Agron. Sustain. Dev. 29 (2009) 339–352
© INRA, EDP Sciences, 2009
DOI: 10.1051/agro/2008065

Available online at:
www.agronomy-journal.org

Agronomy for Sustainable Development

Research article

Statistical modelling of nitrogen use efficiency of dairy farms in Flanders

Elena Ramírez*, Dirk Reheul

Department of Plant Production, Faculty of Bioscience Engineering, Ghent University, Coupure Links 653, 9000 Ghent, Belgium

(Accepted 10 November 2008)

Abstract – In the past decade it has repeatedly been shown that agriculture is a significant source of ground- and surface water pollution. Nitrogen losses and nitrogen use efficiency (NUE) are major concerns in agricultural practice and of policy-makers. Rapid intensification of livestock production, a result of the focus on increasing productivity from the 1950s onwards, has contributed to a large increase in nutrient surpluses. Here, we performed a quantitative analysis of the variables influencing the nitrogen use efficiency in Flemish grassland-based farming systems. The analysis was based on the large dataset of the Farm Accountancy Data Network, holding technical and economic data of Flemish farms. A statistical model is proposed by performing multiple regression with several variable selection procedures. Many combinations of variables were studied in 92 models and different criteria were taken into account to select the most adequate combination of variables. This approach focuses on a deep statistical analysis and interpretation of the model. The novelty of this research is the quantification and comparison of the influence of different inputs and other variables in nitrogen use efficiency at the farm level. Our results show that, contrary to current knowledge, a higher nitrogen use efficiency was observed for farms with a higher number of 'Dairy cows ha^{-1}', holding the rest of the N inputs constant. A higher stocking density is compatible with a higher agricultural sustainability. It is demonstrated that the amount of milk N produced by added cows is higher than the decrease in milk N produced by each single cow due to a reduced input of feed N per cow. The dairy cow becomes more efficient in the use of N, increasing the farm-gate NUE and the farm sustainability. In the dataset of this study, the variable 'Dairy cows ha^{-1}' is more relevant than suggested by previous studies: 1.4 times more relevant than the variable 'kg of N in fertilisers ha^{-1}', which is 1.5 times more relevant than 'kg of N in concentrates ha^{-1}'. According to previous knowledge, the N input variables present a negative sign. Decreasing the N input in fertilisers, concentrates and by-products are recommended actions to increase the NUE. Unexpected interaction effects were found.

sustainability / nitrogen use efficiency (NUE) / statistical model / dairy farming / dairy cows ha^{-1}

1. INTRODUCTION

In the past decade it has repeatedly been shown that agriculture is a significant source of ground- and surface water pollution (Heathwaite et al., 1996; Yadav et al., 1997; Carpenter et al., 1998; Hadas et al., 1999). In Flanders, agriculture is still a major contributor to nitrate contamination of groundwater (Van Gijseghem et al., 2002) and to acidifying emissions such as NO_x and NH_3, and as in other European regions, N losses and N use efficiency are major concerns in agricultural practice and of policy-makers (Nevens et al., 2006). An urgent problem related to nutrient pollution from the environment due to dairy farming is the increasing concentration of greenhouse gases in the atmosphere and the consequent global warming (Oenema et al., 2001).

The ecological quality of many surface waters is poor and nitrate concentration of groundwater exceeds 50 mg of nitrate per litre of groundwater in several areas. This is mainly due to relatively high discharges of nitrogen (N) and phosphate from agriculture (Oenema et al., 2005). The eutrophication of surface and marine waters, partly caused by agriculture, has also become a major concern (Carpenter et al., 1998; Rejesus and Hornbaker, 1999). Rapid intensification of livestock production, a result of the focus on increasing productivity from the 1950s onwards, has contributed to a large increase in nutrient surpluses (Oenema et al., 1998). The surplus expresses the potential loss from the system both in terms of volatilisation (e.g. in the process of handling manure in barns, during storage and in the fields) and in terms of denitrification and leaching from the soil (Borsting et al., 2003).

* Corresponding author: Elena.RamirezFanlo@UGent.be

To bring sustainability into practice many farms focus on improving efficiency wherever possible (Van Passel, 2007). Efficient use of resources, particularly of nutrients, is one of the major assets of sustainable agricultural production systems. Inefficient nutrient use not only results in excessive and potentially harmful losses to the environment, it also negatively affects the economic performance of production systems (Oenema and Pietrzak, 2002).

Efficiency is an important research topic because it plays an important role within the ecological aspects of grassland-based farming systems because it has a substantial impact both on the source and sink side. Considering the major share of 54% grassland and forages in the total utilised agricultural area in Flanders, working on enhanced sustainability on grasslands and forages is very effective in increasing the sustainability of the agricultural sector as a whole (Nevens, 2003).

Introducing nutrient balances on farms increases awareness of nutrient flows in the farming system and the information can serve as a guideline for improvements in nutrient management (Ondersteijn, 2002; Goodlass et al., 2003; Hanegraaf and den Boer, 2003; Swensson, 2003) and to study the efficiency of nitrogen use (NUE). However, the nutrient balance sheet does not tell what decisions a farmer should take to reduce surpluses and it does not quantify the effects of particular decisions (Kuipers and Mandersloot, 1999).

This research introduces statistical modelling of the effects of the variation in the nitrogen flows on the NUE on Flemish dairy farms. To our knowledge, there are very few multiple regression analyses reported to quantify the effects of managerial aspects on the NUE of farms. Several authors tried to model decision-making and many models have been created, but either they do not quantify the effect of each independent variable (they are not statistical models), they are not designed and validated to use in Flemish conditions or they do not focus on the general N surplus but on the nitrate leaching or volatilisation.

The statistical approach is chosen because it is appropriate when there are uncertainties surrounding the systems under study. Statistical modelling plays an important role in modern control practice, particularly in assisting in higher level decision-making, process monitoring, data analysis and in statistical process control. Statistically-based models can yield useful information in a relatively short time frame and in a cost-effective manner (Wells and cole, 2001).

Much agricultural scientific research, in which multiple regression is performed, does not comment on important issues such as, e.g., multicollinearity and interactions limiting the overall use of these studies. From 537 studies published in scientific journals about agriculture and biology since 1990, only 30 of them comment on multicollinearity; 240 of them comment on interactions; only 21 comment both on interactions and on multicollinearity, and only two of them centred the models in order to interpret the coefficients when interaction terms were present. The objective was to develop a regression equation including as many independent variables as possible to study their relevance and relative importance.

2. MATERIALS AND METHODS

2.1. Characterising the dataset

Data used in this research was stored in the Farm Accountancy Data Network (FADN), which holds technical and economic data of Flemish farms. From this dataset, 233 specialised dairy farms were selected. Specialised dairy farms were defined as farms where dairy produced at least 95% of the farm income. The farms were followed during the period 1989–2001. Not all farms were monitored during all years: some farms disappeared from the Network, others joined the Network during the monitoring period. In total, 1 298 observations were available. Since several farms were recorded in subsequent years, a random effect for 'Farm' was introduced in the models in order to avoid the negative impact of correlated observations. The selected farms did not buy forage maize and had no N input by fixation since they had neither leguminous crops on their farms nor white clover in their grassland. They did buy straw and by-products from the food industry. Characteristics of the selected dairy farms are presented in Nevens et al. (2006).

2.2. Conceptualising a farm-gate balance and calculating the NUE

The NUE use (%) was calculated as:

$$NUE = 100 \times \frac{N\ output}{N\ input}$$

Total N input is the sum of N in purchased concentrates, by-products, straw (or sawdust), animals, mineral fertiliser and manure, and in atmospheric deposition. Total N output is the total amount of N in exported milk, animals, manure and crops. All inputs and all outputs are expressed in kg N ha^{-1} of the total utilised farm area. A detailed description of the calculation of the balance is presented in Nevens et al. (2006).

2.3. Statistical analysis

The development of the statistical analysis of farm data was done in different steps.

First step: several multiple regression methods (ascendant and descendant substantive knowledge method, stepwise methods and best-subsets methods) were used in order to find as many candidate models as possible. In total, 92 models were developed, from which 51 were developed with the descendant substantive knowledge method. The final models were found with the descendant knowledge method. This method starts by including in the equation all the candidate independent variables and proceeds by eliminating step by step the independent variables that are not relevant enough. The criteria for eliminating variables are the researcher's knowledge based on previous literature and the partial correlation of the independent variables with the dependent variable NUE. An overview of the candidate independent variables is given in Table I.

Table I. Candidate independent variables for inclusion in the multiple regression including descriptive statistics.

Group	Candidate variables	Type of variable	Minimum	Maximum	Range	Mean	Std dev	Variance	Distribution
Group 1: Other inputs	‘kg of N in purchased straw ha^{-1}’	Metric, continuous	0	26.9	26.9	1.6	2.8	7.9	Other
	‘kg of N in purchased forage maize ha^{-1}’	Metric, continuous	0	150.7	150.7	2.9	12.9	168.1	Other
	‘kg of N from deposition ha^{-1}’	Metric, discrete	42.4	53.5	11.1	47.8	2.9	8.7	Other
Group 2: Fertilisers	‘kg N in fertilisers ha^{-1}’	Metric, continuous	0	543.8	543.8	194.6	76.4	5845	Normal
	‘kg of N in fertilisers per ha grassland’	Metric, continuous	0	604.7	604.7	262.1	101.3	10261	Normal
Group 3: Concentrates	‘kg of N in dairy cow concentrates dairy cow^{-1}’	Metric, continuous	0	109.3	109.2	44.2	16.7	277	Normal
	‘kg of N in dairy cow concentrates ha^{-1}’	Metric, continuous	0	282.5	282.5	77.3	35.7	1271	Other
	‘kg of N in dairy cow concentrates per ha grassland’	Metric, continuous	0	1519.0	1519.0	126.6	84.2	7087	Other
	‘kg of N in concentrates LU^{-1}’	Metric, continuous	0	75.4	75.4	29.7	9.7	94	Normal
	‘kg of N in concentrates ha^{-1}’	Metric, continuous	0	282.5	307.2	92.9	42.1	1770	Other
Group 4: By-products	‘kg of N in by-products ha^{-1}’	Metric, continuous	0	140.0	140.0	22.8	27.4	748	Other
	‘kg of N in by-products LU^{-1}’	Metric, continuous	0	36.4	36.4	6.5	7.0	49	Other
Group 5: % dairy cows	‘% dairy cows’	Metric, continuous	32.1	99.9	67.8	66.8	10.3	105	Normal
Group 6: Stocking density	‘LU ha^{-1}’	Metric, continuous	1.24	8.06	6.82	3.07	0.77	0.59	Normal
	‘LU per ha grassland’	Metric, continuous	1.50	27.76	26.26	4.94	1.95	3.81	Other
	‘LU dairy cows per ha grassland’	Metric, continuous	0.84	22.05	21.21	3.29	1.41	2.00	Other
	‘LU dairy cows ha^{-1}’	Metric, continuous	0.66	5.15	4.49	2.04	0.55	0.30	Normal
	‘ha LU^{-1}’	Metric, continuous	0.12	0.80	0.68	0.34	0.08	0.01	Normal
	‘ha dairy cow^{-1}’	Metric, continuous	0.25	1.64	1.39	0.61	0.15	0.02	Normal
	‘ha per LU dairy cows’	Metric, continuous	0.19	1.52	1.33	0.53	0.15	0.02	Other
Group 7: Land distribution	‘% grassland’	Metric, continuous	17.1	100.0	82.9	65.4	13.5	183.2	Normal
	‘% non-grass forages’	Metric, continuous	0	82.9	82.9	33.4	13.2	174.1	Normal
	‘% arable crops’	Metric, continuous	0	10.0	10.0	0.9	2.3	5.1	Other
Group 8: Dummy variables	‘Dummy variable % arable crops’	Metric, discrete							

Second step: the verifications of the assumptions of multiple regression were conducted:

1. The relationship between the independent variable and the dependent variables is linear, at least approximately;
2. The error term e has zero mean and constant variance;
3. The errors are uncorrelated and normally distributed.

Gross violations of the assumptions may yield an unstable model, in the sense that a different sample could lead to a totally different model with opposite conclusions (Montgomery et al., 2001). Third step: the dummy variable '% of arable crops' was tested in order to check if it was necessary to split the dataset of 1300 farms into two datasets: one dataset comprising farms without arable crops and another dataset comprising farms with arable crops. Fourth step: the degree of multicollinearity (correlation among the independent variables, which is not desirable) was checked in order to assure a correct interpretation of the regression coefficients. Fifth step: interaction effects were checked. To interpret the interaction terms, the models were centred. When a model is centred, the degree of multicollinearity decreases and the regression coefficients are easier to interpret.

To centre a model, the mean is subtracted from each observation. The regression coefficients are different in the centred and in the uncentred model. In the presence of interactions, the coefficients of an uncentred model represent the variation in the dependent variable when a particular independent variable varies, keeping the others constant at zero value. This has little sense because the independent variable 'Dairy cows ha^{-1}' cannot equal zero in a dairy farm. The coefficients of a centred model represent the variation in the dependent variable when a particular independent variable varies and the rest of the independent variables are kept constant at their average value. This situation is more common in reality. Therefore, it makes more sense to interpret the regression coefficients of the centred model.

Two final models were retained by eliminating the models which did not fulfil the assumptions of multiple regression. Finally, the influential observations were studied, and the models were validated and interpreted.

Two different statistical software packages were used: SPSS 16.0 and S-PLUS 6.1. SPSS gives the most complete information about the multicollinearity diagnostics (tolerance values, variance inflation factors (VIF), condition indices and the proportion of the variance). S-PLUS is the most adequate program to test the functional form of the relationships between the dependent and independent variables. The stepwise methods are most practically performed with this software. The program also allows a fast interaction analysis.

3. RESULTS AND DISCUSSION

The aim of this research was to quantify and to compare the influence of the different inputs and other variables on nitrogen use efficiency (NUE), specifically the magnitude and sign of the regression coefficients, at the farm level. Many combinations of variables were studied in 92 multiple regression models and different criteria were taken into account to select the most adequate combination of variables.

3.1. Final models

The dummy variable '% of arable crops' was significant, which means that applying the same regression equation to a group of farms with arable crops and to a group of farms without arable crops, the difference in the calculated average NUE is significant for both groups. As a consequence, two different regression models were selected: the dataset of Model 1 includes farms without arable crops. The dataset of Model 2 includes farms with arable crops. Both models present interaction terms and both models fulfil the assumptions of multiple regression.

The final uncentred models had the following equations:
Model 1 (uncentred)

$$\begin{aligned}
&NUE = 9.379 - 0.056\, kg\, of\, N\, in\, fertilisers\, ha^{-1}\\
&- 0.052\, kg\, of\, N\, in\, concentrates\, ha^{-1}\\
&- 0.086\, kg\, of\, N\, in\, by{-}products\, ha^{-1}\\
&+ 15.688\, dairy\, cows\, ha^{-1}\\
&- 0.058\, kg\, of\, N\, in\, by-products\, ha^{-1} \times dairy\, cows\, ha^{-1}\\
&+ 0.00029\, kg\, of\, N\, in\, fertilisers\, ha^{-1}\\
&\times kg\, of\, N\, in\, concentrates\, ha^{-1}\\
&+ 0.00063\, kg\, of\, N\, in\, fertilisers\, ha^{-1}\\
&\times kg\, of\, N\, in\, by-products\, ha^{-1}\\
&+ 0.00032\, kg\, of\, N\, in\, concentrates\, ha^{-1}\\
&\times kg\, of\, N\, in\, by-products\, ha^{-1}\\
&- 0.013\, kg\, of\, N\, in\, fertilisers\, ha^{-1} \times dairy\, cows\, ha^{-1}\\
&- 0.028\, kg\, of\, N\, in\, concentrates\, ha^{-1} \times dairy\, cows\, ha^{-1}\\
&- 0.0000031\, kg\, of\, N\, in\, fertilisers\, ha^{-1}\\
&\times kg\, of\, N\, in\, concentrates\, ha^{-1}\\
&\times kg\, of\, N\, in\, by-products\, ha^{-1}\\
&+ 0.00027\, kg\, of\, N\, in\, concentrates\, ha^{-1}\\
&\times kg\, of\, N\, in\, by-products\, ha^{-1} \times dairy\, cows\, ha^{-1} \qquad (1)
\end{aligned}$$

Model 2 (uncentred)

$$\begin{aligned}
&NUE = 12.628 - 0.065\, kg\, of\, N\, in\, fertilisers\, ha^{-1}\\
&- 0.061\, kg\, of\, N\, in\, concentrates\, ha^{-1}\\
&+ 0.116\, kg\, of\, N\, in\, by{-}products\, ha^{-1}\\
&+ 13.586\, dairy\, cows\, ha^{-1} - 0.084\, kg\, of\, N\, in\, by\\
&- products\, ha^{-1} \times dairy\, cows\, ha^{-1} \qquad (2)
\end{aligned}$$

Tables II, III and IV show a resume of the main characteristics of both models (uncentred and centred).

3.2. Comparison of the two models

Model 1 represents farms without arable crops. Model 2 represents farms with arable crops.

Table II. Main characteristics of Model 1 (uncentred). Std. Dev: Standard deviation of the independent variables; R^2: Coefficient of determination; Std dev: Standard deviation; B: unstandardised regression coefficient; * Significance for $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; Std Err: Standard error of the unstandardised regression coefficient; Beta: Standardised regression coefficients; Tol: Tolerance; VIF: Variance Inflation Factor; Cond Index: Condition Index.

Model			Dataset		Characteristics of the independent variables								
Number	Dataset	Independent variables	Mean	Std Dev	R^2	Std dev		Unstd coef		Std Coef	Tol	VIF	Cond
						Intercept	residual	B	Std Err	Beta			Index
1 uncentred	1078 observations without arable crops				0.772	1.47	2.01						309
		Constant						9.379***	1.460				
		kg of N in fertilisers ha^{-1}	206.87	70.98				–0.056***	0.006	–0.776	0.009	115	
		kg of N in concentrates ha^{-1}	92.46	37.27				–0.052***	0.014	–0.382	0.008	121	
		kg of N in by-products ha^{-1}	19.03	21.53				–0.086*	0.040	–0.365	0.005	182	
		Dairy cows ha^{-1}	2.05	0.50				15.688***	0.839	1.533	0.013	76	
		kg of N in by-products ha^{-1} × Dairy cows ha^{-1}						–0.058***	0.014	–0.706	0.008	131	
		kg of N in fertilisers ha^{-1} × kg of N in concentrates ha^{-1}						2.94E-4***	4.3E-5	0.702	0.003	308	
		kg of N in fertilisers ha^{-1} × kg of N in by-products ha^{-1}						6.32E-4***	1.34E-4	0.645	0.008	129	
		kg of N in concentrates ha^{-1} × kg of N in by-products ha^{-1}						3.24E-4	3.07-4	0.210	0.003	290	
		kg of N in fertilisers ha^{-1} × Dairy cows ha^{-1}						–0.013***	0.003	–0.538	0.004	282	
		kg of N in concentrates ha^{-1} × Dairy cows ha^{-1}						–0.028***	0.005	–0.690	0.003	320	
		kg of N in fertilisers ha^{-1} × kg of N in concentrates ha^{-1} × kg of N in by-products ha^{-1}						–3.09E-6**	9.5E-7	–0.509	0.006	181	
		kg of N in concentrates ha^{-1} × kg of N in by-products ha^{-1} × Dairy cows ha^{-1}						2.69E-4**	8.9E-5	0.558	0.006	166	

Table III. Main characteristics of Model 1 (centred). Std. Dev: Standard deviation of the independent variables; R^2: Coefficient of determination; Std dev: Standard deviation; B: unstandardised regression coefficient; * Significance for $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; Std Err: Standard error of the unstandardised regression coefficient; Beta: Standardised regression coefficients; Tol: Tolerance; VIF: Variance Inflation Factor; Cond Index: Condition Index.

Model			Dataset		Characteristics of the independent variables								
Number	Dataset	Independent variables	Mean	Std Dev	R^2	Std dev		Unstandardised coefficients		Std Coef	Tol	VIF	Cond Index
						Intercept	residual	B	Std Err	Beta			
1 centred	1078 observations without arable crops				0.772	1.47	2.01						7
		Constant						19.047***	0.146				
		kg of N in fertilisers ha^{-1}	0	70.98				–0.049***	0.001	–0.682	0.770	1.298	
		kg of N in concentrates ha^{-1}	0	37.27				–0.044***	0.004	–0.319	0.456	2.195	
		kg of N in by-products ha^{-1}	0	21.53				–0.053***	0.006	–0.223	0.398	2.516	
		Dairy cows ha^{-1}	0	0.50				9.790***	0.282	0.957	0.443	2.248	
		kg of N in by-products ha^{-1} × Dairy cows ha^{-1}						–0.033***	0.008	–0.402	0.248	4.027	
		kg of N in fertilisers ha^{-1} × kg of N in concentrates ha^{-1}						2.3E-4***	3.59E-5	0.549	0.409	2.445	
		kg of N in fertilisers ha^{-1} × kg of N in by-products ha^{-1}						3.5E-4***	6.72E-5	0.357	0.398	2.510	
		kg of N in concentrates ha^{-1} × kg of N in by-products ha^{-1}						2.4E-4	1.32E-4	0.156	0.194	5.165	
		kg of N in fertilisers ha^{-1} × Dairy cows ha^{-1}						–0.013***	0.003	–0.538	0.430	2.326	
		kg of N in concentrates ha^{-1} × Dairy cows ha^{-1}						–0.023**	0.004	–0.562	0.322	3.104	
		kg of N in fertilisers ha^{-1} × kg of N in concentrates ha^{-1} x kg of N in by-products ha^{-1}						–3.09E-6***	9.46E-7	–0.509	0.246	4.068	
		kg of N in concentrates ha^{-1} × kg of N in by-products ha^{-1} × Dairy cows ha^{-1}						2.69E-4**	8.93E-5	0.558	0.171	5.834	

Table IV. Main characteristics of Model 2 (uncentred and centred). Std. Dev: Standard deviation of the independent variables; R^2: Coefficient of determination; Std dev: Standard deviation; B: unstandardised regression coefficient; * Significance for $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; Std Err: Standard error of the unstandardised regression coefficient; Beta: Standardised regression coefficients; Tol: Tolerance; VIF: Variance Inflation Factor; Cond Index: Condition Index.

Model Number	Dataset	Independent variables	Dataset Mean	Dataset Std Dev	R^2	Std dev intercept	Std dev residual	Unstandardised coefficients B	Unstandardised coefficients Std Err	Std Coef Beta	Tol	VIF	Cond Index
2 uncentred	220 observations with arable crops				0.761	2.01	1.92						21
		Constant						12.628***	1.258				
		kg of N in fertilisers ha^{-1}	163.95	48.16				–0.065***	0.004	–0.595	0.933	1.072	
		kg of N in concentrates ha^{-1}	76.62	37.59				–0.061***	0.008	–0.434	0.546	1.830	
		kg of N in by-products ha^{-1}	9.95	13.71				0.116*	0.065	0.300	0.042	23.573	
		Dairy cows ha^{-1}	1.75	0.40				13.586***	0.798	1.040	0.445	2.246	
		kg of N in by-products ha^{-1} × Dairy cows ha^{-1}						–0.084*	0.029	–0.516	0.038	26.346	
2 centred	220 observations with arable crops				0.761	2.01	1.92						3
		Constant						20.738***	0.291				
		kg of N in fertilisers ha^{-1}	0	48.16				–0.065***	0.004	–0.594	0.933	1.072	
		kg of N in concentrates ha^{-1}	0	37.59				–0.061***	0.008	–0.434	0.546	1.830	
		kg of N in by-products ha^{-1}	0	13.71				–0.031*	0.020	–0.081	0.429	2.331	
		Dairy cows ha^{-1}	0	0.40				12.752***	0.744	0.976	0.510	1.963	
		kg of N in by-products ha^{-1} × Dairy cows ha^{-1}						–0.084*	0.029	–0.134	0.560	1.787	

The main statistical difference between these two models is the larger number of interaction terms of Model 1. The degree of multicollinearity of the uncentred Model 1 is very high (Tab. II). The uncentred Model 2 has a lower degree of multicollinearity but it is still high (Tab. IV). The interpretation of the independent variables is more meaningful in both centred models.

The two centred models present an acceptably low degree of multicollinearity (tolerance values are higher than the accepted cut-off value of 0.100 (Hair et al., 1998), variance inflation factor values are lower than the accepted cut-off value of 10 (Hair et al., 1998; Von Eye and Schuster, 1998) and condition indices are not higher than the recommended cut-off values of 15 or 30 (Hair et al., 1998), or 100 (Montgomery et al., 2001)) (Tabs. III, IV).

The statistical power achieved in both models is very similar: Model 1 has a R^2 of 0.772 and Model 2 has a R^2 of 0.761. The standard deviation of the intercept and of the residual are also similar in both models (see Tabs. III, IV). Both models have the same independent variables except the interaction terms.

3.3. Interpretation of the regression coefficients of centred Model 1

The interpretation of Model 2 is not shown but very similar.

3.3.1. Unstandardised coefficients of centred Model 1

The unstandardised regression coefficients of the centred model show how much a variation in one unit of an independent variable influences the dependent variable NUE when the other independent variables are around their mean value.

As indicated in Table III, a change of 1 'kg of N in fertilisers ha^{-1}', 1 'kg of N in by-products ha^{-1}' and 1 'kg of N in concentrates ha^{-1}' causes an average change of approximately 0.05% NUE in the opposite direction; a higher NUE was observed for farms with a higher number of LU per ha when the other independent variables are at their mean. The positive sign of this independent variable is unexpected according to previous literature studies which suggested a negative correlation (e.g. Nevens et al., 2006). However, this result was based on simple regression or observations and straight comparisons among farms. This positive sign is not due to multicollinearity, because the centred model does not suffer from multicollinearity, as mentioned in Section 2.3, but due to the inclusion of the relevant variables in the regression equation. The causes, explanation and demonstration of this positive sign are explained in Section 3.4.

3.3.2. Standardised (Beta) coefficients of centred Model 1

The value of a beta coefficient expresses how many standard deviations the dependent variable NUE changes on average when varying an independent variable with one standard deviation. They show the relative importance of each independent variable in the model, when the other independent variables are at their mean (Tab. III). The beta coefficients are influenced by the different standard deviations of each independent variable. For variables measured in the same units and with very similar unstandardised coefficients ('kg of N in fertilisers ha^{-1}', 'kg of N in concentrates ha^{-1}' and 'kg of N in by-products ha^{-1}'), the ones which have a higher standard deviation will have a higher beta coefficient. The beta coefficients of the centred Model 1 show that, contrary to current knowledge, the most relevant variable is 'Dairy cows ha^{-1}': it has a beta coefficient of 0.957 (Tab. III). So this variable has the highest priority if one wants to improve the NUE. 'kg of N in fertilisers ha^{-1}' has the second largest beta coefficient, –0.682. The interpretation is that an increase in one standard deviation of 'kg of N in fertilisers ha^{-1}' (70.97 units) will cause an average decrease of 0.682 standard deviations of the NUE. 'kg of N in concentrates ha^{-1}' (beta coefficient –0.319) is about half as relevant as 'kg of N in fertilisers ha^{-1}' and the relevance of 'kg of N in by-products ha^{-1}' (beta coefficient –0.223) is about one-third as relevant as 'kg of N in fertilisers ha^{-1}'.

3.3.3. Interactions and simple slopes of Model 1

The simple slopes of an independent variable indicate how much the dependent variable will change on average when varying by one unit that independent variable and keeping the other independent variables constant at certain levels: low, mean and high. If the simple slopes change during these variations there are interaction effects.

3.3.3.1. Simple slopes of the variable 'kg of N in fertilisers ha^{-1}'

All the simple slopes of the variable 'kg of N in fertilisers ha^{-1}' are significant and negative, independently of the level of the other variables (Figs. 1–3). The NUE decreases as the 'kg of N in fertilisers ha^{-1}' increases.

The conclusion is that farms with low levels of 'kg of N in concentrates ha^{-1}' and 'kg of N in by-products ha^{-1}' and high levels of 'Dairy cows ha^{-1}' will experience a relatively higher average increase in the NUE (0.073) when decreasing the 'kg of N in fertilisers ha^{-1}' than the farms with high levels of 'kg of N in concentrates ha^{-1}' and 'kg of N in by-products ha^{-1}' and low levels of 'Dairy cows ha^{-1}'.

3.3.3.2. Simple slopes of the variable 'kg of N in concentrates ha^{-1}'

All the simple slopes of the variable 'kg of N in concentrates ha^{-1}' are significant and negative, independently of the level of the other variables. The NUE decreases as the 'kg of N in concentrates ha^{-1}' increases. Values were very similar to those shown in Figures 1–3 (data not shown). The conclusion is that farms with low levels of 'kg of N in fertilisers ha^{-1}' and 'kg of N in by-products ha^{-1}' and high levels of 'Dairy

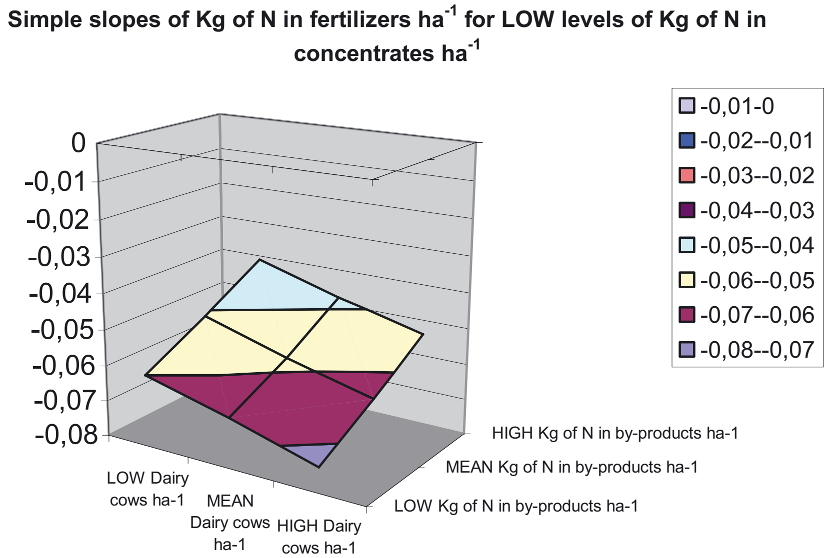

Figure 1. Simple slopes of the variable 'kg of N in fertilisers ha^{-1}' for LOW levels of 'kg of N in concentrates ha^{-1}' and different levels of the other variables.

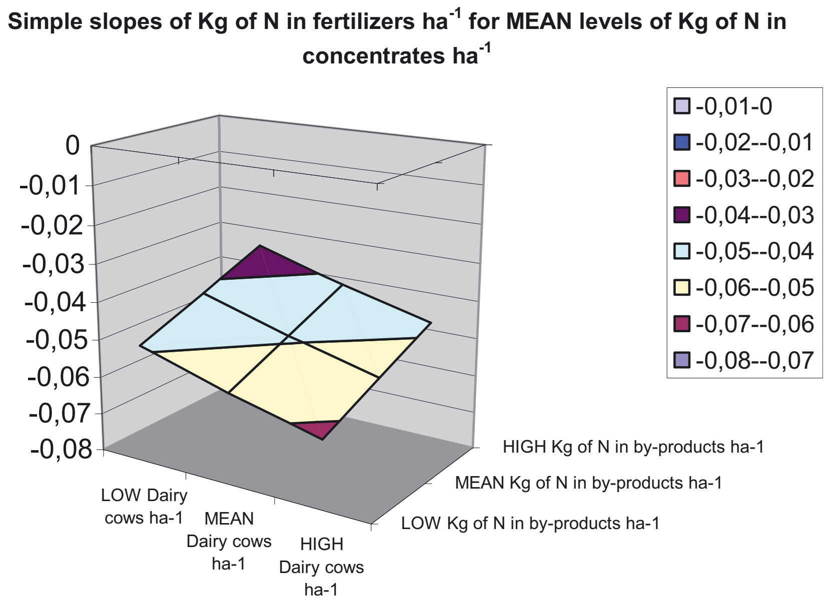

Figure 2. Simple slopes of the variable 'kg of N in fertilisers ha^{-1}' for MEAN levels of 'kg of N in concentrates ha^{-1}' and different levels of the other variables.

cows ha^{-1}' will experience a relatively higher increase in the NUE when decreasing 'kg of N in concentrates ha^{-1}' than farms with high levels of 'kg of N in fertilisers ha^{-1}' and 'kg of N in by-products ha^{-1}' and low levels of 'Dairy cows ha^{-1}'.

3.3.3.3. Simple slopes of the variable 'kg of N in by-products ha^{-1}'

All the simple slopes of the variable 'kg of N in by-products ha^{-1}' are significant and negative, independently of the level of the other variables. The average NUE decreases as the 'kg of N in by-products ha^{-1}' increases. Values were very similar to those shown in Figures 1–3 (data not shown). The conclusion is that farms with low levels of 'kg of N in fertilisers ha^{-1}' and 'kg of N in concentrates ha^{-1}' and high levels of 'Dairy cows ha^{-1}' will experience a relatively higher increase in the NUE when decreasing 'kg of N in by-products ha^{-1}' than farms with high levels of 'kg of N in fertilisers ha^{-1}' and 'kg of N in concentrates ha^{-1}' and low levels of 'Dairy cows ha^{-1}'.

3.3.3.4. Simple slopes of the variable 'Dairy cows ha^{-1}'

All the simple slopes of the variable 'Dairy cows ha^{-1}' are positive, independently of the level of the other variables

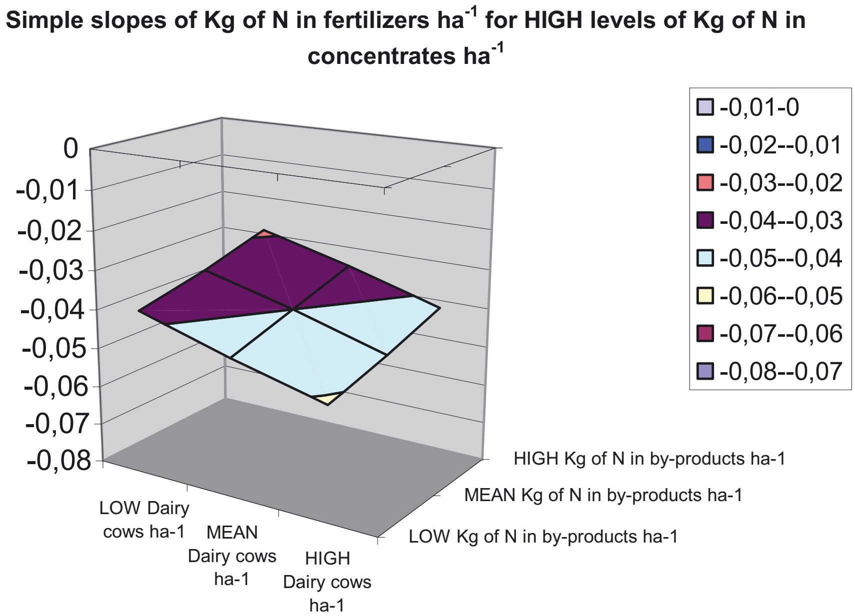

Figure 3. Simple slopes of the variable 'kg of N in fertilisers ha^{-1}' for HIGH levels of 'kg of N in concentrates ha^{-1}' and different levels of the other variables.

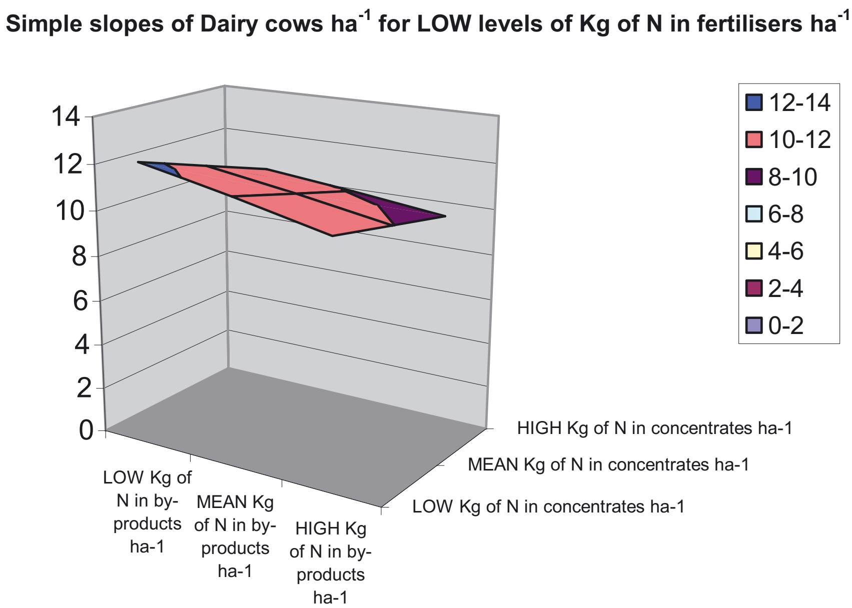

Figure 4. Simple slopes of the variable 'Dairy cows ha^{-1}' for LOW levels of 'kg of N in fertilisers ha^{-1}' and different levels of the other variables.

(Figs. 4–6). A higher NUE was observed for farms with a higher number of 'Dairy cows ha^{-1}' holding the N inputs constant.

The conclusion is that farms with low levels of 'kg of N in fertilisers ha^{-1}', 'kg of N in concentrates ha^{-1}' and 'kg of N in by-products ha^{-1}' will experience a relatively higher increase in the NUE when increasing 'Dairy cows ha^{-1}' than farms with high levels of 'kg of N in fertilisers ha^{-1}', 'kg of N in concentrates ha^{-1}' and 'kg of N in by-products ha^{-1}'. The results are similar in Model 2 (data not shown).

3.3.4. Percentage coefficients of Model 1

Percentage coefficients (Tab. V) may be more useful in practice than the beta coefficients since they are intuitively easier to understand because they relate changes in percentages

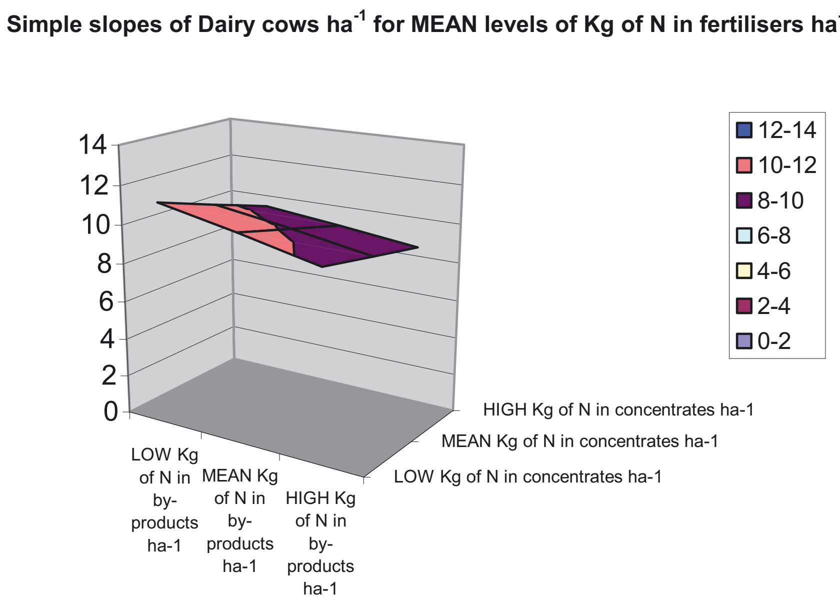

Figure 5. Simple slopes of the variable 'Dairy cows ha^{-1}' for MEAN levels of 'kg of N in fertilisers ha^{-1}' and different levels of the other variables.

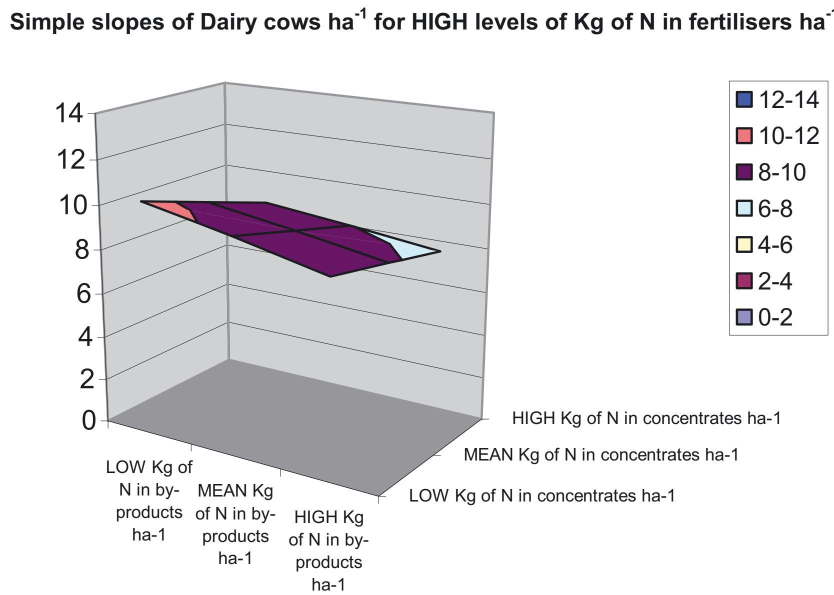

Figure 6. Simple slopes of the variable 'Dairy cows ha^{-1}' for HIGH levels of 'kg of N in fertilisers ha^{-1}' and different levels of the other variables.

between the independent and the dependent variables and not standard deviations like the beta coefficients.

So, as an example, a 10% change in the variable (1) will cause an average variation of 6.1% in the NUE.

3.4. Interpretation of the positive sign of the independent variable 'Dairy cows ha^{-1}'

As known from the literature, the more intensive the management of a dairy farm, the lower the NUE might be. The literature usually does not mention clearly if the variable 'Dairy cows ha^{-1}' is considered in a simple regression or in a multiple regression. Because an increase in 'Dairy cows ha^{-1}' is usually considered as an element of the high intensity of the management, it was surprising to find a positive correlation between the number of dairy cows ha^{-1} and the NUE. In the next paragraph we explain the veracity of this statistical result.

Potential causes of an unexpected sign are: (1) the dataset is too small; (2) the variance of the independent variables is too small; (3) the degree of multicollinearity is high; (4) computational errors have been made, and (5) important independent

Table V. Percentage coefficients of Model 1 indicating the average percentage change in the dependent variable by changing an independent variable by 1%.

Independent variables			
kg of N in fertilisers ha^{-1}	kg of N in concentrates ha^{-1}	kg of N in by-products ha^{-1}	Dairy cows ha^{-1}
(1)	(2)	(3)	(4)
0.628	0.200	0.048	1.080

variables have been included in the model. The cause of the unexpected sign is the inclusion of important independent variables in the model (data not shown).

The unexpected positive correlation of 'Dairy cows ha^{-1}' with the dependent variable NUE means that the higher (up to a certain limit) the stocking density of dairy cows ha^{-1}, the higher the NUE will be, provided that N feed and fertiliser inputs in the farm and cows are kept at a constant level. If the feed has to be divided over a higher number of cows, one expects a lower milk production and a lower N output in the milk per cow. Here, we prove that the N output in the milk of the extra cows is higher than the sum of the decreases in the milk of the cows of the original herd. In other words, the herd becomes more efficient in N use.

Kebreab et al. (2001) presented an equation for the relationship between the daily input of N per cow and the N output in the milk per cow (Eq. (3)). The research was performed with 50 cows which were fed amounts of N from 0.300 to 0.600 kg cow^{-1} day^{-1}, in different combinations of concentrates and forages.

The equation is:

$$g\ of\ milk\ N/cow\ day = 0.19 \times g\ of\ N\ intake/cow\ day + 38.2 \quad (3)$$

This is a linear relationship and not a curve. This interval of feed nitrogen (300–600 g) is just a part of the whole curve that ranges from 0 to 1000. The curve has been linearised for this interval, which is the interval representing the feed level of many dairy herds.

The cow efficiency is:

$$Dairy\ cow\ Efficiency = \frac{g\ of\ milk\ N/cow\ day}{g\ of\ N\ intake/cow\ day} \times 100 \quad (4)$$

substituting g of milk N / cow day by equation (3) results in:

$$Dairy\ cow\ Efficiency = \frac{0.19 \times g\ of\ N\ intake/cow\ day + 38.2}{g\ of\ N\ intake/cow\ day} \times 100 \quad (5)$$

$$Dairy\ cow\ Efficiency = \left(0.19 + \frac{38.2}{g\ of\ N\ intake/cow\ day}\right) \times 100 \quad (6)$$

So, the higher the N intake/cow day, the lower the efficiency of the dairy cow.

The definition of NUE is:

$$NUE = 100 \times \frac{N\ output}{N\ input}$$

Two different scenarios are tested:

Scenario 1: The total feed N input is fixed as input$_1$ and fed to n_1 cows.

$$NUE_1 = 100 \times \frac{N\ output_1}{N\ input_1}$$

$$Output_1 = Milk\ Output_1 + Meat\ Output_1$$

Milk Output$_1$ is:

$$Milk\ Output_1 = n_1 \times milk\ N\ production/cow_1$$

$$Milk\ N\ production/cow_1 = N\ intake/cow_1 \times 0.19 + 38.2$$

$$N\ intake/cow_1 = \frac{Total\ feed\ N\ input_1}{n_1}$$

$$Milk\ N\ production/cow_1 = \frac{Total\ feed\ N\ input_1}{n_1} \times 0.19 + 38.2$$

Scenario 2: The total feed N input stays constant = feed N input$_1$, the number of cows increases to n_2.

$$NUE_2 = 100 \times \frac{N\ output_2}{N\ input_1}$$

$$Output_2 = Milk\ Output_2 + Meat\ Output_2$$

Milk Output$_2$ is:

$$Milk\ Output_2 = n_2 \times milk\ N\ production/cow_2$$

$$Milk\ N\ production/cow_2 = N\ intake/cow_2 \times 0.19 + 38.2$$

$$N\ intake/cow_2 = \frac{Total\ feed\ N\ input_1}{n_2}$$

$$Milk\ N\ production/cow_2 = \frac{Total\ feed\ N\ input_1}{n_2} \times 0.19 + 38.2$$

To prove the veracity of the hypothesis we proceed as follows:
If:

$$NUE_1 < NUE_2$$

$$100 \times \frac{N\ output_1}{N\ input_1} < 100 \times \frac{N\ output_2}{N\ input_1}$$

$$\frac{N\ output_1}{N\ input_1} < \frac{N\ output_2}{N\ input_1}$$

$$Output_1 < Output_2$$

$$Total\ Milk\ N\ production_1 < Total\ Milk\ N\ production_2$$

$$n_1 \times milk\ N\ production/cow_1 < n_2 \times milk\ N\ production/cow_2$$

$$n_1 \times \left(\frac{Total\ feed\ N\ input_1}{n_1} \times 0.19 + 38.2\right) < n_2 \times \left(\frac{Total\ feed\ N\ input_1}{n_2} \times 0.19 + 38.2\right)$$

$$\frac{n_1 \times Total\ feed\ N\ input_1 \times 0.19}{n_1} + 38.2 \times n_1 < \frac{n_2 \times Total\ feed\ N\ input_1 \times 0.19}{n_2} + 38.2 \times n_2$$

$$Total\ feed\ N\ input_1 \times 0.19 + 38.2 \times n_1 < Total\ feed\ N\ input_1 \times 0.19 + 38.2 \times n_2$$

$$n_1 < n_2$$

which proves the veracity of the results indeed.

The following step is to calculate how much the number of dairy cows (or livestock units) can increase without jeopardising the minimum intake limit of 0.300 kg N cow^{-1} day^{-1}. The potential increase in the number of cows depends on the initial number of cows and on the initial N intake cow^{-1} day^{-1}. The larger the initial herd and the higher the initial intake cow^{-1} day^{-1}, the more "extra cows" can be included. We have not studied potential effects on animal health and fertility when the N supply to the dairy cows is very sharp. Within our dataset, a small increase in the dairy cow herd does not necessarily impair the economic performance of a dairy farm.

4. CONCLUSION

Two statistical models were developed, which include the most relevant variables influencing the nitrogen use efficiency (NUE) of Flemish dairy farms. One model is developed for farms without arable crops and the other model is developed for farms with arable crops. One major result is that, contrary to current knowledge, 'Dairy cows ha^{-1}' is the most relevant variable, being 1.4 times more relevant than 'kg of N in fertilisers ha^{-1}', 3.1 times more relevant than 'kg of N in concentrates ha^{-1}' and 4.6 times more relevant than 'kg of N in by-products ha^{-1}'. Hence, when the aim is to increase the NUE and the farm sustainability, the variable 'Dairy cows ha^{-1}' should be modified in the first place. Another major result is that, according to previous knowledge, the N input variables present a negative correlation (sign) with NUE. However, the variable 'Dairy cows ha^{-1}' shows a positive correlation (sign) with 'NUE'. At a constant N feed and fertiliser input to the farm, the increase in 'Dairy cows ha^{-1}' implies an increase in the total farm N output, which causes an increase in the average NUE and in the farm sustainability. In other words, if 'Dairy cows ha^{-1}' is increased, and if the farm N feed and fertiliser input is constant, the sum of the decreases in N in the milk production of each cow is lower than the extra N produced by the added cows. Therefore, the total farm milk N output will be higher and the NUE as well. Hence, a higher stocking density does not imply a lower agricultural sustainability provided other factors are kept constant.

REFERENCES

Borsting C.F., Kristensen T., Misciattelli L., Hvelplund T., Weisbjerg M.R. (2003) Reducing nitrogen surplus from dairy farms. Effects of feeding and management, Livest. Prod. Sci. 83, 165–178.

Carpenter S.R., Caraco N.F., Correll D.L., Howarth R.W., Sharpley A.N., Smith V.H. (1998) Non-point pollution of surface waters with phosphorus and nitrogen, Ecol. Appl. 8, 559–568.

Goodlass G., Halberg N., Verschuur G. (2003) Input output accounting systems in the European community – an appraisal of their usefulness in raising awareness of environmental problems, Eur. J. Agron. 20, 17–24.

Hadas A., Sagiv B., Haruvy N. (1999) Agricultural practices, soil fertility management modes and resultant nitrogen leaching rates under semi-arid conditions, Agr. Water Manage. 42, 81–95.

Hair J.F., Anderson R.E., Tatham R.L., Black W.C. (1998) Multivariate data analysis, Fifth edition, Prentice-Hall International, New Jersey.

Hanegraaf M., den Boer D.J. (2003) Perspectives and limitations of the Dutch minerals accounting system (MINAS), Eur. J. Agron. 20, 25–31.

Hansen B., Kristensen E.S., Grant R., Hogh-Jensen H., Simmelsgaard S.E., Olesen J.E. (2000) Nitrogen leaching from conventional versus organic farming systems – a systems modelling approach, Eur. J. Agron. 13, 65–82.

Hanson J.D., Skiles J.W., Parton W.J. (1988) A multi-species model for rangeland plant communities, Ecol. Model. 44, 89–123.

Heathwaite A.L., Johnes P.J., Peters N.E. (1996) Trends in nutrients, Hydrol. Process. 10, 263–293.

Kebreab E., France J., Beever D.E., Castillo A.R. (2001) Nitrogen pollution by dairy cows and its mitigation by dietary manipulation, Nutr. Cycl. Agroecosys. 60, 275–285.

Kuipers A., Mandersloot F. (1999) Reducing nutrient losses on dairy farms in The Netherlands, Livest. Prod. Sci. 61, 139–144.

Montgomery D.C., Peck E.A., Vining G.G. (2001) Introduction to linear regression analysis, John Wiley & Sons, New York.

Nevens F. (2003) Nitrogen use efficiency in grassland, silage maize and ley/arable rotations, PhD Thesis, Ghent University, Belgium.

Nevens F., Verbruggen I., Reheul D., Hofman G. (2006) Farm gate nitrogen surpluses and nitrogen use efficiency of specialized dairy farms is Flanders: Evolution and future goals, Agr. Syst. 88, 142–155.

Oenema O., Boers P.C.M., Van Eerdt M.M., Fraters B., Van der Meer H.G., Roest C.W.J., Schröder J.J., Willems W.J. (1998) Leaching of nitrate from agriculture to groundwater: the effect of policies and measures in the Netherlands, Environ. Pollut. 102, 471–478.

Oenema O., Koskamp G.J., Galama P.J. (2001) Guiding commercial pilot farms to bridge the gap between experimental and commercial dairy farms; the project 'Cows and Opportunities', Neth. J. Agr. Sci. 49, 277–296.

Oenema O., Pietrzak S. (2002) Nutrient management in food production: achieving agronomic and environmental targets, Ambio 31, 159–168.

Oenema O., Van Liere L., Schoumans O. (2005) Effects of lowering nitrogen and phosphorus surpluses in agricultura on the quality of groundwater and surface water in the Netherlands, J. Hydrol. 304, 289–301.

Ondersteijn C.J.M. (2002) Nutrient management strategies on Dutch dairy farms: an empirical analysis, PhD Thesis, Wageningen University, The Netherlands.

Rejesus R.M., Hornbaker R.H. (1999) Economic and environmental evaluation of alternative pollution-reducing nitrogen management practices in central Illinois, Agr. Ecosyst. Environ. 75, 41–53.

Swensson C. (2003) Analyses of mineral element valances between 1997 and 1999 from farms in the south of Sweden, Eur. J. Agron. 20, 63–69.

Van Gijseghem D., Overloop S. (2002) Milieu- en natuurrapport Vlaanderen, Achtergronddocument 2002, 2.12 Vermesting. Vlaamse Milieumaatschappij, 75 p.

Van Passel S. (2007) Sustainability performance of farms: an efficiency approach, PhD Thesis Ghent University, Belgium.

Von Eye A., Schuster C. (1998) Regression analysis for social sciences, Academic Press, San Diego.

Wells S.A., Cole T.M. (2001) TMDLS: Statistical correlations or mechanistic modelling? The National TMDL Science and Policy Conference, November 13-16, Phoenix, Arizona.

Yadav S.N., Peterson W., Easter K.W. (1997) Do farmers overuse nitrogen fertilizer to the detriment of the environment? Environ. Resour. Econ. 9, 323–340.