



## A multivariate and stochastic approach to identify key variables to rank dairy farms on profitability

A. S. Atzori,\*<sup>1</sup> L. O. Tedeschi,† and A. Cannas\*

\*Dipartimento di Agraria, Sezione di Scienze Zootecniche, Università di Sassari, 07100 Sassari, Italy

†Department of Animal Science, Texas A&M University, College Station 77843-2471

### ABSTRACT

The economic efficiency of dairy farms is the main goal of farmers. The objective of this work was to use routinely available information at the dairy farm level to develop an index of profitability to rank dairy farms and to assist the decision-making process of farmers to increase the economic efficiency of the entire system. A stochastic modeling approach was used to study the relationships between inputs and profitability (i.e., income over feed cost; IOFC) of dairy cattle farms. The IOFC was calculated as: milk revenue + value of male calves + culling revenue – herd feed costs. Two databases were created. The first one was a development database, which was created from technical and economic variables collected in 135 dairy farms. The second one was a synthetic database (sDB) created from 5,000 synthetic dairy farms using the Monte Carlo technique and based on the characteristics of the development database data. The sDB was used to develop a ranking index as follows: (1) principal component analysis (PCA), excluding IOFC, was used to identify principal components (sPC); and (2) coefficient estimates of a multiple regression of the IOFC on the sPC were obtained. Then, the eigenvectors of the sPC were used to compute the principal component values for the original 135 dairy farms that were used with the multiple regression coefficient estimates to predict IOFC (dRI; ranking index from development database). The dRI was used to rank the original 135 dairy farms. The PCA explained 77.6% of the sDB variability and 4 sPC were selected. The sPC were associated with herd profile, milk quality and payment, poor management, and reproduction based on the significant variables of the sPC. The mean IOFC in the sDB was  $0.1377 \pm 0.0162$  euros per liter of milk (€/L). The dRI explained 81% of the variability of the IOFC calculated for the 135 original farms. When the number of farms below and above 1 standard deviation (SD) of the dRI were

calculated, we found that 21 farms had  $dRI < -1$  SD, 32 farms were between  $-1$  SD and 0, 67 farms were between 0 and  $+1$  SD, and 15 farms had  $dRI > +1$  SD. The top 10% of the farms had a dRI greater than 0.170 €/L, whereas the bottom 10% farms had a dRI lower than 0.116 €/L. This stochastic approach allowed us to understand the relationships among the inputs of the studied dairy farms and to develop a ranking index for comparison purposes. The developed methodology may be improved by using more inputs at the dairy farm level and considering the actual cost to measure profitability.

**Key words:** principal component analysis, ranking index, decision making unit, modeling

### INTRODUCTION

The competitiveness of dairy farms often depends more on the improvement of technology and efficiency than on the size of the farms (Tauer, 2001; Cabrera et al., 2010). Differences in input combinations can change farm efficiency; hence, adequate technical strategies must be carried out to increase the efficiency of the farm production process (Stokes et al., 2007). Managerial capacity and approaches are highly associated with farm performance and efficiency (Solano et al., 2006). As highlighted by Solís et al. (2009) and Cabrera et al. (2010), the study of the sources of inefficiency at the farm level could be helpful to farmers who want to improve their farm performance and also for policy-makers who want to promote actions and interventions to improve productivity and economics.

The improvement of dairy farm efficiency has frequently been obtained by benchmarking partial indicators, such as the amount of a known input used per unit of output (Fraser and Cordina, 1999). However, several studies demonstrated that the use of partial indicators of technical efficiency is an inadequate approach to improve productivity, mainly because they are very dependent on the input used and one cannot assess if factors affected the numerator or the denominator of the partial indicator (Stokes et al., 2007). Furthermore, maximizing one indicator at a time may directly or in-

Received October 11, 2012.

Accepted January 5, 2013.

<sup>1</sup>Corresponding author: [asatzori@uniss.it](mailto:asatzori@uniss.it)

directly affect other correlated inputs or outputs (Fraser and Cordina, 1999). The improvement of technical efficiency is usually driven by the goal of maximizing the annual profit (Huirne et al., 1997). Conversely, the increase in annual profit by only maximizing technical efficiency can impair the economic efficiency in some cases (St-Pierre, 2001). Hence, farm efficiency should be evaluated by considering technical performance and economic outputs concurrently.

In certain production conditions, the optimal combination of inputs available to farmers can be identified by studying the profitability of the best farms (i.e., the most efficient decision making units). For this reason, it is important to develop a comprehensive indicator of profitability and efficiency to rank the decision making units and to benchmark their economic efficiency. The income over feed cost (**IOFC** = milk and meat direct revenues – feed costs) has often been used as an indicator of profitability of dairy farms when information on fixed costs is not available, and is also highly associated with short-term farm choices (Oleggini et al., 2001; St-Pierre, 2001; Ely et al., 2003; Bailey et al., 2005; Cabrera, 2010).

A multivariate approach can be used to manage the large number of farm inputs that are routinely collected by the farmer association, the milk plant, or the farm management software. The principal component analysis (**PCA**; Pearson, 1901) can be used to reduce the dimensionality of the data by taking into account the mean square error (Jolliffe, 1986; Jackson, 1991; Fodor, 2002). The PCA has been used to identify indicators of lactation persistency (Macciotta et al., 2006) and in farm management studies (Enevoldsen et al., 1996; Fahey et al., 2002; Tozer et al., 2003; Tauer and Mishra, 2006; Dechow et al., 2011). Kristensen et al. (2008) studied the effect of key performance indicators on the long-term financial performance of dairy farms by aggregating input variables based on empirical relationships developed on previous studies, instead of multivariate analysis, then classifying the key performance indicators (i.e., low, medium, and high) and regressing them on gross margin. In our case, the multivariate analysis could help to identify the main synergy among input variables of the dairy farm system, considering that no previous studies were carried out in this system to identify aggregation criteria. In addition, a stochastic approach combined with PCA could overcome the shortage of measured values in real dairy farms, by considering combinations that might be better than the existing ones and that might be useful to set new targets for the best farms as well.

The objectives of this work were (1) to find a comprehensive index of profitability based on information routinely available at the farm level, and (2) to evaluate

existing and new combinations of inputs and outputs to provide useful information to assist in the short-term decision-making process and to improve the economic efficiency of dairy farms.

## MATERIALS AND METHODS

### Database and Calculations

The data used in this work were from 135 dairy cattle farms located in Arborea (Sardinia, Italy) from October 2009 to September 2010. Eighteen technical and economic variables (original variables) were collected monthly by either the Italian Farmer Association (AIA; i.e., counts and performances of primiparous, multiparous, replacement cattle, dry cows, and culled cows) or the cooperative milk processing plant to which the farmers were associated (i.e., amount of milk sold, milk composition, and milk price). Table 1 has the list of the original variables and Figure 1 has the sequence and graphical representation of the calculations described below.

In the first step, the annual average of these 18 variables was calculated for each farm and used to build a development database (**dDB**). Then, their probability distributions and the Spearman correlations among the variables were obtained with @Risk 5.5 software (Palisade Corporation, Ithaca, NY), and the best distribution for each variable was chosen based on the  $\chi^2$  statistic. The descriptive statistics and the selected distributions for each variable in the dDB are in Table 1.

In the second step, to obtain a broader number of variable combinations than that observed in the studied farms, a synthetic database (i.e., artificially simulated database; **sDB**) was created with 5,000 synthetic dairy farms, in which a specific combination of the 18 variables was used. The 18 variables were randomly sampled based on their distributions and Spearman correlations. The simulation was performed with the Monte Carlo technique using the @Risk 5.5 software to maintain the same statistical characteristics of the dDB. This stochastic approach was used to study the relationships among these variables used as input variables in computing profitability.

In the third step, the income over feed cost for the sDB (**sIOFC**) was calculated as shown in Equation [1]:

$$\begin{aligned} \text{sIOFC} = & (\text{milk revenue} + \text{value of male calves} \\ & + \text{culling revenue}) - \text{herd feed costs}, \end{aligned} \quad [1]$$

where sIOFC is income over feed cost for the sDB, in euros (€); milk revenue is liters of milk sold  $\times$  milk price (€/L); value of male calves is number of male

**Table 1.** Descriptive statistics of input variables and computed income over feed cost (IOFC) obtained from the development data set (n = 135 dairy farms)

Variable	Unit	Mean	SD	Range (maximum to minimum)	Distribution
Potential production level <sup>1,2</sup>	L at 305 DIM	10,294	946	13,139 to 6,968	Normal
Primiparous <sup>1</sup>	Number	42.1	29.8	180.0 to 5.6	Log logistic
Multiparous <sup>1</sup>	Number	72.4	37.9	310.3 to 21.3	Gamma
Dry cows <sup>1</sup>	Number	19.0	11.3	80.0 to 4.6	Inverse Gauss
Replacement per year <sup>1</sup>	Number	52.0	28.6	226.8 to 14.0	Gamma
Calves <sup>1</sup>	Number	120.3	68.6	544.0 to 21.8	Log normal
Age at first calving <sup>1</sup>	Months	28.2	3.1	69.3 to 23.4	Pearson
Age at culling <sup>1</sup>	Months	68.4	11.3	169.7 to 51.7	Inverse Gauss
Culled cows <sup>1</sup>	Number	37.5	25.0	17.5 to 0.7	Extreme value
Days open <sup>1</sup>	Days	164.4	26.9	259.2 to 115.1	Weibull
Milk sold <sup>3</sup>	1,000 L/yr	974.3	569.5	3,972.5 to 241.8	Inverse Gauss
Milk fat content <sup>3</sup>	%	3.78	0.14	4.28 to 3.33	Gamma
Milk protein content <sup>3</sup>	%	3.35	0.08	3.76 to 3.16	Gamma
Milk price <sup>3</sup>	€/1,000 L	353.26	13.70	397.66 to 316.19	Weibull
Bonus fat <sup>3</sup>	€/1,000 L	5.46	4.78	28.19 to -14.93	Log logistic
Bonus protein <sup>3</sup>	€/1,000 L	6.25	4.06	34.22 to -0.63	Pearson
Bonus SCC <sup>3</sup>	€/1,000 L	5.53	8.14	17.04 to -17.04	Triangular
Bonus other <sup>3</sup>	€/1,000 L	10.12	3.95	14.20 to -4.30	Betagenereal
Farm profitability, IOFC <sup>4</sup>	€/L	0.138	0.162	0.175 to 0.066	Betagenereal

<sup>1</sup>From farmer association.

<sup>2</sup>It represents the phenotypic mean of the genetic level of the herd.

<sup>3</sup>From the milk processing plant.

<sup>4</sup>Our calculation.

calves × 50 €/calf; culling revenue is number of culled cows × 300 €/cow; and herd feeding costs is herd energy requirements (Mcal) × 0.158 (€/Mcal).

Regarding the value of calves and culled cows, the studied farms are part of a cooperative service that collects: (a) all male calves few days after birth, paying the same price to all farmers, and (b) all culled cows, paying a fixed price per kilogram of live weight. The income of the farmers might vary depending on the weight of culled cows, but no data of animal weights of each farm were available. Thus, we used the average price paid by the cooperative for culled cows.

In Eq. [1], the IOFC was assumed to be an indicator of profitability because no information about fixed costs was available for the studied farms. The cooperative milk processing plant to which the studied farms were associated calculated the milk price on the basis of the quality of the milk sold, fat and protein content, SCC, bacterial count, cryoscopy, and farm involvement in quality and traceability programs. Feeding costs were assumed to be equal in all farms and were calculated as proportion of total energy requirements, considering the recommendations of the NRC (2001) for maintenance, lactation, pregnancy, and growth of the cattle in the herd.

### Statistical Analyses

All statistical analyses conducted with the dDB and sDB were performed using SAS Version 9.3 (SAS Institute, Cary, NC).

**Analyses with the Synthetic Database.** A PCA was performed in the sDB, excluding the sIOFC, with the PROC COMP. The principal components (**PC**) of the sDB (**sPC**; n = 5,000 for each selected PC) that explained at least 75% of variance were selected. Subsequently, each selected PC was matched with managerial and technical aspects of the farm. Then, a multiple regression analysis of sIOFC on selected sPC scores was performed with the PROC REG to estimate the coefficients of the relationship between profitability (i.e., IOFC) and the input variables as shown in Equation [2]:

$$\text{sIOFC} = a + \text{sPC1 } X_1 + \text{sPC2 } X_2 + \dots + \text{sPCn } X_n + e, \quad [2]$$

where a is the intercept that represents the average of the sIOFC; sPC1 is the score of the first PC of the sDB;  $X_1$  is the regression coefficient of PC1 and so forth; and e is the random error.

A ranking index for the sDB (**sRI**) was computed using the coefficients estimated with Equation [2] as a proxy for the farm profitability as follows (Equation [3]):

$$\text{sRI} = a + \text{sPC1 } X_1 + \text{sPC2 } X_2 + \dots + \text{sPCn } X_n. \quad [3]$$

To evaluate the adequacy in predicting sIOFC with the sRI, adequacy statistics were obtained with the Model Evaluation System (Tedeschi, 2006; <http://nutrition-models.tamu.edu/mes.html>), including model precision

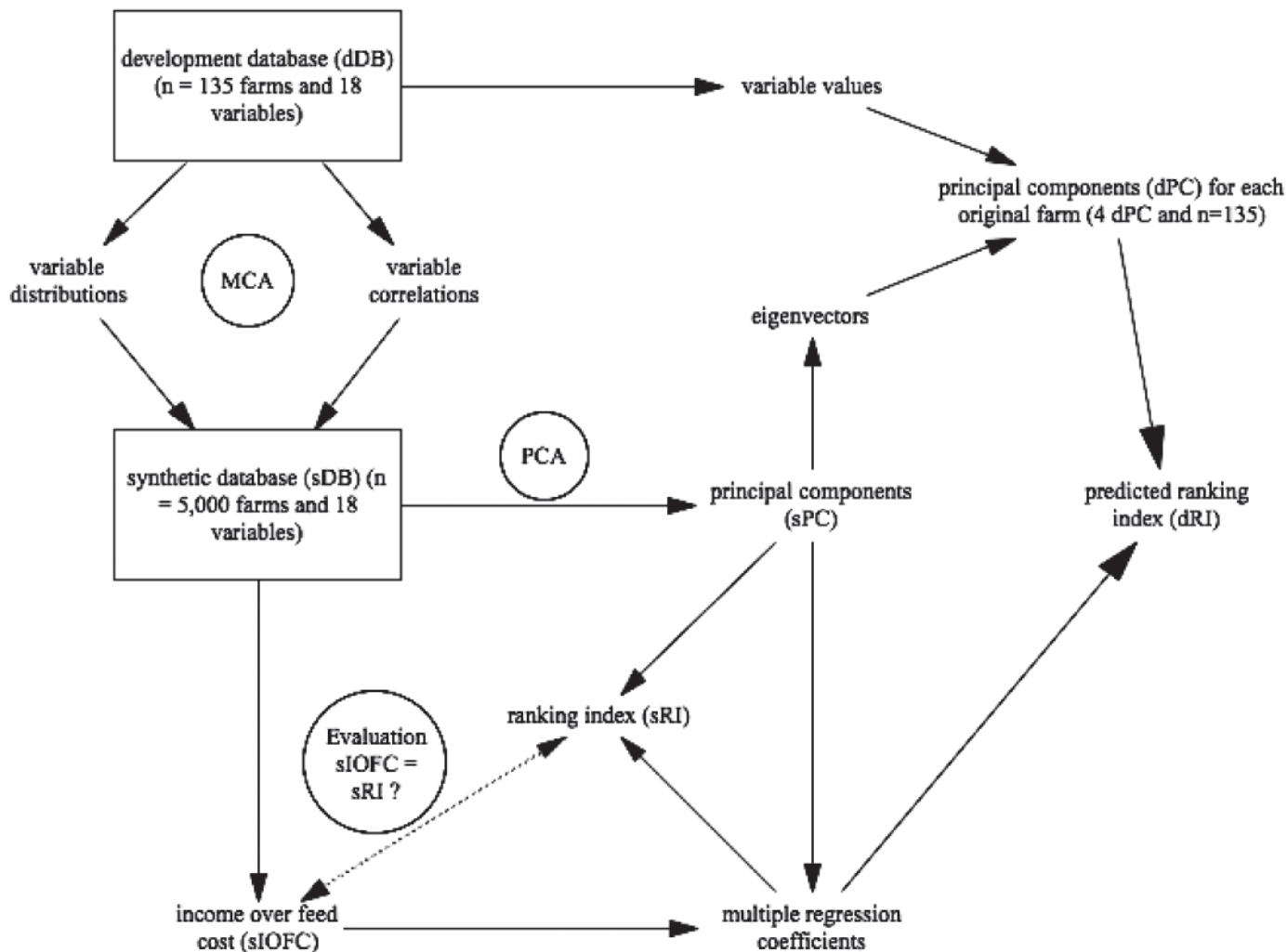


Figure 1. Schematic representation of the calculations. MCA = Monte Carlo analysis; PCA = principal component analysis; dotted line represents the evaluation process.

( $R^2$ ; Neter et al., 1996), accuracy (Lin, 1989), mean bias (MB; Cochran and Cox, 1957), and mean square error of prediction (MSEP; Bibby and Toutenburg, 1977) and its decomposition in 3 components (Theil, 1961).

**Analyses with the Development Database.** The eigenvectors of the sPC (Table 2) were used to estimate the PC scores of the dDB (dPC;  $n = 135$  for each selected PC). Similarly, the regression coefficients estimated with Equation [2] (i.e.,  $X_1, X_2, \dots, X_n$ ) were multiplied by the estimated dPC to obtain the ranking index for the 135 original farms (dRI) as shown in Equation [4]:

$$dRI = a + dPC1 X_1 + dPC2 X_2 + \dots + dPCn X_n, \quad [4]$$

where  $a, X_1, X_2,$  and  $X_n$  were defined in Equation [2]. Finally, each term of this equation was calculated for

each farm and reported as the percentage of the average dRI (i.e., the estimated value for  $a$ ).

## RESULTS

The studied dairy farms had on average 115 cows, ca. 95% of them being Italian Holsteins and the remaining Italian Brown cows. About 64% of the cows were multiparous and 36% were primiparous. On average, farms produced 1.05 calves/yr and approximately 8,500 L of milk/yr per cow, with 3.78% of fat and 3.35% of protein in the milk (Table 1). The equivalent mature cow production at 305 DIM was 10,294 L/yr per cow. The productive life of the cows was on average 2.7 lactations, considering the average calving interval reported by the Italian Farmer Association for the cows in the province of Oristano (AIA, 2011).

Four PC were selected by the PCA performed in the sDB. Based on their significance in accounting for the

**Table 2.** Eigenvectors of the principal component (PC) obtained with the PC analysis of the synthetic database (n = 5,000 synthetic dairy farms)<sup>1</sup>

Input variable	PC1 (herd profile)	PC2 (milk quality and payment)	PC3 (poor management)	PC4 (reproduction)
Primiparous	0.35	0.04	0.05	-0.04
Multiparous	0.37	0.03	0.15	-0.03
Dry cows	0.35	0.02	0.17	0.12
Replacement per year	0.36	0.03	0.13	-0.08
Calves	0.37	0.02	0.10	-0.03
Culled cows	0.35	-0.01	0.12	0.00
Milk sold	0.38	0.04	0.05	-0.01
Milk fat content	-0.09	0.41	0.14	-0.18
Milk protein content	-0.09	0.36	0.23	-0.10
Milk price	0.02	0.48	-0.16	0.11
Bonus fat	-0.09	0.40	0.15	-0.18
Bonus protein	-0.07	0.38	0.19	-0.06
Bonus other	0.08	0.32	-0.27	0.17
Potential production level	0.12	0.03	-0.40	0.22
Age at first calving	-0.11	-0.04	0.38	0.12
Age at culling	-0.11	-0.08	0.40	0.26
Bonus somatic cell count	0.10	0.22	-0.38	0.32
Days open	-0.07	0.06	0.26	0.79
Explained variability, %	37.4	20.2	15.6	4.4

<sup>1</sup>Values greater than 0.3 and smaller than -0.3 were used to associate the PC with a specific managerial and technical aspect (within parentheses below the PC) of the dairy farms.

total variation, they could be matched to 4 managerial and technical aspects of the farms: herd profile (**HP**), milk quality and payment (**MQP**), poor management (**PM**), and reproduction (**RE**) for sPC1, sPC2, sPC3, and sPC4, respectively. The total variance of the sDB was equal to the number of original standardized variables ( $\sigma^2 = 18$ ; which corresponds to 324,373,789,543 obtained from the variance - covariance matrix) and the selected PC explained 77.6% of the original variance, with an eigenvalue equal to 13.96 (Table 2). The HP was positively associated with the number of animals for each physiological stage and with the amount of milk sold. The MQP was positively associated with the milk fat and protein contents and their economic values. The PM was associated negatively with the potential production level of the herd and SCC, and positively with the age at first calving (i.e., presence of old cows), which is an inadequate management goal. The RE was positively associated with days open, which was the only reproductive variable available in the original database.

Equation [5] shows the coefficient estimates for the  $X_1$ ,  $X_2$ ,  $X_3$ , and  $X_4$  parameters that were subsequently used to compute sRI of Equation [3]. It also has the estimated average for the sIOFC ( $a = \text{€}0.1377 \pm 0.0162$ ).

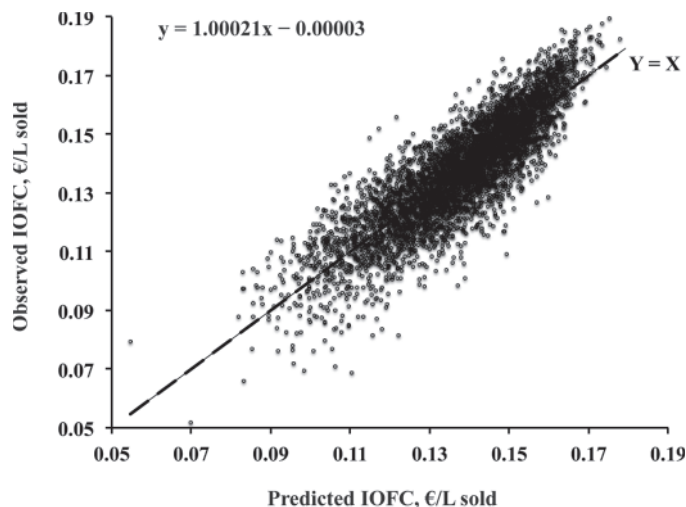
$$\begin{aligned} \text{sRI} &= 0.1377 + 0.00131 \times \text{sPC1} + 0.00542 \\ &\times \text{sPC2} - 0.00681 \times \text{sPC3} + 0.00410 \times \text{sPC4}; \\ &(\text{r}^2 = 0.745, n = 5,000, \text{MSE} = 0.0002), \end{aligned} \quad [5]$$

where sRI is ranking index predicted for the sDB; sPC1, sPC2, sPC3, and sPC4 are the selected PC that represent herd profile, milk quality and payment, poor management, and reproduction, respectively.

The sRI obtained with Equation [5] was very accurate (correlation bias = 0.99) and highly precise ( $R^2 = 0.74$ ) (Figure 2). It had a MB (€/L) of nearly zero ( $P = 0.99$ ) and the root of MSEP was -0.00034 €/L. The decomposition of the MSEP consisted of 0.3% for MB, 9.7% for systematic bias, and 90% for random errors. Therefore, the sRI was assumed to be an acceptable index (high precision and accuracy) to represent profitability (i.e., IOFC).

When Equation [5] was used to compute dRI of the original variables in the dDB (n = 135), 81% of the variability of the calculated IOFC for the dDB was explained. The dRI of the ranked farms ranged from +0.024 €/L to -0.082 €/L of the mean IOFC (Figure 3). Considering the number of farms below and above 1 SD of the dRI, 21 farms had dRI less than -1 SD, 32 farms were between -1 SD and 0, 67 farms were between 0 and +1 SD, and 15 farms had dRI greater than +1 SD. The top 10% of the farms had a dRI greater than 0.170 €/L, whereas the bottom 10% farms had a dRI lower than 0.116 €/L.

Figure 4 shows the individual contribution of the selected PC to the calculated dRI of each farm in the dDB (n = 135). The MQP and PM represented the most important managerial and technical aspects that contributed to the IOFC expressed as dRI (Figure 4); in particular, HP contributed with 2.5% (ranging from



**Figure 2.** Regression of the income over feed cost (sIOFC) on the ranking index (sRI) for the simulated database ( $n = 5,000$ ). Statistics of model adequacy were  $C_b = 0.99$ ;  $R^2 = 0.74$ ; mean bias (€/L) nearly zero ( $P = 0.99$ ); and the root of mean square error of prediction (MSEP) =  $-0.00034$  €/L. The decomposition of the mean square error of prediction consisted of 0.3% for mean bias, 9.7% for systematic bias, and 90% for random errors. The dotted line represents the linear regression and the solid line represents the  $Y = X$ .

$-3.5$  to  $14.3\%$ ) of the dRI, MQP contributed with  $7.5\%$  (ranging from  $-18.2$  to  $28.3\%$ ) of the dRI, PM contributed with  $8.4\%$  (ranging from  $-46.9$  to  $15.4\%$ ) of the dRI, and RE contributed with  $2.6\%$  (ranging from  $-10.2$  to  $7.6$ ) of the dRI.

## DISCUSSION

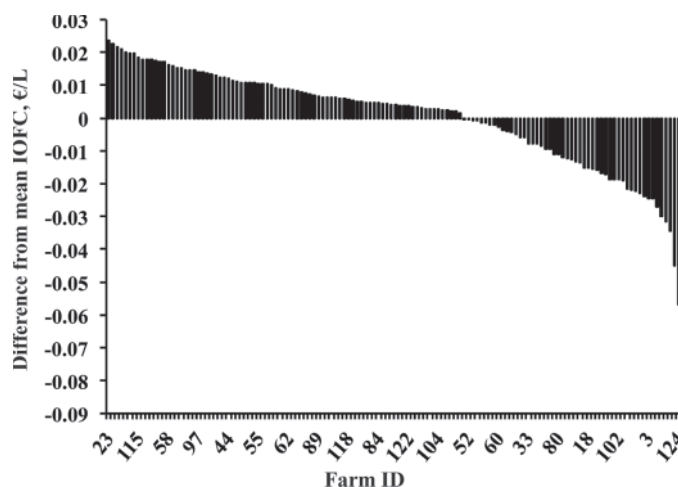
Our dDB was representative of the dairy farms from the province of Oristano, used as a case study, because it included more than  $75\%$  of the intensively managed dairy cows of the province (AIA, 2011). Although the original database (dDB) had enough information to compute the IOFC values for each farm, the calculation of the IOFC by using the selected PC allowed us to match them to different managerial and technical aspects of the dairy farms (i.e., HP, MQP, PM, and RE), which can be used to evaluate profitability. Furthermore, the selected principal components are not correlated among themselves because they are, by definition, orthogonally related to each other.

The interpretation of the PC (i.e., HP, MQP, PM, and RE) affecting IOFC is illustrated for 5 specific farms in Figure 5. It is interesting to note that the 2 farms (farms 1 and 2) that had positive average standardized profitability (dRI above the mean, line at zero value) had positive performances only in certain areas and negative performances in others, suggesting that they can target the negative components of their dRI

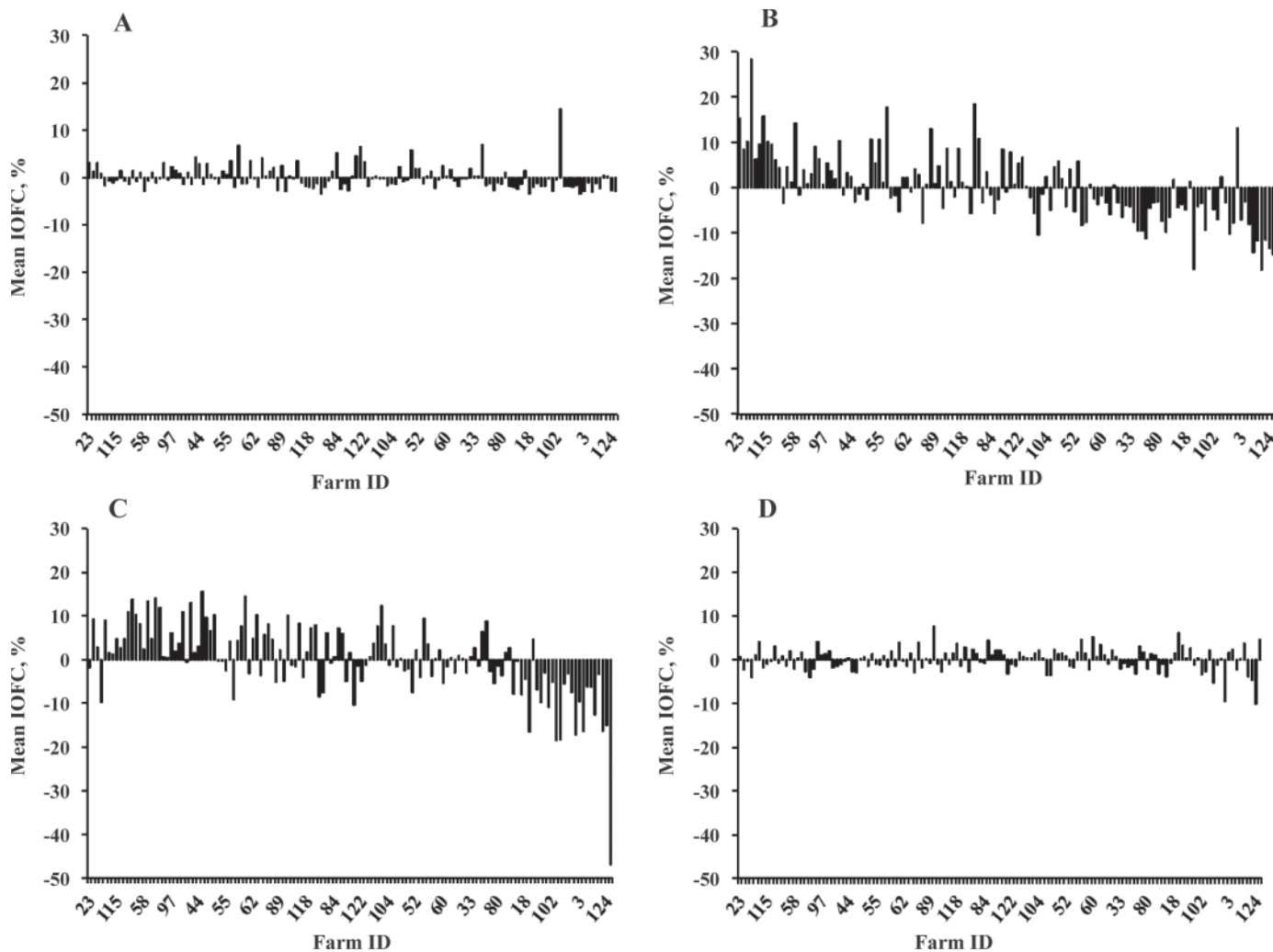
to improve their overall profitability (Figure 5). The contribution of the individual identified technical areas to profitability is presented as percentage of the overall mean ( $\text{€}0.1377/\text{L}$ ; Figures 4 and 5). In an alternative application of this method, it could be calculated as the percentage of the total dRI of a reference farm. In this case, the reference farm should be the same across years, so that the relative technical and managerial improvements would be consistent among the dairy farms of the region.

Although the principal component HP explained the greatest amount of variability in the database (PC1 in Table 2), it gave the smallest contribution in computing the dRI, as indicated by the regression coefficients of Equation [5] and as shown in Figure 4. This finding suggests that herd profile may have a small effect when measuring profitability per unit of product ( $\text{€}/\text{L}$  of milk). However, because the same feed cost per megacalories of net energy was applied to all farms and all animal physiological stages, the relative contribution of herd profile in explaining the variation in the profitability might have been reduced.

The PM (PC3) and the MQP (PC2) were the PC that showed the largest influence on the profitability (Equation [5]). The PM had the greatest effect on the profitability and was highly variable among the dairy farms (Figure 4). The production level and the quality of the milk of the MQP component had similar effects on the profitability; in fact, farms with high milk quality and low milk yield can achieve similar profitability of farms with high milk yield and low milk quality. This depends on the revenue compensation due to the milk payment method used in the studied area. When



**Figure 3.** Sorting of the ranking index (dRI) of the original dairy farms ( $n = 135$ ). The results are shown as the difference of each dairy farm from the mean income over feed cost (IOFC) observed in the synthetic database ( $0.138$  €/L).

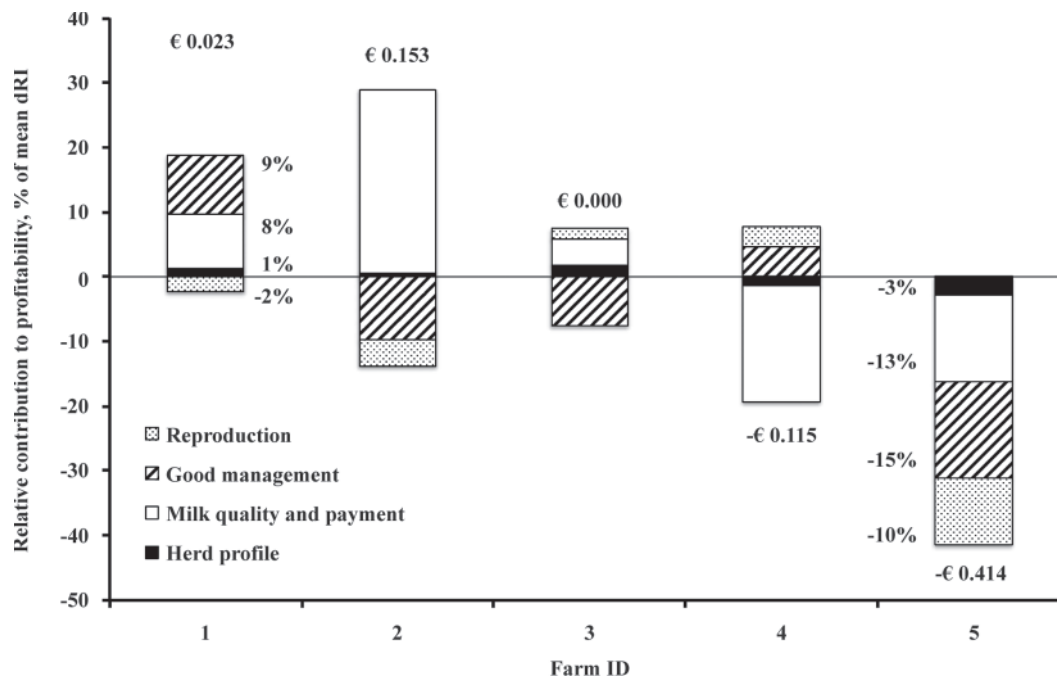


**Figure 4.** Contribution to the profitability (% of mean IOFC; 0.1377 €/L) of the 4 components associated with managerial and technical aspects of the farms in the development database: herd profile (A), milk quality and payment (B), “good” management (C; computed as  $-1 \times$  poor management estimates), and reproduction (D). IOFC = income over feed cost.

milk fat and protein increased, the milk price increased faster than the energy content of the milk. These results may explain the positive correlation between RE (i.e., days open) and IOFC per liter. Indeed, in previous calculations carried out to assess the effect of milk payment method for these farms, we found that cows with greater DIM, although reaching lower revenues per cow per day, produced greater IOFC per liter of milk compared with cows in early lactation [A. S. Atzori A. Cannas, and G. Pulina (Dipartimento di Agraria, Sezione di Scienze Zootecniche, Università di Sassari, Sassari, Italy), unpublished data]. This fact has also been reported by several papers evaluating cows with different milk yields and calving intervals (Arbel et al., 2001; Auldust et al., 2007).

The fourth PC (RE) was matched with days open and explained the least amount of the total variability

in the database (Table 2 and Figure 4). This is in contrast with papers that reported large variability among farms in reproductive performances of the cows (Caraviello et al., 2006; Olynk and Wolf, 2008). In our database, the only variable associated with reproduction efficiency was days open, which is not considered an exhaustive variable to describe the fertility status of a herd and its relative costs and benefits (LeBlanc, 2005). Nevertheless, in our study the PCA assigned a high effect of RE on profitability, agreeing with several papers that evaluated the economic importance of reproduction in dairy farms (Arbel et al., 2001; LeBlanc, 2005; De Vries, 2006). Probably, in our results, part of the variability due to RE was absorbed by the herd profile component that included milk produced and sold. Indeed, milk production is always strongly affected by RE. This fact is in agreement with Giordano



**Figure 5.** Relative contribution of selected principal components (bar patterns) to profitability in 5 dairy farms. Farms (1 to 5) are ranked for decreasing income over feed cost (IOFC). The bar patterns represent the relative contribution of the principal components to the mean dRI (profit, €/L). “Good” management was computed as  $-1 \times$  poor management estimates. dRI = ranking index from development database.

et al. (2012), who evaluated the economic benefits of various reproduction programs at farm level and observed that RE had a large effect on milk production, which in turn accounted for most of the economic benefits in programs with superior economic performance. In the same study, they also reported that in some cases replacement cost could greatly affect profitability, whereas calf value had the lowest effect; similar results were also obtained by Giordano et al. (2011).

The results obtained in our analysis depended on the fact that we considered only the IOFC as profitability indicator. In contrast, Wilson (2011) analyzed total costs of dairy farms and divided them into 4 classes based on performance quartiles of net margin (ranging from 4.68 to  $-6.96$  €/L). The increase in milk production was the most important and labor was the second most important factor influencing the variation of economic performance among farms. The same author reported small differences in gross margin when only variable costs were considered (ranging from 15.80 to 10.53 €/L). The assumption of equal feed cost per unit of requirements does not represent the actual situation of the studied area, but, unfortunately, actual feed costs were not available for each farm. This limitation, however, has a positive drawback because it allows highlighting differences in efficiency among farms when feed market prices and value of calves and culled cows are the same, which is often true for homogeneous production areas, such as that considered in the study. In

addition, it would have been beneficial to have a larger number of input variables to account for other factors (e.g., reproduction and sanitary cost, land, facilities, labor, field operation, energy, machinery, liabilities) to broaden the scope of our analysis regarding profitability and other areas of intervention. However, these factors were not available for all farms included in our database.

Our approach was aimed at finding a comprehensive indicator of profitability that could be obtained from variables available at the farm level and could provide practical guidelines to compare and improve the profitability of dairy farms. The lack of other variables associated to the production process (e.g., land, labor, reproduction) often limits the ability to estimate the absolute efficiency of the farms. Absolute efficiency is based on the minimization of the amounts of used inputs and the quantification of the slack of inputs per unit of output (Stokes et al., 2007). Absolute efficiency can be estimated with other techniques, such as the nonparametric data envelopment analysis (Aigner et al., 1977; Lawson et al., 2004; Stokes et al., 2007) or the parametric stochastic frontier analysis (Charnes et al., 1979; Fraser and Cordina, 1999), which are largely used to study technical and economic profitability and production efficiency (Tone, 2001; Theodoridis and Psychoudakis, 2008). These methods have different objectives from those of our study and, therefore, are not comparable with our approach.



Some studies adopted similar approaches to ours. For instance, Windig et al. (2005) used PCA and Enevoldsen et al. (1996) and Fahey et al. (2002) used factor analysis to identify key variables related with managerial aspects of dairy farms, but they did not connect the multivariate output with profitability or other indicators of farm efficiency. Others studies accounted for profitability, but are not comparable with our approach because they used a different statistical analysis from ours, such as the path analysis method used by Rougoor et al. (1997). On the other hand, the work of Kristensen et al. (2008) was very similar to ours, accounted for many input variables, and used a meta-model approach to simulate the effects of technical variables on the long-term financial performance of dairy farms. Kristensen et al. (2008) identified key performance indicators by aggregating farm-collected input variables based on empirical, personal experience and studied their effects on gross margin. Although our approach identified key variables using PCA, our results were consistent with those of Kristensen et al. (2008) in the sense that their key performance indicators were similar to our PC and the ranking of importance to explain profitability related to managerial and technical aspects of the dairy farms was also similar.

Our approach accounted for the variability of the input variables through PC. It allowed us to use farm level inputs to identify inefficient management areas and to assist in the identification of managerial goals for each farm to improve profitability. The stochastic approach helped us to understand the relationship between the input variables of the studied area and to develop a ranking index for broader application. From this point of view, a new farm in the same area that has the same 18 input variables used in our analyses could calculate its own ranking index by using the estimated PC scores reported in Table 2 and by multiplying them by the regression coefficients estimated with Eq. [3] and shown in Eq. [5]. In this way, this new dairy farm could be compared with the other 135 dairy farms of our database. Therefore, the results of our study could be applied to different farms within the study zone, whereas the proposed approach could be replicated in other zones, by following it step-by-step and, if possible, considering additional original variables. Additionally, it may be improved by using other inputs (e.g., reproduction and sanitary cost, land, facilities, labor, field operation, energy, machinery, liabilities) collected at the dairy farm level and considering the actual feed costs to measure farm profitability.

#### ACKNOWLEDGMENTS

The authors are grateful to Ana Helena Dias Francesconi, from the University of Sassari, for helping in

reviewing and editing the manuscript. The research was supported by data and information provided by the Cooperativa 3A Latte Arborea of Arborea (Sardinia, Italy), the Associazione Provinciale Allevatori di Oristano (Sardinia, Italy), and the J-Service Web Agency of Arborea (Sardinia, Italy).

#### REFERENCES

- AIA. 2011. Controlli della produttività del latte in Italia. Statistiche Ufficiali 2011 della Associazione Italiana Allevatori (Italian Farmer Association, annual report). Accessed Jun 3, 2012. <http://www.aia.it/aia-website/it/attivita/area-tecnica/ufficio-sviluppo/bollettino-dei-controlli-della-produttivita-del-latte>.
- Aigner, D. J., C. A. K. Lovell, and P. Schmidt. 1977. Formulation and estimation of stochastic frontier production function models. *J. Econom.* 6:21–37.
- Arbel, R., Y. Bigun, E. Ezra, H. Sturman, and D. Hojman. 2001. The effect of extended calving intervals in high-yielding cows on milk production and profitability. *J. Dairy Sci.* 84:600–608.
- Auld, M. J., G. O'Brien, D. Cole, K. L. Macmillan, and C. Grainger. 2007. Effects of varying lactation length on milk production capacity of cows in pasture-based dairying systems. *J. Dairy Sci.* 90:3234–3241.
- Bailey, K. W., C. M. Jones, and A. J. Heinrichs. 2005. Economic returns to Holstein and Jersey herds under multiple component pricing. *J. Dairy Sci.* 88:2269–2280.
- Bibby, J., and H. Toutenburg. 1977. Prediction and Improved Estimation in Linear Models. John Wiley and Sons, Berlin, Germany.
- Cabrera, V. E. 2010. A large Markovian linear program to optimize replacement policies and dairy herd net income for diets and nitrogen excretion. *J. Dairy Sci.* 93:394–406.
- Cabrera, V. E., D. Solís, and J. del Corral. 2010. Determinants of technical efficiency among dairy farms in Wisconsin. *J. Dairy Sci.* 93:387–393.
- Caraviello, D. Z., K. A. Weigel, P. M. Fricke, M. C. Wiltbank, M. J. Florent, N. B. Cook, K. V. Nordlund, N. R. Zwald, and C. L. Rawson. 2006. Survey of management practices on reproductive performance of dairy cattle on large US commercial farms. *J. Dairy Sci.* 89:4723–4735.
- Charnes, A., W. W. Cooper, and E. Rhodes. 1979. Measuring the efficiency of decision-making units. *Eur. J. Oper. Res.* 3:339.
- Cochran, W. G., and G. M. Cox. 1957. *Experimental Design*. Wiley, New York, NY.
- De Vries, A. 2006. Economic value of pregnancy in dairy cattle. *J. Dairy Sci.* 89:3876–3885.
- Dechow, C. D., E. A. Smith, and R. C. Goodling. 2011. The effect of management system on mortality and other welfare indicators in Pennsylvania dairy herds. *Anim. Welf.* 20:145–158.
- Ely, L. O., J. W. Smith, and G. H. Oleggini. 2003. Regional production differences. *J. Dairy Sci.* 86(E. Suppl.):E28–E34.
- Enevoldsen, C., J. Hindhede, and T. Kristensen. 1996. Dairy herd management types assessed from indicators of health, reproduction, replacement, and milk production. *J. Dairy Sci.* 79:1221–1236.
- Fahey, J., K. O'Sullivan, J. Crilly, and J. F. Mee. 2002. The effect of feeding and management practices on calving rate in dairy herds. *Anim. Reprod. Sci.* 74:133–150.
- Fodor, I. K. 2002. A survey of dimension reduction techniques. LLNL Technical Report, 2002. UCRL-ID-148494. US Department of Energy Office of Scientific and Technical Information, Oak Ridge, TN.
- Fraser, I., and D. Cordina. 1999. An application of data envelopment analysis to irrigated dairy farms in Northern Victoria, Australia. *Agric. Syst.* 59:267–282.
- Giordano, J. O., P. M. Fricke, M. C. Wiltbank, and V. E. Cabrera. 2011. An economic decision-making support system for selection of reproductive management programs on dairy farms. *J. Dairy Sci.* 94:6216–6232.

- Giordano, J. O., A. S. Kalantari, P. M. Fricke, M. C. Wiltbank, and V. E. Cabrera. 2012. A daily herd Markov-chain model to study the reproductive and economic impact of reproductive programs combining timed artificial insemination and estrus detection. *J. Dairy Sci.* 95:5442–5460.
- Huirne, R. B. M., S. B. Harsh, and A. A. Dijkhuizen. 1997. Critical success factors and information needs on dairy farms: The farmer's opinion. *Livest. Prod. Sci.* 48:229–238.
- Jackson, J. E. 1991. *A User's Guide to Principal Components*. John Wiley and Sons, New York, NY.
- Jolliffe, I. T. 1986. *Principal Component Analysis*. Springer-Verlag, New York, NY.
- Kristensen, E., S. Østergaard, M. A. Krogh, and C. Enevoldsen. 2008. Technical indicators of financial performance in the dairy herd. *J. Dairy Sci.* 91:620–631.
- Lawson, L. G., J. Bruun, T. Coelli, J. F. Agger, and M. Lund. 2004. Relationships of efficiency to reproductive disorders in Danish milk production: A stochastic frontier analysis. *J. Dairy Sci.* 87:212–224.
- LeBlanc, S. 2005. Using DHI records on-farm to evaluate reproductive performance. *Adv. Dairy Technol.* 17:319–330.
- Lin, L. I.-K. 1989. A concordance correlation coefficient to evaluate reproducibility. *Biometrics* 45:255–268.
- Macciotta, N. P. P., D. Vicario, and A. Cappio-Borlino. 2006. Use of multivariate analysis to extract latent variables related to level of production and lactation persistency in dairy cattle. *J. Dairy Sci.* 89:3188–3194.
- Neter, J., M. H. Kutner, C. J. Nachtsheim, and W. Wasserman. 1996. *Applied Linear Statistical Models*, 4th ed. McGraw-Hill Publishing Co., Boston, MA.
- NRC. 2001. *Nutrient Requirements of Dairy Cattle*. 7th rev. ed. Natl. Acad. Sci., Washington, DC.
- Oleggini, G. H., L. O. Ely, and J. W. Smith. 2001. Effect of region and herd size on dairy herd performance parameters. *J. Dairy Sci.* 84:1044–1050.
- Olynk, N. J., and C. A. Wolf. 2008. Economic analysis of reproductive management strategies on US commercial dairy farms. *J. Dairy Sci.* 91:4082–4091.
- Pearson, K. 1901. On lines and planes of closest fit to systems of points in space. *Philos. Mag.* 2:559–572.
- Rougoor, C. W., A. A. Dijkhuizen, R. B. M. Huirne, F. Mandersloot, and Y. H. Schukken. 1997. Relationships between technical, economic and environmental results on dairy farms: An explanatory study. *Livest. Prod. Sci.* 47:235–244.
- Solano, C., H. León, E. Pérez, L. Tole, R. H. Fawcett, and M. Herrero. 2006. Using farmer decision-making profiles and managerial capacity as predictors of farm management and performance in Costa Rican dairy farms. *Agric. Syst.* 88:395–428.
- Solís, D., B. E. Bravo-Ureta, and R. E. Quiroga. 2009. Technical efficiency among peasant farmers participating in natural resource management programmes in Central America. *J. Agric. Econ.* 60:202–219.
- St-Pierre, N. 2001. Economic factors affecting nutrient balance on dairy farms. *Adv. Dairy Technol.* 13:263–281.
- Stokes, J. R., P. R. Tozer, and J. Hyde. 2007. Identifying efficient dairy producers using data envelopment analysis. *J. Dairy Sci.* 90:2555–2562.
- Tauer, L. W. 2001. Efficiency and competitiveness of the small New York dairy farm. *J. Dairy Sci.* 84:2573–2576.
- Tauer, L. W., and A. K. Mishra. 2006. Dairy farm cost efficiency. *J. Dairy Sci.* 89:4937–4943.
- Tedeschi, L. O. 2006. Assessment of the adequacy of mathematical models. *Agric. Syst.* 89:225–247.
- Theil, H. 1961. Economic forecasts and policy. Pages 6–48 in *Contributions to Economic Analysis*. 2nd ed. R. Strotz, J. Tinbergen, P. J. Verdoorn, and H. J. Witteveen, ed. North-Holland, Amsterdam, the Netherlands.
- Theodoridis, A. M., and A. Psychoudakis. 2008. Efficiency measurement in Greek dairy farms: Stochastic frontier vs. data envelopment analysis. *Int. J. Econ. Sci. Appl. Res.* 1:53–67.
- Tone, K. 2001. A slacks-based measure of efficiency in data envelopment analysis. *Eur. J. Oper. Res.* 130:498–509.
- Tozer, P. R., F. Bargo, and L. D. Muller. 2003. Economic analyses of feeding systems combining pasture and total mixed ration. *J. Dairy Sci.* 86:808–818.
- Wilson, P. 2011. Decomposing variation in dairy profitability: The impact of output, inputs, prices, labour and management. *J. Agric. Sci.* 149:507–517.
- Windig, J. J., M. P. L. Calus, and R. F. Veerkamp. 2005. Influence of herd environment on health and fertility and their relationship with milk production. *J. Dairy Sci.* 88:335–347.