Development of a validity test for survey data on milk-from-grass from German dairy farms

Becker T.H., Blume L., Kayser M. and Isselstein J.
Georg-August University Göttingen, Department of Crop Sciences, Germany;
talea.becker@agr.uni-goettingen.de

Abstract

Questionnaires are a frequently used instrument to analyse the productivity of farms. As surveys might include wrong or incorrect data, there is a need for validity testing. A validity test aims at generating an adjusted, reliable data-set with fewer outliers. Large data sets require an automated approach. We conducted a survey on 47 German dairy farms to evaluate the role of grassland in milk production. The farms are located all over Germany with a focus on Lower Saxony and Hesse. In a first step, we developed a generally applicable validity test for assessing the milk yield directly related to grassland. Several simple, directly measurable parameters were defined which correlate with important parameters with a high error rate. These relations were put into a formula and applied to the data set. We found that out of the 47 data sets three had to be excluded from further analysis because of large deviations from the defined confidence limits. The experience with the validity test did not only result in a more reliable data set, but helped to optimize the questionnaire for future surveys. The farms in this survey produced 4,916 l grassland-milk ha^{-1}.

Keywords: survey, validity test, grassland-milk, dairy

Introduction

For evaluating the productivity of farms, questionnaires are a frequently used instrument. However, with surveys there is always a high risk for misunderstandings and miscalculations. The validity and plausibility of the data for the most important parameters should be tested before a final analysis takes place. Ideally, this is done in a systematic way based on confidence limits and following a protocol. For extensive surveys, faulty data-sets are then excluded while for small surveys it can make sense to check doubtful data and eventually correct obvious mistakes by hand.

In order to determine the amount of milk based on grass (grassland-milk), 47 German dairy farmers were interviewed. We developed a protocol to test the validity and plausibility of answers to important parameters and applied it to the survey data. We hypothesize that this will improve data quality and the reliability of the following analysis.

Material and methods

In 2012 and 2013, 47 dairy farmers in Germany were interviewed by students from the Faculty of Agricultural Sciences of the Georg-August University Göttingen. Since it is not possible to ask directly for the amount of grassland-milk produced on the farm, this parameter has to be calculated from the following basic parameters which were provided by the farmers from farm records: amount of concentrates and maize in the ration; number of cows; milk yield per cow; and amount of grassland used for milk production. Some of these basic parameters were also used for the validity test. In a first step the data for these basic parameters have to fall within defined confidence limits and are then combined in a formula to calculate parameters that cannot directly be measured, e.g. the amount of grass (grass + grass-silage + hay) in the feed ration (Table 1). The results obtained from these formulas (final parameters) have to match previously defined confidence limits as well. Some additional parameters, which were not asked in the survey, were taken from the literature (Table 2). The overview of the protocol of the study is given in Table 3.

Table 1. Basic formulas for determining the final parameters and the amount of grassland-milk per hectare.

Calculation of the amount of grass in the daily ration of one cow

- Maintenance requirement per cow per day (MJ NEL) + energy requirement for lactation per cow per day (MJ NEL) = Energy requirement per cow per day (MJ NEL)
- Energy requirement per cow per day (MJ NEL) – energy from maize per cow per day (MJ NEL) – energy from concentrates per cow per day (MJ NEL) = Energy from grass (grass + grass-silage + hay) per cow per day (MJ NEL)
- Energy from grass per cow per day (MJ NEL) / energy content of silage (MJ NEL kg^{-1} DM) = Amount of grass in feed ration (kg DM per cow per day)

Calculation of the milk from grass per cow per year

- Milk yield per cow per day (l) / energy requirement per cow per day (MJ NEL) × energy from concentrates (MJ NEL) = Milk from concentrates per cow per day (l)
- Milk yield per cow per day (l) / energy requirement per cow per day (MJ NEL) × energy from maize (MJ NEL) = Milk from maize per cow per day (l)
- Milk yield per cow per day (l) – milk from maize per cow per day (l) – milk from concentrates per cow per day (l) × 305[a] = Milk from grass per cow per year (l)

Calculation of the milk from grass per ha grassland per year

- Milk (l) from grass per cow per year × cows per farm = Milk from grass per farm per year (l)
- Milk (l) from grass per farm per year per ha grassland per farm = Milk from grass per ha grassland per year (l)

Calculation of the yield of grass per ha grassland per year

- Amount of grass in feed ration (kg per cow per day) × cows per farm × 305 days of lactation = Need for grass (for milk production) per farm per year (kg DM)
- Need for grass (for milk production) per farm per year (kg DM) ha^{-1} grassland per farm = Yield of grass (for milk production) per ha per year (kg DM)
- Yield of grass (for milk production) per ha per year (kg DM) / 60%[b] = Total yield of grass per ha per year (kg DM)

[a] Lactation duration.

[b] Amount of grass yield ha^{-1} that is used for milk production.

Table 2. Basic parameters and final parameters which are needed for the validity test.

Basic parameters	**Confidence limits**	**Source[1]**
Amount of concentrates in the ration per cow per day	0-13 kg DM[a]	KTBL, 2009
Amount of maize in the ration per cow per day	0-13 kg DM[a]	KTBL, 2009
Milk-yield per cow per year	0-12,000 l[a]	KTBL, 2009
Energy content of roughage	4-7 MJ NEL kg^{-1} DM[a]	KTBL, 2009
Final parameters for validity test	**Confidence limits**	
Calculated amount of grass (grass + grass-silage + hay) in feed ration per cow per day	0-19 kg DM[c]	KTBL, 2009
Calculated dry matter intake per cow per day	15-24 kg[c]	KTBL, 2009
Calculated yield of grass per year ha^{-1}	0-11 Mg DM[c]	KTBL, 2009
Basic-parameters taken from literature	**Values**	
Live weight of a dairy cow	650 kg[c]	KTBL, 2009
Lactation duration	305 days[d]	Loeffler (2002)
Maintenance requirement per cow	37.7 MJ NEL day^{-1} [c]	KTBL, 2009
Power requirement per liter milk	3.28 MJ NEL[c]	KTBL, 2009
Energy density of maize	6.7 MJ NEL kg^{-1} DM[c]	KTBL, 2009
Energy density of concentrates	7.6 MJ NEL kg^{-1} DM[c]	KTBL, 2009
Amount of grass yield ha^{-1}, that is used for milk production	60%	-[2]

[1] Confidence limit based on the source.

[2] Assumption based on common agricultural knowledge of farming practice.

Table 3. Protocol for executing the validity test.

Extensive surveys	Small surveys
• Defining basic parameters and final parameters. Determining confidence limits for the parameters.	• Defining basic parameters and final parameters. Determining confidence limits for the parameters.
• Checking the basic parameters.	• Checking the basic parameters.
• Data-sets with one or several basic parameters that do not match the confidence limits are excluded.	• Data-sets with one or several basic parameters that do not match the confidence limits are marked.
• Applying the formulas for calculating the values of the final parameters.	• Applying the formulas for calculating the values of the final parameters.
• Data-sets with one or several final parameters that do not match the confidence limits are excluded.	• Data-sets, with one or several final parameters that do not match the confidence limits are marked.
	• Search for mistakes in marked data-sets. Eventually correction of the mistakes.
	• Exclusion of data-sets which do not pass the validity tests after correction.
• Final analysis of the data from the remaining data-sets.	• Final analysis of the data from the remaining data-sets.

Table 4. The effect of applying a validity test on some parameters for evaluating the importance of grassland for milk production.

	Before validity-testing (n=47)		After validity-testing (n=44)	
	Mean	SD	Mean	SD
Grassland used for milk production (ha)	64	47	54	36
Milking cows	95	70	98	71
ha grassland per cow	0.7	0.4	0.6	0.3
Milk yield per cow per year (l)	9,104	1,410	9,117	1,225
Grassland-milk ha^{-1} (kg)	4,206	4,442	4,916[a]	2,053[a]
Grassland-milk per cow per year	2,572	2,137	2,294	1,393
Calculated yield of grassland (t ha^{-1} per year) [b]	4.7	5.1	5.7	2.3
Calculated grass-intake (kg per cow per day) [b]	6.0	5.7	7.1	3.8

[a] Weighted mean.

[b] Parameter of validity test.

Results and discussion

In our survey of 47 dairy farms, the data sets from four farms did not match the basic parameters. Three of these had also not matched the final parameters. Besides these three, there were nine other data-sets which did not match the final parameters. In small surveys, it often makes sense to have a closer look at the faulty data sets to find out why they failed the test.

When we did that, mistakes in ten data-sets could be corrected, and they passed the test in a second run. Most errors were caused by a confusion of dry matter and fresh matter of feed or by miscalculating the grassland area. Three data-sets, that are three farms, did not pass the test even after an intensive search for mistakes and the correction. As it was not possible to contact the farmers for clarification, these data-sets had to be excluded from further analysis. Data from the remaining farms were then used for calculating the amount of grassland-milk. The validity testing did not lead to drastic changes of the means but resulted in much lower standard deviations. Before the test, standard deviations of grassland-milk ha^{-1}, grassland-milk per cow, grass intake per cow per day and of yield of grass ha^{-1} were in the range of the mean, but were reduced to 50% of the means after applying the test (Table 4).

References

Kuratorium für Technik und Bauwesen in der Landwirtschaft (KTBL) (2009) *Faustzahlen für die Landwirtschaft* 14, Darmstadt (GER), pp. 382-404, 659-670.

Loeffler K. (2002) *Anatomie und Physiologie der Haustiere* 10, Ulmer, Stuttgart (GER), 340.