Journal of Agricultural Machinery
Vol. 12, No. 1, Spring 2022, p. 43-54

نشریه ماشین‌های کشاورزی
جلد ۱۲، شماره ۱، بهار ۱۴۰۱، ص ۵۴-۴۳

Full Research Paper

Evaluating the Efficiency of Sugarcane Harvesting Units Using a Combined Approach to Data Envelopment Analysis and Data Mining

N. Monjezi[1*]

Received: 11-07-2020
Accepted: 22-12-2020

Abstract

Every organization needs an evaluation system in order to be aware of the level of performance and desirability of its units. It is more important for agricultural companies, including agro-industries. In this study, 20 sugarcane harvesting units were selected. After modeling based on input-oriented CCR and BCC models, efficiency values for sugarcane harvesting units were calculated and the CART decision tree was used to extract rules to predict the efficiency of these units. The results of a study of 20 sugarcane harvesting units in the CCR model showed that 6 units had an efficient score and 14 units had an inefficient score, and their technical efficiency score was in the range of 0.73-0.95. The results of the BCC model study also showed that out of a total of 20 sugarcane harvesting units, 8 units had efficient scores. As can be seen, in the BCC model, more units are introduced as efficient units and there is less dispersion between inefficient units. Also, the distribution of efficient units in the BCC model is less than the CCR model. The average technical efficiency, pure technical efficiency, and scale efficiency were 93%, 88%, and 93%, respectively. Also, the accuracy of the decision tree model for technical efficiency and pure technical efficiency was 86% and 93%, respectively.

Keywords: CART, DEA, Harvest, Sugarcane

Introduction

Harvesting crops and transporting them from two dimensions has a great impact on sugarcane production and the income of an agro-industry. In terms of cost, which accounts for the majority of sugarcane production costs, and in terms of income, due to the amount of sugarcane waste at harvest time and the quality of the product sent to the factory and the amount of damage to the farm and sugarcane stump, it has a great impact on the amount of yield in the same year, ratooning[2] years and eventually agro-industry income. In Iran, sugarcane is industrially produced by Khuzestan Sugarcane and By-Products Development Company, Haft Tappeh, Karun and Mian Ab agro-industries (about 135,000 hectares). In these agro-industries, there are four harvesting groups for sugarcane harvesting operations. Each harvesting team consisted of six harvesters. There are four active and stationary in the group and two as reserve units. There are also 20 transporters and 24 drivers (mechanic operators) of the harvester and the transporter in each group, along with an expert in the group (one person), mechanic technician (2 persons), electricians (3 persons), and service workers (3 persons) are used.

Data envelopment analysis (DEA) model is a useful tool in measuring the efficiency of several organizational units with the same structure. In other words, the DEA model minimizes the ratio of inputs to outputs. This study is an attempt to determine the efficiency of harvesting units located in sugarcane production companies to identify inefficient or less efficient units. On the other hand, due to the increase in data volume and the need to analyze information and predict variables, data mining, especially the decision tree, can be helpful. The decision tree is one of the data mining methods. A decision tree is a diagram that shows a classification system or a

---

1- Assistant Professor, Biosystems engineering Dept., Faculty of Agriculture, Shahid Chamran University of Ahvaz, Ahvaz, Iran
(*- Corresponding Author Email: n.monjezi@scu.ac.ir)
DOI: 10.22067/jam.2020.58695.0
2- Sugarcane is a perennial plant and its operation does not end in a year. The sugarcane fields in the following years of operation are called ratoon.

predictive model and is a way to display a series of rules that lead to a category or value (Jenhani *et al.*, 2008). In recent years, a wide range of researchers in various fields have used DEA model, data mining model, and a combination of the two methods (Rahman *et al.*, 2019; Liu *et al.*, 2019; Li *et al.*, 2018). Chiang *et al.* (2017) evaluated the performance of the information and communication technology (ICT) industry in Taiwan using a combination of DEA and decision tree and examined 16 relevant companies. Toloo *et al.* (2009), in a paper by using a combination of data mining and DEA evaluated different rules for evaluating performance. Wanke and Barros (2016) investigated the role of heterogeneity in the insurance sector. This study focused on predicting the performance of Brazilian insurance companies by proposing an integrative, two-stage approach, which involved the determination of a DEA meta-frontier in the first stage and the use of several data mining techniques in the second. Wu (2009) presented a hybrid model using DEA, decision trees, and neural networks (NNs) to assess supplier performance. The model consists of two modules: Module 1 applies DEA and classifies suppliers into efficient and inefficient clusters based on the resulting efficiency scores. Module 2 utilizes firm performance-related data to train decision trees, NNs model, and apply the trained decision tree model to new suppliers. Raorane and Kulkarini (2012) discussed the role of data mining as an effective tool for yield estimation in the agricultural sector. As crop production depends on geographical, biological, political, and economic factors, data mining can solve the challenge of extracting knowledge from this raw data and estimate the amount of crops production. Ferraro *et al.* (2009) analyzed a large production database describing crop yield patterns. They studied the influence of several factors controlling sugarcane productivity in one of the most important areas of sugarcane production in Argentina. They proposed using a data mining technique called classification and regression tree (CART) to identify the dependence of sugarcane yield on the variation of both environmental and management factors. Ramesh and Vardhan (2013) predicted agricultural products yield using different data mining techniques such as K-Means, K-Nearest Neighbor, Support Vector Machines, and Artificial Neural Networks. They wanted to find a model with high accuracy and ability for prediction of the yield of agricultural products. Jeysenthil *et al.* (2014) designed and predicted a support system for a database of sugarcane soil using the data mining clustering technique (k-means). Everinghama *et al.* (2009) in Australia and Fernandes *et al.* (2011) in Brazil have estimated the yield of sugarcane farms using data mining techniques. Medar and Rajpurohit (2014) presented various crop yield prediction methods using data mining techniques. Different Data Mining techniques such as K-Means, K-Nearest Neighbor (KNN), Artificial Neural Networks (ANN), and Support Vector Machines (SVM) for very recent applications of data mining techniques in the agriculture field.

This study examines how to evaluate the efficiency of sugarcane harvesting units using a combination of DEA and decision trees models and as a case study, sugarcane harvesting units in Amirkabir sugarcane agro-industry in Khuzestan province, Iran are considered. Because DEA evaluates the efficiency of harvesting units and a decision tree to predict their effectiveness, this research enables managers to use the results to improve performance in their future decisions.

Materials and Methods

Study area

The data for this study were collected from Amirkabir agro-industry Company of Khuzestan province in Iran. This agro-industry has 480 farms with 25.5 hectare area for each farm. It covers an area of about 14000 ha. Figure 1 shows the position of Khuzestan province in Iran and a map of Amirkabir agro-industry farms. The data is obtained for the years from 2015 to 2020. The study area is located in Khuzestan Province which is a major agricultural region in Iran. The geographical location of the study area is between latitudes 31° 15′ to 31° 40′ North, and longitudes 48° 12′ to 48° 30′ East. The average elevation of the study area is 8 m above sea level. Mean annual rainfall within the study area is 147.1 mm, mean annual temperature is approximately 25 °C, and mean soil temperature at 50 cm depth is 21.2 °C.

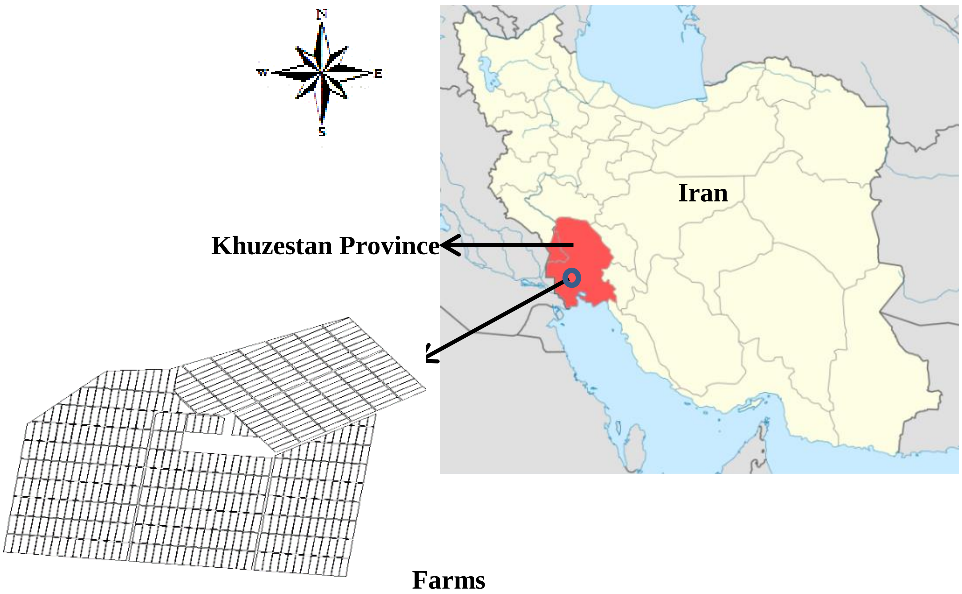

Fig.1. Amirkabir Agro-Industry position

Criteria for measuring the efficiency of sugarcane harvesting units

In this study, to measure the efficiency of sugarcane harvesting units and the best way to identify which unit has the best harvesting performance, nine indicators including the actual amount of hydraulic oil consumed during harvest after the overhaul of the harvester (litre), fuel consumption (litre), repair costs and the costs of consumable parts of harvesters (Rials), the amount of crop harvested (ton), harvest time (day), the amount of area harvested (hectare), factory no-cane hours, the amount of trash sent to the factory (percentage), and the amount of sugarcane waste on the farm (kilogram). To analyze the data, the DEA model has been used to measure the efficiency of sugarcane harvesting units, and the classification and regression trees (CART) model has been used in modeling and predicting the efficiency of these units. Software used also includes IBM SPSS MODELER 14.2 and DEA SOLVER. The research steps are shown in Figure 2.

Data envelopment analysis (DEA)

The DEA has four main models: Constant Return to Scale-CRS, Variable Return to Scale-VRS, Increase Return to Scale–IRS, and Decrease Return to Scale-DRS. Each of the above models has two directions of output-oriented and input-oriented (Liu *et al.*, 2019). In DEA, an inefficient DMU can be made efficient either by reducing the input levels while holding the outputs constant (input-oriented); or symmetrically, by increasing the output levels while holding the inputs constant (output-oriented). Data analysis will be performed with two models CCR[1] and BCC[2]. The CCR DEA model assumes constant returns to scale. It measures the technical efficiency by which the DMUs are evaluated for their performance relative to other DMUs in a sample. On the other hand, the BCC DEA model assumes variable returns to scale conditions. Choosing the right DEA model

1- Charns, Cooper and Rhodes
2- Banker, Charns and Cooper

depends on the level of control over the inputs and outputs; by making each one more controllable, the appropriate model is selected based on that. In this study, because the increase and decrease of inputs are more practical, CCR and BCC input-oriented models are used (Equations 1 and 2). In both models, efficient and inefficient units are identified and the types of technical efficiency, pure technical efficiency, and scale efficiency were calculated .

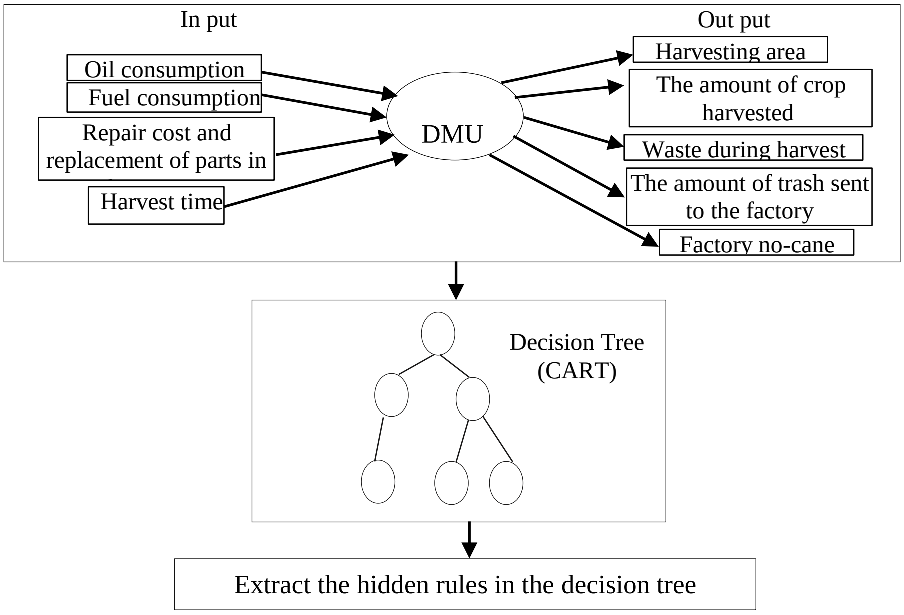

Fig.2. Research model

$$\max E_p = \sum_{r=1}^{r=s} U_r Y_{rp}$$
$$\sum_{i=1}^{i=m} V_i X_{ip} = 1 \qquad (1)$$
$$\sum_{r=1}^{r=s} U_r Y_{rj} - \sum_{i=1}^{i=m} V_i X_{ij} \le o,\ j = 1,2,...,n$$
$$V_i \ge \varepsilon, U_r \ge \varepsilon$$

$$\max E_p = \sum_{r=1}^{r=s} U_r Y_{rp} + w$$
$$\sum_{i=1}^{i=m} V_i X_{ip} = 1 \qquad (2)$$
$$\sum_{r=1}^{r=s} U_r Y_{rj} - \sum_{i=1}^{i=m} V_i X_{ij} + w \le o,\ j = 1,2,...,n$$
$$U_r \ge \varepsilon, V_i \ge \varepsilon, w\ free$$

In the above Equations: j= 1,2,…, n, n: the number of DMUs, s: the number of outputs, m: the number of inputs, X_{iP}: the amount of input ith for DMUp, Y_{rP}: the amount of output rth for DMUp, U_k and V_j, respectively, the weight of the outputs, The weight of the inputs and E_p the efficiency of the unit ith. (Banker *et al.*, 1984.)

The relationship between technical efficiency (TE), pure technical efficiency (PTE) (managerial efficiency), and scale efficiency (SE) is defined as Equation 3:

$$\frac{\text{TE}}{\text{PTE}} = \text{SE} \qquad (3)$$

The scale efficiency will not be more than one. The efficiency of the CCR model is called total technical efficiency because it is not affected by scale and size. On the other hand, the BCC model shows pure technical efficiency under variable returns to scale. The above relationship shows the efficiency analysis, which shows the efficiency sources. Which determines whether the inefficiency is due to managerial inefficiency or is due to conditions that indicate the scale efficiency or to both factors.

Data mining

A CART tree is a binary decision tree that is constructed by splitting a node into two child nodes repeatedly, beginning with the root node that contains the whole learning sample. Each node of a tree has two branches, which are related to the outcome of a test on one of the contextual variables. Data used in this study include 11 variables obtained from 20 sugarcane harvesting units during the years 2015-2020. The variables used were divided into two categories: predictive variables and target variables. The variables of pure technical efficiency and technical efficiency were considered as target variables (dependent variable) and other variables were considered as predictive variables (independent variable). In the CART model, the input data includes: harvesting area, the amount of crop harvested (ton), waste during harvest, fuel consumption, oil consumption, repair cost and replacement of parts in harvesters, harvest time, amount of trash sent to the factory (percentage), factory no-cane hours, technical efficiency and pure technical efficiency units of sugarcane harvesters. In this algorithm, 70% of the data were used for training and 30% of the data were used for testing. The Gini index was used to build the tree. Decision tree algorithm try to minimize diversity in the nodes. This non-uniformity in nodes can be measured using the impurity measure, the most important and most widely used method is the Gini index (Yoneyama *et al.*, 2002). Figure 3 depicts the data modeling process using IBM SPSS Modeler 14.2.

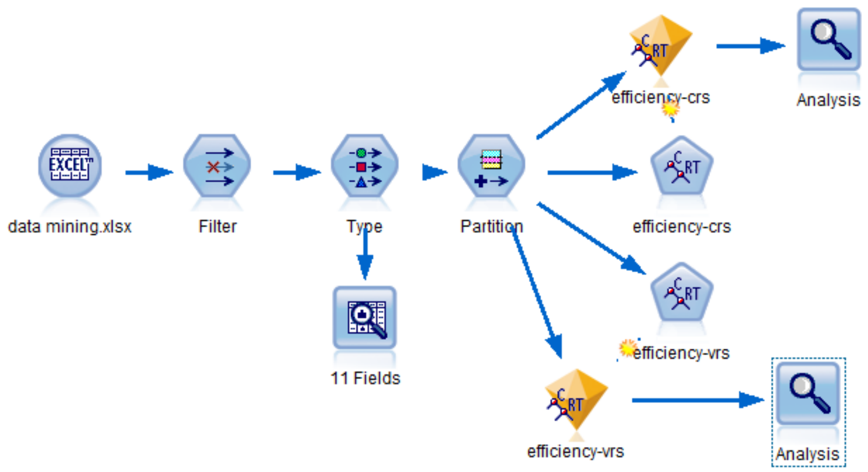

Fig.3. Data modeling process using IBM SPSS Modeler 14.2

Results and Discussion

Determining efficient and inefficient sugarcane harvesting units

The results of a study of 20 sugarcane harvesting units in the CCR model showed that 6 units had an efficient score (30%) and 14 units (70%) had an inefficient score, and their technical efficiency score was in the range of 0.73-0.95 (Table 1). The results of the BCC model study also showed that out of a total of 20 sugarcane harvesting units, 8 units (40%) had efficient scores (Table 1). On the other hand, the remaining harvest units, which were 12 units (60%), scored less than 100 points and were inefficient. In the BCC model, more units are introduced as efficient units and there is less dispersion between inefficient units. Also, the distribution of efficient units in the BCC model is less than the CCR model. The average values of pure technical efficiency, technical efficiency, and scale efficiency for all 20 units are 0.93, 0.88, and 0.93, respectively (Table 2). Ullah *et al.* (2019) analyzed the efficiency of different sugarcane production systems of Thailand. The result showed that the average efficiency score of sugarcane production systems is approximately 52%. The efficiency analysis indicates a huge potential for the improvement in efficiency through a reduction in the current pattern of farm inputs. The efficiency can also be improved by providing good management practices for sugarcane farms. Kaab *et al.* (2019) reported mean scores of technical efficiency, scale efficiency, and pure technical efficiency in sugarcane production in Iran as 0.91, 0.98, and 0.93, respectively. Khai and Yabe (2011) reported TE for paddy production in Vietnam as 0.816. Elhami *et al.* (2016) computed that TE, PTE, and SE for chickpea production in Isfahan province of Iran were 0.94, 0.99, and 0.94, respectively.

The average technical efficiency of inefficient units is 0.83, which shows that by using 83% of inputs and remaining the same output of them, these units can reach the efficiency limit and save 17% of inputs by increasing their efficiency (Table 1). The BCC analysis results in this table show that units 1, 5, 8, 11, 12, 13, 15, and 16 are efficient. The efficiency of the units means that each unit must be able to reduce its consumption of inputs by (1-θ)% without reducing the amount of production. The efficiency of unit 2 is 0.81%. This means that unit 2 should be able to reduce its consumption from all inputs by 19% to reach the efficiency limit. In Table 1, the ranking results of the harvest units are based on the BBC and CCR input-oriented models, which rank the inefficient units, and all efficient units are prioritized in the rankings on the inefficient units and assigned the first rank. Ranking of efficient units will be based on the number of inefficient units to which reference is made. The ranking of inefficient units is based on the value of efficiency points they have earned. According to the CCR model, after six efficient units, unit No. 14 is ranked first among inefficient sugarcane harvesting units and 7th among all harvest units, followed by harvesting units number 18, 6, 12, 20, 17, etc., respectively. In the BBC model, after eight efficient units, the inefficient unit No. 14 ranks first among the inefficient harvesting units and 9th among the total units, and then the harvesting units number 18, 6, 17, 20, 19, etc., respectively, have the next ranks. If a unit is quite efficient in terms of the BBC model but has a low efficiency of the CCR model, then it is relatively efficient but not overall efficiency. So basically it is reasonable to determine the efficiency of the scale of a unit by these two functions. The technical efficiency of a unit with a constant return to scale (CRS) is obtained from the CCR model. However, in the case of variable efficiency compared to the variable scale (VRS) of the BCC model, technical efficiency can be calculated. The relationship between technical efficiency, pure technical efficiency (management), and scale efficiency is defined in Equation 3. The scale efficiency will not be more than one. The efficiency of the CCR model is called total technical efficiency, because it is not affected by scale and size. On the other hand, BCC shows pure technical efficiency under variable returns to scale. Relationship 3 shows the efficiency analysis that shows the relationship between the sources of inefficiency, i.e. it determines whether the inefficiency is due to managerial inefficiency or is due to the conditions that indicate the scale efficiency or from both factors. According to the results obtained in Table 1, units 12 and 13 operate locally efficiently (pure technical efficiency = 1) and total inefficiency (1 <total efficiency) is due to scale inefficiency. The inefficiencies of units 2, 3, 4, 6, 7, 9, 10, 14, 17, 17, 18, 19, and 20 are due to management inefficiency and also due to the condition of the units (scale inefficiency).

The structure of the CART tree was created to predict the technical efficiency and pure technical efficiency of sugarcane harvesting units is shown in Figures 3 and 4. In the technical efficiency model, this tree consists of 4 nodes. Three of these nodes are the final nodes.

Table 1- Evaluation results of sugarcane harvesting units

DMU NO.	DMU Name	Technical efficiency score (CRS)	Ranking CCR	Pure technical efficiency score (VRS)	Ranking BCC	Scale efficiency score	RTS
1	2015-2016...1	1	1	1	1	1	Constant
2	2015-2016…2	0.81	15	0.86	17	0.94	Constant
3	2015-2016…3	0.73	18	0.88	16	0.82	Constant
4	2015-2016…4	0.89	13	0.91	15	0.97	Constant
5	2016-2017...1	1	1	1	1	1	Constant
6	2016-2017…2	0.93	9	0.96	11	0.96	Increasing
7	2016-2017…3	0.75	16	0.78	20	0.96	Constant
8	2016-2017…4	1	1	1	1	1	Constant
9	2017-2018...1	0.70	19	0.86	18	0.81	Constant
10	2017-2018…2	0.68	20	0.79	19	0.86	Constant
11	2017-2018…3	1	1	1	1	1	Constant
12	2017-2018…4	0.91	10	1	1	0.91	Constant
13	2018-2019...1	0.74	17	1	1	0.74	Constant
14	2018-2019…2	0.95	7	0.97	9	0.97	Decreasing
15	2018-2019…3	1	1	1	1	1	Constant
16	2018-2019…4	1	1	1	1	1	Constant
17	2019-2020...1	0.90	12	0.95	12	0.94	Decreasing
18	2019-2020…2	0.94	8	0.96	10	0.97	Decreasing
19	2019-2020…3	0.88	14	0.92	14	0.95	Decreasing
20	2019-2020…4	0.91	11	0.92	13	0.98	Decreasing

Source: Own calculation

Table 2- Efficiency average of harvesting units

Efficiency	Average	Standard deviation
Pure technical efficiency (efficiency in BCC model)	0.93	0.07
Technical efficiency (efficiency in CCR model)	0.88	0.11
Scale efficiency	0.93	0.07

CART algorithm

The information and description of the input data in the CART model are given in Table 3.

Table 3- Description of variables used for this study

Variable name	Unit	Average	Minimum amount	Maximum amount	Standard deviation
Harvesting area	ha	2311	2017	2679	166.71
The amount of crop harvested	ton	179321	140942	215033	28501.01
Waste during harvest	kg	11856.5	8341	15032	1755.52
Fuel consumption	lit	364150	301000	448000	39205.63
Oil consumption	lit	10965.90	8600	13984	2003.51
Repair cost and replacement of parts in harvesters	Rials	7668500000	5800000000	12530000000	1763173798.02
Harvest time	day	113.60	98	140	14.37
The amount of trash sent to the factory (percentage)	ton	6.24	5.51	6.80	0.42
Factory no-cane hours	hr	23.95	0	59	17.17

Source: Own calculation

The first variable selected to create a branch in the tree is the amount of fuel consumed with the Gini index, 0.002. The next division is using the cost of repairs and replacement of parts in harvesters, which is divided into two branches, costs more than 1,200,000,000 Rials and less than 1,200,000,000 Rials. Also, in the pure technical efficiency model, the tree consists of 2 nodes. The variable selected to create a branch in the tree is the amount of fuel consumed with the Gini index, 0.001. The tree is divided into two branches, the amount of fuel consumed is more than 348,000 liters and less than 348,000 liters.

Efficiency-CRS

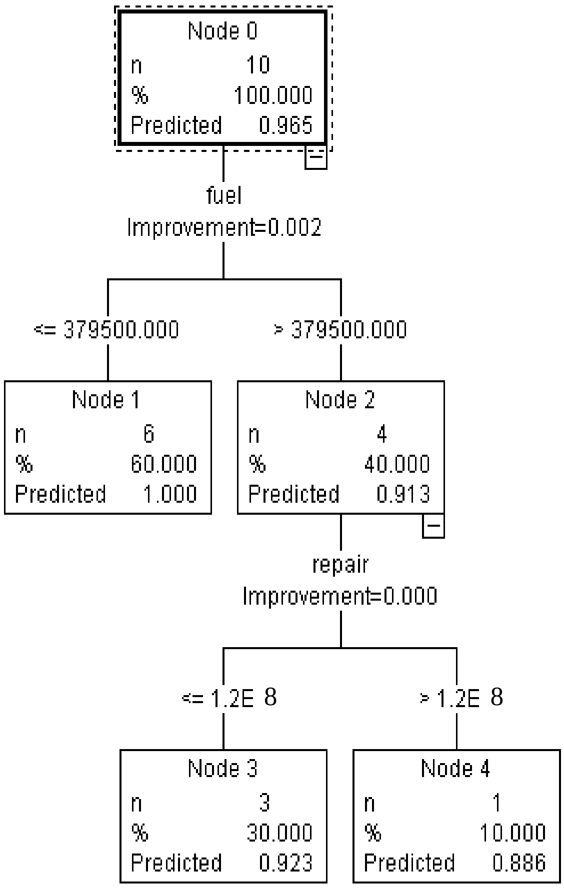

Fig.3. CART model predicting technical efficiency of sugarcane harvesting units

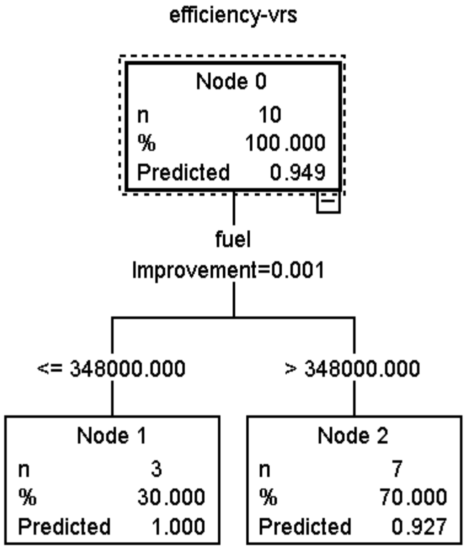

Fig.4. CART model predicting pure technical efficiency of sugarcane harvesting units

Modeling accuracy

Using the linear correlation relationship between the results predicted by the model and the actual results, it can be seen to what extent the resulting model has been successful in predicting the technical efficiency and pure technical efficiency of sugarcane harvesting units. According to Table 4, the accuracy of the CART model for predicting technical efficiency and pure technical efficiency was 86% and 93%, respectively.

Table 4- Results for CART algorithm

Variable	Technical efficiency	Pure technical efficiency
Minimum Error	-0.036	-0.031
Maximum Error	0.077	0.017
Mean Error	0.008	-0.003
Mean Absolute Error	0.013	0.005
Standard Deviation	0.027	0.01
Linear correlation	0.86	0.93

Conclusions

In this study, using the DEA method, technical efficiency, pure technical efficiency, and scale efficiency in sugarcane harvesting units in Khuzestan province, in Iran were investigated. The results of CCR and BCC input-oriented models showed that the average technical efficiency, pure technical efficiency, and scale efficiency in the sugarcane harvesting units were 93%, 88%, and 93%, respectively. Since both CCR and BCC models used for calculating the technical and pure technical efficiencies are unit invariant, the extended models are recommended for future study. In the next step, using the CART decision tree method, the efficiency of sugarcane harvesting units was predicted. In the CART model, the input data includes 11 variables. The results showed that the fuel consumption variable in the CART model for predicting technical efficiency and pure technical efficiency has emerged as the most important independent variable in modeling. This study results showed that the accuracy of the CART model for the variables of technical efficiency and pure technical efficiency were 86% and 93%, respectively. Therefore, it can be clearly seen that the decision tree model has a very good accuracy in estimating the values of technical efficiency and pure technical efficiency of sugarcane harvesting units. Finally, the results of this study show that the use of predictive methods, by providing an accurate picture of the situation of sugarcane harvesting in Amirkabir agro-industry, allows increasing the productivity of inputs.

References

1. Banker, R. D., A., Charnes, and W. W. Cooper. 1984. Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis. Management Science 30 (9): 1078-1092.
2. Charnes, A., W. W., Cooper, and E. Rhodes. 1978. Measuring the efficiency of decision making units. European Journal of Operational Research 2: 429-44.
3. Chiang, T. C., P. Y., Cheng, and F. Y. Leu. 2017. Prediction of technical efficiency and financial crisis of Taiwan's information and communication technology industry with decision tree and DEA. Soft Computing 21 (18): 5341-5353.
4. Elhami, B., A. Akram, and M. Khanali. 2016. Optimization of energy consumption and environmental impacts of chickpea production using data envelopment analysis (DEA) and multi objective genetic algorithm (MOGA) approaches. Information Processing in Agriculture 3: 190-205.
5. Everinghama, Y. L., C. W., Smyth and N, G. Inman-Bamber. 2009. Ensemble data mining approaches to forecast regional sugarcane crop production. Agricultural and Forest Meteorology 149: 689-696.
6. Ferraro, D. O., D. E. Rivero and C. M. Ghersa. 2009. An analysis of the factors that influence sugarcane yield in Northern Argentina using classification and regression trees. Field Crops Research 112: 149-157.
7. Jenhani, I., N. B. Amor, and Z. Elouedi. 2008. Decision Trees as Possibilistic Classifiers. International Journal of Approximate Reasoning 48 (3): 784-807.
8. Jeysenthil, K. M. S., T. Manikandan and E. Murali. 2014. Third Generation Agricultural Support System Development Using Data Mining. International Journal of Innovative Research in Science, Engineering and Technology 3 (3): 9923-9930.
9. Kaab, A., M. Sharifi, H. Mobli, A. Nabavi-Pelesaraei, and K. Chau. 2019. Use of optimization techniques for energy use efficiency and environmental life cycle assessment modification in sugarcane production. Energy 181: 1298-1320.
10. Khai, H., and M. Yabe. 2011. Technical efficiency analysis of rice production in Vietnam. Journal of the International Society for Southeast Asian Agricultural Sciences 17: 135-146.
11. Li, N., Y. Jiang, H. Mu, and Z. Yu. 2018. Efficiency evaluation and improvement potential for the Chinese agricultural sector at the provincial level based on data envelopment analysis (DEA), Energy 164: 1145-1160.
12. Liu, X., P. Guo, and S. Guo. 2019. Assessing the eco-efficiency of a circular economy system in China's coal mining areas: emergy and data envelopment analysis, Journal of Cleaner Production 206: 1101-1109.
13. Medar, R. A., and V. S., Rajpurohit. 2014. A survey on Data Mining Techniques for Crop Yield Prediction. International Journal of Advance Research in Computer Science and Management Studies 2 (9): 59-64.
14. Rahman, M. T., R., Nielsen, M. A., Khan, and M., Asmild. 2019. Efficiency and production environmental heterogeneity in aquaculture: A meta-frontier DEA approach. Aquaculture 509: 140-148.
15. Ramesh, D., and B. Vishnu Vardhan. 2013. Data Mining Techniques and Applications to Agricultural Yield Data. International Journal of Advanced Research in Computer and Communication Engineering 2 (9): 3477-3480.

16. Raorane, A. A., and R. V. Kulkarni. 2012. Data Mining: An effective tool for yield estimation in the agricultural sector. International Journal of Emerging Trends and Technology in Computer Science 1 (2): 75-79.
17. Toloo, M., B. Sohrabi, and S. Nalchigar. 2009. A New Method for Ranking Discovered Rules from Data Mining by DEA. Journal of Expert Systems with Applications 36 (4): 8503-8508.
18. Torgo, L. 2011. Data Mining with R: Learning with Case Studies. CRC Press, Boca Raton.
19. Ullah, A., T. Silalertruksa, P. Pongpat, and S. H. Gheewala. 2019. Efficiency analysis of sugarcane production systems in Thailand using data envelopment analysis. Journal of Cleaner Production 238: 1-11.
20. Wanke, P., and C. P., Barros. 2016. Efficiency drivers in Brazilian insurance: A two-stage DEA meta-frontier-data mining approach. Economic Modelling 53: 8-22.
21. Yoneyama, Y., S., Suzuki, R., Sawa, K., Yoneyama, G. G. Power, and Araki, T. 2002. Increased plasma adenosine concentrations and the severity of preeclampsia. Obstetrics & Gynecology 100 (6): 1266-1270.
22. Wu, D. 2009. Supplier selection: A hybrid model using DEA, decision tree and neural network. Expert Systems with Applications 36: 9105-9112.

Journal of Agricultural Machinery
Vol. 12, No. 1, Spring 2022, p. 43-54

نشریه ماشین‌های کشاورزی
جلد ۱۲، شماره ۱، بهار ۱۴۰۱، ص ۵۴-۴۳

مقاله پژوهشی

ارزیابی کارآیی واحدهای برداشت نیشکر با استفاده از یک رویکرد ترکیبی از تحلیل پوششی داده‌ها و داده‌کاوی

نسیم منجزی[1*]

تاریخ دریافت: ۱۳۹۹/۰۴/۲۱
تاریخ پذیرش: ۱۳۹۹/۱۰/۰۲

چکیده

هر سازمانی به‌منظور آگاهی از میزان عملکرد و مطلوبیت فعالیت واحدهای خود به یک نظام ارزشیابی جهت سنجش این مطلوبیت نیاز دارد و این موضوع برای موسسات کشاورزی از جمله کشت و صنعت‌ها اهمیت بیشتری دارد. در تحقیق حاضر، ۲۰ واحد برداشت نیشکر انتخاب گردید. پس از مدل‌سازی بر مبنای مدل‌های CCR و BCC ورودی محور، مقادیر کارایی برای واحدهای برداشت نیشکر محاسبه گردید و با استفاده از درخت تصمیم CART به استخراج قوانین برای پیش‌بینی کارایی این واحدها پرداخته شد. نتایج مطالعه ۲۰ واحد برداشت نیشکر در مدل CCR نشان داد که ۶ واحد دارای امتیاز کارآمد و ۱۴ واحد دارای امتیاز ناکارآمد بودند و امتیاز کارایی فنی آن‌ها در محدوده ۰/۷۳ تا ۰/۹۵ بود. نتایج مطالعه مدل BCC همچنین نشان داد که از مجموع ۲۰ واحد برداشت نیشکر، ۸ واحد دارای امتیاز کارآمد بودند. همان‌طور که مشاهده می‌شود، در مدل BCC، واحدهای بیشتری به‌عنوان واحدهای کارآمد معرفی می‌شوند و پراکندگی کمتری بین واحدهای ناکارآمد وجود دارد. همچنین، توزیع واحدهای کارآمد در مدل BCC کمتر از مدل CCR است. میانگین کارایی فنی، کارایی فنی خالص و کارایی مقیاس به‌ترتیب برابر ۹۳، ۸۸ و ۹۳ درصد به‌دست آمد. همچنین دقت مدل درخت تصمیم تولید شده برای کارایی فنی و کارایی فنی خالص نیز به‌ترتیب برابر ۸۶ و ۹۳ درصد به‌دست آمد.

واژه‌های کلیدی: CART، برداشت، تحلیل پوششی داده‌ها، نیشکر

۱- استادیار گروه مهندسی بیوسیستم، دانشکده کشاورزی، دانشگاه شهید چمران اهواز، اهواز، ایران
(*- نویسنده مسئول: Email: n.monjezi@scu.ac.ir)

DOI: 10.22067/jam.2020.58695.0