Scientific Papers Series Management, Economic Engineering in Agriculture and Rural Development
Vol. 14, Issue 4, 2014
PRINT ISSN 2284-7995, E-ISSN 2285-3952

# RESEARCH ON REGRESSION MODELING OF PROFIT RELATED TO MILK YIELD IN DAIRY FARMING

**Agatha POPESCU, Livia DAVID**

University of Agricultural Sciences and Veterinary Medicine Bucharest, 59 Marasti, District 1, Zip code 011464, Bucharest, Romania, Phone: +40 213182564/232, +40722462647, +40722949405, Fax:+40213182888, Emails:agatha_popescu@yahoo.com, livia_david@hotmail.com

***Corresponding author***: agatha_popescu@yahoo.com

***Abstract***

*The paper aimed to establish a corresponding regression model reflecting the relationship between profit, as the main barometer of economic efficiency and milk yield in dairy farming using a sample of 8 farms operating in the Southern Romania. Two regression models were compared: the linear regression and the quadratic fit. Average milk yield registered 6,088.57 kg/cow and had just 9.24 % variation among farms. Profit per cow recorded Lei 2,096.57 in average with a very high variation from a farm to another ( 46.20%). The correlation coefficient between milk yield and profit per cow, $r_{xy}$= 0.921, reflected a strong positive link between the two economic indicators. The regression model had the form Y=1.584x - 7,508.66 with the standard error $S_{est}$= 407.370428 and the parabolic fit was Y=0.006$x^2$ -5.7649 + 14,250 having a higher standard error $S_{est}$= 18,786.96. From this comparison, the linear regression model proved to be the most suitable one to reflect the relationship between profit per cow and milk yield with the highest accuracy. According to this model, it was estimated that for an annual 500 kg gain in milk yield, profit per cow could be higher by Lei 792 per year with a deep impact on farm profitability.*

**Key words**: *dairy farming, regression modeling, profit, milk yield*

## INTRODUCTION

Milk yield has a deep impact on the economic results in dairy farming being in its turn influenced by technological aspects, mainly regarding cow feeding, but also breeding herd structure, reproduction and natural conditions.
Miron and Lup (2013) affirmed how important is dairy farm size in determining economic efficiency . [8]
Kopecek in 2002, using the cost function, noticed that the costs for marketed milk for one feeding day of a cow have a lower growth compared to milk yield and using the expense function he determined the maximum profit per litre of marketed milk, maximum profit per cow and year and the interval of profitability for milk production. [6]
Grigoroiu (2008) proposed a new method for establishing the threshold average marketed milk production for assuring profitability in dairy farms in Romania. [4]
Between milk yield, variable cost and gross margin per cow is a close relationship with major results for farm profitability as mentioned by Pirvutoiu and Popescu Agatha in 2012. [10]
The correlation between milk yield and profit per cow is a strong positive one, as found by Popescu Agatha and Gyeresi Stefan in 1989 [15]
Economic performance in terms of milk yield and the financial performance in terms of gross margin are closely related to farm size as affirmed by Popescu Agatha in 2009 and 2010. [11,12]
Profit variation depends on marketed milk and production cost as mentioned by Kopecek in 2002 and Popescu Agatha in 2014. [6,14] A gain of 100 kg in milk production could increase milk cost by Lei 0.02 per milk kilogram and farmers profit by Lei 118.9 per cow and year as obtained by Popescu Agatha in 2014 in Romania. [14]
The relationship between milk cost, return and profitability in dairy farming was put into evidence by Popescu Agatha (2014b) [15]

The relationship between milk yield and economic indicators in dairy farming was studied using various modeling techniques. Popescu Agatha in 2010 used linear regression function in gross margin forecast based on milk yield.[12,13]

Murphy et al. in 2014 used nonlinear auto-regressive model with exogenous input, a static neural network and a multiple linear regression model for predicting daily milk yield per herd over various forecast horizons. He concluded that the non linear auto-regressive model with exogenous input is the more accurate solution to conventional regression techniques for short-term prediction of milk yield. [8]

Ramsbottom et al. in 2011 used linear and quadratic models to determine the correlation between dairy cow genetic merit in terms of economic breeding index, milk yield, fat and protein content, calving interval and financial indicators: income per cow, cost per cow, and profit in commercial spring calving dairy farms. [18]

Hansen et al. in 2005 measured the financial performance in dairy farms in close relationship with milk production. [5]

The correlation coefficient is largely used to reflect the relationship between various variables as affirmed by Colton in 1974, [2], Pearson in 1985, [9], Sponaugle et al. in 2014, [20].

Linear regression is used to reflect the relationship and evolution trend between different variables which depend one to another as affirmed by Murphy et al. in 2014 [6], Popescu Agatha in 2010 [12], Sokal et al. in 1995 [19].

Quadratic models are also used to characterize the link between different variables and their dependence one to another as mentioned by Popescu Agatha and Gyeresi Stefan in 1989 [16], Ramsbottom et al. in 2011 [18].

Standard error of linear regression, quadratic fit and other mathematical models assures the highest accuracy of the prediction [19].

For this reason, it is commonly used in the decision what mathematical model should be chosen to reflect the best way the link between various indicators or variables and to predict their future evolution with the highest precision.

Also, the determination coefficient or R squared is used to explain how much of the total variation of the dependent variable is given by the independent variable as mentioned by Bolboaca [1], Dufour et al. in 2011 [3].

In this context, the present paper aimed to test two mathematical models: linear regression and quadratic fit in the analysis of milk yield impact on profit in dairy farming in Romania in order to establish the most suitable modeling technique in profit prediction based on average milk production.

## MATERIALS AND METHODS

In order to set up this paper, the primary data were collected from a sample of 8 dairy farms situated in the Southern Romania in the year 2013.

The economic indicators taken into consideration were divided into two categories of variables: (i) independent variables, in this case X-milk yield/cow/year, and (ii) dependent variables symbolized Y, as follows: $Y_1$-milk production/farm, $Y_2$-production expenses/cow/year, $Y_3$-milk cost, $Y_4$-profit/cow/year, and $Y_5$-profit/farm.

For each economic indicators, the following statistical parameters have been determined:

(a)Average of the variable**,** $\overline{X}$ , using the well known formula:

$$\overline{X} = \frac{X_1 + X_2 + ...X_n}{n} \qquad (1)$$

(b)Variance of variable**,** $S^2$, according to the formula:

$$S^2 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})}{n-1} \qquad (2)$$

(c)Standard Deviation**,** S, based on the formula:

$$S = \left(\sqrt{\frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1}}\right) \quad (3)$$

(d)Variation Coefficient, $V_{\%}$, using the formula:

$$V_{\%} = \frac{S}{\overline{X}} \text{ x } 100 \quad (4)$$

The correlation coefficients between milk yield and each of the other economic indicators: milk production/farm, production expenses/cow/year, milk cost, profit/cow/year and profit/farm were determined using Pearson product-moment correlation coefficient, $r_{xy}$, whose mathematical formula is:

$$r_{xy} = \frac{\sum(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum(x_i - \overline{x})^2 \sum(y_i - \overline{y})^2}} \quad (5)$$

The economic optimization was based on the estimation of profit per cow/year, the dependent variable, $Y_4$, reflecting the best manner the results of the activity in dairy farming, related to milk yield, X, considered the independent variable.

In order to establish which mathematical model is suitable to the evolution of this pair of indicators, two regression functions were tested:

- ***Linear regression***, whose formula is:

y = ax +b, (6)

where y = the dependent variable and x = the independent variable,

***Least Square Method of Regression Analysis*** allowed to calculate the values of the coefficients a and b, solving the system of the two normal equations given below:

$$\sum y = a\sum x + nb$$
$$\sum xy = a\sum x^2 + b\sum x \quad (7)$$

and using the formulas:

$$a = \frac{\sum x \sum y - n\sum xy}{(\sum x)^2 - n\sum x^2} \quad (8)$$

$$b = \frac{\sum x \sum xy - \sum y \sum x^2}{(\sum x)^2 - n\sum x^2} \quad (9)$$

***The standard error of the estimate***, $S_{est}$, was also required to be determined as a measure of the accuracy of predictions, using the formula given below:

$$S_{est} = \sqrt{\frac{\sum(Y_i - Y_{icalc})^2}{N-2}} \quad (10)$$

where Y is the actual value, $Y_{ical}$ is a predicted value, and N is the number of pairs of values.

Also, ***the determination coefficient or R squared***, $R^2$, was calculated using the formula:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(Y_i - Y_{icalc})^2}{\sum_{i=1}^{n}(Y_i - \overline{Y})^2} \quad (11)$$

***-Quadratic or Parabolic Fit,*** whose formula is:

$y = ax^2 + bx + c$, (12)

where y = the dependent variable and x = the independent variable.

***Least Square Method of Regression Analysis*** allowed to calculate the values of the coefficients a, b and c solving the system of the normal equations given below:

$$\sum_i x_i^2 y_i = a\sum_i x_i^4 + b\sum_i x_i^3 + c\sum_i x_i^2$$
$$\sum_i x_i y_i = a\sum_i x_i^3 + b\sum_i x_i^2 + c\sum_i x_i$$
$$\sum_i y_i = a\sum_i x_i^2 + b\sum_i x_i + nc \quad (13)$$

where $x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$ (14)

The formulas for the parameters a, b and c are given below: (15)

$$a = \frac{[\sum_i x_i \sum_i x_i^3 - (\sum_i x_i^2)^2]\sum_i y_i - [n\sum_i x_i^3 - \sum_i x_i \sum_i x_i^2]\sum_i x_i y_i + [n\sum_i x_i^2 - (\sum_i x_i)^2]\sum_i x_i^2 y_i}{\Delta}$$

$$b = \frac{[\sum_i x_i^2 \sum_i x_i^3 - (\sum_i x_i \sum_i x_i^4]\sum_i y_i - [(\sum_i x_i^2)^2 - n\sum_i x_i^4]\sum_i x_i y_i + [\sum_i x_i \sum_i x_i^3 - n\sum_i x_i^3]\sum_i x_i^2 y_i}{\Delta}$$

$$c = \frac{[\sum_i x_i^2 \sum_i x_i^4 - (\sum_i x_i^3)^2]\sum_i y_i - [\sum_i x_i \sum_i x_i^4 - \sum_i x_i^2 \sum_i x_i^3]\sum_i x_i y_i + [\sum_i x_i \sum_i x_i^3 - (\sum_i x_i^2)^2]\sum_i x_i^2 y_i}{\Delta}$$

where $\Delta = [\sum_i x_i^2 \sum_i x_i^4 - (\sum_i x_i^3)^2]n - [\sum_i x_i \sum_i x_i^4 - \sum_i x_i^2 \sum_i x_i^3]\sum_i x_i + [\sum_i x_i \sum_i x_i^3 - (\sum_i x_i^2)^2]\sum_i x_i^2$

and $\sum_{i=1}^{n} x_i = \sum_i x_i$ . (16)

Also, the standard error and R squared were calculated for the parabolic function.

The most adequate regression function was chosen to forecast profit per cow using the regression model having the lowest standard error for assuring the highest prediction and a 500 kg/cow gain interval for milk yield.

## RESULTS AND DISCUSSIONS

**The economic indicators** characterizing each farm are presented in Table 1.

Table 1.The economic indicators by farm in the year 2013

| Farm | Number of dairy cows | Milk yield kg/cow | Milk production per farm Kg/farm | Production Expenses per cow Lei/cow | Milk cost Lei/kg | Profit/cow Lei/cow | Profit per farm Lei/farm |
|---|---|---|---|---|---|---|---|
| F1 | 25 | 5,840 | 146,000 | 7,030 | 1,20 | 1,135 | 28,375 |
| F2 | 30 | 5,950 | 178,500 | 6,696 | 1,13 | 1,505 | 45,150 |
| F3 | 18 | 6,370 | 114,660 | 6,762 | 1,06 | 3,062 | 55,116 |
| F4 | 22 | 6,600 | 145,200 | 7,096 | 1,08 | 3,089 | 67,958 |
| F5 | 28 | 6,730 | 188,440 | 7,235 | 1,08 | 3,079 | 86,212 |
| F6 | 34 | 5,500 | 187,000 | 7,034 | 1,28 | 1,089 | 37,026 |
| F8 | 50 | 5,630 | 268,000 | 7,198 | 1,34 | 1,717 | 85,850 |
| F9 | 20 | 5,115 | 102,300 | 7,129 | 1,40 | 900 | 18,000 |

Source: Farm bookkeeping in the Southern Romania, 2013 [22]

**The average, variance, standard deviation and variation coefficient.**

***Milk yield*** registered an average of 6,088.57 kg/cow with a variation from the minimum 5,115 kg/cow in case of F8 and 6,730 kg/cow in case F5. The variation coefficient was 9.24 % reflecting a low variation from a farm to another.(Table 2).

Table 2. Average and variation coefficient of each economic indicator

| Indicator | MU | $\overline{X}$ | $S^2$ | S | V% |
|---|---|---|---|---|---|
| Milk yield | Kg/cow/year | 6,088.57 | 317,035.27 | 563.06 | 9.24 |
| Milk production per farm | Kg/farm | 175,400 | 2,718,625,421.63 | 52,140.44 | 29.72 |
| Production expenses per cow | Lei/cow | 7,007.29 | 38,193.14 | 195.43 | 2.78 |
| Milk production cost | Lei/kg | 1.19 | 0.02 | 0.13 | 11.07 |
| Profit per cow | Lei/cow | 2,096.57 | 938,399.14 | 968.71 | 46.20 |
| Profit per farm | Lei/farm | 57,955.29 | 651,221,174.13 | 25,519.04 | 44.03 |

Source: Own calculations

***Milk production per farm*** recorded 175,400 kg in average, varying from the lowest level in case of F5, 102,300 kg, and the highest level in case of F8, 268,000 kg. The variation of this indicator among farms was very high, 29.72 %, being determined by the number of cows raised in farms and their milk yield.

***Production expenses per cow*** accounted for Lei 7,007.29 per dairy cow, ranging between Lei 6,696 in case of F2, the minimum level, and Lei 7,198 in case of F8. The variation coefficient reflected a small variation among farms regarding this economic indicator, 2.78 %.

***Milk production cost*** was Lei 1.19 per milk kilogram, varying between Lei 1.06 in case of F3, the lowest level and Lei 1.40 in case of F8, the highest level. The variation of milk cost among farms was a middle one, as the variation coefficient indicated, 11.07 %.

***Profit per cow*** recorded Lei 2,096.57 in average with a variation between Lei 900, the lowest level registered by F9, and Lei 3,089, the highest value registered by F4. This indicator had a very high variation from a farm to another, as the variation coefficicient confirmed, 46.20%.

***Profit per farm coming from milk*** was Lei 57,955.29 in average, varying between Lei 18,000 in case of F9, the lowest value, and Lei 86,212 in case of F5, the highest value. The variation regarding this indicator was very large, the coefficient of variation being 44.03 %.(Table 2)

**The the correlation coefficients** between milk yield and the other five economic indicators taken into consideration are presented in Table 3.

Table 3. Pearson product-moment correlation coefficients between milk yield and the other economic indicators

| Pairs of economic indicators taken into account | PPMCC, correlation coefficient, $r_{xy}$ |
|---|---|
| Milk yield x Milk production per farm | -0.020999472 |
| Milk yield x Production expenses per cow | -0.060981647 |
| Milk yield x Milk production cost | -0.92375365 |
| Milk yield x Profit per cow | 0.921096492 |
| Milk yield x Profit per farm coming from milk | 0.6475586 |

Source: Own calculations

A negative weak correlation, $r_{xy}$= -0.020 was found between milk yield and milk production per farm.

This could be explained by the fact that milk production is influenced by the number of dairy cows which varied from a farm to another and also by milk consumption for calves up to weaning.

Between milk yield and production expenses is was found a negative low correlation, $r_{xy}$= -0.060, reflecting a large variety of factors influencing production costs per cow, besides average milk production.

The correlation coefficient between milk yield and profit per cow, $r_{xy}$= 0.921, reflected a strong positive link between the two economic indicators. Therefore, the higher milk yield, the higher profit per cow.

Also, between milk yield and profit per farm coming from milk, it was found a positive high correlation, $r_{xy}$= 0.647, showing that a higher milk yield could lead to a higher profit.

Taking into account that the strongest positive correlation was found betwen milk yield and profit per cow, $r_{xy}$= 0.921096492, it was considered that profit per cow is the main economic indicator which should be optimized in close relation to average milk production per cow.

**Comparative results for the linear regression and quadratic fit** regarding the formulas, the values for the parameters a, b and c and, the standard error and the R squared in case of profit per cow related to milk yield are presented in Table 4.

Table 4.Comparative analysis between linear regression and quadratic fit

| | Linear regression $y = ax + b$ | Quadratic Fit $y = ax^2 + bx + c$ |
|---|---|---|
| Regression Model | $Y=1.584x$ - 7,508.66 | $Y=0.006x^2$ - 5.7649 + 14,250 |
| Standard error | 407.370428 | 18,786.96 |
| Pearson product-moment correlation coefficient | 0.921086315 | 0.936589558 |
| $R^2$ (R squared) | 0.84841848 | 0.8772 |
| a coefficient | 1.584693005 | 0.0006 |
| b coefficient | -7,508.665076 | -5.7649 |
| c coefficient | | 14,250 |

Source: Own calculations

As one can easily see, the lowest standard error was registered in case of the linear regression model, $S_{est}$ =407.370428 compared to $S_{est}$ = 18,786.96 recorded in case of the quadratic fit.

Therefore, only the linear regression model assures the highest accuracy in predicting the profit related to milk yield.

The graphical representation of the two regression models is shown in Fig.1. and, respectively, Fig.2.

**Forecast of profit per cow based on milk yield for a 500 kg gain.**

Taking into account the linear regression model, Y=1.584x - 7,508.66, assuring the lowest standard error, that is the highest precision, it was estimated profit per cow for an interval of milk yield gain of 500 kg/cow. The results are presented in Table 5 and showed that for an increased milk yield by 500 kg per year, profit per cow will grow by Lei 792/year.

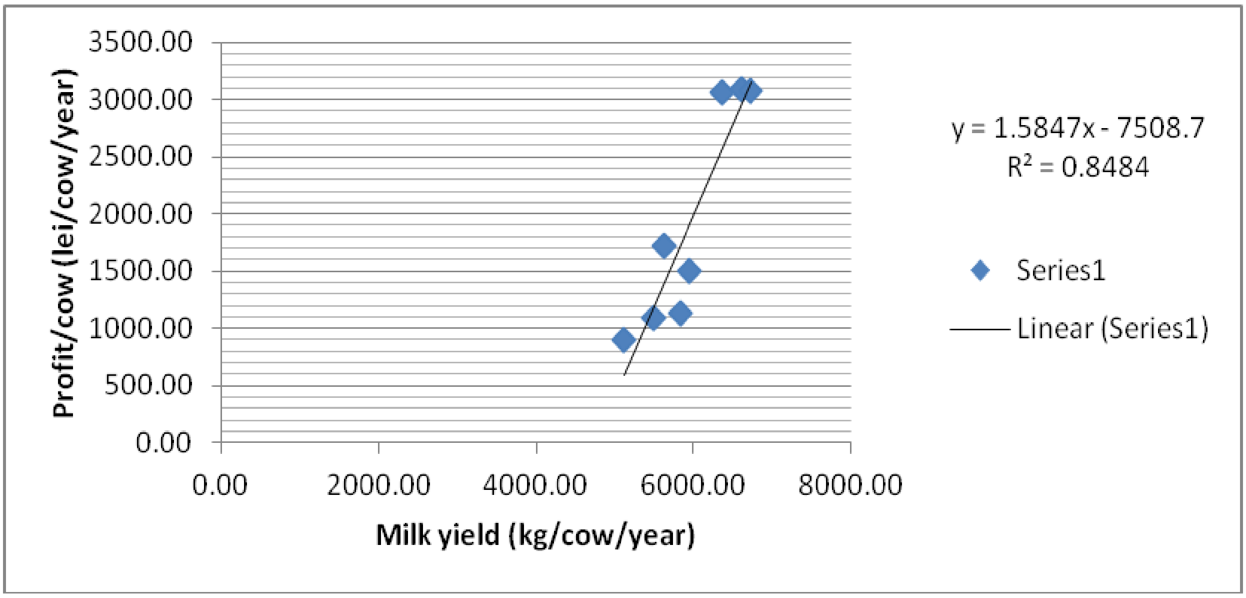


Fig.1.Linear regression between milk yield and profit per cow

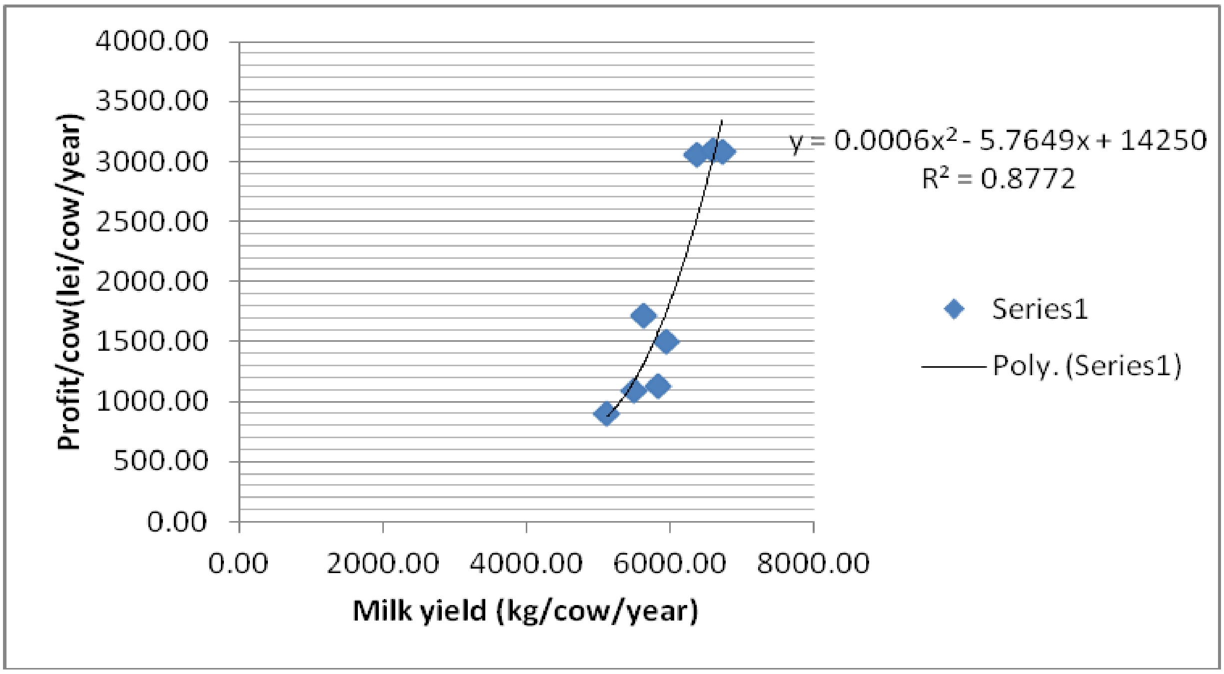


Fig.2. Parabolic fit between milk yield and profit per cow

Table 5.Estimated profit per cow based on milk yield using the linear regression model Y=1.584x - 7,508.66

| Milk yield, X | Estimated Profit per cow, $Y_{est}$ |
|---|---|
| 5.500 | 1,203.34 |
| 6,000 | 1,995.34 |
| 6,500 | 2,787.34 |
| 7,000 | 3,579.34 |
| 7,500 | 4,371.34 |
| 8,000 | 5,163.34 |
| 8,500 | 5,955.34 |
| 9,000 | 6,747.34 |
| 9,500 | 7,539.34 |
| 10,000 | 8,331.34 |

Source: Own calculations

## CONCLUSIONS

Average milk yield registered 6,088.57 kg/cow and had a reduced variation among farms ( 9.24 %).

Profit per cow recorded Lei 2,096.57 in average with a very high variation from a farm to another (46.20%).

Profit per cow is deeply influenced by milk yield as Pearson product-moment correlation coefficient proved ( $r_{xy}$ = 0.921), existing a strong positive relationship between these two economic indicators.

The comparison between the two regression models: the linear regression and the quadratic fit had the following mathematical representation: Y=1.584x - 7,508.66 with the standard error $S_{est}$= 407.370428 and the parabolic fit was $Y=0.006x^2 - 5.7649 + 14,250$ with the standard error $S_{est}$= 18,786.96.

The linear regression model proved to be the most suitable one to reflect the relationship between profit per cow and milk yield with the highest accuracy as its standard error was the lowest one.

For this reason, based on the linear regression model, it was estimated that for an increase of 500 kg in milk yield, profit per cow could grow by Lei 792 per year.

As a conclusion, the most important indicators with a deep impact on farm profitability in dairy farming are milk yield and profit per cow and the most adequate mathematical model for reflecting the link between them is the linear regression.

## ACKNOWLEDGMENTS

The authors thank the eight dairy farmers for their kindness to provide the production and financial data as this research work to be carried out.

## REFERENCES

[1]Bolboaca Sorana, Correlation and Linear Regression, http://sorana.academicdirect.ro

[2]Colton, T., Statistics in Medicine, 1974, Little Brown and Company, New York

[3]Dufour Jean-Marie, 2011, Coefficients of determination, CIRANO, McGill University, Montreal

[4]Grigoroiu, E, 2006, Method for establishing the average threshold marketed milk production in dairy farms, Scientific Papers Series Management, Economic Engineering in Agriculture and Rural Development, Vol.6, pp.277-279

[5]Hansen, B. G. and G. Stokstad (2005). Measuring financial performance on Dairy Farms, Acta Agricultural Scandinavica, Section C-Food Economics, June 2005, vol.2, Issue 2: 99-109

[6]Kopecek Petr, 2002, Analysis of the milk yield effect on the economics of milk production, Agric. Econ., *48* (10): 473–479

[7]Miron Liliana, Lup Aurel, 2013, Analysis of the structure of the agricultural farms by size and the share of the livestock among them. An argument for more helping for the small farms, Vol.13(2):

[8]Murphy, M.D., Mahony, M.J.O., Shallo, I., French, P., Upton, J., 2014, Comparison of modelling techniques for milk-production forecasting, Journal of Dairy Science, Vol.97)(6):3352-3363

[9]Pearson Karl, 1985, Notes on regression and inheritance in the case of two parents, Proceedings of the Royal Society of London, 58: 240–242

[10]Pirvutoiu, Ion, Popescu, Agatha, 2012, Research Concerning Standard Gross Margin Depending on Yield in Dairy Farming, Scientific Papers: Animal Science & Biotechnologies, Vol. 45(2):339

[11]Popescu Agatha, 2009, Analysis of milk production and economic efficiency in dairy farms, Scientific Papers Animal Science and Biotechnologies, Timisoara, vol. 42 (1):508-512

[12]Popescu Agatha, 2010a, Research Concerning the Use of Regression Function in Gross Margin Forecast, Bulletin UASVM Horticulture, 67(2):197-202

[13]Popescu Agatha, 2010b, Research concerning gross margin forecast based on milk yield using the least square method of regression analysis, Vol.10(2):175-180

[14]Popescu Agatha, 2014a, Research on profit variation depending on marketed milk and production cost in dairy farming, Scientific Papers Series

Management, Economic Engineering in Agriculture and Rural Development, Vol. 14(2):223-229

[15]Popescu Agatha, 2014b, Research on milk cost, return and profitability in dairy farming, Scientific Papers. Series Management, Economic Engineering in Agriculture and rural development, Vol. 14(2): 219-222

[16]Popescu Agatha, Gyeresi Stefan, 1989, Research regarding the harmonization of technical optimum with economic optimum in cow milk production. Novelties in the Technology of Domestic Animals, Vol.XV, Cluj Napoca, pp.78-83

[17]Quadratic Least Square Regression, http://www.azdhs.gov/lab/documents/license/resources/calibration-training/12-quadratic-least-squares-regression-calib.pdf

[18]Ramsbottom, G., Cromie, A. R., Horan, B., Berry, D. P. 2011, Relationship between dairy cow genetic merit and profiton commercial spring calving dairy farms, Animal (2012), 6:7, pp 1031–1039

[19]Sokal, R.R., Rohlf, F.J., 1995, Biometry, Freeman, New York

[20]Sponaugle Britani, 2014, Pearson Product-Moment Correlation: A Relationship Measurement Tool, https://www.udemy.com/blog/pearson-product-moment-correlation/

[21]Stats Tutorial - The Regression Line/Stats Tutorial - Errors in the Regression Equation, http://www.chem.utoronto.ca/coursenotes/analsci/Stats Tutorial/ErrRegr.html

[22]Valuable bookkeeping data for the year 2013 from 8 dairy farms situated in the Southern Romania