

**Modeling cow somatic cell count using sensor data as input to generalized additive models**

**Dorota Anglart<sup>1, 2,\*</sup>, Charlotte Hallén-Sandgren<sup>1</sup>, Patrik Waldmann<sup>3</sup>, Martin Wiedemann<sup>4</sup>  
and Ulf Emanuelson<sup>2</sup>**

<sup>1</sup>DeLaval International AB, PO Box 39, SE-147 21, Tumba, Sweden

<sup>2</sup>Swedish University of Agricultural Sciences, Department of Clinical Sciences, PO Box 7054, SE-750 07 Uppsala, Sweden

<sup>3</sup>Swedish University of Agricultural Sciences, Department of Animal Breeding and Genetics, PO Box 7023, SE-750 07 Uppsala, Sweden

<sup>4</sup>DeLaval GmbH, Wilhelm-Bergner-Str. 5, 21509 Glinde, Germany

## **Abstract**

This research paper investigated if sensor data from an automatic milking rotary could be used to model cow level somatic cell count (CMSCC). CMSCC is valuable for udder health monitoring and the accuracy of individual cow udder health surveillance could be improved by predicting CMSCC between routine samplings. Data regularly recorded in the automatic milking rotary and stored in the database of the herd management system, in one German herd, were collected for analysis. The cows (Holstein-Friesian, n=372) were milked twice daily and sampled once weekly for 8 weeks for CMSCC. From the potential explanatory variables, quarter conductivity, milk flow, blood, kick-offs, not milked quarters and incomplete milkings, new variables that combined quarter data were created. To investigate how useful a variable could be before the actual CMSCC sampling event, past period values, i.e., lags of up to seven days, were added to most of the variables. Univariable Generalized additive models (GAM) were used to screen through the data to select potential explanatory variables. Furthermore, several multivariable GAM were fitted in order to compare the importance of the potential explanatory variables and to explore how the model performance would be affected by using data from various days before the CMSCC sampling event. The result of the model selection, showed that the best explanation of CMSCC was by the model incorporating all significant variables from the variable screening for the seven preceding days, including the day of the CMSCC sampling event. However, using data from only three days before the CMSCC sampling event is suggested to be sufficient to model CMSCC. Conductivity variables from combined quarters together with quarter conductivity are suggested to be important in describing CMSCC. We conclude that CMSCC can be modeled with a relatively high degree of explanation using the information routinely recorded by the milking robot.