Detection of abnormal milk and mastitis using sensor measurements of automatic milking machines

C. Kamphuis, D. Pietersma and H. Hogeveen
Epidemiology and Economics group, Department of Farm Animal Health, Utrecht University, the Netherlands

Introduction
Since the introduction of AM systems in 1992, an approximate 4% of the Dutch dairy farmers have implemented such a system. Van der Vorst (2002) showed that the bulk tank SCC (BTSCC) of dairy producers using such a system, reached a level comparable to that of dairy producers milking conventionally at 1.5 years after implementation. This indicates that dairy producers using AM systems are able to manage mastitis successfully.

However, improvements are still possible regarding some udder health and milk quality issues. Firstly, increasing the sensitivity of mastitis detection would improve the timely detection of mastitis cases that require treatment, prevent production losses and prevent the spread of infection. Secondly, reducing the number of false positives on the mastitis attention lists would decrease the work-load of the farmer and improve the farmers’ trust in those lists. In addition, improvement of the sensitivity of abnormal milk detection, given an acceptable specificity level, is required. Currently, abnormal milk is not always separated causing a decrease in bulk tank milk quality. Finally, pathogen detection using sensor measurements has not been implemented yet. A proper prediction of the most likely mastitis causing pathogen might improve treatment plans and will reduce costs of antibiotics.

PhD research proposal
The objectives of this PhD research project are to improve the detection of mastitis, abnormal milk, and mastitis causing pathogens. This will be achieved by improving the use of data of existing sensors. Data will be collected for a period of 2 years, from 10 commercially
dairy herds with an AM system. The collected data is expected to include more than three million quarter milkings, and approximately one billion sensor measurement records. This huge data set will contain missing values, outlier values, and mislabelled records. Data mining techniques will be used to deal with these constraints as well as to develop classification models. Data mining can be defined as an iterative process to analyze large data sets to discover new patterns that are useful and understandable. The process of data mining is depicted in Figure 1. An important step in data mining is data pre-processing, with the development of potentially predictive variables as one aspect of this pre-processing. These developed variables can be used in a later stage of the data mining process to develop classification models.

![Figure 1. The process of data mining.](image)

An existing data set was used in order to start with creating potentially predictive variables. The data involved a single farm using automatic milking and covered a 2-week period. This period included 8000 quarter milkings from 65 cows. Sensor data consisted of EC, colour (red, green, and blue), and milk flow. Quarter milkings had been scored for homogeneity using 6 levels (increasing from normal milk to milk with a complete loss of its normal character).

Sensor data patterns of series of quarter milking were visualized in S-PLUS, to explore how changes in the measurement patterns relate to changes in homogeneity score. Variables were created to represent different characteristics of the sensor data patterns, such as the pattern level, variability or shape.

**References**


