

BIG DATA may be useful to dairy herd managers – after washing and drying

Peter Løvendahl

Center for Quantitative Genetics and Genomics, Department of Molecular Biology and Genetics, Aarhus University, AU-Foulum, Denmark

Peter.Lovendahl@MBG.AU.DK

Great promises are often made by providers of newly developed sensor based decision support systems. After getting the device paid, installed, connected and switched on, the decision support is expected to work flawless ever after. But, it will not, unless a number of maintenance and control points are attended to. In some cases, the equipment will run but produce wrong results.

General quality parameters like sensitivity and specificity may be available in technical data, but they are insufficient in describing common errors and stability issues. A simple “dashboard style” control panel would be a great help to verify that sensors are working, data is flowing and calibrations are within expected range. Also “oil lamps” and “fuel gauges” offer good help to view when service is needed. Control data from these “lamps and gauges” should be stored automatically in order to verify long-term performance and record service interventions of any kind. It would also be an advantage if equipment of various kinds were tested according to official guidelines provided by an international body such as ICAR. The guidelines should provide a list of control parameters; specify how they are to be measured and calculated; and give latitudes for acceptable function. The test results should be openly available to anyone. Testing for approval to known guidelines has been successful in many industries, including cars, electric hardware, safety equipment, and milk recording devices. During development of new sensor based decision support, with or without internet, BIG data solutions should therefore include routines for quality and performance control, as well as filters to remove faulty data. We have investigated these issues during our work with a mastitis detection and monitoring tool based on frequently recorded somatic cell count obtained from automatically milked cows. Our findings were clear: algorithms will usually work good with good clean data and bad with faulty and biased data. A modular build of software encourages guided sensor maintenance, cleaning or filtering and adjusting data, that was of utmost importance to obtain reliable processed data.

Acknowledgements

This article is based upon work from COST Action FA1308 DairyCare, supported by COST (European Cooperation in Science and Technology, www.cost.eu). COST is a funding agency for research and innovation networks. COST Actions help connect research initiatives across Europe and enable scientists to grow their ideas by sharing them with their peers. This boosts their research, career and innovation.