EFFICIENCY ANALYSIS OF DAIRY FARMS IN THE NORTHERN GREAT PLAIN REGION USING DETERMINISTIC AND STOCHASTIC DEA MODELS

Tímea Gál

University of Debrecen CAAES Coordination Centre for Logistics
138. Böszörményi str. Debrecen 4033 E-mail: galtimea@agr.unideb.hu

Abstract: Running any dairy enterprise is a risky activity: the profitability of the enterprise is affected by the price fluctuation of feed and animal health products from inputs, as well as by the fluctuation of end-product prices. Under these circumstances, it is essential for the cattle breeders, in order to survive, to harness the reserves in management as effectively as possible.

In this research the efficiency and risk of 32 sample dairy farms were analysed in the Northern Great Plain Region from the Farm Accountancy Data Network (FADN) by applying classical Data Envelopment Analysis (DEA) and stochastic DEA models. The choice of this method is justified by the fact that there was not such an available reliable database by which production functions could have been defined, and DEA makes possible to manage simultaneously some inputs and outputs, i.e. complex decision problems. By using DEA, the sources that cause shortfall on inefficient farms can be identified, analysed and quantified, so corporate decision support can be reinforced successfully.

A disadvantage of the classical DEA model is that the stochastic factors of farming cannot be treated either on the side of inputs or outputs; therefore, their results can be adopted with reservations, especially in agricultural models. This may have been because we could not discover that many agricultural applications. Considering the price of inputs and outputs as probability variables, 5000 simulation runs have been done in this research. As a result, it can be stated that at which intervals of the input and output factors can become competitive and the fluctuation of these factors can cause what level of risk at each farm.

Keywords: deterministic DEA, stochastic DEA, efficiency, dairy enterprise, risk.

INTRODUCTION

Material flow processes of animal husbandry constitute a compound system which involves procurement, stocking and marketing tasks. Material flow processes are closely linked with resource management, operation and demand management. An optimal production and logistics strategy can only be determined in the light of these circumstances. These processes are much more complex in animal husbandry than in a conventional enterprise, since biological laws must also be considered in the timing of processes.

The share of agricultural enterprises from GDP is about 4%. Within this share, the bovine sector gives one-fourth of the GDP in animal husbandry, thus it is of the third largest volume animal enterprise. Based on the database of the Central Statistical Office (CSO), domestic bovine herds have decreased by 200 thousand head, cows by 100 thousand in just the last 10 years. These numbers have considerable influence on keepers not being able to recognize the costs of keeping in their selling prices. The number of bovine farms has also decreased to one-third within the last seven years, according to the CSO's Farm Structure Survey of 2007. Since such a drastic decrease was not followed by the bovine herd's decrease, this suggests that a concentration happened among the farms, i.e. there are fewer farms, but bigger herds are being kept. On the basis of the statements of the Dairy Board and Interbranch Organization, in 2010, Hungary could fulfil only 80% of its milk quota l, which represents a continuous 10% decrease in the last 3–4 years. This statistic also confirms that the bovine sector is in a long-term crisis; its profitability has been falling. Farms must try to make their farming as efficient as possible with every available tool, so as to avoid the disposal of their herds and closure. The system approach application of logistics can be such an instrument in the processes of animal husbandry. However, for the improvement of efficiency, the exact level – and the input and output parameters – must be known that are to be changed to reach a more expeditious farming. A tool for efficiency analysis is the application of DEA models.

LITERATURE REVIEW

Deterministic DEA models

The idea of *Data Envelopment Analysis* (hereinafter DEA) method was originated by FARREL (1951), who wanted to develop a method that is more suitable for measuring productivity. However, in 1978, CHARNES et al. reformed this as a mathematical programming problem. This technique is a relatively new *"data-oriented" process*, which can be applied for measuring the performances of decision making units (DMU's) producing from several inputs several outputs (COOPER et al. 2004a). In recent years, the method of DEA has been used in many applications for performance measurement. It has been used for measuring the efficiency of a service's internal quality (SOTERIOU and STAVRINIDES, 2000; BECSER, 2008), efficiency measurement of banks (SHERMAN and LADINO, 1995; TÓTH, 1999), of educational (TIBENSZKYNÉ, 2007) and other public bodies, and also for measuring the efficiency of business parks (FÜLÖP and TEMESI, 2000). However, its application in agricultural practice was not significant. The efficiency analysis of animal farms and agricultural production processes can be carried out by *simulation methods* (SZŐKE et al. 2009; KOVÁCS and NAGY, 2009); however, the quality of available database does not always allow the full mapping of technological processes. In these cases, DEA is a more efficient tool.

The DEA process has two known approaches: *input-oriented* (cost-oriented) and *output-oriented* (result-oriented). In the case of the input-oriented approach, we examine how much and at which proportion the inputs should be used to minimize a cost at the same emission level. In the output-oriented approach, we determine the partial increase of outputs without changing the quantity of inputs (FARREL, 1957; CHARNES, et al. 1978).

This is complicated by the fact that we must take into consideration in our efficiency measurement that not every input benefits an enterprise in the same way: if we calculate with the intake on the same level *Constant Return to Scale* (CRS) is counted, if not, then *Variable Return to Scale* (VRS) (COOPER et al. 2004a). KOVÁCS and EMVALOMATIS (2011) applied a VRS output-oriented model to analyse the efficiency of dairy enterprises in Germany, Hungary and The Netherlands.

DEA is a non-parametric multiple statistic method by which we can determine a unit's efficiency of transforming inputs into outputs; therefore, it is suitable to determine the unit (e.g., farm, university or restaurant) with the "best practice" (ALBRIGHT and WINSTON, 2007). Thus, the DEA process gives the marginal efficiency and, knowing this marginal efficiency curve, the parameters of non-efficient units can be detected. By improving these parameters, optimality can be reached (TOFALLIS, 2001; BUNKÓCZI and PITLIK, 1999).

Stochastic DEA models

By applying DEA models, units of 100% efficiency can be chosen; however, the results are valid for past data: decisions are made for the future. Although bottlenecks – i.e. the factors to be changed so that a decision-making unit (DMU) will be efficient – can be identified by the basic deterministic model, it is not sensitive enough. There are many input and output factors which can be defined as probability variables, so these will be built into the model. Probability variables can be described by different functions: distribution function, density, characteristic and generator function.

Researchers have begun the practical application of stochastic DEA models from the beginning of the 1990s. The comparison of stochastic and deterministic DEA models was published by several researchers (COOPER et al., 2004b; SEIFORD and ZHU, 1998; TSIONAS, 2003; BRUNI et al., 2009). Stochastic DEA models were applied on many fields: to measure the efficiency of libraries (LOTFI et al., 2007), textile factories (KHODABAKHSHI and ASGHARIAN, 2009) and oil companies (SUEYOSHI, 2000).

BARÁTH et al. (2007) applied stochastic DEA model to analyse the total factor productivity change in Hungarian agriculture.

MATERIAL AND METHODS

The database of regional analyses was given by the Farm Accountancy Data Network (FADN) of the Research Institute of Agricultural Economics (Hungarian abbreviation: AKI). FADN is a representative information system in the European Union that measures the financial position and the assets and liabilities of farms. For this research, the data of 32 sample dairy farms was used in the Northern Great Plain Region. For the calculations, the examined year was 2010.

From the examined 32 sample farms, there are 22 individual and 10 corporate farms. Dairy herds of the individual farms number 1,187 cows; the corporate farms' total herds number 3,716 head, the average for one farm is 371 dairy cattle. The herd of the examined 32 farms numbers 4,903 cows, which is 6.22% of the population in the Northern Great Plain Region. The produced milk yield was more than 35.5 thousand tons in 2010.

Description of deterministic DEA model

In the course of the operation of an enterprise, the question of how efficient its units are working often arises. Investment analysts are interested in the efficiency of competing participants within an industrial enterprise. DEA is a linear programming application by which the above-mentioned problems can be solved. In the course of DEA analysis, we get the result of at what efficient level the inputs are transformed into outputs, so it is suitable to find the unit (e.g., a plant, university or restaurant) which has the "best-practice" (ALBRIGHT and WINSTON, 2007). I apply the method of DEA to determine the frontier efficiency by the efficiently operating units (TOFALLIS, 2001; BUNKÓCZI and PITLIK, 1999).

Efficiency can be measured by output/input indices, thus:

$$E_i = \frac{\sum_{j=1}^{n_o} O_{ij} w_j}{\sum_{j=1}^{n_i} I_{ij} v_j} \quad \text{where} \tag{1.}$$

E_i : the efficiency of the unit *i*

O_{ij} : the value of the unit *i*'s output factor *j*

n_o : number of outputs

w_j : the evaluation of one unit of output *j*

I_{ij} : the value of the unit *i*'s input factor *j*

n_j : number of inputs

v_j : the evaluation of one unit of input *j*

Objective function of the model:

$$\sum_{j=1}^{n_o} O_{ij} w_j \Rightarrow MAX! \tag{2.}$$

For every examined unit, we solve a separate LP exercise, by which the economic content of the objective function is the same; namely, my aim is to maximize the value of the units' weighted outputs. After having solved all LP models, we get the best evaluation (input and output weights) as a result (RAGSDALE, 2007).

Constraints:

1. The efficiency of any unit cannot be higher than 100%.

$$\sum_{j=1}^{n_o} O_{kj} w_j \leq \sum_{j=1}^{n_i} I_{kj} v_j \tag{3.}$$

(k=1,2,..., the number of units to be taken under the analysis) that is

$$\sum_{j=1}^{n_o} O_{kj} w_j - \sum_{j=1}^{n_i} I_{kj} v_j \leq 0 \tag{4.}$$

2. For the sake of the calculations, input prices should be scaled in a way that the input cost of economic unit *i* shall be 1 (RAGSDALE, 2007).

$$\sum_{j=1}^{n_I} I_{ij} v_j = 1 \tag{5.}$$

After choosing the non-efficient farms, we can quantify which parameters should be changed on the farms – one by one – to reach the optimal values of the "composite farm" that is 100% efficient. For this, shadow prices can be applied. On those farms that were 100% efficient, the difference of weighted output and input is zero, so it stands on the threshold i.e. it has a shadow price. The given farm's optimal value can be calculated as the scalar product of the vectors for shadow prices and each parameter value.

Description of stochastic DEA model

Basically, the stochastic DEA model is a stochastic linear programming model series. Stochastic LP can be applied if the probability of different events is known or statistically is to be defined.

Obviously, in this DEA model the input and output parameters must be taken into consideration as probability variables. In this research these variables were treated as beta (Milk production for 305 days, milk fat and turnover) or normal distribution (milk protein, on-farm and bought-in feed costs, labour cost, direct costs) based on my previous analyses about the AKI database. The density function of beta-distribution:

$$f(x) = \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1} = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, \quad x \in [0,1] \tag{6.}$$

and f(x)=0 otherwise. In this formula, Γ(x) is the gamma-function, B(α, β) is the beta-function and α and β are positive. Specially, if α = 1 and β = 1, X follows a uniform distribution in the interval [0,1]. The graph of beta-distribution density function can have various shapes. In this case, the values of α and β for the chosen probability variables are listed in Table 1. These values were set according to the practice presented by dairy farmers.

Table 1. Parameters of probability variables with beta distribution

	The likeliest	Minimum	Maximum	Alpha	Beta
Milk production for 305 days (liter)	farm value	2000	10000	1,75	1,9
Milk fat (%)	farm value	2,80%	4,34%	10	6
Turnover (without subsidy) (thousand HUF/cow)	farm value	200	1250	3	5

Source: own calculation

The probability value for beta distribution was determined by random number generator and based on this the inverse of beta distribution function was calculated, which is exactly the value of the probability variable by given α and β parameters that can vary within my own estimated limits (minimum and maximum). These calculated beta distribution variables will be put into this DEA model.

Milk protein, the own and purchased feed costs, labour costs and direct costs were treated as normal distribution probability variables. The density function of normal distribution is:

$$f(x) = \frac{1}{\sqrt{2\Pi \cdot \sigma}} \cdot e^{\frac{-(x-\mu)^2}{2\sigma^2}}. \tag{7.}$$

The parameters of probability variables with normal distribution are shown in Table 2. In the course of the calculations, the INVERZ.NORM function was used in Excel, which gives the value of the normal distribution function's inverse by given expected value and standard deviation. Values can be calculated according to the

Table 2. Parameters of probability variables with normal distribution

	Average	Coefficient of variance
Milk protein (%)	farm value	4,10%
On-farm grain feed cost (thousand HUF/cow)	farm value	50%
On-farm fodder cost (thousand HUF/cow)	farm value	50%
Bought-in grain feed cost (thousand HUF/cow	farm value	50%
Bought-in fodder cost (thousand HUF/cow)	farm value	50%
Labor cost (thousand HUF/cow)	farm value	80%
Direct cost (thousand HUF/cow)	farm value	40%

Source: own calculation

probability for distribution, the distribution's mean (in this case average) and coefficient of variation. The probability value for distribution was determined by random number generator. These calculated variables will be put into this model.

RESULTS

Efficiency analysis of dairy farms in Northern Great Plain region by deterministic DEA model

I analysed the efficiency of 32 dairy sample farms in the Northern Great Plain Region in the deterministic version of DEA. I classified these farms according to the size categories of the Hungarian Central Statistical Office (Table 3). Sixteen per cent of the examined farms have 3–9 cows, which is only 1% of the total herd. Sample farms with 20–99 cows are 39%, which gives only 11% of the population. Only 6 farms have more than 300 cows, but 61% of the total herd belongs to them.

Table 3. Classification of the examined sample tests based on herd size

Central Statistical Office size category (cow)	Size category code	Number of farms	Distribution of farms by categories (%)	Herd by categories (cow)	Distribution of herd by categories (%)
1-2	1	0	0%	0	0%
3-9	2	5	16%	27	1%
10-19	3	3	9%	41	1%
20-29	4	4	13%	105	2%
30-49	5	4	13%	160	3%
50-99	6	4	13%	308	6%
100-199	7	3	9%	493	10%
200-299	8	3	9%	781	16%
300-499	9	4	13%	1598	33%
500-	10	2	6%	1389	28%
Total		**32**	**100%**	**4902**	**100%**

Source: own calculation

Herd size has medium concentration (Figure 1). Based on the calculations, the average difference is 490 cows, which was determined by Gini's formula:

$$G = \frac{2 \cdot \sum_{i=1}^{n} (X_i - X_j)}{n^2} = 489{,}94 \, . \tag{8.}$$

The degree of concentration was calculated by the concentration coefficient which is the quotient of Gini's formula and twice the mean:

$$Ke = \frac{G}{2\overline{\overline{X}}} = 0{,}504 \, . \tag{9.}$$

The value of concentration coefficient is 0.504, which means that the concentration is medium. The coefficient value can take a number between 0 and 1, so the nearer to 1, the stronger the concentration is. Concentration examination was made based on the calculations of LORENZ (1905).

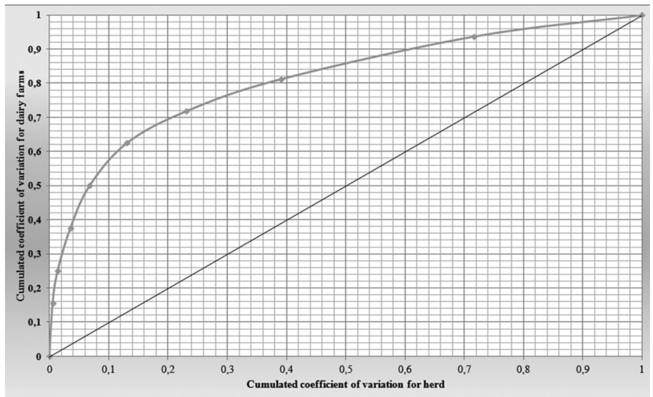

Figure 1. Concentration of herd on the examined dairy farms by Lorenz-curve
Source: own calculation

Classification of farms was executed by cluster analysis. These farms were analysed by two types of cluster analysis: hierarchical and non-hierarchical k-means analysis. The result was the same by both method, 3 groups were established by 11 characteristics: 23 farms are in the first cluster, 4 farms in the second and 5 farms are in the third one. Farms in Cluster 1 constitutes the group of so-called smaller or medium-sized farms, their average herd is 52 cows. Legally 4% of them are corporate farms, the rests are individual farms. The average herd in Cluster 2 is 276 cows, which means that these constitute the group of large sized farms. In Cluster 3, there are 5 corporate farms with an average herd of 521 cows; actually, these are the classical industrial large farms.

In the efficiency model arable land (ha/cow), herd size, on-farm and bought-in grain feed and fodder costs (thousand HUF/cow), labour cost (thousand HUF/cow) and direct costs were taken into account as input factors. Milk production for 305 days, average milk fat and protein from milk quality parameters and turnover with subsidies and without subsidies were set into the model as outputs. The aim of the analysis is to examine the farms' efficiency, to explore the critical factors in cases of non-efficient farms and to determine the direction of further analyses. After solving the model it can be stated that considering the given input and output constraints 20 farms (63%) from 32 operate in an efficient way, the others (12 farms, 37%) does not (i.e. DEA efficiency value is less than 1).

Among the efficient farms, there are 3 corporate and 17 individual farms, while among the non-efficient ones, there are 5 individual and 7 corporate farms. Therefore, according to the examinations, 30% of the corporate farms and 77% of the individual farms work in an efficient way. Consequently, it can be stated that medium and large sized individual dairy farms work more efficiently in the Northern Great Plain Region than the industrial large corporate farms. The classification of efficient farms by legal status, cluster and size category is shown in Table 4. The herd size of corporate efficient farms is 386 cows; cows on the individual farms number 1,239 altogether, which gives 25% of the examined farms' total herd. 95% of the efficient farms belong to Cluster 1, while only 5% are in Cluster 2. Neither of the farms in Cluster 3 was efficient.

Table 4. Efficient farms by legal status, cluster and size category

Farm code	Legal status 1= individual 2= corporate	Number of cluster	CSO size-category code[1]
1	2	1	2
2	1	1	7
3	1	1	8
5	2	2	9
6	1	1	7
8	1	1	4
9	1	1	6
11	1	1	2
12	1	1	5
13	1	1	2
15	1	1	4
17	2	3	9
18	1	1	3
19	1	1	5
20	1	1	2
25	1	1	3
26	1	1	5
28	1	1	2
31	1	1	3
32	1	1	4

([1] Farm size categories in the CSO databases: 1=1-2 cows, 2=3-9, 3=10-19, 4=20-29, 5=30-49, 6=50-99, 7=100-199, 8=200-299, 9=300-499, 10= more than 500 cows)
Source: own calculation

If the size category for efficient farms is analysed, it can be stated that 8 farms have less than 20 cows, 7 farms have 20-100 cows, 3 farms have 100 and 300 cows and only 2 farms keep more than 300 cows.

The model analysis shows that those farms are efficient, for which direct costs (409 thousand HUF/cow on efficient and 620 thousand HUF/cow on non-efficient farms) are much more lower compared to their turnover (541 thousand HUF/cow without subsidy, 667 thousand HUF/cow with subsidy) and produced less milk per cow (5365 kg/cow) but with better quality parameters. In cases of the efficient farms the feed cost was lower (151 thousand HUF/cow) than on the non-efficient ones (193 thousand HUF/cow).

42 per cent of the non-efficient farms are individual farms (Table 5). It can be stated that in Cluster 1 five farms are efficient from 23 (22%). There is only 1-1 farm in Cluster 2 and also in Cluster 3, which is efficient. On the non-efficient farms, 3,167 cows are kept, which is 65% of the examined total herd.

DEA efficiency of non-efficient farms is shown in Figure 2. The average efficiency is marked with blue colour (73.53%). Based on this, we can see that 58% of these 12 non-efficient farms have efficiency above average. Sorting the efficiency values in descending order, Farms 17 and 21

Table 5. Non-efficient farms by legal status, cluster, size category and DEA efficiency

Farm code	Legal status 1= individual 2= corporate	Number of cluster	CSO size-category code[1]	DEA efficiency
22	1	1	6	92%
21	2	2	8	88%
14	1	1	4	84%
7	1	1	5	81%
24	2	3	9	76%
4	1	1	6	76%
30	2	2	7	75%
10	1	1	6	71%
16	2	3	10	67%
23	2	3	10	62%
27	2	3	9	60%
29	2	3	8	51%
Average				73,53%

([1] Farm size categories in the CSO databases: 1=1-2 cows, 2=3-9, 3=10-19, 4=20-29, 5=30-49, 6=50-99, 7=100-199, 8=200-299, 9=300-499, 10= more than 500 cows)
Source: own calculation

reached almost 90% efficiency. Although Farms 23 and 16 have the largest herds (638 and 751 cows), their efficiency is among the lowest (638 and 751 cows).

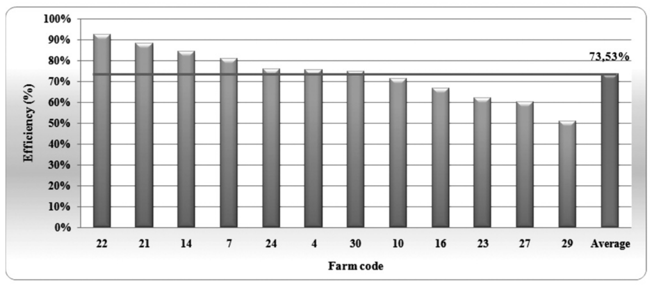

Figure 2. Efficiency of the non-efficient farms
Source: own calculation

In order to make a sounder analysis about the parameters of non-efficient farms, sensitivity analyses and the examination of shadow prices are needed to be made. Using the efficient farms' shadow prices an input and output vector can be created that concern to a complex, hypothetical farm. The input and output parameters of this composite farm can be compared to the present values of the non-efficient farms, thus the defects that decrease efficiency can be explored. In this research, the non-efficient farms' parameters were compared to the hypothetical farms' factors created by the shadow prices, so I could determine which value should be improved to reach a good practice. As an example, the calculation of the composite farm for Farm 4 will be presented. The shadow prices listed in Table 6 were calculated as a solution after solving the LP for Farm 4 in the sensitivity analysis. In the first row of the table (Input Farm 4), DEA efficiency can be found in the column of Shadow price as a dual solution. The weighted input was set with equality in the model, so here the shadow price means that 1% of input change results 0,756% output change. In the other rows, the differences of inputs and outputs are evaluated. The left and right hand side of the constraint is calculated in every case and, if the left hand side value is equal with the right, the shadow price is to be found. This match means that where the difference is zero, that is an efficient farm and the shadow price shows the weight of considering the given farm in the further efficiency analysis.

Table 6. Shadow prices after solving the LP model of Farm 4

Name	Left hand side of the constraint	Shadow price	Right hand side of the constraint
Input Farm 4	1	0,755635466	1
Farm 1 Difference	-0,789828871	0	0
Farm 2 Difference	-3,9968E-15	0,016921441	0
Farm 3 Difference	-0,626195297	0	0
Farm 4 Difference	-0,244364534	0	0
Farm 5 Difference	-0,409044056	0	0
Farm 6 Difference	-0,122871971	0	0
Farm 7 Difference	-0,499653851	0	0
Farm 8 Difference	-0,154663417	0	0
Farm 9 Difference	-5,88418E-15	0,688754471	0
Farm 10 Difference	-0,654209858	0	0
Farm 11 Difference	-0,227827927	0	0
…	…	…	…
Farm 31 Difference	-1,55431E-15	0,327050034	0
Farm 32 Difference	-0,609197277	0	0

Source: own calculation

If the shadow prices are weighted with the farms' parameters, a composite farm for Farm 4 will be created which were compared to the original values of the given farm (Table 7). In the Difference column, those values are listed by which the parameter value is to be modified for the farm to be efficient. This calculation series was made in every case.

Table 7. Parameters of the composite farm created by shadow prices and Farm 4 and the difference of the values

Parameters \ Farm			Parameters of the composit farm	Farm 4	Difference
Outputs	**Milk for 305 days**	kg/cow	4 451,23	4 451,23	**0,00**
	Milk protein	%	3,10%	3,10%	- **0,00**
	Milk fat	%	3,75%	3,38%	**0,37%**
	Turnover without subsidy	thousand HUF/cow	397,76	372,27	**25,49**
	Turnover with subsidy	thousand HUF/cow	434,12	404,17	**29,95**
Inputs	**Arable land**	ha/cow	0,66	1,18	- **0,52**
	Dairy herd	cows	65	87	- **22**
	On-farm grain feed cost	thousand HUF/cow	45,80	60,61	- **14,81**
	On-farm fodder cost	thousand HUF/cow	42,02	68,63	- **26,61**
	Bought-in grain feed cost	thousand HUF/cow	19,78	47,34	- **27,56**
	Bought-in fodder cost	thousand HUF/cow	0,10	2,28	- **2,18**
	Labor cost	thousand HUF/cow	29,49	50,45	- **20,96**
	Direct costs	thousand HUF/cow	434,85	640,99	- **206,14**

Source: own calculation

It is understood that on the non-efficient farms, the specific produced milk is not to be changed either. On the farms with good practice, the value of milk protein was higher on average by 0.22%; however, on five farms, this value does not need to be modified. At the largest degree, on Farm 23, the rate of milk protein should be increased by 0.85%. Considering the milk fat, we can see that an averaged 0.55% should be improved on these 12 farms. The rate of milk fat is only adequate on Farm 30. The highest improvement is needed on Farm 23 again (+1.23%). This led to the conclusion that quality parameters of milk are efficiency increasing factors.

Turnover should be increased by 110 thousand HUF/cow on average on the non-efficient farms. The smallest modification is need on Farm 29 (+17 thousand HUF/cow).

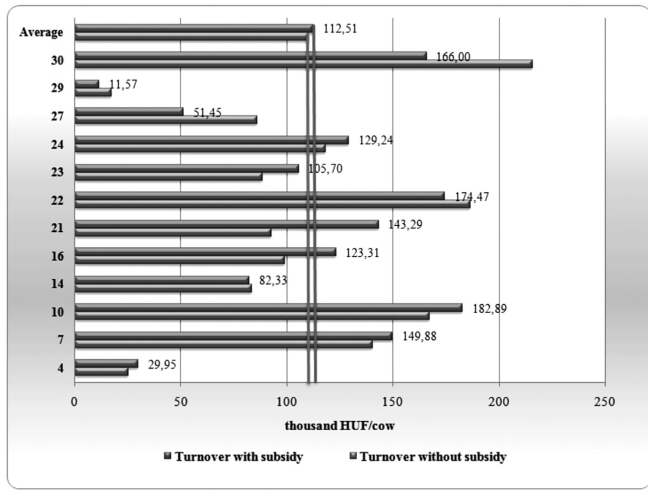

Figure 3. Turnover differences of the non-efficient farms compared to the composite farms
Source: own calculation

Considering the inputs, the size of arable land should be enlarged almost on every farm by an average 2.35 hectares/cow. The largest enlargement should be done on Farm 23 (6.07 ha/cow), the smallest on Farm 7 (0.12 ha/cow).

The size of herd should be decreased by 128 cows, on average. Among the farms of bad practice, on Farm 16, the herd should be reduced by 531 cows, but this farm had the largest livestock (751 cows). The herd of Farm 23 is almost to be halved: the livestock of 638 cows should be reduced by 387 cows.

Considering feed costs, it is to be concluded that both the cost of on-farm and bought-in feeds should be cut by 105 thousand HUF on average in the cases of almost all the farms (Figure 4). Grain feed cost should be reduced by 20 thousand HUF/cow on average, the values fluctuate from 1.3 (Farm 22) to 57 thousand HUF/cow (Farm 7). On-farm fodder cost is to be diminished by 53 thousand HUF on average on all farms. The slightest decrease of this cost is needed on Farm 22 (1.98 thousand HUF/cow), the largest is on Farm 29 (205.6 thousand HUF/cow). Bought-in grain feed cost is not to be modified on only 3 farms (Farm 14, 21, 30), but it should be decreased by 10 thousand HUF/cow on the others. The cost of bought-in fodder is to be moderated by 22 thousand HUF/cow on average. Among feed costs, this cost has the most extreme fluctuation (131%).

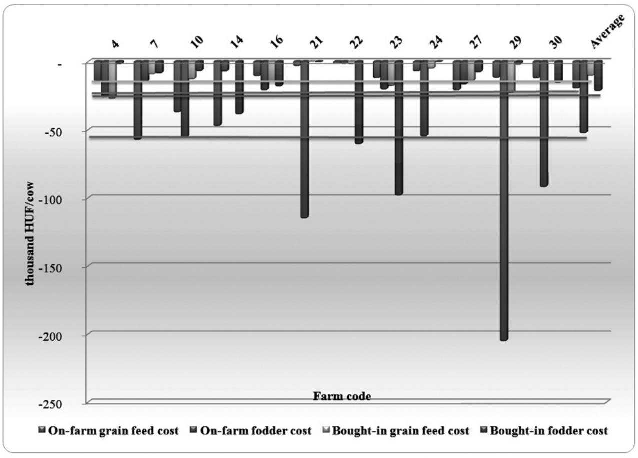

Figure 4. Feed cost differences of the non-efficient farms compared to the composite farms
Source: own calculation

Analysing the differences of labour costs, it can be stated that labour costs should be reduced by 65 thousand HUF/cow on average on all farms. The highest degree of labour cost reduction should be reached on Farm 29 (163 thousand HUF/cow). This cut can be obtained by decreasing the labour hours or the hourly rates.

The level of direct cost is also higher than the level of efficient farms, by 446 thousand HUF/cow on average (Farm 5).

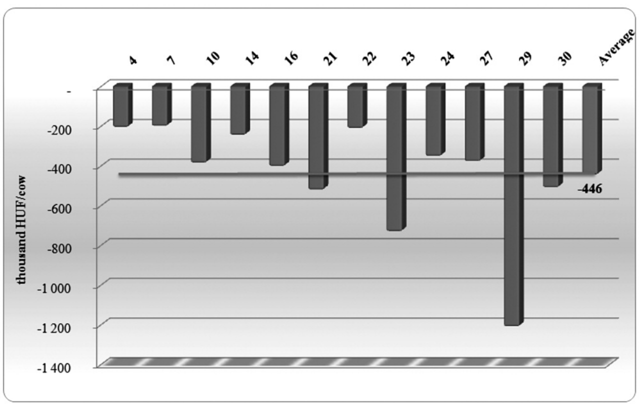

Figure 5. Direct cost differences of the non-efficient farms compared to the composite farms
Source: own calculation

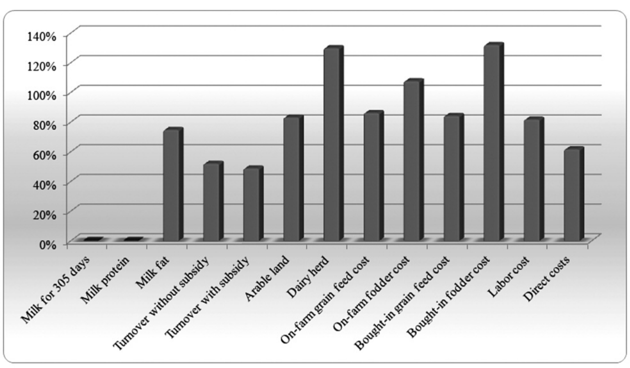

Figure 6. Coefficient of variation values for inputs and outputs
Source: own calculation

Since the values for milk production and milk protein are not needed to be modified, their coefficient of variation is 0% (Figure 6). Coefficient of variation for turnover and direct cost is around 50–60%, which implies an extreme fluctuation. The highest variation of coefficient has the herd size and bought-in fodder costs (around 120%). Coefficient of variation for arable land, on-farm and bought-in grain feed costs fluctuates around 80%.

Efficiency analysis of dairy farms in the Northern Great Plain region by stochastic DEA model

During the evaluation of the results for the stochastic DEA model, I applied the index of the **rate of efficiency (%)**. We can calculate this if the number of simulation runs that were 100% efficient is divided by the total number of simulation runs. The number of simulation tests was 5000. This large number was justified in order that the simulated values for variables shall cover the range of observation in the given parameter intervals based on the distributions, because this way, the results represent better all situations for the future after running the simulation.

After running the stochastic model, it was found that 9 farms are efficient in consideration of the risks, which is half of the result for deterministic DEA. (The considered risk factors are mentioned in the description of the stochastic DEA model of this paper.) The rate of efficiency is better on those 11 farms that were efficient according to the deterministic version, but the rate of efficiency is above 50% in case of only 3 dairy farms (Figure 7). This rate fluctuates around 10–50% on all other farms, which indicates that if the conditions change a bit to a less favourable way, these farms will not meet the criteria of the farms with good practice in the region.

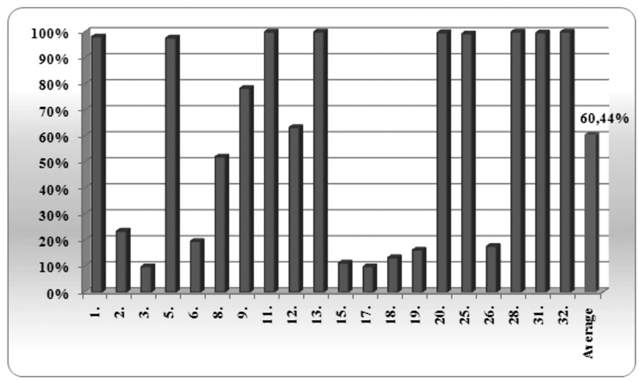

Figure 7. Stochastic DEA efficiency of efficient farms based on the deterministic DEA model
Source: own calculation

The average rate of efficiency in cases of non-efficient farms was 8.31% (Figure 8) compared to the value of 60.44% for efficient ones. This definitely implies that the results of deterministic model are reliable, because the chance of reaching such an input-output combination, in order to be considered 100% efficient, is very small on those farms which were non-efficient.

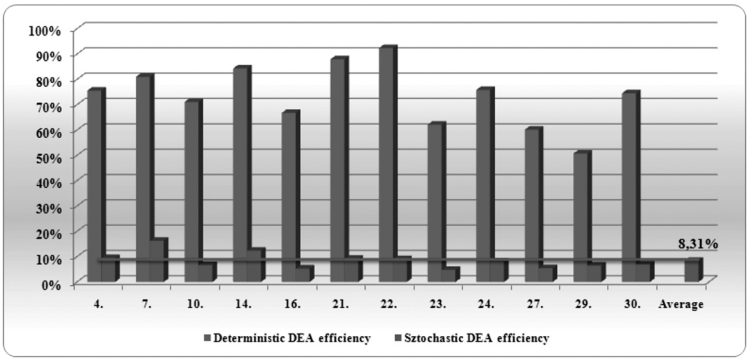

Figure 8. Non-efficient farms' deterministic and stochastic DEA efficiency values
Source: own calculation

According to the deterministic DEA there were 20 efficient farms (Table 8). After 5000 simulation runs, in the most favourable cases, the number of efficient farms was 22 and the minimum value was 7. The distribution is presumed to be close to symmetric, because mean and median are close to each other in every category. Median in the efficient category is 13 farms, which means that in 50% of the cases (in 2500 cases) 13 farms were efficient from 32. This is better than the presented 9 farms at the rate of efficiency, but the constraints are much more solid as well.

In the further columns of Table 8, I decreased the efficiency values by 0.1 and presented the cumulated statistical indices to the given categories. The value of DEA efficiency shows the extent of interventions to be made for the best practice. The lower the efficiency value, the more drastic action is to be made. In the case of deterministic DEA, all 32 farms show values above 0.5; moreover, 24 farms – ¾ of dairy farms – have efficiency above 0.8. If the results of stochastic DEA is analysed, we can see that the median value of category for >0.5 is 24, which means that for these 24 farms the chance of having efficiency above 0.5 is 50%. The median of category above 0.8 is 15, so we can expect efficiency above 0.8 at less than half of the number of farms in 2500 runs.

These results are in fully compatible with the consequences drawn at the rate of efficiency, so it can be stated that at present most of the dairy farms (62.5%) in the Northern Great Plain Region have good practice according to the deterministic version of DEA analysis. However, if the risks of inputs and outputs are also considered, it is found that even most of the farms with 100% efficiency can correct a small split of the present balance with difficulties.

Table 8. Some statistical indices of farms in different DEA categories

Unit: number of farms

		y	>0,9	>0,8	>0,7	>0,6	>0,5
Deterministic		20	21	24	28	31	32
Stochastic	MAX	22	24	25	28	31	31
	MIN	7	7	8	8	9	9
	MED	13	14	15	18	22	24
	Mean	13,1	13,9	15,3	17,4	20,8	23,7
	Standard deviation	2,4	2,7	3,0	3,5	3,9	3,3

Source: own calculation

As the next step, the influence of input and output factors was analysed on the efficiency. In the literature it is meant that we analyse the effect of factors with random variable on the forecasted values. The simplest way to do this is to make a regression analysis.

At multiple linear regression calculations, one of the most frequent problems is the narrow observation range. In this case it was not a problem, because 5000 runs provided data of proper quality and quantity.

The other most frequent problem is multicollinearity. In this case, there is a strong correlation between the two factors, stronger than with the dependent variable. In case of multicollinearity, the definition of partial regression coefficients for the given factors is inexact; therefore, one of the two factors of strong correlation must be left out of the model, so the estimation for the other parameter will be exact.

The partial regression coefficients in the regression analysis show the absolute effect of the influencing factors. The measurement unit and the order of magnitude for input and output factors in this DEA model are significantly differ, so it is practical to apply the standardized regression coefficient, the ß-coefficient (EZÉKIEL–FOX, 1970) in the comparison.

The calculation of ß-coefficient:

$$\beta_i = \frac{b_i \cdot S_i}{S_Y} \qquad (10.)$$

where
b_i: the partial regression coefficient,
S_i: standard deviation for independent variable i,
S_Y: standard deviation of the dependent variable.

Variables of the regression analysis:
- DEA efficiency (dependent variable)
- independent variable:
  - milk production,
  - milk protein,
  - milk fat,
  - turnover (without subsidy),
  - on-farm grain feed cost,
  - bought-in grain feed cost,
  - on-farm fodder cost,
  - bought-in fodder cost,
  - labour costs,
  - direct costs.

Milk production, milk fat, turnover and direct costs were involved in the model; the other variables were eliminated because of multicollinearity.

Based on beta-weights, milk fat has the greatest effect on DEA efficiency (Table 9). This factor was the most important at 2/3 of the farms. According to the ranks, milk production is on the second place, direct cost is the third and turnover is the last one. Analysing the average beta values, we can see that on the second and third place, milk production and direct costs change places, which implies that direct cost has smaller influence, but on those farms where its significance is higher, its effect is stronger on DEA efficiency.

Table 9. Effect of beta weights on DEA efficiency rank

Name	Place in ranking				Average beta	Average place
	1	2	3	4		
Milk production	4	8	13	5	1,32	2,63
Milk fat	20	5	3	2	2,87	1,57
Turnover	3	6	7	14	1,15	3,07
Direct cost	3	11	7	9	1,68	2,73

Source: own calculation

If the rank of influencing factors is analysed by farms, it can be stated that beta values are substantially higher at the farms with lower DEA efficiency, so less efficient farms are more sensitive to changes. In Table 10, farms were sorted by their rates of efficiency and the date show that farms with lower efficiency levels have higher beta values: it is concluded that there are weak-medium correlation among them (the correlation coefficient is between 0.5-0.6)

Table 10. Rank of farms by beta values

Farm	Milk production		Milk fat		Turnover		Direct cost		DEA efficiency
	Beta	Rank	Beta	Rank	Beta	Rank	Beta	Rank	
23.	8,294	10	-13,142	22	-6,159	8	10,569	19	4,74%
16.	2,264	21	-5,383	12	0,671	25	1,906	12	5,28%
27.	-0,314	13	-0,432	10	0,462	9	0,29	17	5,44%
29.	3,509	7	-7,191	8	-0,827	23	3,968	11	6,44%
10.	-1,313	23	0	20	5,908	26	-5,003	27	6,66%
30.	1,971	6	-4,882	4	0,868	13	1,47	3	6,92%
24.	1,115	19	-4,269	21	1,502	27	1,053	22	8,08%
22.	1,126	9	-2,609	19	0,851	16	0,248	15	8,96%
21.	0,438	11	-2,46	29	1,221	2	0,32	2	9,14%
4.	1,711	20	-4,247	23	0,359	12	1,72	18	9,40%
17.	-1,239	14	0	18	3,82	21	-3,095	14	9,78%
3.	1,206	27	-3,473	26	1,083	28	0,994	28	9,82%
15.	-0,125	18	-2,696	9	2,341	24	0,311	9	11,20%
14.	0,855	29	-3,708	13	0,317	4	2,013	24	12,36%
18.	2,676	4	-9,307	5	-0,663	18	0,702	10	13,26%
19.	1,614	12	-3,374	29	-0,726	3	2,122	5	16,18%
7.	1,794	3	-5,725	2	-0,859	19	4,362	20	16,26%
26.	-0,182	8	0,097	11	0,136	17	-0,066	8	17,60%
6.	0,566	22	-1,427	24	0,307	22	0,236	21	19,54%
2.	0,672	25	-2,723	15	0,309	7	1,542	23	23,46%
8.	0,823	16	-1,109	14	-0,303	14	0,402	26	52,00%
12.	1,203	1	-2,062	1	-0,46	1	1,309	1	63,26%
9.	1,44	17	-1,915	7	-0,751	5	1,279	16	78,32%
1.	1,337	24	-0,828	17	-1,167	10	0,781	6	98,14%
25.	0,471	28	-2,081	28	-0,907	29	2,55	29	99,26%
31.	-1,13	26	2,196	25	1,272	20	-2,266	25	99,64%
20.	0,642	30	-0,624	27	-0,409	30	0,426	30	99,72%
11.	-0,705	2	0,741	3	0,866	15	-0,888	4	100%
13.	-0,289	5	0,226	6	0,214	11	-0,17	13	100%
28.	-0,011	15	0,108	16	-0,048	6	-0,049	7	100%

Source: own calculation

CONCLUSIONS, RECOMMENDATIONS

According to the deterministic DEA efficiency analysis, almost 2/3 of the examined dairy farms have “good practice”. The model analysis shows that those farms are efficient which direct costs (409 thousand HUF/cow on efficient and 620 thousand HUF/cow on non-efficient farms) are much more lower compared to their turnover (541 thousand HUF/cow without subsidy, 667 thousand HUF/cow with subsidy) and produced less milk per cow (5365 kg/cow), but with better quality parameters (milk fat and protein). In cases of the efficient farms, the feed cost was lower (151 thousand HUF/cow) than on the non-efficient ones (193 thousand HUF/cow). If the input and output parameters are considered to be probability variables, risk is then also considered. After 5000 simulation runs, this rate reduces to 1/3, which implies that even half of the efficient farms is quite sensitive to the unfavourable change of conditions. Based on the ß-weights, milk fat is the most important factor among the risk factors, which is followed by milk production, direct cost and turnover. The other involved input and output factors had to be eliminated from the model because of multicollinearity.

REFERENCES

1. **ALBRIGHT S.C.–WINSTON W.L. (2007):** Management Science Modeling, Revised Third Edition. Thomson, 184–192. p.
2. **BARÁTH L.-HOCKMANN H.-KESZTHELYI Sz.-SZABÓ G. (2007):** A TFP változásának forrásai a magyar mezőgazdaságban. Statisztikai Szemle, 87. évfolyam 5. szám. 471–492. p.
3. **BECSER N. (2008):** Szolgáltatásminőség modellek, BCE Vállalatgazdaságtani Intézet, 89. sz. Műhelytanulmányok, 2008. január, 19.
4. **BRUNI, M.E.-CONFORTI, D.-BERALDI P.-TUNDIS, E. (2009):** Probablistical constrained models for efficiency and dominance. In: Int. J. Production Economics 117. 219–228. p.
5. **BUNKÓCZI L.–PITLIK L. (1999):** A DEA (Data Envelopment Analysis) módszer alkalmazási lehetőségei üzemhatékonyságok mérésére, IA’99, Debrecen, 1999. augusztus 26.
6. **CHARNES A.-COOPER W.-RHODES E. (1978):** Measuring the efficiency of decision-making units. European Journal of Operational Research vol. 2, 429–444. p.
7. **COOPER W. W.–SEIFORD L. M.–ZHU J. (2004a):** Data Envelopment Analysis, International Series in Operations Research & Management Science. vol. 71. Chapter 1. 1–39. p.
8. **COOPER W. W.- DENG H. – HUANG, Z. – LI, S. X. (2004b):** Chance constrained programming approaches to congestion in stochastic data envelopment analysis. European Journal of Operational Research 155, 487–501. p.

9. **EZÉKIEL M.-FOX K. A. (1970**): Korreláció- és regresszióanalízis, lineáris és nem-lineáris módszerek. Közgazdasági és Jogi Könyvkiadó, Budapest.
10. **FARREL, M.J. (1957):** The Measurement of Productive Efficiency. Journal of the Royal Statistical Society vol. 120, 253–281. p.
11. **FÜLÖP J.–TEMESI J. (2000):** A Data Envelopment Analysis (DEA) alkalmazása ipari parkok hatékonyságának vizsgálatára. Szigma 32. kötet 3–4. füzet, 85–110. p.
12. **KHODABAKHSHI M. – ASGHARIAN M. (2009):** An input relaxation measure of efficiency in stochastic data envelopment analysis, 2010–2023.
13. **KOVÁCS K.–EMVALOMATIS G. (2011):** Dutch, Hungarian and German dairy farms technical efficiency comparison. APSTRACT Vol. 5. Agroinform Publishing House. Budapest.
14. **KOVÁCS S.–NAGY L. (2009):** An application of Marcov chain Monte Carlo simulation. Internal Congress on the aspects and visions of applied economics and informatics (AVA4), Debrecen, 1333–1338. p.
15. **LORENZ, M. O. (1905):** Methods of measuring the concentration of wealth. Publications of the American Statistical Association. Vol. 9. No. 70. June 1905. 209–219. p.
16. **LOTFI F. H., – JAHANSHAHLOO, G. R. – ESMAEILI M. (2007):** Sensitivity analysis of efficient units in the presence of non-discretionary inputs. Applied Mathematics and Computation 190, 1185–1197. p.
17. **RAGSDALE C.T. (2007):** Spreadsheet Modeling & Decision Analysis: A Prectical Introduction to Management Science, Fifth Edition. Thomson, 102–113. p.
18. **SEIFORD L. M. – ZHU J. (1998):** Stability regions for maintaining efficiency in data envelopment analysis. European Journal of Operational Research 108, 127–139. p.
19. **SHERMAN H.D. – LADINO G. (1995):** Managing Bank Productivity Using Data Envelopment Analysis (DEA). Interfaces 25, no. 2, 60–80. p.
20. **SOTERIOU A.C.–STAVRINIDES Y. (2000):** An internal customer service quality data envelopment analysis model for bank branches. International Journal of Bank Marketing, Vol. 18., No. 5. 246–252. p.
21. **SUEYOSHI T. (2000):** Stochastic DEA for restructure strategy: an application to Japanese petroleum company. Omega 23, 385–398.
22. **SZŐKE Sz. – NAGY L. – KOVÁCS S. – BALOGH P. (2009):** Examination of pig farm technology by computer simulation. Internal Congress on the aspects and visions of applied economics and informatics (AVA4), Debrecen. 1317–1325. p.
23. **SZŰCS I. (2002):** Alkalmazott statisztika. Agroinform Kiadó, Budapest.
24. **Tej Terméktanács:** 2010. Tájékoztató anyag. Kézirat. pp. 1–17.
25. **TIBENSZKYNÉ F. K. (2007):** Az oktatás hatékonyságának mérése a ZMNE 2006-ban végzett hallgatóin Data Envelopment Analysis (DEA) módszer használatával. Doktori Ph.D. értekezés, ZMNE, 149–165. p.
26. **TOFALLIS C. (2001):** Combining two approaches to efficiency assessment. Journal of the Operational Research Society 52 (11), 1225–1231. p.
27. **TÓTH Á. (1999):** Kísérlet a hatékonyság empirikus elemzésére. MNB füzetek 1999/2, 11-33. p.
28. TSIONAS E. G. (2003): Combining DEA and stochastic frontier models: An empirical Bayes approach. European Journal of Operational Research 147, 499-510. p.