Distributed computing as a new opportunity for stakeholders in dairy cattle management and breeding

N. Gengler
Numerical Genetics, Genomics and Modelling Group, Gembloux Agro-Bio Tech, University of Liège, B-5030 Gembloux, Belgium
nicolas.gengler@gmail.com

Abstract

Currently the uses of on-farm computers and of centralized performance-recording based tools are considered as two opposite models for dairy management. However, dividing the problem into complementary tasks, each of which is optimally solved, is an opportunity that should also be considered by stakeholders. Recent development solved data exchange issues allowing the use of adapted distributed computing algorithms. As example milk yield and composition are given. Different research projects and several commercial companies are focussing on the development of on-farm tools, mostly Near InfraRed (NIR) based, other projects are developing and implementing tools based on Mid InfraRed (MIR), available only through performance recording. Both are complementary, as NIR based measurements are easier and less expensive, available at every milking, but MIR based measurements are more precise, however only obtained every 4 weeks. Numerous advantages arise when combining both types of measurements. It will be shown that statistical theory exists to base advanced modelling on, using optimally the longitudinal data generated by this type of setting. This will open different novel opportunities to optimize currently used on-farm and off-farm management and breeding tools.

Keywords: Milk recording, Data exchange, Central database, Distributed computing, Hierarchical modelling

Introduction

The use of direct milk yield meters and similar sensors in robotic milking units and fully computerized milking parlours associated to farm computers running adapted on-farm herd management system, is often seen as a classical case of “precision livestock farming” in dairy cattle. In opposition to many other species and production systems, dairy farming has also another, well-developed historic dairy herd management approach which relies on classical performances recording, mostly supervised by technicians, on centralized milk testing and on centralized data bases. This data is then used to generate herd and cow reports that, until recently on paper nowadays often by the Internet, are send back to help manage associated herds. Basic principles of milk recording are harmonized across countries by the International Committee for Animal Recording (ICAR, 2012). This data is also the primary source of dairy cattle data used in animal breeding (INTERBULL, 2012). Currently the uses of on-farm computers based systems and of centralized performance-recording based tools are considered as two opposite “models” for dairy cattle management. The objective of the present report is to show an alternative integrative approach currently under development presenting the different layers and how by dividing the problem into complementary tasks. By optimizing each task and achieving its optimal solving at the adequate level, dairy herd management can then be optimized. Additionally the statistical background of this approach will be explained and some practical examples will be given.

Material and Methods

Current status of interaction between on-farm and off-farm systems

Use of sensors available on-farm is increasing. This includes a wide range from classical milk yield meters and conductivity sensors, over sensors directly measuring on the animal as pedometers up to very advanced sensors as those for LDH (Lactate Dehydrogenase), urea, BHB (Beta Hydroxy Butyrate) and progesterone becoming commercially available in the Herd Navigator (developed by Danish Lattec A/S, a jointure venture company of Delaval International and FOSS Analytical: Mazeris, 2010). Classically all these tools were conceived as stand-alone products or potentially linked by a common on-farm infrastructure but only if provided by the same manufacturer. Despite some efforts especially from milking parlour manufacturers common standard “languages” for data exchange are still sparse (i.e., TAURUS Standard Interface).

Very early with the first sensors (classical milk yield meters) becoming available performance recording agencies have started to develop ways to import data. Basically two strategies were pursuit. The first strategy is the development of own on-farm management systems, the PCDART program (Dairy Records Management Systems, Raleigh, NC, USA) being an example. Unfortunately this limits the choice of herd owners and is considered not necessarily optimal by them because of their choice of other systems. Therefore manual transfer of data was often needed. A second strategy was to develop methods to export the data from the farms to central databases independently from the manufacturers of the different on-farm systems. Again the natural limit that appeared was the need or, unfortunately, the lack of common exchange standards. An innovative idea was developed by Valacta (Dairy Production Center of Expertise Quebec-Atlantic, Canada) in its Trans-D software that was from the beginning on multi-manufacturer and pluged-in directly into on-farm data bases (Saunier *et al*, 2012). Based on this principle, in collaboration France Conseil Elevage (FCEL) and Valacta developed Ori-Automate a bi-directional interface tool that links farm management software to performance-recording databases. Only this interface is installed additionally on-farm. The XML format is then used to communicate with the central milk recording databases (Saunier *et al*, 2012). In the Walloon Region of Belgium the Walloon Breeding Association (AWE) is currently implementing Ori-Automate. In other countries alternative tools are under development or already deployed. There are two other hidden advantages in a bi-direction approach. First all on-farm sensor-based tools need to access basic animal data in order to operate. By linking up with the recording agencies farmers no longer need to enter this information, potentially even several times, as it is readily available in the central databases. Also as described above current on-farm systems when provided by different manufacturer are seldom designed to exchange data. By communicating with Ori-Automate or similar systems the exchange between on-farm tools is, indirectly, established.

The next step: distributed computing

Having this exchange architecture in place, the development of distributed systems is the next logical step. With the steady increase of computing power current desktop PC, as routinely used on-farm, are underused. The quantity and quality of data produced by on-farm sensors is also increasing, potentially overwhelming centralized data bases. A way to address these issues is by developing adapted distributed computing algorithms dividing the problem into complementary tasks. By optimizing each task and achieving its optimal solving at the adequate level, dairy herd management (and breeding) can then be optimized. The quantity and quality of milk produced will be used as practical example to show the statistical and modelling background of this approach.

Results and Discussions

Modelling

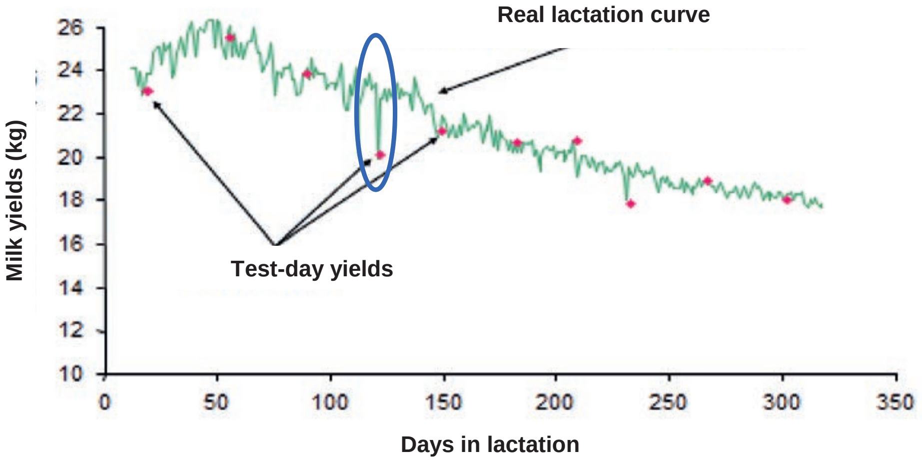

Figure 1: Evolution of the real lactation curve (daily yields) and yields recorded at specific test-days (for clarity only three indicated by an arrow); an outlier is highlighted.

A typical lactation curve for milk yield is shown in Figure 1. Numerous mathematical models were developed for this type of data (e.g., Wood, 1967), an overview was given by Gengler (1996). The natural limitation of all method was the limitation of data to few (between 8 and 11) test-day records available over a standard lactation period of 305 day yields. More advanced computational methods became available as Vallait Concept implemented by AWE based on the study by Mayeres et al (2004). One of the specific features of this method is that the lactation curve modelling included all information known of this cow including the specificity of her herd, her breed and her genetics. The used equations were set-up as mixed model equations (Henderson, 1984; Robinson, 1991), but can also be interpreted in a Bayesian setting (Blasco, 2001). Other mathematical approaches were proposed as the Kalman Filter (Van Bebber *et al*, 1999), but these were only useful when applied to daily data. With on-farm meters the limitation on the availability of data was replaced by that of finding adequate methods to limit the burden at a central database level. An adequate solution is to consider this as a two step process (Gengler, 2002) adopting adequate solving algorithms (Gengler *et al*., 2000) where herd-individual and population levels are seperated. Statistically this can be written in a more formal context using a Bayesian hierarchical modelling approach (e.g., Jamrozik *et al*, 2001). Also extending this to use more data on a farm level exchanging only specific lactation parameters (e.g., Gengler, 2002) is also straightforward.

<u>New opportunities</u>

An example for current developments leading to new opportunities is the detection and use of fine milk composition. Different research projects (e.g., MILKINIR: Nguyen *et al*., 2011) and several commercial companies are focussing on the development of on-farm tools for milk quality that are mostly based on Near InfraRed (NIR). Other projects (e.g., RobustMilk: Soyeurt *et al*, 2012; OptiMIR: Massart, 2011) are developing and/or implementing tools based on Mid InfraRed (MIR), available only through performance recording milk laboratories. Both are complementary, as NIR based measurements are easier and less expensive, available at every milking, but MIR based measurements are more precise, however only obtained every 4 weeks. If both sources of information are combined MIR will also help improve quality of NIR measurement, which can be used to make management decision on a short term. Advanced modelling of this longitudinal data generated by this type of setting combining optimal both levels: population for the performance recording and herd-individual on a herd level. This type of interactions will open numerous opportunities to optimize currently used tools. A few examples are first on an on-farm level that these tools would be near real-time during or shortly after milking but still allowing benchmarking and comparison to other farms. On an off-farm level, it would also allow quality control on the performance recording side as outliers (cf Figure 1) linked to a specific event (e.g. heat) could be more easily detected. Also, following the proposal by Gengler (2002) alternative parameters could be transmitted to the central database. These parameters would then generate new phenotypes (i.e., stability of records as an indicator of animal robustness). Boichard & Brochard (2012) gave many other opportunities created by the linking on Precision Livestock Farming and performance recording, especially for animal breeding and genetics.

Conclusions

Recently and also through the generalisation of Internet access and the development of "cloud computing ", the development of distributed systems is becoming a reality. For Precision Livestock Farming, especially after the solving of the data exchange problem this gives numerous new opportunities on-farm but also off-farm. In particular, theory exists to develop distributed computing. It is therefore a real new opportunity for stakeholders in dairy cattle management and breeding. Given the positive interaction between performance recording agencies and manufacturers of equipment recently establish for data exchange, by adding adapted research the practical development of such systems will become a reality. This paper gave first indications what can be done, but many other opportunities exist.

References

Blasco, A. 2001. The Bayesian controversy in animal breeding. *Journal of Animal Science* **79**(8) 2023–2046.

Boichard, D., and Brochard, M. 2012. New phenotypes for new breeding goals in dairy cattle. *Animal* **6**(4) 544–550.

Gengler, N. 1996. Persistency of lactation yields: A review. *Interbull Bulletin*, **12** *http://www-interbull.slu.se/ojs/index.php/ib/article/download/254/254*. Accessed May 16, 2013.

Gengler, N. 2002. Considerations on the path from test-day records to national and international genetic evaluations and its consequences. In *Proc. of the 7th World Congress on Genetics Applied to Livestock Production,* Montpellier, France, August, 2002. Session 1. (pp. 1-4). Institut National de la Recherche Agronomique (INRA).

Gengler, N., Tijani, A., and Wiggans G.R. 2000. Use of sequential estimation of regressions and effects on regressions to solve large multitrait test-day models. *Journal of Dairy Science* **83**(2) 369.e1–369.e18.

Henderson, C.R. 1984. Applications of linear models in animal breeding.

ICAR, 2012. Aims and main objectives of the Committee. *http://icar.org/pages/aims.htm*. Accessed May 16, 2013.

INTERBULL, 2012. Genetic Evaluations: Information on evaluations for production, conformation, udder health, longevity, calving, female fertility traits and workability. *http://www-interbull.slu.se/eval/framesida-genev.htm*. Accessed May 16, 2013.

Jamrozik, J., Gianola, D., and L. R. Schaeffer, L.R. 2001. Bayesian estimation of genetic parameters for test day Records in dairy cattle using linear hierarchical models. *Livestock Production Science* **71**(2) 223–240.

Massart, X. 2011. OptiMIR: Un projet européen innovant pour la durabilité des exploitations laitières [OptiMIR : An innovative European projet to improve the sustainability of dairy farms]. *Wallonie Elevage* 5 26-27. http://www.optimir.eu/files/20110501_Article%20_Wallonie_Elevage.pdf. Accessed May 16, 2013.

Mayeres, P., Stoll, J., Bormann, J., Reents, R. and Gengler, N. 2004. Prediction of daily milk, fat, and protein production by a random regression test-day model. *Journal of Dairy Science* **87**(6) 1925–1933.

Mazeris, F. 2010. DeLaval Herd Navigator: Proactive Herd Management. In *Proceedings of First North American Conference on Precision Dairy Management (pp. 26-27).* http://www.precisiondairy.com/proceedings/s1mazeris.pdf Accessed May 16, 2013.

Nguyen, H.N. , Dehareng, F. , Hammida, M. , Baeten, V. , Froidmont, E. , Soyeurt, H. , Niemöller, A. and Dardenne, P. (2011). Potential of near infrared spectroscopy for on-line analysis at the milking parlour using a fibre-optic probe presentation. *NIR news*, **22**(7) 11-13.

Robinson, G.K. 1991. That BLUP is a good thing: The estimation of random effects. *Statistical Science*. **6**(1) 15-32.

Saunier, D., Clyde, G., and Moore, R. 2012. New interface to exchange data on the farm: Ori-Automate by FCEL and Valacta. In *Proceedings of the 38th ICAR Session, Cork (Ireland).* http://www.icar.org/Cork_2012/Manuscripts/Published/Saunier.pdf. Accessed May 16, 2013.

Soyeurt, H., McParland, S., Berry, D., Wall, E., Coffey, M., Gengler, N., Dehareng, F., and Dardenne, P. (2012). Mid-infrared predictions of fatty acids in bovine milk: final results of the RobustMilk project. In *Book of Abstracts of the 63rd Annual Meeting of the European Federation of Animal Science.* Wageningen Academic Publishers.

Wood, P.D.P. 1967. Algebraic model of the lactation curve in cattle. *Nature*, **216** 164-165.