the region and realize the innovation drive. Domestic and foreign scholars have achieved rich results on the evaluation of resource allocation efficiency and the factors affecting it, but the systematic research on the efficiency of scientific and technological resource allocation in the digital economy environment is still insufficient, especially the relative lack of research based on the group perspective to explore the path of the integration of digital and real development. Existing studies mostly analyze resource allocation efficiency from the perspective of a single factor, ignoring the complex causal relationship of multi-factor combinations, which makes it difficult to comprehensively reveal the mechanism of resource allocation efficiency enhancement in the context of digital economy. This study focuses on the issue of S&T resource allocation efficiency in the digital economy, constructs a scientific and reasonable index system, and adopts the data envelopment analysis method to comprehensively measure the S&T resource allocation efficiency of China's eight major economic zones and 30 provinces. On this basis, the fuzzy set qualitative comparative analysis method is applied to explore the multiple paths of the integration of digital economy and real economy from the group perspective. The study first conducts static measurement based on the DEA-BCC model to analyze the differences in the technical efficiency, pure technical efficiency and scale efficiency of S&T resource allocation in each region; secondly, it adopts the DEA-Malmquist model to conduct dynamic measurement to reveal the evolutionary trend of the efficiency of S&T resource allocation; finally, it applies the fsQCA method to identify the key combinations of conditions for the development of high-level digital-real integration and puts forward targeted Finally, the fsQCA method is used to identify the key conditions for the development of high-level digital-real integration and propose targeted policy recommendations. This study not only enriches the research on digital economy theory and resource allocation efficiency, but also provides practical guidance for promoting the deep integration of digital economy and real economy and optimizing the efficiency of S&T resource allocation.

II. Resource allocation efficiency measurement model based on data envelopment analysis

In order to realize the optimization and efficiency improvement of resource allocation under the digital economy, this chapter selects scientific and technological resources as the specific research object, and constructs the resource allocation efficiency measurement model based on data envelopment analysis.

II. A. Data Envelopment Analysis Methods

II. A. 1) Basic concepts of the DEA methodology

Data Envelopment Analysis (DEA) method [14], as a non-parametric statistical analysis method, does not require conditional assumptions and the establishment of regression models, but only needs to set good weighting coefficients according to input and output indicators and establish relevant linear programming models, so as to obtain the evaluation of the relative effectiveness of the research object. This paper adopts the DEA related model to study the efficiency of science and technology resource allocation in China under the digital economy.

(1) Efficiency

Efficiency generally refers to a proportional relationship between total inputs and total outputs. In data envelopment analysis, efficiency mainly includes: comprehensive technical efficiency, pure technical efficiency and scale efficiency; numerically, comprehensive technical efficiency is the product of pure technical efficiency and scale efficiency.

Combined technical efficiency is the ratio of actual output to desired output, assuming that the inputs to the decision-making unit do not change during the production activity. Generally, the integrated technical efficiency takes a value between 0 and 1. If the efficiency value is equal to 1, it means that the decision-making unit maximizes output at the current level of inputs and is technically efficient. If the efficiency value is less than 1, it indicates that there is still a gap between the actual and ideal outputs of the decision-making unit, and the ratio of inputs needs to be improved. In this paper, the allocation efficiency of science and technology resources refers to the ratio of science and technology-related outputs obtained by enterprises, universities, research institutions and other scientific and technological innovation subjects in the operation process to the total input of scientific and technological resources.

(2) Decision Making Unit

In general, research objects with inputs and outputs can be called decision-making units (DMUs). Only decision-making units of the same type have practical significance for efficiency evaluation, i.e., such decision-making units of the same type must have common goals and tasks, the same external decision-making environment, and the same number of dimensions of inputs and mandatory output indicators.

(3) Possible sets of production for the data envelopment analysis approach

If in the process of a certain production activity, the decision-making unit carried out m input, which can be expressed in the form of a vector as: $x=(x_1,x_2,x_3,\cdots,x_m)^T$, and got s outputs, which can also be expressed in the form of a vector as: $y=(y_1,y_2,y_3,\cdots,y_s)^T$. Therefore, (x,y) can be regarded as the inputs and outputs of the decision-making unit in the whole production process of the project, which completes the construction of the production possibility set. The decision-making process input and output indicators are shown in Figure 1.

Definition 1: Set T is said to be the production possibility set in a production activity, assuming that set $T=\{(x,y)\}$ consists of all possible situations in the actual production activity and that output variable y can be produced by input variable x.

When using DEA-related models to evaluate the efficiency of production activities, it can generally be assumed that all production possibility sets simultaneously satisfy the following four theorems.

Theorem 1: $(x_j,y_j)\in T$, $j=1,2,3,\cdots,n$, i.e., for the process of production activities, the basic activity (x_j,y_j) of the decision unit with input variables x_j and output variables y_j is a mode of production.

Theorem 2: Suppose that set T is a convex set, $(x_j,y_j)\in T$, $j=1,2,3,\cdots,n$, and that there exists $\lambda_i\geq 0$ that also satisfies $\sum_{j=1}^{n}\lambda_j=1$, then there is $\left(\sum_{j=1}^{n}\lambda_j x_j,\sum_{j=1}^{n}\lambda_j y_j\right)\in T$.

Theorem 3: Assume $(x,y)\in T$, $\hat{x}\geq x$, $\hat{y}\leq y$, then there is $(\hat{x},\hat{y})\in T$.

Theorem 4: For the set T to be a cone, assume $(x,y)\in T$, then $(kx,ky)\in T$ holds whenever $k>0$.

The set T can be expressed as if the production of the minimal set of possible sets can satisfy the above four theorems at the same time:

$$T=\left\{(x,y)\left|\sum_{j=1}^{n}x_j\lambda_j\leq x,\sum_{j=1}^{n}y_j\lambda_j\geq y,\lambda_j\geq 0,j=1,2,3,\cdots,n\right.\right\} \quad (1)$$

$$\begin{array}{ccc|cccccc|cc}
 & & & 1 & 2 & \cdots & j & \cdots & n & & \\
v_1 & 1\rightarrow & & x_{11} & x_{12} & \cdots & x_{1j} & \cdots & x_{1n} & & \\
v_2 & 2\rightarrow & & x_{21} & x_{22} & \cdots & x_{2j} & \cdots & x_{2n} & & \\
\vdots & \vdots & & \vdots & \vdots & & \vdots & & \vdots & & \\
v_m & ms\rightarrow & & x_{m1} & x_{m2} & \cdots & x_{mj} & \cdots & x_{mn} & & \\
\hline
 & & & y_{11} & y_{12} & \cdots & y_{1j} & \cdots & y_{1n} & \rightarrow 1 & u_1 \\
 & & & y_{21} & y_{22} & \cdots & y_{2j} & \cdots & y_{2n} & \rightarrow 2 & u_2 \\
 & & & \vdots & \vdots & & \vdots & & \vdots & \vdots & \vdots \\
 & & & y_{s1} & y_{s2} & \cdots & y_{sj} & \cdots & y_{sn} & \rightarrow s & u_s
\end{array}$$

Figure 1: Input indicators and output indicators in the decision-making process

(4) Benefits of Scale and Technical Effectiveness of DEA Methods

(1) Return to scale

The ratio of a decision unit's individual output indicators to its individual input indicators over the same period is known as the gain in scale for that decision unit, and is expressed by the formula:

$$k = \frac{\Delta y}{y} \Big/ \frac{\Delta x}{x} \tag{2}$$

If $k > 1$, it indicates that the returns to scale for this decision unit are gradually increasing, and consideration can be given to increasing inputs to improve outputs. If $k < 1$, it means that the returns to scale of the decision unit are gradually decreasing, and it can be considered to reduce inputs. If $k = 1$, then the returns to scale for the decision unit are stable, i.e., the decision unit is efficient in scale at this time.

(2) Technically valid

Assuming any decision unit $(x, y) \in T$, which satisfies the output result that there is no $y' > y$, decision unit $(x, y) \in T$ is considered to have achieved technical validity.

II. A. 2) Basic DEA model

(1) CCR model [15].

It is assumed that the research object has n decision making unit (DMU), while each decision making unit has m inputs and s outputs, which are expressed in the form of vectors as:

$$X_j = \left(x_{1j}, x_{2j}, \cdots, x_{mj}\right)^T, Y_j = \left(y_{1j}, y_{2j}, \cdots, y_{mj}\right)^T \ And \ j = 1, 2, 3, \cdots, n \tag{3}$$

In the production process, appropriate weights must be assigned to each input variable as well as to each output variable. The vector of weights for m input and s outputs is expressed accordingly:

$$v = \left(v_1, v_2, \cdots, v_m\right)^T, u = \left(u_1, u_2, \cdots, u_s\right)^T \tag{4}$$

Then the combined input and output value of the j st decision unit can be expressed as:

$$\sum_{i=1}^{m} v_i x_{ij} \ And \ \sum_{r=1}^{s} u_r y_{rj} \tag{5}$$

Therefore, the combined efficiency evaluation index of the j st decision unit can be expressed as follows:

$$h_j = \frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \tag{6}$$

Setting the weights of inputs and outputs in the production activity, the efficiency evaluation index of the decision unit is measured as h_j, and the maximum efficiency evaluation index of the decision unit is 1. When $h_j = 1$, it indicates that the production of the decision unit is efficient with respect to other decision units. And when $h_j < 1$, it indicates that the production of the decision unit is non-efficient relative to other decision units.

The relative efficiency of the j th decision unit is calculated by the formula:

$$\max h_{j0} = \frac{\sum_{i=1}^{s} u_r y_{rj_0}}{\sum_{i=1}^{m} v_i x_{ij_0}}$$
$$s.t. \begin{cases} \dfrac{\sum_{i=1}^{s} u_r y_{rj_0}}{\sum_{i=1}^{m} v_i x_{ij}} \le 1, j = 1, 2, 3, \cdots, n \\ v = (v_1, v_2, \cdots, v_m)^T \ge 0 \\ u = (u_1, u_2, \cdots, u_s)^T \ge 0 \end{cases} \tag{7}$$

To facilitate the optimal solution of the model, let:

$$t = \frac{1}{\sum_{i=1}^{m} v_i x_{ij_0}}, u_r = tu_r, w_i = tv_i \tag{8}$$

To facilitate the optimal solution of the model, let:

$$\max h_{j0} = \sum_{r=1}^{x} u_r y_{rj_0}$$
$$s.t.\begin{cases} \sum_{r=1}^{x} u_r y_{rj} - \sum_{i=1}^{m} w_i x_{ij} \le 0, j = 1,2,\cdots,n \\ \sum_{i=1}^{m} w_i x_{ij_0} = 1 \\ u_r, w_i \ge 0, i = 1,2,\cdots m \end{cases} \tag{9}$$

To wit:

$$\max h_{j0} = u^T Y_0$$
$$s.t.\begin{cases} u^T Y_j - w^T X_j \le 0 \\ w^T X_0 = 1 \quad j = 1,2,3,\cdots,n \\ w \ge 0, u \ge 0 \end{cases} \tag{10}$$

$$\min \theta$$
$$s.t.\begin{cases} \sum_{j=1}^{n} \lambda_j x_j \le \theta x_0 \\ \sum_{j=1}^{n} \lambda_j y_j \ge y_0 \\ \lambda_j \ge 0, j = 1,2,3,\cdots,n \\ \theta \text{ Unrestricted } s^+ \ge 0, s \ge 0 \end{cases} \tag{11}$$

where s^+ represents the slack variable and s^- represents the residual variable, so the above inequality can be transformed into:

$$\min \theta$$
$$s.t.\begin{cases} \sum_{j=1}^{n} \lambda_j y_j - s^+ = \theta x_0 \\ \sum_{j=1}^{n} \lambda_j y_j - s^- = y_0 \\ \lambda_j \ge 0, j = 1,2,3,\cdots,n \\ \theta \text{ Unrestricted}, s^+ \ge 0 \end{cases} \tag{12}$$

The optimal solution θ, s^+, s^- can be found.

a) When $\theta = 1$, and $s^+ = 0$, $s^- = 0$, indicating that the inputs and outputs at this time are optimal, that is, the integrated technical efficiency is the highest, that is, the decision unit to achieve the DEA effective.

b) When $\theta = 1$, s^+ and s^- are not all 0 and at least one of them is greater than 0, then the decision unit is weakly effective in DEA.

c) When $\theta < 1$, the decision unit is ineffective, the combined efficiency of inputs and outputs is not maximized, and it is necessary to improve the allocation of input and output variables in production activities.

d) BCC model [16].

On the basis of DEA-CCR model with constant returns to scale, operations researchers proposed DEA-BCC model based on variable returns to scale.BCC model decomposes the comprehensive technical efficiency in CCR model into pure technical efficiency and scale efficiency, so the value of comprehensive technical efficiency can be understood computationally as the product of pure technical efficiency and scale efficiency.

Assuming that there is n decision cell, each of which has m inputs as well as s outputs, the input and output vectors of the j th decision cell can be represented as vectors respectively:

$$x_j = \left(x_{1j}, x_{2j}, \cdots, x_{mj}\right)^T, y_j = \left(y_{1j}, y_{2j}, \cdots, y_{sj}\right)^T, x_j \geq 0, y_j \geq 0 \tag{13}$$

The weight vectors of their inputs and outputs can be expressed separately:

$$v = (v_1, v_2, \cdots, v_m), u = (u_1, u_2, \cdots, u_s) \tag{14}$$

Then the output-oriented DEA-BCC model can be expressed as:

$$\min \theta - \varepsilon\left(\sum_{i=1}^{m} s^- + \sum_{r=1}^{s} s^+\right)$$
$$s.t.\begin{cases} \sum_{j=1}^{n} \lambda_j x_j + s^- = \theta x_0 \\ \sum_{j=1}^{n} \lambda_j y_j - s^+ = y_0 \\ \sum_{j=1}^{n} \lambda_j = 1 \\ \lambda_j \geq 0, s^- \geq 0, s^+ \geq 0 \\ j = 1, 2, \cdots, n \end{cases} \tag{15}$$

When $\theta = 1$, $s^- = s^+ = 0$, the decision cell realizes DEA validity; when $\theta = 1$, s^- & s^+ are not all 0, the decision cell is DEA weakly valid; when $\theta < 1$, the decision cell is invalid.

II. B.DEA-Malmquist model and indicator selection

II. B. 1) DEA-Malmquist modeling

In the empirical literature related to the use of DEA models, the approaches commonly used to measure allocative efficiency are divided into two categories, namely constant returns to scale (CCR model) and variable returns to scale (BCC model). Comparison of these two models reveals that the BCC model considers the production state of DMUs more comprehensively, and believes that decision-making units may be in the state of increasing or decreasing returns to scale, so it decomposes the pure technical efficiency (TE) in the CCR model into two types of efficiency expressions: pure technical efficiency (PTE) and scale efficiency (SE), and thus the BCC model is relatively more complete and better able to reflect the real situation. , more able to reflect the reality. However, for different decision-making units, its efficiency level will also produce changes and differences, BCC model can not solve this problem, so scholars finally improved the DEA-Malmquist model after practice [17]. DEA-Malmquist index can make up for the shortcomings of the CCR and BCC model, and measure the changes between the decision-making units.

In this paper, the Malmquist index of productivity change is shown in the following form:

$$M(x^{t+1}, y^{t+1}, x^t, y^t) = \frac{D^{t+1}(x^{t+1}, y^{t+1})}{D^t(x^t, y^t)} \times \left[\left(\frac{D^t(x^{t+1}, y^{t+1})}{D^{t+1}(x^{t+1}, y^{t+1})}\right)\left(\frac{D^t(x^t, y^t)}{D^{t+1}(x^t, y^t)}\right)\right]^{\frac{1}{2}} \tag{16}$$

Among them:

$$Effech = \frac{D^{t+1}(x^{t+1}, y^{t+1})}{D^{t}(x^{t}, y^{t})} \tag{17}$$

$$Tech = \left[\left(\frac{D^{t}(x^{t+1}, y^{t+1})}{D^{t+1}(x^{t+1}, y^{t+1})}\right)\left(\frac{D^{t}(x^{t}, y^{t})}{D^{t+1}(x^{t}, y^{t})}\right)\right]^{\frac{1}{2}} \tag{18}$$

The Malmquist index of productivity change (TFP) refers to the magnitude of change in the total productivity of a decision unit from time t to time $t+1$, and can be divided into two indices: the index of change in technical efficiency (Effech) and the index of technological progress (Tech).Tech is used to measure the efficiency of technological change in the production of a particular decision unit from time t to time $t+1$. , which is mainly a reflection of the change in the technological level, if $Tech > 1$, it means that the decision unit has been upgraded in the technological level or the level of innovation has increased. If it is a measure of whether a decision-making unit has been improved at the level of organizational management, the Effch is used to show that when $Effech > 1$, it indicates that this decision-making unit measured has been improved compared to the technical efficiency of the previous period, i.e., the technical efficiency has been improved. It can be expressed by the formula:

$$TFP = Effech \times Tech \tag{19}$$

This paper refers to the way of research in the related literature, and adopts the Malmquist productivity change index (TFP) to represent the efficiency of science and technology resource allocation, $TFP > 1$ indicates that the efficiency of science and technology resource allocation has been improved compared with that of the previous year, and $TFP < 1$ indicates that the efficiency of science and technology resource allocation has been reduced compared with that of the previous year.

II. B. 2) Selection of input and output variables

This paper refers to the relevant empirical research literature, based on the current scientific and technological resource allocation efficiency research related to the selection of indicators, based on the availability of data, the input part of the input only consider the scientific and technological human resources, scientific and technological financial resources two kinds of resources, and the output part of the output to consider the output of scientific and technological achievements and scientific and technological achievements marketization. Finally, this paper constructs a system of indicators for measuring the efficiency of S&T resource allocation, including two primary indicators, four secondary indicators and six tertiary indicators, as shown in Table 1.

Table 1: Input-output indicators for the efficiency of resource allocation

First-level indicator	Secondary indicators	Third-level indicators	Variable name
Input indicator	Input of scientific and technological human resources	Full-time equivalent of R&D personnel	X1
	Financial investment in science and technology	Internal expenditure on R&D funds	X2
Output indicators	Output of scientific and technological achievements	The number of scientific and technological papers indexed by SCI, CPCI-S and EI	Y1
		The number of authorized invention patents	Y2
	Marketization of scientific and technological achievements	Transaction volume of technology market contracts	Y3
		Sales revenue of new products in high-tech industries	Y4

II. B. 3) Basis for the delineation of the eight integrated economic zones

Great Southwest Comprehensive Economic Zone“, ”Great Northwest Comprehensive Economic Zone“, ”Middle Yellow River Comprehensive Economic Zone“, ”Middle Yangtze River Comprehensive Economic Zone", “North Coastal Comprehensive Economic Zone”, “South Coastal Comprehensive Economic Zone” and “East Coastal Comprehensive Economic Zone”, and analyzing the “Eight Comprehensive Economic Zones”, where the level of economic development itself differs greatly. We analyze whether there is a big difference in the level of economic development of the “eight comprehensive economic zones” in the level of science and technology resource allocation efficiency, and whether the impact of the digital economy on the efficiency of science and technology resource allocation will also be different because of the different economic zones. The division of the eight integrated economic zones is shown in Table 2.

Table 2: The division of the eight major comprehensive economic zones

Economic zone	Contained area
Northeast Comprehensive Economic Zone	Liaoning, Jilin, Heilongjiang
Northern Coastal Comprehensive Economic Zone	Beijing, Tianjin, Hebei, Shandong
Eastern Coastal Comprehensive Economic Zone	Shanghai, Jiangsu, Zhejiang
Southern Coastal Comprehensive Economic Zone	Fujian, Guangdong, Hainan
The Comprehensive Economic Zone of the Middle Reaches of the Yellow River	Shaanxi, Shanxi, Henan, Inner Mongolia
The Comprehensive Economic Zone of the Middle Reaches of the Yangtze River	Hubei, Hunan, Jiangxi, Anhui
Southwest Comprehensive Economic Zone	Yunnan, Guizhou, Sichuan, Chongqing, Guangxi
Comprehensive Economic Zone of Northwest China	Gansu, Qinghai, Ningxia, Tibet, Xinjiang

II. C.Allocation Efficiency Measurement with China's S&T Resources as an Example

In this section, using China's S&T resources as an example, the DEA-BCC model and the DEA-Malmquist model were used to measure the efficiency of S&T resource allocation.

II. C. 1) Data sources

The data of the study mainly includes 30 provincial-level administrative regions in mainland China, excluding Tibet and Hong Kong, Macao and Taiwan. The method of dividing the eight economic zones by the Development Research Center of the State Council of China is used, but since Tibet is excluded from the measurement, the efficiency value of the Northwest region in this part is calculated without Tibetan data. The data are from China Science and Technology Statistical Yearbook (2016-2024) and China Statistical Yearbook (2016-2023). Meanwhile, in order to exclude the influence of price changes and inflation factors, the data of internal expenditure on R&D funding, the amount of technology market contract transactions, and the sales revenue of new products in high-tech industries were deflated by using the GDP deflator with 2015 as the base period.

II. C. 2) Static measurement based on DEA-BCC modeling

(1) National level

From the national level, the technical efficiency, pure technical efficiency and scale efficiency of the allocation of scientific and technological resources increase year by year from 2016 to 2020, and the technical efficiency and pure technical efficiency show a decreasing trend from 2020 to 2022, while the scale efficiency remains stable. And in the same period, scale efficiency is higher compared with technical efficiency and pure technical efficiency.

(2) Regional level

a) Technical efficiency

The technical efficiency of S&T resource allocation in China's eight economic zones is shown in Figure 2. In the figure, “Northeast Comprehensive Economic Zone”, “North Coastal Comprehensive Economic Zone”, “East Coastal Comprehensive Economic Zone”, "South Coastal Comprehensive Economic Zone Economic Zone”, “Middle Yellow River Comprehensive Economic Zone”, “Middle Yangtze River Comprehensive Economic Zone”, “Greater Southwest Comprehensive Economic Zone” and “Greater Northwest Comprehensive Economic Zone”. The “Great Northwest Comprehensive Economic Zone” is denoted by the letters A, B, C, D, E, F, G, and H, respectively, and the same applies to the latter.

It can be seen that: the technical efficiency of the Northeast region is higher than the national average in all years, while the technical efficiency of the North Coast, the Middle Yellow River and the Middle Yangtze River is lower than the national average in all years. The Eastern Coastal Comprehensive Economic Zone and the Southern Coastal Comprehensive Economic Zone are higher than the national average in all years except for 2022, which is slightly lower than the national average. The Greater Southwest Comprehensive Economic Zone is above the national average in all years except 2016, 2017, and 2018. The Greater Northwest Comprehensive Economic Zone is below the national average in all years except 2020, 2021, and 2022.

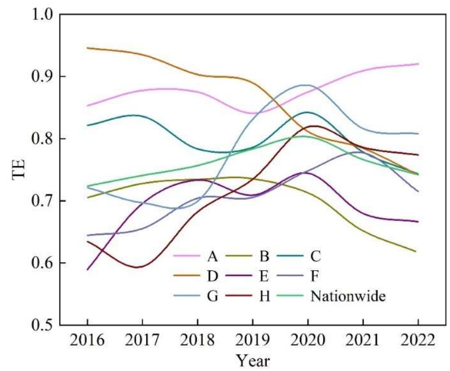

Figure 2: The efficiency of scientific and technological resource allocation (TE)

b) Pure technical efficiency

The results of measuring the pure technical efficiency of S&T resource allocation in China's eight economic regions are shown in Figure 3. It can be seen that: the pure technical efficiency of the Northeast, the East Coast, the South Coast, and the Great Northwest are all higher than the national average in all years, while the pure technical efficiency of the North Coast, the Middle Yellow River, and the Middle Yangtze River are all lower than the national average. The Great Southwest is above the national average in all years except 2016, 2017 and 2018.

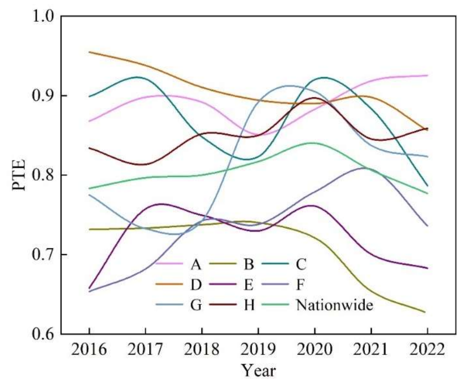

Figure 3: The efficiency of scientific and technological resource allocation (PTE)

c) Scale efficiency

The results of measuring the scale efficiency of S&T resource allocation in China's eight major comprehensive economic zones are shown in Figure 4. It can be seen that: the scale efficiency of the Northeast, the northern coast of the region, and the Great Southwest are all higher than the national average in all years. The scale efficiency of the East Coast and the Great Northwest in all calendar years is lower than the national average. The middle reaches of the Yangtze River are above the national average in all years except 2019. The South Coast is above the national average in all years except 2020, 2021 and 2022. The middle reaches of the Yellow River are above the national average in all years except 2016, 2017 and 2018.

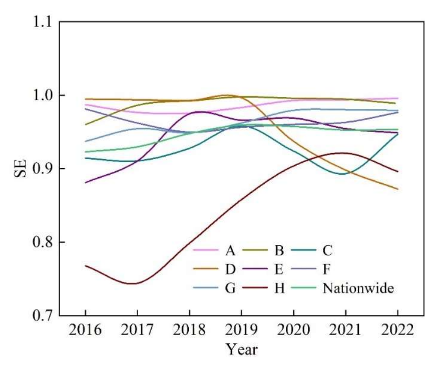

Figure 4: The efficiency of scientific and technological resource allocation (SE)

II. C. 3) Dynamic measurement based on DEA-Malmquist modeling

(1) National level

Changes in China's STI resource allocation efficiency from 2016 to 2022 are shown in Figure 5. As can be seen from Figure 5:

a) The efficiency of S&T resource allocation increases year by year except for 2021. Tfpch average value is 1.074>1, and overall the efficiency of S&T resource allocation increases. Specifically, tfpch is greater than 1 in all periods except 2021/2022.

b) Technical efficiency has an inflection point in 2020. Effch mean value is 1.011>1, and the overall is characterized by an increase. However, the two periods of 2020/2021 and 2021/2022 are less than 1, and all other periods are greater than 1, which indicates that technical efficiency increases year by year from 2016-2019 and decreases after 2020.

c) Technology fluctuates and progresses. Techch average is 1.067, which is greater than 1 in all periods except 2017/2018 which is less than 1. There are fluctuations, but overall and especially in recent years, it is in a state of progress.

d) Pure technical efficiency will see an inflection point in 2020. China pech average value of 1.004, the overall is characterized by the promotion. But after 2020 there is a declining inflection point, especially in recent years, pure technical efficiency decline, which is consistent with the performance of technical efficiency.

5) Scale efficiency improves year by year except for 2020. Sech average value of 1.011, the overall rise, except for 2020/2021 sech is less than 1, the other periods are greater than 1.

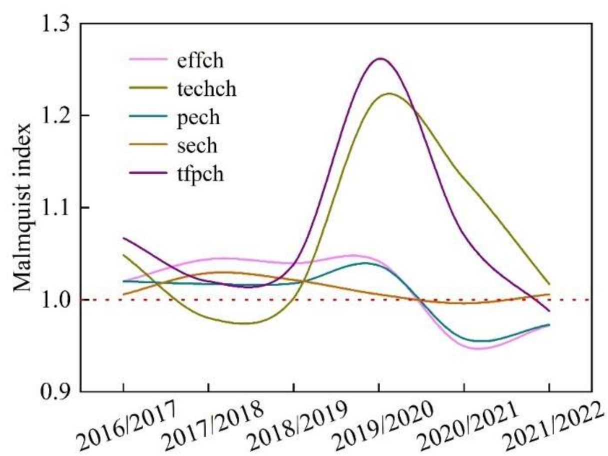

Figure 5: The variation of resource allocation efficiency

(2) Regional level

The results of the Malmquist index decomposition of S&T resource allocation efficiency in the eight economic comprehensive regions of China are shown in Table 3. As can be seen from the table: the efficiency of S&T resource allocation is improving in most regions. Except for the Southern Coastal Comprehensive Economic Zone (D) tfpch is equal to 1, all other regions are greater than 1. Combined with the decomposition index to further analyze the reasons for the change in efficiency:

a) On the east coast and north coast, the total factor productivity increase in the allocation of science and technology resources comes from technological progress and scale efficiency. This is because both techch and sech are greater than 1, while effch and pech are less than 1.

b) Technical progress is offset by technical efficiency, pure technical efficiency and scale efficiency in the southern coastal integrated economic zone, so that the total factor productivity of S&T resource allocation remains unchanged. effch, pech, sech are all less than 1, and only techch is greater than 1.

c) The increase in total factor productivity of S&T resource allocation in other regions comes from the increase in technical progress, technical efficiency, pure technical efficiency and scale efficiency.

Table 3: Malmquist exponential decomposition of resource allocation efficiency

Region	effch		techch		pech		sech		tfpch	
	Value	Ranking	Value	Ranking	Value	Ranking	Value	Ranking	Value	Ranking
A	1.03	5	1.07	6	1.03	3	1.01	6	1.08	4
B	0.99	7	1.09	4	0.99	8	1.03	5	1.07	7
C	0.99	6	1.09	3	0.99	7	1.03	4	1.07	6
D	0.96	8	1.06	8	0.99	6	0.99	8	1.00	8
E	1.05	2	1.09	2	1.03	5	1.04	2	1.09	2
F	1.04	4	1.06	7	1.04	1	1.00	7	1.08	5
G	1.04	3	1.07	5	1.04	2	1.03	3	1.09	3
H	1.06	1	1.09	1	1.03	4	1.05	1	1.14	1
Nationwide	1.03	-	1.08	-	1.02	-	1.03	-	1.09	-

(3) Provincial level

The results of Malmquist index demarcation of S&T resource allocation efficiency in Chinese provinces are shown in Table 4. From the table, it can be seen that: except for Yunnan, Tianjin and Hainan, the efficiency of S&T resource allocation in other provinces of China has been improving and developing well. This is because the total factor productivity (TFP) index of S&T resource allocation in Yunnan, Tianjin and Hainan is less than 1, while the TFP indexes of the other 27 provinces are all greater than 1.

Table 4: Changes in the efficiency of resource allocation in 2016-2022

Region	effch	techch	pech	sech	tfpch	Region	effch	techch	pech	sech	tfpch
Guangxi	1.156	1.077	1.156	1.004	1.221	Shanxi	1.009	1.04	1.006	1.003	1.062
Qinghai	1.045	1.148	1.004	1.052	1.205	Guangdong	1.000	1.047	1.004	1.002	1.044
Ningxia	1.115	1.041	1.007	1.104	1.159	Inner mongolia	0.942	1.114	0.936	0.995	1.054
Henan	1.114	1.031	1.072	1.036	1.143	Hubei	1.023	1.031	1.014	1.000	1.054
Shanxi	1.042	1.089	1.021	1.031	1.136	Jilin	1.000	1.042	1.000	1.000	1.046
Anhui	1.05	1.072	1.06	0.995	1.118	Chongqing	1.000	1.052	1.000	1.000	1.048
Xinjiang	1.022	1.075	1.036	0.987	1.109	Shanghai	0.965	1.082	0.947	1.013	1.036
Heilongjiang	1.018	1.082	1.014	1.006	1.087	Jiangsu	0.973	1.061	0.985	0.998	1.024
Jiangxi	1.033	1.041	1.049	1.005	1.088	Hunan	0.978	1.058	0.982	0.998	1.028
Liaoning	1.03	1.059	1.032	0.998	1.081	Gansu	0.988	1.028	0.988	0.997	1.023
Guizhou	1.025	1.045	1.025	1.013	1.072	Fujian	0.955	1.054	0.963	1.002	1.004
Shandong	0.996	1.085	0.989	1.001	1.077	Yunnan	0.931	1.069	0.909	1.032	0.992
Zhejiang	1.023	1.052	1.006	1.015	1.087	Tianjin	0.935	1.038	0.926	1.008	0.973
Sichuan	1.012	1.066	1.009	1.000	1.072	Hainan	0.908	1.05	0.979	0.927	0.954
Beijing	1.000	1.077	0.998	1.000	1.072	Nationwide	1.009	1.057	1.003	1.009	1.074
Hebei	1.011	1.077	0.987	1.013	1.072						

Further, through the Malmquist decomposition index, we analyze the reasons for the dynamic changes in the allocation efficiency of scientific and technological resources in each province.

a) Yunnan, Tianjin and Hainan: the total factor productivity index is less than 1, but Yunnan and Tianjin are due to the technical efficiency and pure technical efficiency offsetting the technical progress and scale efficiency, which makes the efficiency of S&T resource allocation lower; while Hainan is due to the technical efficiency, pure technical efficiency and scale efficiency offsetting the technical progress, and ultimately the efficiency of S&T resource allocation lower.

b) The remaining 27 provinces: the total factor productivity index is greater than 1, but it should be pointed out that the reasons for the improvement in efficiency vary from province to province. Anhui, Xinjiang and Liaoning provinces have scale efficiency change indexes less than 1, and other decomposition indexes are all greater than 1. The increase in the efficiency of scientific and technological resource allocation in the three provinces comes from the improvement of technical progress, pure technical efficiency and technical efficiency. The indexes of technical efficiency and pure technical efficiency in Shandong, Hebei, Shanghai and Fujian are less than 1, indicating that the improvement of the efficiency of S&T resource allocation in these four provinces comes from the improvement of technical progress and scale efficiency. The technical efficiency index, pure technical efficiency index and scale efficiency index of Inner Mongolia, Jiangsu, Hunan and Gansu are less than 1. The improvement of the efficiency of S&T resource allocation in these four provinces comes from technical progress. The improvement of the efficiency of S&T resource allocation in other provinces comes from the growth of technical progress, technical efficiency, pure technical efficiency and scale efficiency together.

c) Looking at the total factor productivity decomposition indices of China's 30 provinces, it can be found that the technological progress indices of China's 30 provinces are all greater than 1, indicating that technological progress exists in all provinces.

III. Research on the development path of the integration of the digital economy and the real economy

This chapter explores the development path of the integration of the digital economy and the real economy on the basis of measuring the efficiency of resource allocation under the digital economy.

III. A. Study design

III. A. 1) Research methodology

The level of development of digital-real fusion is influenced by multiple factors with intricate driving paths. fsQCA methodology, with its group-based logic, integrates qualitative and quantitative analyses, and is particularly suitable for dealing with interdependencies and complex causality among conditional variables. In addition, the fsQCA method has multiple concurrency and is able to deal with multiple antecedent condition change scenarios at the same time, and thus is able to effectively cope with multi-factor combinations and their interactions that affect the number-reality fusion. This is in line with the reality of the development of number-reality fusion. fsQCA method also has asymmetry, which can explore the nonlinear relationship in causality. By fully exploring multiple concurrent and asymmetric causal relationships, fsQCA is able to effectively differentiate between core conditions and edge conditions, deeply excavate the causal chain behind the efficiency of number-reality fusion, and provide practical paths to improve the quality and efficiency of fusion.

Table 5: Evaluation index system

<table>
<tr><th>First-level indicator</th><th>Secondary indicators</th><th>Third-level indicators</th><th>Variable name</th></tr>
<tr><td rowspan="8">The level of digital economic development</td><td rowspan="2">Digital infrastructure</td><td>Mobile phone penetration rate</td><td>X1</td></tr>
<tr><td>Internet broadband access user rate</td><td>X2</td></tr>
<tr><td rowspan="3">Digital industrialization</td><td>Operating income of the electronic information manufacturing industry</td><td>X3</td></tr>
<tr><td>Total volume of telecommunication services</td><td>X4</td></tr>
<tr><td>Revenue from information technology services</td><td>X5</td></tr>
<tr><td rowspan="3">Industrial digitalization</td><td>The number of websites per hundred enterprises</td><td>X6</td></tr>
<tr><td>The proportion of enterprises with e-commerce transaction activities</td><td>X7</td></tr>
<tr><td>E-commerce sales volume</td><td>X8</td></tr>
<tr><td rowspan="6">The development level of the real economy</td><td rowspan="3">Development scale</td><td>The number of scientific and technological papers indexed by SCI, CPCI-S and EI</td><td>V1</td></tr>
<tr><td>The number of authorized invention patents</td><td>V2</td></tr>
<tr><td>Per capita regional GDP</td><td>V3</td></tr>
<tr><td>Development benefits</td><td>Main business income of industrial enterprises above designated size</td><td>V4</td></tr>
<tr><td rowspan="2">Development potential</td><td>R&D expenditure of industrial enterprises above designated size</td><td>V5</td></tr>
<tr><td>The number of valid invention patents of industrial enterprises above designated size</td><td>V6</td></tr>
</table>

III. A. 2) Variable Selection and Measurement

(1) Outcome variables

Development level of digital-real integration: the index is the level of digital-real coupling coordination of each region in 2024 measured by the coupling coordination degree model based on the entropy value method. In order to construct a comprehensive evaluation system, taking into account the key role of digital infrastructure in the development of digital economy, the practical significance of industrial digitization and digital industrialization in that development process, this paper constructs an evaluation index system for the level of development of digital economy in each province in China from the aspects of digital infrastructure, digital industrialization and industrial digitization. In assessing the development level of the real economy, this study starts from the three dimensions of real economy development scale, development efficiency and development potential. Among them, the scale of the real economy is selected to be quantified by three indicators: gross output value minus the financial and real estate industries, total retail sales of consumer goods, and total import and export of goods. In view of the core position of industry in the real economy and its key role in promoting the development of digital-real integration, this study selects the relevant indicators of industrial enterprises above the scale as the evaluation indicators of the real economy's efficiency and potential. The evaluation index system of the development level of digital-real integration is specifically shown in Table 5.

In view of the discrete degree of each indicator, the entropy value method is first used in this study to standardize and assign weights to the indicators in order to calculate the comprehensive evaluation index of the two systems, digital economy and real economy. In addition, in order to reveal the degree of integration and development of the two, this paper constructs a coupling coordination degree model with the following formula:

$$O = \frac{2 \times \sqrt{Digital \times Real}}{Digital + Real} \tag{20}$$

$$Z = \beta_1 Digital \times \beta_2 Real \tag{21}$$

$$XT = \sqrt{O \times Z} \tag{22}$$

In the formula, $Digital$ represents the development level of digital economy calculated by entropy value method, Real represents the development level of real economy calculated by entropy value method, and β_1 and β_2 represent the degree of contribution of the two to the coupling coordination degree, respectively. In this paper, the parameter value is set to 0.5, which indicates that the digital economy and the real economy play equally important roles in the improvement of the coupling coordination degree. XT indicates the measured level of coupling coordination between the digital economy and the real economy, and its value range is $[0,1]$.

(2) Conditional variables

Marketization of data elements: reflected by the data trading platforms set up in each region. If 30 provincial-level administrative regions in China have set up data trading platforms in 2023 and before, the indicator will be 1, otherwise it will be 0.

Digital human capital: measured by the number of people working in the information transmission, software and information technology services industry, and the computer, communication and other electronic equipment manufacturing industry per 10,000 employed people.

Government Intervention: Measured by the intensity of government intervention in terms of fiscal expenditure per capita.

Financial support: the intensity of financial support is measured by the ratio of the loan balance of financial institutions to the output value of the secondary industry at the end of the year.

Market Demand: The level of market demand is reflected by per capita disposable income.

Small and medium-sized enterprise (SME) development: the level of SME development is measured by the comprehensive SME prosperity index of each province in the China SME Prosperity Index Report (2023).

III. A. 3) Research sample and data sources

In this chapter, 30 provincial-level administrative regions in China are used as research cases, which are widely and evenly selected to cover regions with higher and lower levels of digital-realistic integration in order to form a comparative analysis. The data used are mainly derived from the China Statistical Yearbook, China Science and Technology Statistical Yearbook, China Industrial Statistical Yearbook, and the Cathay Pacific database. When dealing with missing data, the interpolation method is used to make up for the missing data to ensure the integrity of the data. In addition, this paper sets a lag period of one year during data analysis, i.e., the conditional variable in 2023 corresponds to the development level of digital-realistic integration in 2024 to ensure the timeliness and accuracy of the analysis results.

III. B. Empirical analysis

III. B. 1) Necessity analysis of individual conditions

The NCA method was applied through R software to analyze whether individual antecedent conditions were necessary for the outcome variable, and before the analysis, the raw values of all conditions were converted to pooled affiliation scores, and the results of the NCA method are shown in Table 6. Among them, the effect size is the ratio of the upper limit region to the range, indicating the degree of constraint on the outcome conditions, with a value range of 0 to 1. Considering that the antecedent conditions are all continuous values, the upper limit regression

(CR) analysis technique was chosen to measure them, and the envelope regression (CE) was selected as the robustness treatment. Based on the judgment criteria of the necessary conditions, the effect size (d) of the necessary conditions represents the effect of the necessary conditions, with $0<d<0.1$ as a low effect, $0.1<d<0.3$ as a medium effect, and $0.3<d<0.5$ as a high effect. This antecedent variable can be considered necessary for the outcome variable when d is greater than 0.1 and significant at the 5% level ($p<0.05$).

As can be seen from Table 6, data factor marketization, digital human capital and financial support are all necessary conditions for a high level of digital-real integration development. This paper chooses the method of judging the necessary conditions from the size of the effect size and finds that the degree of importance of the three necessary conditions is financial support, data factor marketization and digital human capital in descending order.

Table 6: Analysis results by NCA method

Conditional variable	Method	Accuracy /%	Upper limit area	Range	d	p
V1	CE	100	0.142	1	0.142	0.000
	CR	84.5	0.253	1	0.253	0.000
V2	CE	100	0.180	0.994	0.182	0.000
	CR	87.9	0.224	0.994	0.227	0.000
V3	CE	100	0.002	1	0.002	0.241
	CR	100	0.002	1	0.002	0.241
V4	CE	100	0.383	1	0.383	0.000
	CR	87.9	0.349	1	0.349	0.000
V5	CE	100	0.094	1	0.094	0.000
	CR	91	0.205	1	0.205	0.205
V6	CE	100	0.031	1	0.031	0.021
	CR	100	0.019	1	0.019	0.027

The bottleneck level analysis of NCA allows for the measurement of the minimum level of individual antecedent conditions that need to be achieved within the range of observations in order to achieve a specific target outcome variable. The bottleneck level allows the analysis of the value of the level at which the antecedent condition needs to be achieved to achieve a specific level of the outcome. The results of the bottleneck level analysis are shown in Table 7, where the values represent the affiliation values of the conditions, and NN means “not necessary”. It can be seen that in order to achieve a high level of digital and real integration development (greater than 0.5), the member affiliation scores of data factor marketization, digital human capital, financial support and market demand should be at least greater than 0.158, 0.145, 0.313 and 0.016, respectively, and there is no bottleneck level for government intervention and SMEs' development at this level.

Table 7: Bottleneck level analysis results

The level of integrated development of digital and real economies	V1	V2	V3	V3	V5	V6
0.00	NN	NN	NN	NN	NN	NN
0.10	NN	NN	NN	NN	NN	NN
0.20	NN	NN	NN	NN	NN	NN
0.30	NN	NN	NN	0.072	NN	NN
0.40	0.021	0.031	NN	0.193	NN	NN
0.50	0.158	0.145	NN	0.313	0.016	NN
0.60	0.284	0.263	NN	0.431	0.173	NN
0.70	0.416	0.378	NN	0.548	0.322	NN
0.80	0.543	0.482	NN	0.664	0.468	NN
0.90	0.675	0.594	NN	0.785	0.616	NN
1.00	0.809	0.709	0.031	0.907	0.765	0.62

III. B. 2) Necessity analysis of configuration conditions

(1) Necessary condition test

The fuzzy set qualitative comparative analysis method requires a necessary condition test to be performed on the cause variables first; if a cause variable is a necessary condition, it indicates that the results cannot be produced without that condition, and thus there is no need to include this variable in the subsequent qualitative comparative analysis. The results of applying the fsQCA software to analyze the necessity condition of whether the causal variable is a high/low number of real integration development levels are shown in Table 8. Under the measure of whether the consistency is higher than 0.9 as a measure of whether the variable has necessity, the necessity of financial support is higher than 0.9, and for this reason this was excluded from the subsequent qualitative comparative analysis. The consistency thresholds for the other antecedent variables are all below 0.9, indicating that none of the other antecedent variables selected, which are not a single condition of necessity, should be retained in the subsequent truth table calculations.

Table 8: Analysis results of necessary conditions

Cause variable	The level of integrated development of the digital and real economies is high		The level of integrated development of the digital and real economies is low	
	Consistency	Coverage rate	Consistency	Coverage rate
~V1	0.8052	0.8417	0.3064	0.3026
V1	0.3348	0.3368	0.8413	0.8042
~V2	0.8396	0.8926	0.2425	0.2364
V2	0.2864	0.2843	0.8946	0.8416
~V3	0.5639	0.6075	0.5291	0.5375
V3	0.5723	0.5637	0.6154	0.5719
~V4	0.9074	0.8942	0.2792	0.2581
V4	0.2385	0.2576	0.8843	0.8985
~V5	0.7961	0.8749	0.2561	0.2649
V5	0.3342	0.3241	0.8824	0.8044
~V6	0.7145	0.7415	0.3715	0.3672
V6	0.3936	0.3975	0.7393	0.7063

(2) Analysis of conditional grouping results

Fuzzy set qualitative comparative analysis calculates all possible groupings or combinations by constructing a truth table and screening the data by establishing a frequency count threshold and consistency criteria. The frequency number describes the number of samples contained in the grouping state in which it is located, and the consistency measures the degree of approximation that the set of samples of a particular grouping state is a subset of the set of samples of the outcome variable. In this paper, the consistency threshold is set to 0.8 with reference to previous studies, and the frequency number is set to 1. The results of the conditional grouping analysis through fsQCA software are shown in Table 9. Where L1~L3, L1a, L1b, and R1~R3 indicate the grouping categories, "●" represents that the core antecedent variable appears, "⊗" represents that the core antecedent variable does not appear, and "√" represents the presence of marginal antecedent variables, "×" represents the absence of marginal antecedent variables, and "blank" areas indicate that the presence or absence of antecedent variables has no effect on the results. The overall consistency of the paths of the high/low number of real integration development level groupings is much higher than 0.9, and the consistency of individual groupings is also higher than 0.9, which makes it possible to consider that the groupings in Table 9 can be regarded as sufficiently conditional combinations of high/low number of real integration development levels.

Three types of solutions can be obtained by fsQCA: complex solutions, intermediate solutions and simplified solutions. The complex solution is not simplified and presents all possible combination conditions, and generally does not need to do path analysis. Both the simple and intermediate solutions are simplified, but the simple solution is simplified while adding a complex counterfactual analysis, and thus there may be an oversimplification problem, so this paper ends up only explaining and analyzing the intermediate solution. During the analysis of the fsQCA solution, the elements that appear in both the intermediate and simple solutions are called core cause variables with more prominent causal relationships, and the elements that appear only in the intermediate solution are called marginal cause variables with weaker causal relationships. To better represent the differences in groupings, differences based on core conditions are distinguished as first-order equivalence, and groupings with the same core conditions are distinguished as second-order equivalence under first-order equivalence.

Table 9: The results of conditional configuration analysis

	The level of integrated development of the digital and real economies is high				The level of integrated development of the digital and real economies is low		
	Configuration L1		Configuration L2	Configuration L3	Configuration R1	Configuration R2	Configuration R3
	L1a	L1b					
V1	√	√	√	×	√	×	√
V2	●	●	●	●	⊗	⊗	⊗
V3		√	×	×	×	●	●
V5	√		●	●	√	×	×
V6	√	√	×		●		●
Consistency	0.992	0.997	0.952	0.971	0.924	0.993	0.945
Original coverage	0.595	0.434	0.151	0.164	0.091	0.342	0.085
Unique coverage	0.142	0.039	0.028	0.029	0.064	0.305	0.054
Overall solution coverage	0.705				0.451		
Consistency of all solutions	0.991				0.982		

In this paper, three high number of real integration development level realization groupings and three low number of real integration development level groupings are obtained, and the consistency and overall consistency of each path are higher than 0.9. The specific analysis of each grouping is as follows:

(1) Analysis of the high number of real integration development level regional grouping states. Grouping pattern L1 has the highest coverage and is the main path of the integration and development of the digital economy and the real economy. In the grouping state L1, high digital human capital is the core condition, and high data factor marketization, high market demand and high SME development level complement each other as the marginal condition, i.e. digital resources play a leading role, digital environment and digital industry play a synergistic role, and together they reach the goal of promoting high-level digital-real integration development, which is named as digital environment-industry-driven under digital resource domination. This grouping condition is further divided into two second-order equivalent paths. The role of government intervention is not significant in the group state Lla, and the market demand is not significant in the group state L1b, but the path of the integrated development of digital economy and real economy shows consistency in general.

In group state L2 and group state L3, high digital human capital and high market demand are the core conditions, and high data factor marketization is complementary as a marginal condition, which together reach the development goal of high-level digital-real integration, and it is named as the digital human-market demand two-way dominant type.

(2) Analysis of regional groupings of low number of real integration development levels. There are three group states that produce low digital real integration development level, and this result does not have a symmetrical relationship with the high digital real integration development level group state, which is consistent with causal asymmetry. The problem of lack of digital force capital exists in the group states R1, R2 and R3. In the group state R1, even though there are rich digital industrial resources, it is difficult to realize technological progress under the support of lack of talents, and it cannot reach the high level of digital-real integration development. In configuration R2, although the government intervention is more reasonable, the lack of the remaining antecedent conditions also indicates that it is difficult for the region to achieve a high level of digital-real integration development without the support of rich digital industries and digital resources. In configuration R3, although the region has made reasonable government interventions and promoted the development of small and medium-sized enterprises (SMEs), and also has a certain level of marketization of data factors, the lack of digital talents also makes it difficult for the region to achieve a high level of digital-realistic convergence development.

IV. Policy recommendations

Based on the conclusions of the previous analysis, this chapter puts forward policy recommendations such as strengthening support for technological innovation, stimulating the potential of market demand, improving the efficiency of resource allocation, increasing policy support, and strengthening international cooperation and exchange.

IV. A. Enhanced support for technological innovation

(1) Implementing a deep integration strategy: actively exploring the path of deep integration between the digital economy and strategic emerging industries, relying on new-generation information technology, such as big data, cloud computing, artificial intelligence, etc., and promoting the transformation of strategic emerging industries into intelligent and green ones.

(2) Promote cross-border integration and innovation: constantly promote cross-border cooperation among enterprises, break the boundaries of industries, and promote the free flow of innovation factors such as knowledge, technology and talents through the establishment of platforms such as innovation consortiums and co-innovation centers, so as to create an open and shared innovation network, stimulate the vitality of industrial innovation, and open up a new path for industrial development.

(3) Improve the science and technology innovation fund system: attach great importance to and give full play to all kinds of science and technology innovation funds, and continuously increase financial input to provide strong financial support for key technological breakthroughs and transformation of achievements.

IV. B. Stimulating market demand potential

(1) Innovation Demand Oriented: Through government-enterprise cooperation, a market information sharing platform has been built to help enterprises quickly capture market dynamics, promote the research and development of personalized and customized products, and meet the increasingly diversified and personalized consumer demand, so as to provide endogenous impetus for the upgrading of strategic emerging industries.

(2) Explore new business models: Encourage enterprises to actively explore new business models such as digital marketing and e-commerce platforms, and continue to optimize the policy environment for e-commerce, strengthen the construction of network infrastructure, and safeguard data security and consumer rights and interests, so as to escort e-commerce development. Utilizing the cross-border e-commerce platform, it helps enterprises to explore the international market of strategic emerging industries, and promotes the integrated development of international and domestic double-cycle.

(3) Implementing a globalized market strategy: encouraging strategic emerging enterprises to actively integrate into the global industrial chain and participate in international competition and cooperation. Enterprises are encouraged to actively integrate into the “Belt and Road” initiative, broaden export channels, participate in the formulation of international standards, enhance international discourse, and pave the way for the internationalization of the industry.

(4) Optimizing market structure: Continuously promoting the development of strategic emerging industry clusters and building globally competitive industrial agglomerations. Increase support for small and medium-sized enterprises, encourage large-scale enterprises to collaborate with small and medium-sized enterprises, and build an open, synergistic and symbiotic industrial ecology. Through policy guidance and market mechanisms, promote the efficient allocation of resources and enhance the overall competitiveness of the industry.

IV. C. Increased efficiency in resource allocation

(1) Enhancing the efficiency of resource allocation: Continuously promoting the construction of digital infrastructure to provide solid support for the efficient flow of resources. It has actively utilized cutting-edge technologies to build an intelligent decision-making support system, realizing accurate matching and efficient utilization of resource information.

(2) Strengthen supply chain synergy: build a digital supply chain system and promote in-depth synergy between upstream and downstream enterprises in the industrial chain through an intelligent management platform. The government and industry associations should play a bridging role to promote cross-industry cooperation and create a more solid and efficient industry chain ecosystem.

(3) Cutting costs and accelerating market flow: further encourage and promote the popularization of digital transaction modes, strengthen data security and personal privacy protection, create a safe and trustworthy digital transaction environment, promote the free flow of resources and commodities, and stimulate market vitality.

(4) Accurate allocation of talent resources: Talent is the first resource for the development of strategic emerging industries. The government and enterprises work together to build a multi-level talent cultivation system, attracting domestic and foreign high-level talents to settle in the country through preferential policies on the one hand, and strengthening cooperation with higher education institutions and vocational training institutions on the other hand, to customize the cultivation of skilled talents urgently needed by the industry, so as to inject intellectual resources for the industrial upgrading in a sustained manner.

IV. D. Increased policy support

(1) Fiscal and tax incentives: Aiming at the characteristics of strategic emerging industries, such as high R&D investment and long return cycle, design a flexible and efficient fiscal and tax policy system, so as to reduce the financial burden of the enterprises and incentivize them to increase R&D investment continuously. Provide preferential loans, credit guarantees and other cost-cutting financial tools for enterprise financing to enhance their innovation ability.

(2) R&D incentive system: build a comprehensive R&D incentive mechanism to give direct economic rewards or honorary recognition to enterprises that have made major technological breakthroughs and successfully realized the transformation of their achievements, so as to stimulate the innovation power of enterprises. Through the establishment of innovation funds and achievement transformation funds, financial support is provided to promote the rapid commercialization of technology and the formation of a virtuous cycle of innovation and industrial development.

(3) Industrial planning-oriented: Combining with the actual industrial development, scientifically plan the layout of strategic emerging industries. Through policy guidance, combined with the market mechanism, resources are optimally allocated to continuously promote the development of industrial clusters, form scale and synergistic innovation effects, and enhance the overall competitiveness of the industrial chain.

(4) Infrastructure and talent strategy: Strengthen digital infrastructure and intelligent transformation to provide strong fundamental support for strategic emerging industries. Through various channels such as higher education, vocational training and introduction of overseas talents, create a multi-level and high-quality talent team to provide intellectual support for industrial development.

IV. E. Strengthening international cooperation and exchange

(1) Enhance the influence of international standards: actively guide and support enterprises and research institutions to participate in the formulation of international standards, especially in advantageous fields such as new energy vehicles, and through cooperation with international standardization organizations, transform technological advantages into international standards, enhance the status of enterprises in the international market, and strengthen the dominance and discourse power in the global industrial chain.

(2) Broaden international cooperation network: Based on the current pattern of economic development, further broaden international cooperation channels, and utilize international cooperation platforms, such as the Belt and Road and RCEP, to deepen cooperation with countries and regions along the routes in the field of emerging industries.

(3) Encourage transnational investment and cooperation: Support enterprises to participate in international production capacity cooperation and global supply chain reconstruction through various forms such as greenfield investment, mergers and acquisitions, and the establishment of overseas R&D centers, so as to realize the simultaneous introduction and output of technology, promote the optimization of the layout of industries globally, and enhance international competitiveness.

(4) Promote international cooperation in science and technology: Make full use of and give full play to the advantages of scientific research, and strengthen joint R&D projects with top international scientific research institutions and universities. At the same time, the international cooperation projects will be subsidized through the guiding fund and venture capital, so as to accelerate the international transfer and transformation of scientific research results and promote the strategic emerging industries to the high-end of the global value chain.

V. Conclusion

Through DEA method and fuzzy set qualitative comparative analysis, a systematic study on the efficiency of S&T resource allocation and the development path of digital-real integration in China's digital economy has been conducted. The data show that the average value of China's total factor productivity index of S&T resource allocation is 1.074 in 2016-2022, with an overall upward trend, and the efficiency of S&T resource allocation is improved in 27 out of 30 provinces, showing a good development trend. The Great Northwest Comprehensive Economic Zone has the highest efficiency of S&T resource allocation, with a tfpch value of 1.14, mainly due to the joint improvement of technological progress, technical efficiency, pure technical efficiency and scale efficiency. There are obvious differences in the allocation efficiency between regions, with regions such as the east coast and the north coast mainly relying on technological progress and scale efficiency to improve the overall efficiency, while the Great Southwest and the Great Northwest rely on the synergistic improvement of multiple factors. fsQCA analysis results show that the marketization of data elements, digital human capital, and financial support are necessary for the development of a high-level digital-realistic convergence with the effect sizes of 0.253, respectively, 0.227 and 0.349. Two effective integration paths, digital environment-industry-driven and digital human-market demand two-waydriven under the domination of digital resources, are identified, which provide a scientific basis for promoting the integration and development of the digital economy and the real economy. It is recommended to strengthen the support of technological innovation, stimulate the potential of market demand, improve the efficiency of resource allocation, increase policy support, and strengthen international cooperation and exchange.

References

[1] Wickham, H. (2016). Data analysis. In ggplot2: Elegant graphics for data analysis (pp. 189-201). Cham: Springer International Publishing.

[2] Heeringa, S. G., West, B. T., Heeringa, S. G., & Berglund, P. A. (2017). Applied survey data analysis. chapman and hall/CRC.

[3] Tsai, C. W., Lai, C. F., Chao, H. C., & Vasilakos, A. V. (2015). Big data analytics: a survey. Journal of Big data, 2, 1-32.

[4] Mölder, F., Jablonski, K. P., Letcher, B., Hall, M. B., Tomkins-Tinch, C. H., Sochat, V., ... & Köster, J. (2021). Sustainable data analysis with Snakemake. F1000Research, 10, 33.

[5] Perez-Sanz, F., Navarro, P. J., & Egea-Cortines, M. (2017). Plant phenomics: An overview of image acquisition technologies and image data analysis algorithms. GigaScience, 6(11), gix092.

[6] Hasan, B. M. S., & Abdulazeez, A. M. (2021). A review of principal component analysis algorithm for dimensionality reduction. Journal of Soft Computing and Data Mining, 2(1), 20-30.

[7] Challoumis, C. (2024, October). Building a sustainable economy-how ai can optimize resource allocation. In XVI International Scientific Conference (pp. 190-224).

[8] Dobruna, J., Fazliu, Z. L., Humar, I., & Volk, M. (2025). On Optimizing Resource Allocation: A Comparative Review of Resource Allocation Strategies in HetNets. Computer Modeling in Engineering & Sciences (CMES), 142(3).

[9] Di, S., & Wang, C. L. (2012). Dynamic optimization of multiattribute resource allocation in self-organizing clouds. IEEE Transactions on parallel and distributed systems, 24(3), 464-478.

[10] Halabian, H. (2019). Distributed resource allocation optimization in 5G virtualized networks. IEEE Journal on Selected Areas in Communications, 37(3), 627-642.

[11] Raghavendar, K., Batra, I., & Malik, A. (2023). A robust resource allocation model for optimizing data skew and consumption rate in cloud-based IoT environments. Decision Analytics Journal, 7, 100200.

[12] Jayaprakash, S., Nagarajan, M. D., Prado, R. P. D., Subramanian, S., & Divakarachari, P. B. (2021). A systematic review of energy management strategies for resource allocation in the cloud: Clustering, optimization and machine learning. Energies, 14(17), 5322.

[13] Tang, L., & Meng, Y. (2021). Data analytics and optimization for smart industry. Frontiers of Engineering Management, 8(2), 157-171.

[14] Chaoyang Zheng, Yushuo Zhang & Dan Wu. (2025). Corporate Governance, Resource Allocation, and Efficiency in Fisheries: A Data Envelopment Analysis Approach. Frontiers in Economics and Management,6(1),147-159.

[15] Mohd. Shamim Ansari, Monika & Irfan Ali. (2023). Data envelopment analysis in efficiency measuring of District Central Cooperative Banks in India: a case study. International Journal of Operational Research,47(4),440-460.

[16] Mette Asmild, Joseph C. Paradi & Jesus T. Pastor. (2006). Centralized resource allocation BCC models.Omega,37(1),40-49.

[17] Li Ruochen, Luo Yuyan, Chen Borui, Huang Huan & Liu Ping. (2023). Efficiency of scientific and technological resource allocation in Chengdu–Chongqing–Mianyang Urban agglomeration: based on DEA–Malmquist index model. Environment, Development and Sustainability,26(4),10461-10483.