56th Annual meeting of the European Association for Animal Production
Uppsala,Sweden, June 5th - June 8th, 2005
skarsten@tierzucht.uni-kiel.de

The use of a database for genetic evaluation and to manage health in dairy cows

S. Karsten[1], E. Stamer[2], W. Junge[1],
W. Lüpping[3] and E. Kalm[1]

[1]*Institute of Animal Breeding and Husbandry, Christian-Albrechts-University Kiel, Olshausen Str. 40, D-24118 Kiel, Germany,skarsten@tierzucht.uni-kiel.de*
[2] *TiDa Tier und Daten GmbH, Brux, Germany,*
[3]*Landwirtschaftskammer Schleswig-Holstein, Kiel, Germany*

Abstract

The quantity of data recorded for each animal will continue to increase in the future. Records computerised in management programs can be used for operational analyses but the information systems are of limited use for research and the evaluation of experiments. Problems will arise especially if analyses are performed across farms. Various software systems and data sources and unstandardised data codes make evaluations difficult.
A database has been developed to organise data of numerous farms. Records of parentage, feeding, fertility, health and individual performance are electronically transferred between management programs and the database. Irregularly occuring data can be entered on-line via an electronic form. An accurate and consistent data retention is guaranteed by validity checks. The data can be compiled, exported and analysed easily because of a uniform coding system. A user interface will be provided for on-line access. Data processing by a statistical program package is also be possible. Health records can be analysed easily for example by control charts. In the moment the database is applied on two research farms. On one farm data of trials like feeding experiments are collected, on the other farm functional traits are measured within the scope of a performance test and evaluated genetically.

Introduction

New process technologies have evolved, for example automated monitoring of individual animal performance and activity, supplying many records on each day. To make detailed recommendations about the herd's health and management, the veterinarian or any other adviser needs additional information, on all aspects of fertility and diseases. Within the framework of the existing legislation more records are to be kept by the farm staff anyway in order to provide for traceability and quality assurance (e.g. medicinal treatments). The management systems for these on-farm records have also improved and the capability for herd management has enhanced. But the indices and codes used in software differ widely, and there is often no definition of the indices used. Therefore, the comparison across farms using different software is problematically.
The use of such data management systems for research and genetic evaluation is limited as well, especially when the analyses are performed across farms. On research and test farms even more data are recorded (e.g. functional traits, metabolic parameters, individual feed intake) and the demand and necessity to manage more and diverse data becomes more obvious.
Many different sources of information have to be integrated into data evaluations. The data used in genetic evaluations originate from the on-farm record systems, the milk recording

schemes, laboratories etc. The validity of the estimated genetic parameters depends on an accurate recording of parentage, production and environmental factors. Up to now a file-oriented data processing has been mostly used. The file structures are closely related to the data structures and the processing routines of the programs, and may therefore differ between farms. Reduncancy and inconsistency in the data are consequences of the file-oriented data processing. In the time of the data analysis when time is sparse, consistency and adaptation problems have to be solved. Especially in evaluations across farms (e.g. BLUP) additional problems arise due to the use of unstandardised codes, different data structure etc. Therefore, the basic requirement for systematic use of field and experimental data is a central database and a standardised coding system in order to reduce the danger of data inconsistency and to increase the productivity of application programmes. This paper deals with the development of a concept for the management of the information from test and research farms.

Sources of information

If an automated milk recording system is used on the farm, the milkings will be usually stored in the local computer programme. The various data on barn environment, management routines and feeding are often recorded only on paper or also in the local herd management system of the farm. The same is applied to the data files created by the milk recording laboratory. The records of administered medicines may be managed by the herd management software, on paper or by other software like for example Excel. On research farms even more sources of information exist: data files with individual feed intakes, metabolic parameters and others. These data of special experiments are usually stored in ASCII-files or Excel sheets. For the data evaluation, these very heterogeneuos sources of information have to be integrated.
The data preparation is usually extensively and time-consuming. Because of the parallel data storing in files and different software the data are not continuously adjusted and the validity of the data is not checked till the data processing. Because of the local data storing, the data have to be transfered to the analyst electronically or by mail.
More problems arise when the data analysis is performed across herds. Several programmes and process technologies may be used on the farms leading to different data structures. The different farm staff and veterinarians may use diverse codes and synonyms resulting in complications at combining the data. For a good evaluation of the health recordings a consistent classification of the diseases and diagnoses in categories is indispensable.

Concept of the database "KUHDAM"

The computer-based data handling described in this paper is realised by a relational data model. The information on objects are stored in tables. Each animal is characterised by a unique individual number that is used in the whole model to identify the animal. At the time of data entry, the cow and bull are specified by their ear tag or herd book number, respectively. This external identification is translated into the internal unique individual number. The same is done with the farms: The name is translated into an individual farm-ID. For all traits a definite reference exists from the relation to the rest of entries for the animal. Consequently, more traits can be analysed for the individual animals (e.g. treatment costs for each animal).
The generic design of the data model allows to handle nearly all occurring traits and parameters. Due to the flexible and comprehensive design, future extensions are possible. The most important complexes in the data model are shown in table 1. Additionally, activity, animal weights, metabolic parameters and others are integrated into the data model.

Table 1
Complexes in the data model and exemplary traits

Milking	**Feeding**	**Health**	**Fertility**	**Individual data**
Milk yield	Feed intake	Diagnosis	Insemination	Ear tag
Average milk flow	Water intake	Disease category	Pregnancy test	Birth date
Peak in milk flow	Feed composition	Treatment	Calving date	Pedigree
Milk constituents	Feed ratio	Drug used	Calving ease	Sex

A standardised coding system is used that works across the herds and that makes evaluations more easy. Diseases entered into the database have to be classified into disease categories (metabolic disorders, fertility, claw and leg disorders, udder) by the veterinarian. In this way, a good analysis of the health data is possible. The data base system as well as the application programmes provide validity checks to remove erroneous or inconsitent records. In this way, inconsistency in the data and invalid values are avoided. The scheme of the database-oriented data processing is shown in figure 1.

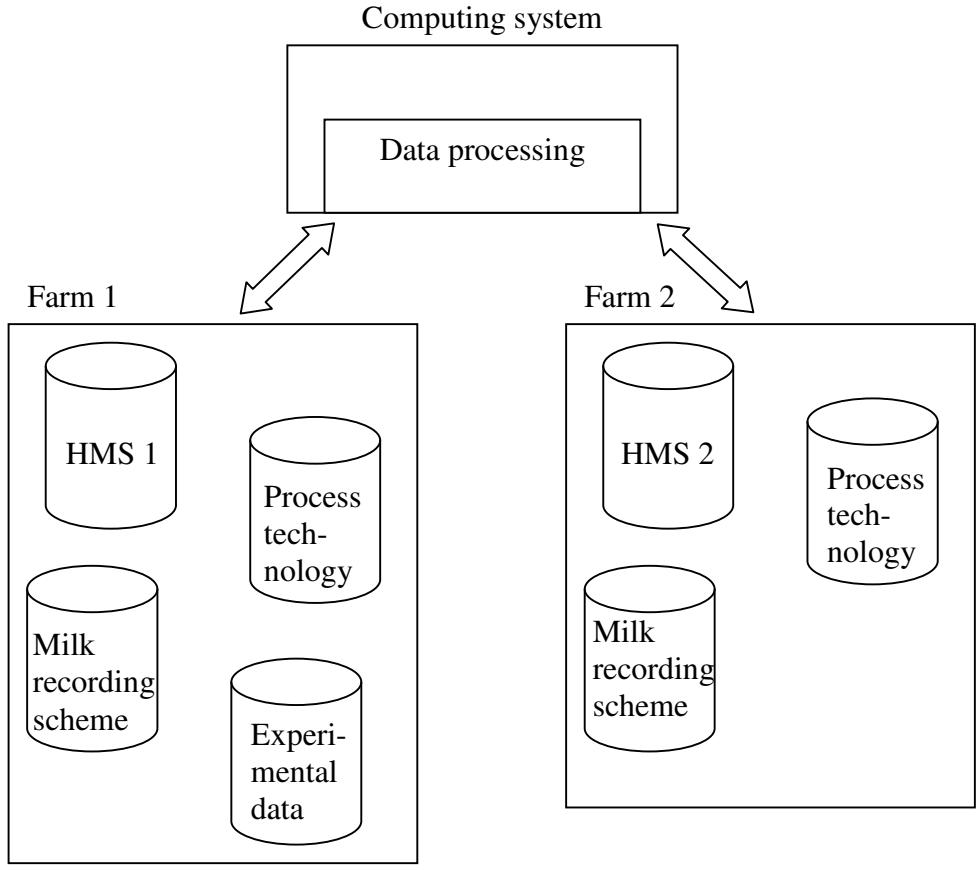

Figure 1
Scheme of the database-oriented data processing
(HMS = Herd management software)

Technical facts

As the database is based on the open source database software 'PostgreSQL', the implementation of the database does not incur any license costs. Access to the data is provided

by standardised interfaces (SQL). The database is implemented on a Linux server. Software for the automatic data transfer and validity checks has been developed in Perl 5.

Data import/ export

For an efficient electronic data transfer between the different sources of information on the farms and the database standardised formats for data exchange are necessary. Lists provided by the herd management software on the farms, the process technology and other sources are imported automatically into the database by the means of an application software programmed in Perl. Irregularly occurring data can be entered directly into the database by the staff at the farm using an online form. Validity checks are performed in the time of the data entry and therefore almost immeditely after data recording. In case of a discrepancy, the problem may be solved easily by contacting the farmer.
Data export out of the database is possible on different ways. On the one hand, data can be extracted from the database and subjected to statistical analysis with software like SAS or Excel. On the other hand, the desired data can be displayed on a web page. Thus, the farm staff is enabled to obtain immediate management information on the herd. Another advantage of the central database is the possibility of individual combination of the desired data depending on the task.

Application of the database

In this stage of the pilot project, the data of two farms are managed by the database. One of them is a bull dam testing station with individual performance testing on milk production and functional traits. The other farm is a research farm performing large-scale experiments, e.g. feeding experiments. The data collection have started. The routine operation includes automatic data transfer of lists exported regularly from the herd management system of the farms including milk recordings and actions like inseminations. The online forms for direct data entry are under development.

Discussion

The data of different farms is managed by one central database system that provides an easy access to the data and a powerful use of the aggregated information. The redundancy is reduced by characterising each animal by an individual id and storing each value only once. The quality of an analysis depends on the quality of the data used. Validity checks are performed in the time of the data import reducing inconsistency in the data and erroneous entries.
Because of the integration of the data of different farms comparative evaluations are easily possible. To make a variable application of the database possible, the data model should be extended by analytical programs as well as action lists. The former means that the effects of changes in management can be monitored, e.g. by graphs and control charts. Additionally, indices may be calculated up to date. Such modules for the evaluation of the data may motivate the farm staff to enter the records accurately and regularly. On the base of a central database the indices of the farms in the database are calculated the same way and in this way a comparison across farms is easily possible. Daily printed action lists can flag up cows that need to be observed by the farm staff or veterinarian. Standardised codes and disease categories enable less time-consuming evaluations. Long-term and especially cross-project evaluations will be more easy.