Available online at http://ijim.srbiau.ac.ir/
Int. J. Industrial Mathematics (ISSN 2008-5621)
Vol. 14, No. 4, 2022 Article ID IJIM-1568, 12 pages
DOI: http://dx.doi.org/10.30495/ijim.2022.65237.1568
Research Article

Developing a Two-Stage Network Data Envelopment Analysis Model with Desirable and Undesirable Outputs (Case Study of Butter Manufacturing Industry)

A. Shirvani *, A. Azizi †‡

Received Date: 2022-01-03 Revised Date: 2022-06-19 Accepted Date: 2022-07-22

Abstract

Determining and measuring efficiency is an important strategy in decision making. Due to the lack of transparency in the efficiency of production stages in Butter production units, the purpose of this study is to design a data envelopment analysis model to evaluate the efficiency of production stages of these units. The contribution of this research is to develop a model for evaluating the efficiency of a two-stage production system with both desirable and undesirable output using the network model of Khalili and Shahmir (2015). The developed two-stage model has been conducted using accurate and certain data of twenty Butter manufacturing factories in 1399, which can evaluate efficiency separately and in an integrated manner. At the end of the first stage, there are two types of desired output. There is a desired output of breakfast cream that does not enter the second stage of production and it is supplied to market for sale, and another desirable output is the pasteurized cream that is transferred to the second stage of production. At the end of the second stage, a desirable output, which is the packaged butter and an undesirable output, which is butter dough are obtained. The efficiency of all twenty units was evaluated and compared. Depending on the efficiency of the whole process, there was only one efficient unit. According to the efficiency of the first stage, there was only one efficient unit out of twenty units, and according to the second stage, there are two efficient units. The conclusion is that this industry has a high inefficiency in the factories located in provinces of Tehran and Alborz on their production process. Thus, it is necessary for the managers of this industry to increase efficiency of their production process.

Keywords : Efficiency; Two-stage production network: Data Envelopment Analysis; Undesirable output.

1 Introduction

Despite the growth and improvement of food production conditions, methods, and techniques, food safety, and health are still significant public health concerns. According to the World

*Department of Industrial Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran.
†Corresponding author. azizi@srbiau.ac.ir, Tel:+98(912)1813261.
‡Department of Industrial Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran.

Health Organization (WHO), up to 30 percent of people in industrialized countries suffer from foodborne illness each year. Another critical issue in the food production sector is the mismatch between production growth and demand. Another issue is that food production must be hygienic and high quality [9].

For maximizing the reliability and efficiency of complex industrial production systems, such as food products, management decisions depend on the experience of the manager and the organization, and success is not achieved with limited or inaccurate information. Therefore, the researchers task is to know the components, variables, steps, and operating times to increase reliability and efficiency [4].

One of the major industries in Iran is the butter industry. More than five hundred tons of cream are transferred daily to butter factories to produce animal butter from the cream. As the production volume is high, the efficiency of these production units to prevent waste of resources should be enhanced. So far in Iran, the efficiency of this industry has not been evaluated and compared; therefore, the inefficiency of these production units has been unknown. If these production units overall efficiency and two-stage efficiency are identified, their production costs can be reduced, and the waste of resources may be prevented.

Nevertheless, how to evaluate the efficiency of these production units? The philosophy of network data envelopment analysis (NDEA) models ignores the internal steps and processes within decision-making units (DMUs) in traditional ways. Conventional DEA models consider each company as a DMU and confine its calculations to its initial inputs and final outputs. A further explanation of DEA is that it is a conventional non-parametric method for evaluating the relative efficiency of DMUs with multiple inputs and outputs. In recent years, DEA models have been developed by many researchers in various fields, such as sustainable supplier selection, power distribution unit ranking, supplier sustainability assessment, and predicting supplier group membership in the sustainable supply chain (Boudaghi & Farzipoor Saen, 2018; Tavassoli & Farzipoor Saen, 2019) [[8].

In the present study, an animal butter factory was selected as a case whose production units have two production stages. The traditional DEA models cannot determine their efficiency; therefore, a model should be developed. The efficiency with the desirable output in the first production stage, the efficiency with the undesirable output in the second production stage, and the overall efficiency can be evaluated. Thus, a deterministic model, in which desirable and undesirable variables are considered, is presented based on the Khalili and Shahmir (2015) NDEA model. The differences and advantages of the proposed model compared to Khalili and Shahmir study can be summarized as follows [5]:

(i) In this proposed model, two production stages are investigated: the production of pasteurized cream and the production of butter;

(ii) At the end of the first production stage, a desirable output, i.e., breakfast cream, leaves the production line. It does not enter the second production stage; and

(iii) Undesirable output is buttermilk, appearing in the second production stage. It has an essential and effective variable in the efficiency of butter production factories.

Therefore, the main research question is, what is the network performance evaluation model of a two-stage production system with desirable and undesirable outputs?

2 Review of literature

The DEA[1] model, developed by Charnes, Cooper, and Rhodes (1978), is a mathematical technique for evaluating the performance of decision-making units with similar inputs and outputs. Traditional DEA models ignore the internal structure of the units and have a black-box view. Then, network DEA (NDEA) models were developed to evaluate structures with more than

[1]Data envelopment analysis

one stage; however, these models are static and cannot assess the units. Dynamic DEA (DDEA) models were developed to assess units over time, considering the unit structure as a black box in each period. Many organizations (including banks), whose continuous activities in successive periods, have a multi-stage process. Thus, using NDEA and DDEA models alone is insufficient to evaluate them.

In network models, Overall efficiency includes the whole system or organization, defined differently depending on the type of structure, application, and model. Since network models for an organization have several parts, each of which contains several DMUs, efficiency is also measured for its subunits.

The proposed models of the primary network are of the radial CCR and BCC type models, which mean that the changes in inputs and outputs occur in the same proportion.

Standard models in data envelopment analysis consider decision-making units as a general system. In many cases, a decision-making system or unit consists of several decision-making processes or subunits. One type is a two-stage system in most of which the efficiency of each part is first calculated, and then the overall efficiency is obtained using the weighted arithmetic mean of divisional efficiencies. Divisional efficiency and overall efficiency always have a value between zero and one. The use of performance measurement models based on NSBM auxiliary network variables has been used in recent years due to very high capabilities in applied fields where the changes of inputs and outputs are not the same ratio [3]. Figure 1 shows a two-stage system in

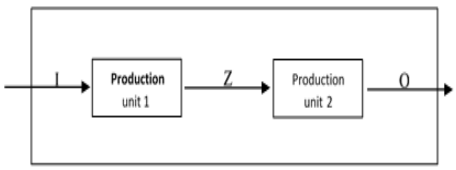

Figure 1: A two-stage network

which I, O, and Z are the decision-making units input and output, and the intermediate relationship between the subunits, respectively. The outputs of the first subunit are the inputs of the second subunit. The second subunit consumes no output, and the first subunit produces no output (Kao & Hwang, 2008). The general shape of a series system is shown in Figure 2 g subunits are related to each other. $x_{ik}(i = 1, \ldots, m)$ and $y_{rk}(r = 1, \ldots, s)$ are the inputs and outputs of $\text{DMU}_k(k = 1 \ldots, n)$, respectively [3] In

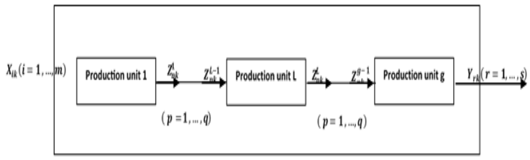

Figure 2: A series network

$Z^t_{pk}(p = 1 \ldots, q)$, P is the second subunit of t from DMU_k. In fact, $Z^t_{pk}(p = 1 \ldots, q)$ is the output of subunit t and the input of subunit $t + 1$. Note that the output of subunit g is the output of the whole system, and the input of subunit 1 is the input of the whole system. The following model for calculating the efficiency of a series system is presented by Kao and Hwang (2008):

$$
\begin{aligned}
E_k = \max & \sum_{r=1}^{s} u_r y_{rk} \\
\text{s.t.} \quad & \sum_{i=1}^{m} v_i x_{ik} = 1, \\
& \sum_{r=1}^{s} u_r y_{rk} - \sum_{i=1}^{m} v_i x_{ij} \leq 0, \quad j = 1, \ldots, n \,(2a) \\
& \sum_{p=1}^{q} w^1_p z^1_{pj} - \sum_{i=1}^{m} v_i x_{ij} \leq 0, \quad j = 1, \ldots, n \,(2b) \\
& \sum_{p=1}^{q} w^t_p z^t_{pj} - \sum_{p=1}^{q} w^{t-1}_p z^{t-1}_{pj} \leq 0, \; t = 2, \ldots, g-1, \\
& \qquad j = 1, \ldots, n \,(2c) \\
& \sum_{r=1}^{s} u_r y_{rk} - \sum_{p=1}^{q} w^{t-1}_p z^{t-1}_{pj} \leq 0, \; j = 1, \ldots, n \,(2d) \\
& u_r, v_i, w^t_p \geq \varepsilon \quad r = 1, \ldots, s, i = 1, \ldots, m, \\
& p = 1, \ldots, q, t = 1, \ldots, g-1
\end{aligned}
\tag{2.1}
$$

Constraint (2a) corresponds to the performance of the whole system, constraint (2b) is related to the efficiency of the first subunit, constraint (2c) is associated with all subunits except the first and last subunit, and constraint (2d) corresponds to the last subunit [3].

Lee et al. (2022) proposed a new data envelopment analysis (DEA) approach to measure the environmental inefficiencies of two-stage structures with adverse intermediate measures and then apply this approach to 30 provinces in China with pollution generation and treatment activities. Consequently they obtained the following results: (1) China executes well in overall efficiency, with an average inefficiency score of

0.3214. (2) Hunan, Hainan, Chongqing, Ningxia and Hebei are the top five least efficient regions, while Xinjiang, Yunnan, Guizhou, Liaoning and Heilongjiang are the top five most inefficient. (3) Most regions have higher pollution treatment inefficiency than the production inefficiency. (4) The eastern region has a better overall efficiency than the central and western regions, and the developed coastal economic zones have a better efficiency than the remote inland regions (5). The eastern region is consistently better than the western region in both production and pollution treatment, while the central region is similar to the western region in terms of production and is superior to the eastern region in pollution treatment.

Lee et al. (2022) examined the efficiency changes and efficiency levels of maternal and child health hospitals in Turkey from 2014 to 2017 using two-stage data envelopment analysis. In the first stage of their study, the efficiency of hospitals was evaluated with input-oriented CCR (Chaarns, Cooper, Rhodes) and Malmquist Productivity Index (MPI) Has been employed. In the second step, Tobit panel regression analysis was used to identify the factors affecting the efficiency scores in the first step. Based on the results of the first stage, the average efficiency values calculated for the years 1393, 1394, 1395 and 1396 were 0.83, 0.90, 0.86 and 0.80, respectively. In the second stage, it was found that the variables of not being in the metropolis and having at least 200 beds had a statistically significant effect on efficiency score ($p < 0.05$) while being a teaching and research hospital had no significant effect ($p > 0.05$).

Omrani et al. (2022) developed a two-stage network data envelopment analysis (NDEA) model with negative input and undesirable outputs. First, a linear model based on the goal programming method was proposed to avoid nonlinear calculations. Then a method was used to change the negative cases to the positive and undesirable outputs to the desired outputs. Finally, the proposed model is developed using αcut fuzzy approach to combine data uncertainty in the linear goal programming (GP) model. To confirm the validity of the proposed model, a real case of 22 insurance companies has been examined.

Via DEA, Kermantzis et al. (2022) used a variety of approaches in one-stage and two-stage contexts to obtain accurate performance in evaluating decision-making units. Their studies are based on a generalized two-stage structure with additional inputs in the second stage and the ratio of the outputs of the first stage as final outputs. They argued that in this context, the models accuracy increases with increasing measures of the strength of the weight model. He states that this aspect develops the knowledge of DEA methodology and has not been considered so far. A new combination of a high self-efficacy model, a minimax second goal model, and the importance of criteria through the criterion method will enhance these aspects and thus achieve a better degree of cooperation between the steps of a DMU and among DMUs. It is suggested that each DMU intends to evaluate the worst-case scenario at its best. Practical applications of this approach could include supporting the identification of training needs in workflow generation or evaluating sustainable supply chains.

Hanum (2021) aimed to promote the use of network data envelopment analysis (NDEA) to measure the performance of manufacturing companies. NDEA can provide a single composite index derived from multidimensional criteria, showing the overall production performance. An NDEA-based proposal is tested in a pharmaceutical company to show how it works in assessing profitability and marketability. We change the existing profitability and marketability framework to a three-stage model: production acquisition, profitability, and marketability, to illustrate the complexities of production operations. The proposed NDEA-based PMS is useful for conducting a performance review to seek inefficiency sources and determine improvement targets.

Wang et al. (2021) suggested a dynamic network data envelopment analysis approach to assess the environmental performance of China's two-stage industrialization during 2015-2010, reflecting the context of production, considering that the previous studies did not consider the feature of being dynamic,. They have provided practicality and achieving comparable efficiency between different periods. The main results of the study are: First, not considering the dynamic features leads to underestimating the two-stage environmental efficiency. Second, the production efficiency was higher than the treatment efficiency and the two-stage environmental efficiency was between them during the study period. East of China had the best efficiency in terms of two-stage environmental productivity, including production and treatment efficiency, and then central and western regions were ranked. Finally, one-fifth of the 30 regions had relatively low values of input and output efficiencies in both production and refining stages based on clustering analysis.

Lee et al. (2020) in their study presented a two-stage data envelopment analysis model with undesirable input to evaluate the efficiency of recycling and reuse of industrial waste, including wastewater, waste gases and solid waste in China. The model was based on the circular structure of the economy and was proposed by the China Institute of Environmental Protection. Data were collected during the implementation of critical environmental policies: Eleventh Five-Year Plan (2006-2010). Waste treatment efficiency for solid waste has been analyzed during the Eleventh and Twelfth Five-Year Plan periods (2011-2015). Unlike previous studies, which have focused primarily on pollution disposal efficiency, the research took into account both pollution characteristics and waste sources. The results show that the integrated efficiency of the industrial waste processing system is stable over five years. It has also been shown that efficiency in the pollution disposal phase is greater than the resource reuse phase, which contributes to the overall efficiency. Finally, it shows that there are three modes of efficient development: coordinated, ecosystem, and moderated development. They have also provided some criteria from the developed countries for different modes of development that can benefit from their research in the future by identifying their development methods and formulating their environmental policy based on these criteria.

Also, the risk of banking activities was considered by considering arrears as undesirable output. The results of stage and network efficiency show that the model provides a better evaluation of banks' efficiency and identifies the source of inefficiency well. The proposed models can be extended to other structures and applications.

Haruna et al. (2020) used the DEA-CRS and DEA-VRS approaches in their study to calculate the efficiency scores of broiler farms in Nigeria. They determined that the average scores of farm technical efficiency with DEA-CRS was 0.83 and with DEAVRS 0.88. Approximately 7% of broiler farms offered fixed-scale returns, 85% of broiler farms offered incremental returns to scale, and 8% decreased returns to scale.

In order to evaluate the performance of banks, Momeni et al. (2017) first developed a radial network data envelopment analysis model and then a non-radial model based on Russell's idea. Banks, like many organizations, have a two-stage network structure. In the next stage, the non-radial model presented in an applied and experimental study was used to evaluate the two-stage structure of 21 banks, including the deposit collection stage and the profitability stage.

Shafiee Nikabadi et al. (2017) investigated that the overall efficiency of the electricity generation and transmission process in sixteen production and transmission areas of the country's electricity industry as decision-making units. Based on Khalili and Shahmir model, a mathematical model was proposed, which includes two stages of electricity generation and transmission. In the proposed model, at the end of the production stage, output variables were considered that no longer enter the next stage of production. In the second stage, in spite of the inputs received from the production stage, other input variables were considered that do not enter the model from the previous stage. In the first stage, the internal consumption variables of power plants and fuel consumption were considered as input variables and the variables of maximum production load, special production, special production, efficiency, nominal power of the power plant, practical power of the power plant were considered as output variables. In the second stage, the variables of maximum production load, special production, special production, capacity of power transmission substations, and length of power transmission lines are considered as input variables and delivered energy variable as desirable output and energy losses as undesirable output variable. Totally, the regions of Azerbaijan, Isfahan, Tehran, Khorasan, Khuzestan, Semnan, Fars and Kerman were recognized as efficient units.

Mohaqer et al. (2016) developed that data envelopment analysis models to measure the efficiency of provincial units of Islamic Azad University as decision-making units in multi-stage systems with parallel network structure in a case study. They designed a model to measure the efficiency of a parallel network structure by simultaneously considering undesirable outputs and uncontrollable inputs. To design such a model, the provincial units of Islamic Azad University in the horizon of 1404 were studied and using data envelopment analysis without orientation to input or output with a preferential structure. Finally, a model for measuring the efficiency of decision-making units and the whole university was designed.

Via a dynamic multi-stage data envelopment analysis (DMS-DEA) approach, Khalili Damghani et al. (2015) designed a model to evaluate energy efficiency in cotton production. In this model, fields that produce cotton using resources such as pesticides, fertilizers, and seeds are used as DMUs. Input variables that are not consumed are passed to the next period. Finally, a DMS-DEA model was used in the real world, in energy consumption, in the cotton industry, to operate all DMUs with dynamic inputs.

3 Modeling

In this study, animal butter production units were investigated. These production units had an undesirable output, i.e., buttermilk. The emergence of this undesirable output is part of the natural and inevitable process of this type of production system, but the amount of this undesirable output can be due to various reasons such as quality of cream received; device life, the quality of construction of machines, devices, and equipment. The higher the amount of these undesirable outputs, the lower the profitability of these production units.

Here, the research method is designed in three stages. The first step selected variables through library and field studies (interviewing experts). In the second step, a questionnaire was prepared, and the questionnaire was employed to assess the significance of input and output indicators separately. In the third step, efficiency is calculated using the NDEA model.
Step 1: To compile performance evaluation indicators of lean production, past studies, documents, laws, instructions, and bylaws, theoretical records, and backgrounds, based on scientific achievements and records and empirical support, were used.
Step 2: A questionnaire was designed. For investigating its content validity, the opinion of experts, i.e., are the founders, managers, and specialists in dairy products, were used; they were asked to comment on the weight of the questionnaire indicators. Following this procedure, finally, the crucial variables that could be used in this study were determined (Table 1).
Step 3: This step, which is to measure the efficiency of the studied units, is discussed in the continuation of this article and the findings section.

To determine the number of workstations in the research conceptual model, first, a visit was made to the factory environment, and a field study was conducted. In the next step, using the opinions of experts, founders, and factory managers, changes were made in the number of workstations. A secondary conceptual model was set up with two workstations.

In each decision-making unit here, i.e., the animal butter production plant, $\mathrm{DMU}_j, (j = 1, \dots, n)$ has m input variables $X_{ij}, (i = 1, \dots, m)$ in step 1. Among these inputs, a number as many as q outputs are produced. They produce a desirable output in the first production stage, i.e., pasteurized cream. It enters the second production stage as the input of the second stage of

Table 1: Final indicators

Description	**Indicators**
Input	Cost of raw materials
Input	maintenance costs
Input	Operator costs
Output	The amount of produced breakfast cream (desirable)
Output	The amount of produced buttermilk (undesirable output)
Output	The amount of packed butter (desirable)

production $Z_{pj}, (p = 1, \dots, q)$. Also, in the first production stage, they will have a desirable output, i.e., breakfast cream, which does not enter the second production stage. It is shown with $y_{gj}, (g = 1, \dots, a)$.

In addition to the variable received from the first production stage, i.e., pasteurized cream, there are other new input variables in the second production stage. They are limited to this stage and have nothing to do with the first stage. These variables are represented by $z^2_{lj}, (l = 1, \dots, c)$. The final output of the system contains a desirable output variable, which is packed animal butter $y_{rj}, (r = 1, \dots, s)$ and has an undesirable output, which is buttermilk $x^2_{bj}, (b = 1, \dots d)$.

Since the production line has two stages, to evaluate the efficiency of this production line, the overall efficiency score (e_0), the efficiency score of the first stage (e_1), and the efficiency score of the second stage (e_2) are calculated separately. The production line process is presented in Figure 3 below.

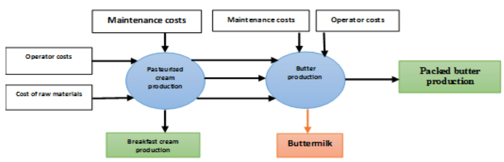

Figure 3: The proposed two-stage model

The model parameters

j: DMU counter

i: Input counter first step

k: Count the intermediate values of the steps

r: Counter of final optimal outputs

p: the embedded unit under investigation

X_{ij}: The input rate of each DMU

y^b_{rj} : The amount of the r^{th} undesirable final output of for j^{th} unit (amount of produced buttermilk)

z_{kj}: The k^{th} average value for j^{th} unit (amount of pasteurized cream)

y''_{rj}: The r^{th} desirable final output of for the j^{th} unit (amount of packed butter)

y'_{rp}: The the r^{th} desirable output of the first stage of for the p^{th} unit (amount of breakfast cream).

Decision variables:

$\theta, \lambda^1_j, \lambda^2_j, \lambda^3_j$: Dual variables in the network data envelopment analysis model

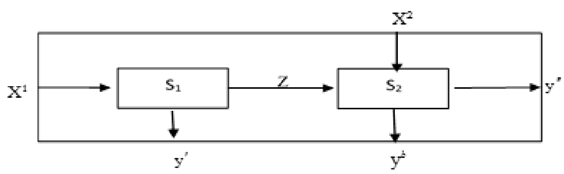

Figure 4: The two-stage model

The overall performance of the NDEA model

$$T_c = \left\{ \begin{pmatrix} x^1 \\ x^2 \\ z \\ y' \\ y'' \\ y^b \end{pmatrix} \middle| \begin{array}{c} x^1 \geq \sum_{j=1}^{n} \lambda_j^1 x_j^1 \\ \& \sum_{j=1}^{n} \lambda_j^2 x_j^2 \leq x^2 \\ \& \sum_{j=1}^{n} \lambda_j^1 z_j \geq \sum_{j=1}^{n} \lambda_j^2 z_j \\ \& \sum_{j=1}^{n} \lambda_j^1 y_j' \geq y' \\ \& \sum_{j=1}^{n} \lambda_j^2 y_j'' \geq y'' \\ \& \sum_{j=1}^{n} \lambda_j^2 y_j^b = y^b \\ \& \lambda^1 \geq 0, \lambda^2 \geq \lambda^2 \geq 0 \end{array} \right\} \tag{3.2}$$

The letter C in T_C indicates the principle of Constant Return To Scale. Evaluating a unit in this set leads to build the model:

$$\begin{aligned} & \min \theta \\ \text{s.t.} \quad & \sum_{j=1}^{n} \lambda_j^1 x_{ij}^1 \leq \theta x_{ip}^1 && \forall i \\ & \sum_{j=1}^{n} \lambda_j^1 z_{kj} \geq \sum_{j=1}^{n} \lambda_j^2 z_{kj} && \forall k \\ & \sum_{j=1}^{n} \lambda_j^1 y_{rj}' \geq y_{rp}' && \forall r \\ & \sum_{j=1}^{n} \lambda_j^2 x_{ij}^2 \leq \theta x_{ip}^2 && \forall i \\ & \sum_{j=1}^{n} \lambda_j^2 y_{rj}'' \geq y_{rp}' && \forall r \\ & \sum_{j=1}^{n} \lambda_j^2 y_{rj}^b \geq y_{rp}^b && \forall r \\ & \lambda_j^1 \geq 0, \lambda_j^2 \geq 0 && \forall j \\ & \theta : \text{ free} \end{aligned} \tag{3.3}$$

Consider the following answer:

$$\begin{aligned} & \lambda_j^1 = \lambda_j^2 = 0, \forall j, \, j \neq 0 \\ & \theta = \lambda_0^1 = \lambda_0^2 = 1 \end{aligned} \tag{3.4}$$

Obviously, Equation (3.4) is a valid answer for Model (3.3) and applies to all constraints of the model. In other words, regardless of the quantitative value of the model inputs, intermediate sizes, and outputs, there will always be a valid answer for model (3.3), which we denote by θ; therefore, model (3.3) is always feasible.

On the other hand, the objective function of model (3.3) is a minimization function with optimal value θ^* whose value, in any case, is less than or equal to the value of the objective function of the given justified answer θ. Therefore, the objective function will have an optimal value, and the objective function will have a relation as $\theta^* \leq \theta \leq 1$. In this case, the optimal value of the objective function of model (3.3) is ≤ 1. As a result, outputs can be reduced without changing the inputs.

Also, a feasible solution such as $(\bar{\theta}\bar{\lambda}_j^1, \bar{\lambda}_j^2)$ is assumed. If $\bar{\theta} = 0$, then from the first, fourth, and fifth constraints of model (3.3) it follows that $\bar{\theta} = 0\bar{\lambda}_j^1 = \bar{\lambda}_j^2$; therefore the value of the output variable of the assessed unit in the second and sixth constraints will be $y_{go} = 0, y_{ro} = 0$, which is contrary to the assumption. The vectors (x_j, y_j) have at least one non-zero component, the result of which is that the optimal value of the objective function of model (3.3) is $\theta^* > 0$. As a result, outputs can be reduced without changing the inputs.

Model (3.5) is the first step in calculating the maximum achievable value to achieve the efficiency of decision-making units.

$$\begin{aligned} & \min(\theta^1 + e_p^* * \lambda_p^1) \\ \text{s.t.} \quad & \sum_{j=1}^{n} \lambda_j^2 x_{ij}^1 \geq \theta^1 xp && \forall i \\ & \sum_{j=1}^{n} \lambda_j^2 y_{rj}' + e_p^* * \lambda_p^1 y_{rp}' \geq y_{rp}' && \forall r \\ & \sum_{j=1}^{n} \lambda_j^2 z_{kj} - \sum_{j=1}^{n} \lambda_j^3 z_{kj} \geq z_{pj} && \forall k \\ & \sum_{j=1}^{n} \lambda_j^3 z_{kj}^2 \leq \theta 0 && \forall k \\ & \sum_{j=1}^{n} \lambda_j^3 x_{ij}^2 \leq 0 && \forall i \\ & \sum_{j=1}^{n} \lambda_j^3 y_{rj}' + e_p^* * \lambda_p^1 y_{rp}' \geq 0 && \forall r \\ & \lambda_j^2 \geq 0, \lambda_j^3 \geq 0 && \forall j \\ & \theta, \lambda_p^1 : \text{ free} \end{aligned} \tag{3.5}$$

Consider the following answer:

$$\lambda_j^2 = 0, \forall j, \; j \neq 0$$
$$\lambda_j^3 = 0, \forall j, \; j \neq 0 \tag{3.6}$$
$$\theta^1 = \lambda_0^1 = \lambda_0^2 = 1$$

Equation (3.6) is a valid answer for Model (3.5) and applies to all model constraints. In other words, regardless of the value of the input values, intermediate sizes, and model outputs, there will always be a valid answer for model (3.5), which we denote by θ^1; therefore, model (3.5) is always feasible.

On the other hand, the objective function of model (3.5) is a minimization function with an optimal value of θ^{1*} whose value is, in any case, less than or equal to the value of the objective function of the assumed justified answer θ^1. Therefore, the objective function will have the optimal value, and the objective function of the assumed justified answer will have the relation; $\theta^{1*} \leq \theta^1 \leq 1$. In this case, it is evident that the optimal value of the objective function of model (3.5) is ≤ 1 and will be finite.

Also, a feasible solution such as $(\bar{\theta}^1, \bar{\lambda}_j^1, \bar{\lambda}_j^2, \bar{\lambda}_j^3)$ is assumed. If $\bar{\theta}^1 = 0$, then the first constraint of model (3.5) show that $\bar{\lambda}_j^2 = 0$, therefore the value of the output variable of the assessed unit in the second constraint will be $y_{go} = 0$, which is contrary to the assumption that the vector (xj, yj) has at least one component. As a result, the optimal value of the objective function of model (3.5) is $\theta^{1*} > 0$.

Model (3.7) is the second stage in calculating the maximum accessible value to achieve the efficiency of decision-making units.

$$\begin{aligned}
&\min \theta^2 \\
&\text{s.t.} \sum_{j=1}^{n} \lambda_j^2 x_{ij}^2 \geq e_p^* \lambda_p^1 x_{ip} && \forall i \\
&\sum_{j=1}^{n} \lambda_j^2 y'_{rj} \geq \lambda_p^1 y''_{rp} && \forall r \\
&\sum_{j=1}^{n} \lambda_j^2 z_{kj} + \theta^2 z_{kj} \geq \sum_{j=1}^{n} \lambda_j^3 z_{kj} && \forall k \\
&e_p^* \lambda_p^1 z_{kj} \sum_{j=1}^{n} \lambda_j^3 z_{kj} \geq \theta^2 z_{kj} && \forall k \\
&e_p^* \lambda_p^1 x_{ij}^2 \sum_{j=1}^{n} \lambda_j^3 x_{ij}^2 \leq \theta^2 x_{ij}^2 && \forall i \\
&\sum_{j=1}^{n} \lambda_j^3 y_{rj}^b \geq \lambda_p^1 y_{rp}^b && \forall r \\
&\lambda_j^1 \geq 0, \lambda_j^2 \geq 0, \lambda_j^3 \geq 0 && \forall j \\
&\theta^2 : \text{free}
\end{aligned} \tag{3.7}$$

Consider the following answer:

$$\lambda_j^1 = 0, \lambda_j^3 = 0, \forall j$$
$$\theta^2 = 1 \tag{3.8}$$
$$\lambda_o^1 = \lambda_o^2 = 0$$

Equation (3.8) is a valid answer for Model (3.7) and applies to all model constraints. In other words, regardless of the value of the input values, intermediate sizes, and model outputs, there will always be a valid answer for model (3.7), which we denote by θ^2; therefore, model (3.7) is always feasible.

Furthermore, the objective function of model (3.7) is a minimization function with optimal value θ^{2*} which is, in any case, less than or equal to the value of the objective function of the assumed justified answer θ^2. Therefore, the objective function will have an optimal value, and the objective function of the assumed justified answer will have a relation; $\theta^{2*} \leq \theta^2 \leq 1$. In this case, it is evident that the optimal value of the objective function of model (3.7) is ≤ 1, and will be finite.

Also, a feasible solution such as $(\bar{\theta}^1, \bar{\lambda}_j^1, \bar{\lambda}_j^2, barlambda_j^3)$ is assumed. If $\bar{\theta}^2 = 0$, then, the fists, fourth, and fifth constraints of model (3.7) show that $la\bar{m}bda_j^3 = \bar{\lambda}_j^2$; therefore the value of the output variable of the unit under evaluation in the sixth constraint will be $y_{ro} = 0$, which is contrary to the assumption that the vector (xj, yj) have at least one non-zero component, the result of which is that the optimal value of the objective function of model (3.7) is $\theta^{2*} > 0$.

Research innovation

The study population in this study are butter production units. These production units are located in Tehran and Alborz provinces. In this research, the efficiency of these production units is evaluated. The innovation of this research is a method used to calculate efficiency in this industry for the first time.

In this method, for the first time using the non-parametric method of data envelopment analysis, in two stages with a desirable output in the first stage and a desirable output and an undesirable output in the second stage, the average efficiency of processes has been calculated.

4 Findings

After selecting the final inputs and outputs, the raw data, separately for each DMU, are entered into MATLAB software to perform calculations to determine the units performance by stochastic analysis of random network data. According to the calculations, the studied units are divided into two groups of efficient and inefficient, and among the units, efficient units are distinguished from inefficient ones. The values of none of the inputs and outputs listed in Table 2 have a negative value. Thus, the model is not built for negative data.

As the whole process efficiency showed, the efficient unit was unit 3 in 2020. According to the first stage, unit 3 was efficient, and units 1 and 3 were efficient according to the second stage.

In the general process, Unit number twenty is the most inefficient unit; that is, it ranked in the twentieth place among the twenty units. It is concluded that the most inefficient butter production unit, among these twenty units, is the twentieth production unit. In general, the conclusion that can be drawn is that this unit has more weakness in the first stage, which is the production of pasteurized cream. If this unit becomes efficient, it can have higher profits and lower costs with a higher overall efficiency score.

5 Conclusion

This study aimed to calculate the average efficiency of twenty units of waste butter production in 2020 in Tehran and Alborz provinces in Iran by a non-parametric DEA model. A model was developed by modeling the Shahmir and Khalili (2015) NDEA model. The model developed in this research evaluates the efficiency of the production process in two stages, with a desirable output in the first stage and a desirable output and an undesirable output in the second stage. This model eliminates the inefficiency of classical data envelopment analysis models that have not been able to calculate the efficiency of the black box. Classical models used the initial input and final output to measure performance and did not evaluate the efficiency of the various production stages. However, this weakness was eliminated in the model developed in this research.

The data were analyzed with MATLAB software. The results showed that production unit 3 is efficient in the first stage, which is the production stage of pasteurized cream, but no other units are efficient. Therefore, it can be concluded that the most efficient desirable output and the most efficient transfer of semi-finished product from the first to the second stage is the production unit (factory) 3.

In the second stage, the butter production stage, units 1 and 3 are the most efficient ones among the twenty units studied, i.e., these two units have the most efficient activity since receiving the semi-finished product from the first stage, which is pasteurized cream and processing. They have it in the second stage and convert it into the final output, i.e., packed butter. Also, according to performance measurement as a general process, unit 3 is the only efficient one; therefore, it is introduced as a model and reference unit for other ones.

The efficiency score of all units in the first stage is better than the second stage, except for units 1 and 3, which received a better score in the second stage. Therefore, it can be concluded that the most significant weakness and inefficiency of units are usually in the second production stage. These units managers are suggested to prioritize the production complexs efficiency over the second stage and use efficient units as a model for their units efficiency.

The conclusion is that the butter industry in Alborz and Tehran provinces has a high inefficiency in the production process and the managers of this industry need to pay attention to increase efficiency.

The researchers suggest developing this model in an inaccurate environment assuming statistical distributions such as the normal distribution for inputs and outputs of decision-making units.

In this study, only experts opinions and judgments were used to collect initial data, and quantitative methods were not employed along with this qualitative method. It is suggested to use nondeterministic methods such as a fuzzy method to receive experts opinions and increase data accuracy.

References

[1] D. Li, M. Q. Wang, C. Lee, The waste treatment and recycling efficiency of industrial waste processing based on two-stage data envelopment analysis with undesirable inputs, *Journal of Cleaner Production* 242 (2020) 118-127.

[2] G. lgn, S. Seda, K. Murat, Y. Birol, Measuring the efficiency of Turkish maternal and child health hospitals: A two-stage data envelopment analysis, *Evaluation and Program Planning* 91 (2022) 102-123.

[3] H. Omrani, A. Emrouznejad, M. Shamsi, P. Fahimi, Evaluation of insurance companies considering uncertainty: A multi-objective network data envelopment analysis model with negative data and undesirable outputs, *Socio-economic planning sciences* 11 (2022) 101-130.

[4] H. Syarifa. Manufacturing enterprise performance using network DEA: a profitability and marketability framework, *International Journal of Business Excellence* 25 (2021) 277-299.

[5] O. Haruna Ezekiel, E. Samuel Elizabeth, T. Ibinayo Blessing, Economic and scale efficiency of broiler production in the federal capital territory, *Abuja Nigeria: a data envelopment analysis approach. RJOAS* 7 (2020) 103-121, http://dx.doi.org/10.18551/rjoas.2020-07.16/

[6] Z. Hashemi, M. Bonyadian, H. Moshtaghi, Evaluation of antibiotic resistance pattern in enterotoxin-producing Staphylococcus aureus bacteria isolated from creamy sweets in Shahrekord city, (1399).

[7] G. Jahanshahloo, F. Hosseinzadeh Lotfi, H. Nikomram, Data envelopment analysis and its applications, *Nafis publication*, (1389).

[8] M. D. Kremantzis, B. Patrick, K. Jonathan, A fairer assessment of DMUs in a generalised two-stage DEA structure, *Expert Systems with Applications* 187 (2022) 115-121.

[9] A. Kumar, G. Negi, S. Pant, M. Ram, S. C. Dimri,Availability-Cost Optimization of Butter Oil Processing System by Using Nature Inspired Optimization Algorithms, *Reliability: Theory & Applications SI* 2 (2021) 188-200.

[10] K. Khalili-Damghani, Z. Shahmir, Uncertain network data envelopment analysis with undesirable outputs to evaluate the efficiency of electricity power production and distribution processes, *Computers & Industrial Engineering* 88 (2015) 131-150.

[11] M. Momeni, H. Safari, M. Rostami, A. Mostafaei, R. Soleimani Damaneh, Design of a non-radial network data envelopment analysis model for performance evaluation, *Journal of Management Studies (Improvement and Transformation)* 26 (1396) 1-23.

[12] A. Mohaghar, H. Safari, A. R. Amirteimoori, M. Soufi,A New Data Envelopment Analysis Model for Efficiency Measurement of Decision Making Units with Parallel Network Structure, *Journal of Operational Research In Its Applications (Applied Mathematics)-Lahijan Azad University* 13 (2016) 9-26.

[13] M. Nikabadi, A. Oveysi Imran, K. Yakideh, Presenting a network data envelopment analysis model with a combination of desirable and undesirable intermediate and final outputs, *Journal of Operations Research in its Applications* 14 (1396) 116-95.

[14] Q. Wang, J. Tang, G. Choi, A two-stage eco-efficiency evaluation of Chinas industrial sectors: A dynamic network data envelopment analysis (DNDEA) approach, *Process Safety and Environmental Protection* 148 (2021) 879-892.

[15] Y. Li, H. Xue, L. Lin, F. Li, L. Liang, G. Kou, Provincial production and pollution treatment performance in China based on a two-stage eco-inefficiency approach with undesirable intermediate outputs, *Journal of Cleaner Production* 331 (2022) 13-26.

Atefeh Shirvani is a phd student of Industrial Engineering, majoring in System Engineering and Operations Research in Islamic Azad University, Science and Research Branch, Department of Industrial Engineering, Tehran, Iran.

AMIR AZIZI is a Assistant professor of Islamic Azad University, Science and Research Branch, Department of Industrial Engineering, Tehran, Iran and Universiti Malaysia Pahang: Kuantan, Pahang, MY.His research focuses on the, Planning and control in lean and flexible production, Total Quality Management, Project and risk management, Comprehensive productivity management, Statistical modeling, Fuzzy and Bayesian inference. He has published in various leading journals such as Procedia Manufacturing, South African Journal of Industrial Engineering and Advances in Decision Sciences.

Table 2: 2:Values of inputs and outputs in 2020

	input	input	input	input	output	input	input	input	output	output
Factory	Cost of raw materials	Maintenance costs1	Operator costs1	The amount of pasteurized cream output	The amount of produced breakfast cream	The amount of pasteurized cream inpu	Operator costs2	Maintenance costs2	The amount of produced buttermil	The amount of packed butter
1	245,718, 000,000	245,718, 000,000	245,718, 000,000	245,718, 000,000	245,718, 000,000	245,718, 000,000	245,718, 000,000	245,718, 000,000	245,718, 000,000	245,718, 000,000
2	163,812, 000,000	163,812, 000,000	163,812, 000,000	163,812, 000,000	163,812, 000,000	163,812, 000,000	163,812, 000,000	163,812, 000,000	163,812, 000,000	163,812, 000,000
3	245,718, 000,000	245,718, 000,000	245,718, 000,000	245,718, 000,000	245,718, 000,000	245,718, 000,000	245,718, 000,000	245,718, 000,000	245,718, 000,000	245,718, 000,000
4	204,832, 320,000	204,832, 320,000	204,832, 320,000	204,832, 320,000	204,832, 320,000	204,832, 320,000	204,832, 320,000	204,832, 320,000	204,832, 320,000	204,832, 320,000
5	204,832, 320,000	204,832, 320,000	204,832, 320,000	204,832, 320,000	204,832, 320,000	204,832, 320,000	204,832, 320,000	204,832, 320,000	204,832, 320,000	204,832, 320,000
6	122,859, 000,000	122,859, 000,000	122,859, 000,000	122,859, 000,000	122,859, 000,000	122,859, 000,000	122,859, 000,000	122,859, 000,000	122,859, 000,000	122,859, 000,000
7	245,718, 000,000	245,718, 000,000	245,718, 000,000	245,718, 000,000	245,718, 000,000	245,718, 000,000	245,718, 000,000	245,718, 000,000	245,718, 000,000	245,718, 000,000
8	147,430, 800,000	147,430, 800,000	147,430, 800,000	147,430, 800,000	147,430, 800,000	147,430, 800,000	147,430, 800,000	147,430, 800,000	147,430, 800,000	147,430, 800,000
9	81,906, 000,000	81,906, 000,000	81,906, 000,000	81,906, 000,000	81,906, 000,000	81,906, 000,000	81,906, 000,000	81,906, 000,000	81,906, 000,000	81,906, 000,000
10	65,524, 800,000	65,524, 800,000	65,524, 800,000	65,524, 800,000	65,524, 800,000	65,524, 800,000	65,524, 800,000	65,524, 800,000	65,524, 800,000	65,524, 800,000
11	189,873, 000,000	189,873, 000,000	189,873, 000,000	189,873, 000,000	189,873, 000,000	189,873, 000,000	189,873, 000,000	189,873, 000,000	189,873, 000,000	189,873, 000,000
12	126,582, 000,000	126,582, 000,000	126,582, 000,000	126,582, 000,000	126,582, 000,000	126,582, 000,000	126,582, 000,000	126,582, 000,000	126,582, 000,000	126,582, 000,000
13	189,873, 000,000	189,873, 000,000	189,873, 000,000	189,873, 000,000	189,873, 000,000	189,873, 000,000	189,873, 000,000	189,873, 000,000	189,873, 000,000	189,873, 000,000
14	158,227, 500,000	158,227, 500,000	158,227, 500,000	158,227, 500,000	158,227, 500,000	158,227, 500,000	158,227, 500,000	158,227, 500,000	158,227, 500,000	158,227, 500,000
15	158,227, 500,000	158,227, 500,000	158,227, 500,000	158,227, 500,000	158,227, 500,000	158,227, 500,000	158,227, 500,000	158,227, 500,000	158,227, 500,000	158,227, 500,000
16	94,936, 500,000	94,936, 500,000	94,936, 500,000	94,936, 500,000	94,936, 500,000	94,936, 500,000	94,936, 500,000	94,936, 500,000	94,936, 500,000	94,936, 500,000
17	189,873, 000,000	189,873, 000,000	189,873, 000,000	189,873, 000,000	189,873, 000,000	189,873, 000,000	189,873, 000,000	189,873, 000,000	189,873, 000,000	189,873, 000,000
18	113,923, 800,000	113,923, 800,000	113,923, 800,000	113,923, 800,000	113,923, 800,000	113,923, 800,000	113,923, 800,000	113,923, 800,000	113,923, 800,000	113,923, 800,000
19	63,291, 000,000	63,291, 000,000	63,291, 000,000	63,291, 000,000	63,291, 000,000	63,291, 000,000	63,291, 000,000	63,291, 000,000	63,291, 000,000	63,291, 000,000
20	50,632, 800,000	50,632, 800,000	50,632, 800,000	50,632, 800,000	50,632, 800,000	50,632, 800,000	50,632, 800,000	50,632, 800,000	50,632, 800,000	50,632, 800,000

Table 3: Overall efficiency

2***situation**	**efficiency E**	2***DMU**
inefficient	0.98	DMU1
inefficient	0.65	DMU2
efficient	1	DMU3
inefficient	0.53	DMU4
inefficient	0.83	DMU5
inefficient	0.63	DMU6
inefficient	0.9	DMU7
inefficient	0.65	DMU8
inefficient	0.57	DMU9
inefficient	0.32	DMU10
inefficient	0.83	DMU11
inefficient	0.31	DMU12
inefficient	0.81	DMU13
inefficient	0.71	DMU14
inefficient	0.75	DMU15
inefficient	0.57	DMU16
inefficient	0.75	DMU17
inefficient	0.61	DMU18
inefficient	0.42	DMU19
inefficient	0.11	DMU20

Table 4: The efficiency of the first stage

2***situation**	**efficiency E**	2***DMU**
inefficient	0.99	DMU1
inefficient	0.84	DMU2
efficient	1	DMU3
inefficient	0.78	DMU4
inefficient	0.92	DMU5
inefficient	0.87	DMU6
inefficient	0.96	DMU7
inefficient	0.86	DMU8
inefficient	0.79	DMU9
inefficient	0.53	DMU10
inefficient	0.97	DMU11
inefficient	0.63	DMU12
inefficient	0.92	DMU13
inefficient	0.9	DMU14
inefficient	0.98	DMU15
inefficient	0.78	DMU16
inefficient	0.91	DMU17
inefficient	0.85	DMU18
inefficient	0.65	DMU19
inefficient	0.41	DMU20

Table 5: The efficiency of the second stage

2***situation**	**efficiency E**	2***DMU**
inefficient	1	DMU1
inefficient	0.78	DMU2
efficient	1	DMU3
inefficient	0.69	DMU4
inefficient	0.91	DMU5
inefficient	0.74	DMU6
inefficient	0.95	DMU7
inefficient	0.77	DMU8
inefficient	0.73	DMU9
inefficient	0.62	DMU10
inefficient	0.86	DMU11
inefficient	0.49	DMU12
inefficient	0.9	DMU13
inefficient	0.79	DMU14
inefficient	0.77	DMU15
inefficient	0.75	DMU16
inefficient	0.83	DMU17
inefficient	0.72	DMU18
inefficient	0.65	DMU19
inefficient	0.28	DMU20