Evaluation of Greenhouse Vegetable Planting Efficiency Based on DEA Method

You Lyu[1], Zhenwei Zhang[2]
[1]*Department of Basic Courses, Beijing University of Agriculture, Beijing 102206, China*
[2]*School of Mathematics, North University of China, Shanxi 030051, China*

Abstract: The modernization of agricultural technology holds an important position in achieving sustainable agricultural development. Practice has shown that greenhouse agriculture will play a significant role in future agricultural development. This study utilized data envelopment analysis (DEA) to evaluate the production efficiency of greenhouse vegetable cultivation. An efficiency evaluation model was established with greenhouse temperature and light radiation intensity as input indicators, and fruit yield, soluble solids content, lycopene content, and sugar-to-acid ratio as output indicators. Based on the environmental and production data collected from the greenhouse, the DEA Frontier software was used to calculate planting efficiency. The optimal production input was determined through result analysis to achieve the highest planting efficiency. Furthermore, optimization was conducted for production methods with lower efficiency, providing reference and basis for technological innovation in greenhouse agriculture.

Key words: Data Envelopment Analysis; Efficiency Evaluation; Greenhouse Agriculture; Input-Output Analysis

1. Introduction

As one of the most populous countries in the world, agriculture in China plays a crucial role in national food security and economic stability. Large connected greenhouses are vital facilities in modern agriculture and horticulture [1]. Typically composed of a series of interconnected greenhouse modules, they are primarily used for cultivating vegetables, fruits, and other crops [2]. the integrated operation of connected greenhouse modules allows for controlled environmental conditions, including temperature, humidity, light, and ventilation, enabling growers to cultivate a variety of fruits and vegetables throughout the year without being constrained by seasonal weather. Simultaneously, environmental control within the greenhouse optimizes crop growth conditions, increasing yields and improving the quality and taste of agricultural products, aiming at achieving efficient agricultural production.

In recent years, large connected greenhouses have flourished in China, with continuously expanding newly constructed areas, and greenhouse technologies becoming increasingly mature. However, greenhouse agriculture faces challenges during its development, notably the issue of energy costs. Maintaining constant environmental conditions within the greenhouse often requires a significant amount of energy [3], particularly in the northern regions of China where winter heating demands are substantial. Therefore, high energy costs may render greenhouse agriculture economically unviable, and the associated energy emissions could pose sustainability challenges. Considering the advantages and disadvantages of connected greenhouse agriculture, achieving high-quality and high-yield crops with minimal energy consumption is a developmental goal. Consequently, scientifically evaluating the production efficiency of greenhouse agriculture and providing valuable insights for the development of large connected greenhouses have significant research value.

DEA (Data Envelopment Analysis) was first proposed by eminent scholars in the field of operations research, including Charnes, Cooper, and Rhodes, in 1978. It is a method for measuring the relative efficiency of decision-making involving multiple inputs and outputs [4, 5]. Its primary computational basis lies in mathematical programming from the field of operations research [6]. As a non-parametric method, DEA does not require knowledge of the functional relationships between inputs and outputs within the evaluated decision-making units but rather allows the "data to speak for themselves" [7].

Today, data envelopment analysis has been widely applied in various fields for efficiency evaluation and decision analysis, continuously expanding and improving in its applications. Lu Shuai et al. [8] used a two-stage DEA method to evaluate fund performance. Quan Wei et al. [9] employed a DEA model to evaluate the efficiency of resource allocation in grassroots healthcare in Chongqing. Li-Jun Sun et al. [10] used DEA to assess the efficiency of corporate financing. In the field of agriculture, Ai-Jun Li et al. [11] studied the efficiency of agricultural supply chains using the DEA method. Xu Qiuyan and Han Hao [12] conducted research on the production efficiency of garlic cultivation in China using a DEA model.

From the current research status, it is evident that most scholars employ DEA models to study the production efficiency of agriculture or specific agricultural products from a macro perspective. Therefore, conducting research on the input-output efficiency of specific vegetable varieties in connected greenhouse cultivation is of significant importance for improving the economic viability of greenhouse agriculture.

2. Theoretical Model and Data Sources

The Data Envelopment Analysis (DEA) method originates from the interdisciplinary fields of operations research, mathematical economics, and management science. the fundamental idea behind measuring efficiency is that if a group of homogeneous decision-making units collectively defines a production possibility set, determining the effectiveness of a decision unit is based on whether it lies on the "production frontier" of this set. the production frontier is an extension in economics of the production function to a multi-output context.

Efficiency for decision units is generally defined as 1 for those on the frontier, while the efficiency measure for inefficient decision units is defined as their relative distance from the frontier. Due to DEA's inherent economic background, models established from input and output data based on DEA theory can be directly utilized for economic analysis. The classical DEA models primarily fall into two categories. One assumes constant returns to scale, known as the CCR model, and the other allows for variable returns to scale, known as the BCC model.

2.1 CCR Model

Assuming there are n decision-making units (DMUs), each DMU having m types of inputs (resource consumption) and s types of outputs (benefits of production), the expression for the CCR programming model is as follows:

$$\begin{cases} \min\left[\theta-\varepsilon(\hat{e}^{\mathrm{T}}S^{-}+e^{\mathrm{T}}S^{+})\right], \\ \sum_{j=1}^{n} X_j\lambda_j + S^{-} = \theta X_0, \\ \sum_{j=1}^{n} Y_j\lambda_j - S^{+} = Y_0, \\ \lambda_j \ge 0. j=1,\cdots,n, \\ S^{-}\ge 0, S^{+}\ge 0. \end{cases} \tag{1}$$

In the formula:

$X_j=(x_1,x_2,\cdots,x_m)^{\mathrm{T}}, j=1,\cdots,n,$

$Y_j=(y_1,y_2,\cdots,y_m)^{\mathrm{T}}, j=1,\cdots,n,$

$\hat{e}=(1,1,\cdots,1)^{\mathrm{T}}\in E^{m},$

$e=(1,1,\cdots,1)^{\mathrm{T}}\in E^{s}.$

Assuming there are *n* decision-making units (DMUs), each DMU having *m* types of inputs (resource consumption) and *s* types of outputs (benefits of production), the expression for the CCR programming model is as follows:

Here, X_j and Y_j represent the input and output vectors of the decision-making unit, respectively, is known data obtained through statistics. λ_j is the weight vector for inputs and outputs, S^{-} is a slack variable, S^{+} is a residual variable, and ε is an infinitesimal quantity according to Archimedean property. θ is the efficiency evaluation index, with a maximum value not exceeding 1. the economic significance of the efficiency evaluation index is as follows: when $\theta=1$, it indicates that the decision-making unit is efficient, maintaining the current input while achieving optimal output; when $\theta<1$, it indicates that the decision-making unit is inefficient, suggesting the existence of input redundancy or insufficient output. For inefficient decision-making units, the closer the efficiency evaluation index is to 1, the relatively higher the production efficiency.

2.2 BCC Model

In CCR model, there is an assumption that scale returns remain constant, meaning the ratio of output increase to input increase remains constant. In actual production, scale effects often vary with changes in input and output scales. To address this, operations researchers such as Banker, Charnes, and Cooper extended the CCR model to establish the BCC model, which allows for variable returns to scale. the programming expression for the BCC model is as follows:

$$\begin{cases} \min\left[\theta-\varepsilon(\hat{e}^{T}S^{-}+e^{T}S^{+})\right], \\ \sum_{j=1}^{n}X_j\lambda_j+S^{-}=\theta X_0, \\ \sum_{j=1}^{n}Y_j\lambda_j-S^{+}=Y_0, \\ \sum_{j=1}^{n}\lambda_j=1, j=1,\cdots,n \\ S^{-}\geq 0, S^{+}\geq 0, \lambda_j\geq 0. j=1,\cdots,n. \end{cases} \tag{2}$$

In BCC model, the definition of the effectiveness of decision-making units aligns with the CCR model. the efficiency value of the same decision-making unit in the BCC model is lower than in the CCR model. Therefore, decision-making units deemed efficient in the BCC model are also efficient in the CCR model. the BCC model represents an enhancement of the CCR model, offering a more flexible consideration of scale elasticity and different input-output orientations while focusing on efficiency assessment. This enables better adaptation in evaluating efficiency across different production scales.

2.3 Data Sources and Indicator Selection

Using the connected greenhouse located in the Daguanyuan area of Xiaotangshan, Beijing, as the experimental site, and focusing on greenhouse tomatoes as the research subject, over ten greenhouses were selected as observation points in each of the six regions. Environmental data, including temperature, humidity, and effective radiation inside the greenhouse, were collected. Growth data, such as yield, quality, and plant development indicators of long-season cultivated tomatoes, were measured. In total, 85 decision-making units were formed.

In conjunction with the research by Han Zequn [13] on identification indicators and screening methods for processing tomato quality traits, soluble solid content, tomato lycopene content, and the sugar-acid ratio were chosen as output indicators for fruit quality. Temperature and effective radiation were selected as input indicators for production. the average values of these indicators within the same observation point were taken as a set of production decision-making units. Data collection for the greenhouse environment in Region 1 is presented in **Table 1**.

3 Empirical Analysis

3.1 CCR Model

To analyze the environmental conditions for the efficient production of high-quality and high-yield tomatoes, greenhouse environmental indicators are used as inputs, and fruit yield indicators are used as outputs. the DEAFrontier software is employed, and through an input-oriented CCR model, the production efficiency of each greenhouse collection point is calculated. Partial calculation results are shown in **Table 2**.

Table 1. Greenhouse Environment and Tomato Growth Data (Region 1)

Region 1 CollectPoint	**Input**		**Output**			
	Temperature /℃	Radiation/ $\mu mol \cdot m^{-2} \cdot s^{-2}$	Yield/kg	Soluble Solid/**%**	lycopene mg/100g	SugarAcidRatio
CollectPoint 1	21.18	128.09	623.88	5.25	9.06	9.28
CollectPoint 2	20.10	99.98	914.70	5.45	8.79	8.58
CollectPoint 3	18.69	133.14	657.42	5.15	8.91	7.63
CollectPoint 4	21.39	98.18	777.23	5.29	9.28	6.92
CollectPoint 5	21.43	105.65	527.07	5.18	7.51	7.61
CollectPoint 6	20.60	114.95	664.92	4.95	6.65	8.13
CollectPoint 7	21.26	103.65	528.34	5.96	9.63	7.53
CollectPoint 8	21.43	90.88	515.80	5.29	8.96	8.48
CollectPoint 9	20.83	107.97	1441.12	5.45	8.23	7.60
CollectPoint 10	28.90	123.19	1429.68	3.24	4.70	7.09
CollectPoint 11	26.02	220.05	1591.93	3.21	3.87	6.42

CollectPoint 12	27.12	133.52	1574.92	3.40	5.27	9.73
CollectPoint 13	21.76	91.19	1422.83	3.68	4.69	5.08
CollectPoint 14	21.71	109.10	1200.62	3.72	4.82	6.72
CollectPoint 15	20.77	96.41	1062.43	3.69	4.84	7.70
CollectPoint 16	20.90	120.30	1312.27	3.69	5.13	8.74

Table 2. CCR Model Result(Region 2)

CollectPoint(Region-Point)	**Efficiency**	**Optimal Multiplier**					
		Temperature/℃	Radiation/ $\mu\mathrm{mol}\cdot\mathrm{m}^{-2}\cdot s^{-2}$	Yield /kg	Soluble Solid/**%**	Lycopen mg/100g	SugarAcid Ratio
CollectPoint 2-1	1.0000	0.00000	0.01369	0.00091	0.00000	0.00000	0.00000
CollectPoint 2-2	1.0000	0.04002	0.00218	0.00000	0.00000	0.12888	0.00000
CollectPoint 2-3	1.0000	0.05050	0.00000	0.00000	0.12213	0.06649	0.00000
CollectPoint 2-4	0.9451	0.04570	0.0000	0.00006	0.00000	0.11372	0.00000
CollectPoint 2-5	0.9632	0.04575	0.00000	0.00006	0.00000	0.11383	0.00000
CollectPoint 2-6	1.0000	0.03551	0.00244	0.00002	0.22405	0.00000	0.00000
CollectPoint 2-7	1.0000	0.00000	0.00983	0.00000	0.00000	0.01227	0.09502
CollectPoint 2-8	0.8708	0.04713	0.00000	0.00000	0.00000	0.09140	0.03812
CollectPoint 2-9	1.0000	0.02622	0.00492	0.00081	0.00000	0.00000	0.00000
CollectPoint 2-10	0.9551	0.04829	0.00000	0.00020	0.15094	0.00000	0.02441
CollectPoint 2-11	0.8320	0.04710	0.00000	0.00042	0.02817	0.00000	0.05713
CollectPoint 2-12	0.9912	0.03833	0.00220	0.00000	0.20853	0.00000	0.01727
CollectPoint 2-13	0.9350	0.04728	0.00000	0.00000	0.16499	0.00000	0.03173

Table 2 presents the production efficiency and indicator weights for each data collection point in Region 2. In the second column of the table, the production efficiency of each decision-making unit is displayed. Six sets of input combinations achieve optimal production efficiency, while the remaining seven decision-making units have efficiencies less than 1. This suggests that, for these seven decision-making units, there is room for improvement in both fruit yield and quality at their current input levels. Alternatively, it could indicate that environmental input should be reduced given the current levels of yield and fruit quality.

Columns three to five of the table represent the weights for each indicator. These weights are determined to maximize the production efficiency of the current decision-making unit. the magnitude of these weights reflects the relative importance of each indicator, providing insights into or aligning with decision-makers' preferences for each indicator. For instance, Collection Points 2-1 and 2-2 both achieve optimal production efficiency. the difference lies in the output aspect: for Collection Point 2-1, the yield weight is the highest, while for Collection Point 2-2, the lycopene weight is the highest. Therefore, if the grower leans towards increasing yield, the input combination for Collection Point 2-1 is more favorable. Conversely, if the focus is on increasing the lycopene content in fruit quality, the input combination for Collection Point 2-2 is more favorable.

Similarly, Collection Points 2-3 and 2-7 both achieve optimal production efficiency. In terms of input, Collection Point 2-3 has a higher temperature weight, while Collection Point 2-7 has a higher effective radiation weight. Consequently, if the cost of maintaining greenhouse temperature is higher, the input combination for Collection Point 2-3 is more favorable. If the cost of radiant energy is higher, then the input combination for Collection Point 2-7 is more favorable.

To comprehensively analyze the characteristics of multiple input combinations across data collection points in each region, the decision-making units that achieve optimal production efficiency and their indicator weights are summarized in **Table 3**.

Table 3 presents the decision units from each data collection point with a production efficiency of 1 (optimal efficiency). It lists the weights of input data and output indicators for each unit. Combining the results allows for a comprehensive analysis:

(1)From Table 3, it can be observed that when the greenhouse temperature is maintained between 20℃ and 22℃:

If the grower prioritizes increasing fruit yield, the optimal production efficiency is achieved with an average photosynthetic radiation intensity of around 73 $\mu\mathrm{mol}\cdot\mathrm{m}^{-2}\cdot s^{-2}$.

If the primary goal is to enhance fruit quality (soluble solids content and lycopene content), the optimal production efficiency is attained with a photosynthetic radiation intensity of approximately 129 $\mu\mathrm{mol}\cdot\mathrm{m}^{-2}\cdot s^{-2}$.

If the focus is on improving fruit texture (sugar-acid ratio), the optimal production efficiency is achieved with a photosynthetic radiation intensity of around 128, $\mu\mathrm{mol}\cdot\mathrm{m}^{-2}\cdot s^{-2}$ and the temperature is controlled at approximately 21°C.

(2)From Table 3, it can be observed that when the greenhouse temperature is maintained between 23℃ and 26℃:

If the primary goal is to increase fruit yield, the optimal production efficiency is achieved with a photosynthetic radiation intensity of approximately 96 $\mu\mathrm{mol}\cdot\mathrm{m}^{-2}\cdot s^{-2}$, and the temperature is controlled at around 25°C. Original production data indicates that at this point, the yield is approximately twice that of the input combination at 20°C with a radiation intensity of 73 $\mu\mathrm{mol}\cdot\mathrm{m}^{-2}\cdot s^{-2}$.

If the focus is on improving taste, the optimal production efficiency is attained with a photosynthetic radiation intensity of about 151 $\mu\mathrm{mol}\cdot\mathrm{m}^{-2}\cdot s^{-2}$, and the temperature is controlled at around 25°C. Production data shows that at this point, the sugar-acid ratio is approximately 1.7 times that of the input combination at 21°C with a radiation intensity of 128 $\mu\mathrm{mol}\cdot\mathrm{m}^{-2}\cdot s^{-2}$.

If the main objective is to enhance fruit quality, the optimal production efficiency is achieved with a photosynthetic radiation intensity of around 113 $\mu\mathrm{mol}\cdot\mathrm{m}^{-2}\cdot s^{-2}$, and the temperature is controlled at approximately 23°C. Additionally, at this level of production input, the fruit's taste is significantly improved.

Table 3. Optimal DMU in CCR Model

CollectPoint (Region-Point)	**Input**		**Output**			
	Temperature/℃	Radiation/ $\mu\mathrm{mol}\cdot\mathrm{m}^{-2}\cdot s^{-2}$	Yield /kg	Soluble Solid/%	lycopene mg/100g	SugarAcidRatio
CollectPoint 1-1	21.18	128.09	0.00000	0.00000	0.00000	0.10772
CollectPoint 1-9	20.83	107.97	0.00046	0.00000	0.04151	0.00000
CollectPoint 1-13	21.76	91.19	0.00070	0.00000	0.00000	0.00000
CollectPoint 2-1	20.30	73.06	0.00091	0.00000	0.00000	0.00000
CollectPoint 2-2	20.17	88.53	0.00000	0.00000	0.12888	0.00000
CollectPoint 2-3	19.80	129.55	0.00000	0.12213	0.06649	0.00000
CollectPoint 2-6	20.75	107.74	0.00002	0.22405	0.00000	0.00000
CollectPoint 2-7	20.73	101.71	0.00000	0.00000	0.01227	0.09502
CollectPoint 3-4	23.23	113.69	0.00000	0.00000	0.08852	0.04071
CollectPoint 3-5	23.98	106.29	0.00000	0.00000	0.12537	0.00000
CollectPoint 4-4	26.19	144.17	0.00000	0.00000	0.03404	0.00000
CollectPoint 4-6	25.27	151.68	0.00000	0.00000	0.00000	0.05917
CollectPoint 4-10	25.04	96.26	0.00047	0.00000	0.00000	0.00000
CollectPoint 4-18	26.83	105.79	0.00029	0.00000	0.00000	0.03976
CollectPoint 5-7	27.03	202.90	0.00021	0.11980	0.00000	0.01132
CollectPoint 5-8	27.27	203.19	0.00027	0.00000	0.02606	0.03861
CollectPoint 5-17	28.63	145.10	0.00048	0.00000	0.00000	0.00000
CollectPoint 5-18	27.55	114.16	0.00052	0.00000	0.04379	0.00000
CollectPoint 3-13	28.17	105.99	0.00077	0.00000	0.00000	0.00000
CollectPoint 3-2	23.11	99.16	0.00016	0.15399	0.02951	0.02763
CollectPoint 5-1	28.92	201.65	0.00005	0.13037	0.00920	0.00582

(3) According to Table 3, the input combination of 23.11°C and 99.16 $\mu\mathrm{mol}\cdot\mathrm{m}^{-2}\cdot s^{-2}$ at Collection Point 3-2, under relatively low input conditions, exhibits positive weights for all output indicators. This indicates that when growers seek a balanced improvement in both yield and quality without excessively emphasizing a single indicator, the production efficiency is optimal under these environmental conditions. Therefore, the

analysis of production efficiency can serve as a reference for production decisions under different types and levels of requirements.

3.2 BCC Model

In the results of the CCR model, multiple decision units achieve optimal production efficiency. To further optimize efficiency and increase output, decision units with a larger sum of output indicator weights are selected. Using fruit yield as the output, an input-oriented BCC model with variable returns to scale is employed for production efficiency calculation. the discussion revolves around how to allocate inputs reasonably to achieve efficiency optimization without reducing output.

Table 4 presents the effective input combinations within the BCC model's efficient decision units, which also happen to be effective input combinations in the CCR model with fruit yield as output. These decision units, while ensuring that fruit yield remains unchanged, already operate at the most energy-efficient levels of input, representing the optimal environmental input levels corresponding to different yield goals. This is also referred to as Pareto optimality.

Table 4. Optimal DMU in BCC Model

CollectPoint (Region-Point)	**Input**		**Output**
	Temperature/℃	Radiation/$\mu mol \cdot m^{-2} \cdot s^{-2}$	Yield/kg
CollectPoint 1-2	20.10	99.98	914.70
CollectPoint 1-9	20.83	107.97	1441.12
CollectPoint 6-14	24.95	158.42	2287.37
CollectPoint 4-10	25.04	96.26	2136.83
CollectPoint 4-11	24.69	97.20	1703.24
CollectPoint 2-1	20.30	73.06	1104.95

Table 5. BCC Model Result

CollectPoint (Region-Point)	**Efficiency**	**Input Target**		**Virtual Unit**					
		Temperature/℃	Radiation/$\mu mol \cdot m^{-2} \cdot s^{-2}$	Benchmark-1	Share	Benchmark-2	Share	Benchmark-3	Share
CollectPoint 1-13	0.98363	21.41	89.69	1-9	0.347	4-10	0.195	2-1	0.458
CollectPoint 1-17	0.97534	20.25	79.08	1-2	0.224	2-1	0.766	——	——
CollectPoint 1-18	0.98701	20.62	94.59	1-9	0.617	2-1	0.383	——	——
CollectPoint 6-2	0.93153	20.47	84.10	1-9	0.316	2-1	0.684	——	——
CollectPoint 6-6	0.89874	20.14	93.83	1-2	0.772	2-1	0.228	——	——
CollectPoint 6-11	0.97457	23.27	55.60	4-11	0.677	2-1	0.323	——	——
CollectPoint 6-13	0.96121	22.98	83.22	4-10	0.507	4-11	0.062	2-1	0.431
CollectPoint 5-6	0.78980	20.10	99.98	1-2	1.000	——	——	——	——
CollectPoint 5-8	0.74084	20.20	86.50	1-2	0.499	2-1	0.501	——	——
CollectPoint 5-17	0.85294	24.42	123.76	1-9	0.138	6-14	0.416	4-10	0.446
CollectPoint 5-18	0.78739	21.69	89.89	1-9	0.310	4-10	0.259	2-1	0.431
CollectPoint 4-18	0.82425	22.11	87.19	1-9	0.162	4-10	0.365	2-1	0.473
CollectPoint 3-13	0.74940	21.11	79.43	1-9	0.074	4-10	0.163	2-1	0.763
CollectPoint 2-13	0.96439	20.52	86.71	1-9	0.389	4-10	0.004	2-1	0.608

According to **Table 5**, the 14 effective decision units in the C^2R model are considered ineffective in the BC^2 model. This is because, under variable returns to scale conditions, these decision units can consume fewer environmental inputs while maintaining their output. In other words, there is input redundancy. the reference baseline for optimizing input to achieve maximum production efficiency is represented by the effective decision units in the BC^2 model, as listed in Table 4. These six input-output combinations collectively form the "production frontier" of a data envelopment model, also known as the data envelopment surface. It provides a direction for improvement for ineffective decision units.

For instance, the decision unit corresponding to Collection Point 5-18 has a production efficiency of approximately 78.7% in the BC^2 model. Its reference baselines on the data envelopment surface are Collection Points 1-9, 4-10, and 2-1, with benchmark shares of 0.31, 0.259, and 0.431, respectively. This means that 31% of the temperature data in the input projection of Collection Point 5-18 comes from the temperature of Collection Point 1-9, 25.9% from Collection Point 4-10, and 43.1% from Collection Point 2-1. the calculation method for the projection of light radiation intensity on the envelopment surface is similar. the input projection obtained by combining these three reference baselines through linear combination for "joint production" is 21.69℃ and 89.89 $\mu\mathrm{mol}\cdot\mathrm{m}^{-2}\cdot s^{-2}$. At this point, input consumption is lower than the original data, and production efficiency can be improved to the optimal level. It is evident from the results that Collection Points 1-9 and 2-1 appear more frequently in the reference baselines of ineffective decision units, reflecting their more prominent efficiency levels in the CCR model with fruit yield as the output indicator.

To compare the consumption of input projections in the CCR model with the consumption of original data from data collection points, an environmental input consumption indicator C is introduced. Its value is the product of temperature value T and light radiation intensity value R, i. e. $C = T \times R$. Figure 1 displays the original input consumption and the consumption of input projections for the decision units as shown in **Table 5**. From **Figure 1**, it is evident that the consumption after projecting onto the data envelopment surface is significantly lower than the consumption of the original inputs. Collection points 5-6 and 5-8, in particular, show savings in input consumption of over 50%. This indicates that by using the optimal reference baseline derived from the CCR model to calculate the best benchmark shares, input costs can be reduced while ensuring no decrease in yield. This effectively enhances the economic and sustainable aspects of crop production.

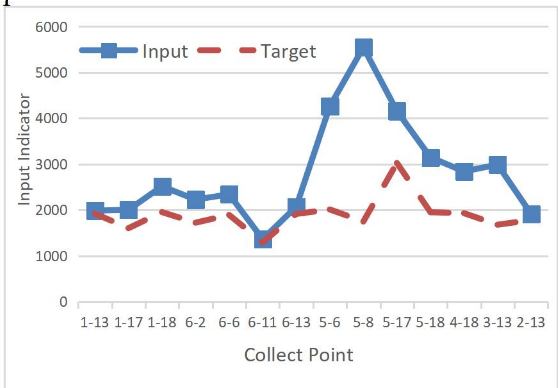

Figure 1. Comparison of Original and Target Input Indicator

4 Conclusion and Outlook

(1) Based on the data envelopment analysis model, taking greenhouse tomatoes as an example, and combining the selection principles of fruit trait identification indicators, we have established a large-scale continuous greenhouse vegetable cultivation efficiency evaluation model. This model utilizes greenhouse environmental conditions as input indicators and fruit yield and quality as output indicators.

(2) Through CCR model calculations and analysis of different greenhouses located in Xiaotangshan, Beijing, the results indicate the following optimal input combinations for improving various aspects:

For increasing fruit yield, the optimal planting efficiency is achieved with temperature at 20°C and light radiation intensity at 73 $\mu\mathrm{mol}\cdot\mathrm{m}^{-2}\cdot s^{-2}$, as well as temperature at 25°C and light radiation intensity at 96 $\mu\mathrm{mol}\cdot\mathrm{m}^{-2}\cdot s^{-2}$.

For enhancing fruit quality, the optimal planting efficiency is attained with temperature at 20°C and light radiation intensity at 129 $\mu\mathrm{mol}\cdot\mathrm{m}^{-2}\cdot s^{-2}$, as well as temperature at 23°C and light radiation intensity at 113 $\mu\mathrm{mol}\cdot\mathrm{m}^{-2}\cdot s^{-2}$.

For improving fruit taste, the optimal planting efficiency is observed with temperature at 21°C and light radiation intensity at 128 $\mu\mathrm{mol}\cdot\mathrm{m}^{-2}\cdot s^{-2}$, as well as temperature at 23°C and light radiation intensity at 113 $\mu\mathrm{mol}\cdot\mathrm{m}^{-2}\cdot s^{-2}$.

(3) In response to the optimal environmental input combinations for planting efficiency in the CCR model, further cultivation efficiency calculations were conducted using the BCC model, with fruit yield as the output indicator. the results indicate that there are 6 sets of temperature and light input combinations where planting efficiency is optimal, while the remaining combinations exhibit input redundancy. the optimal 6 input combinations for planting efficiency can serve as reference benchmarks for other combinations. According to the model results, input combinations constructed based on the benchmark shares derived from the model can reduce input costs without decreasing the original yield, thereby achieving optimal cultivation efficiency.

(4) Through the evaluation and analysis of greenhouse cultivation efficiency, valuable insights are provided for production decisions in greenhouse agriculture. This aims to

enhance agricultural supply capacity and meet diverse food demands. Large continuous greenhouses play an increasingly important role in agricultural production, with potential for further improvement in planting efficiency. Analyzing the results of input-oriented data envelopment analysis, one key pathway to enhancing overall greenhouse planting efficiency is by improving input redundancy. From an input perspective, the results obtained from the model calculations can offer targeted guidance for greenhouse environmental control at different cost levels, particularly in response to fluctuations in the energy market. On the other hand, planting efficiency also quantifies output gaps, providing reference and scientific basis for improving and optimizing resource utilization. This, in turn, promotes innovation in planting technologies and management methods, thereby enhancing the stability and economic viability of greenhouse agriculture. In summary, establishing and refining the greenhouse cultivation efficiency evaluation mechanism is crucial for reducing costs, increasing efficiency, driving agricultural technological innovation, and achieving sustainable agricultural development.

Acknowledgment

This research was supported by the Young Faculty Research and Innovation Capacity Enhancement Program of Beijing University of Agricultural (QJKC-2022037).

References

[1] Wang Binghua, Liu Xiangwei, Wang Yanfang, et al. Modern Large-Scale Greenhouse Tomato Factory Cultivation Management Techniques [J]. Chinese Vegetables, 2021, (04):104-109.

[2] Wang Mingmei, Zhang Yuefeng, Li Shengnan, et al. Current Development Status of Domestic Large-Scale Greenhouse Vegetable Factory Production [J]. Chinese Vegetables, 2023, (05):13-19.

[3] Chen Fan, Sun Weituo, Yu Wenya, et al. Analysis of the Light and Thermal Environment, and Insulation Performance of External Insulation Continuous Greenhouses [J]. Transactions of the Chinese Society of Agricultural Engineering, 2023, 39(06):194-203.

[4] Cai Sanfa, Zhang Guoqiang, Wang Wan. Innovation Performance Evaluation of University Knowledge Economy Circles Based on Data Envelopment Analysis [J]. Journal of Tongji University (Natural Science), 2023, 51(05):682-686.

[5] Ma Zhanxin, Hou Pengbo. DEA Efficiency Measurement and Projection of Effective Decision Units [J]. Systems Engineering: Theory and Practice, 2023, 43(06):1852-1874.

[6] Ji Yuqing. Measurement of Agricultural Financial Support Efficiency in Rural China and Analysis of Influencing Factors [D]. Jilin University, 2023.

[7] Luo Danqi. Research on Production Efficiency and Ecological Efficiency of Alfalfa Planting Enterprises in Gansu Province [D]. Lanzhou University, 2023.

[8] Lu Shuai, Li Shouwei, He Jianmin. Multi-stage Efficiency Considering the Network Position of Funds and Its Prospective Predictive Ability [J]. Chinese Journal of Management Science, 2023, 31(09):22-34.

[9] Quan Wei, Wang Shuai, Luo Yongjun. Efficiency and Input Redundancy Analysis of Primary Healthcare Resource Allocation in Chongqing [J]. Health Economics Research, 2023, 40(07):45-49.

[10] Sun Lijun, Meng Xianwei, Li Xintong. Research on Financing Efficiency and Influencing Factors of Cultural Enterprises under the Background of Industrial Upgrading [J]. Economic and Management Review, 2023, 39(04):134-145.

[11] Li Aijun, Li Na, Wang Chengwen. Research on the Efficiency of Agricultural Product Supply Chains Based on DEA and Malmquist Index Model [J]. Statistics and Decision, 2017, (11):42-45.

[12] Xu Qiuyan, Hao Han. Research on Production Efficiency of China's Garlic Industry Based on DEA Model [J]. Northern Horticulture, 2020, (17):153-159.

[13] Han Zequn, Jiang Bo. Identification Indicators and Selection Methods for Processing Tomato Quality Traits [J]. Hubei Agricultural Sciences, 2014, 53(16):3812-3816.

[14] Joe Zhu. Data envelopment analysis-let the data speak for themselves [M]. Beijing: Science Press, 2016.