Detection of diseases in dairy cows based on water and feed intake measurements

R.M. de Mol[1], M.A.R.M. Oomen[2], R.L.G. Zom[1], A.H. Ipema[1] and P.H. Hogewerf[1]
[1]*Wageningen UR Livestock Research, P.O. Box 65, 8200 AB Lelystad, The Netherlands*
[2]*Wageningen University, Farm Technology Group, P.O. Box 317, 6700 AH Wageningen, The Netherlands*
rudi.demol@wur.nl

Abstract

Automated detection of diseases like mastitis and lameness in dairy cows based on automated measurements of milk yield, activity and concentrate intake is possible nowadays. Automated measurement of water intake might also be useful for this purpose but this is not customary. A data set of 40 cows with measurements of water intake, concentrate intake and roughage intake and milk yield during 102 days was available to test this hypothesis.
For each cow on each day, variable measurements were compared with the expected value (the running average over the four preceding days). A univariate alert was given when the difference was outside the 80% confidence interval. A combined alert was given when two or more variables were alerted on the same day.
Eight disease cases (three mastitis, three lameness and two other disease) occurred in the experimental period. For most cases there were one or more univariate alerts in the six-days period from five days before up to and including the day of disease diagnose. The block sensitivity (percentage of detected cases) was 100% for detection based on water intake and lower when based on another variable or on combined alerts. The specificity (percentage of non-alerted healthy cow-days) was around 86% for detection based on water intake and at the same level or lower when based on other variables or combined alerts. Disease detection based on water intake has good prospects.

Keywords: Dairy cows, water intake, feed intake, monitoring

Introduction

Economic losses associated with diseases like mastitis and lameness and a decrease of reproductive performance negatively influence the profitability of the dairy farm. So it is helpful to detect diseased dairy cows in an earlier state, which makes the treatment shorter and more effective. In addition of restrict economic losses, an earlier treatment is also positive for the cow's welfare. The number of cows per available amount of labour on dairy farms increases. This trend results in less time available per animal. In order to ensure that this development is not at the expense of animal health, welfare and sustainability, the Smart Dairy Farm project (www.smartdairyfarming.nl) attempts

to develop real time decision supporting systems that recognizes diseased cows and reports them to the farmer. This project is a collaboration between research institutes, technological companies, veterinarians, practical farmers and the founders of the project, FrieslandCampina, Agrifirm and CRV. The goal of the project is to increase the number of lactations per cow in 2015 with two and increase the lifetime milk production per cow with 20,000 kg. This finally should result in an increase in animal health and sustainability of dairy farms. To realise this goal the research is focussed on three main topics, namely animal health, nutrition and fertility.
This paper is restricted to the topic animal health. The objective of this research was to investigate the perspectives of the application of automated measurements of water intake for detection of diseases in dairy cows by developing and testing a detection model based on elementary data processing.
Water intake is important for dairy cows. A sufficient supply of clean water is essential to prevent negative effects on animal health, performance and welfare (Meyer *et al.*, 2004). The main water intake (83%) is drinking. The other part comes from water included in feedstuffs and by water originated from metabolic oxidation of body tissues (Meyer *et al.*, 2004). Total water intake is often calculated by the sum of drinking water intake and ingestion of water contained in feed (neglecting the metabolic water). A dairy cow consumes on average 84 litres per day (Cardot *et al.*, 2008). Some studies concluded that water intake was a difficult variable to detect diseases, because water intake was influenced by many factors. This resulted in wide variation of intake levels under apparently similar conditions (Winchester and Morris, 1956). In contrast a study over 70 dairy cows housed in a tie stall showed that water intake had the potency to detect diseases or oestrus, because of the strong correlation with dry matter intake (Lukas *et al.*, 2008).

Material and methods

Experimental data

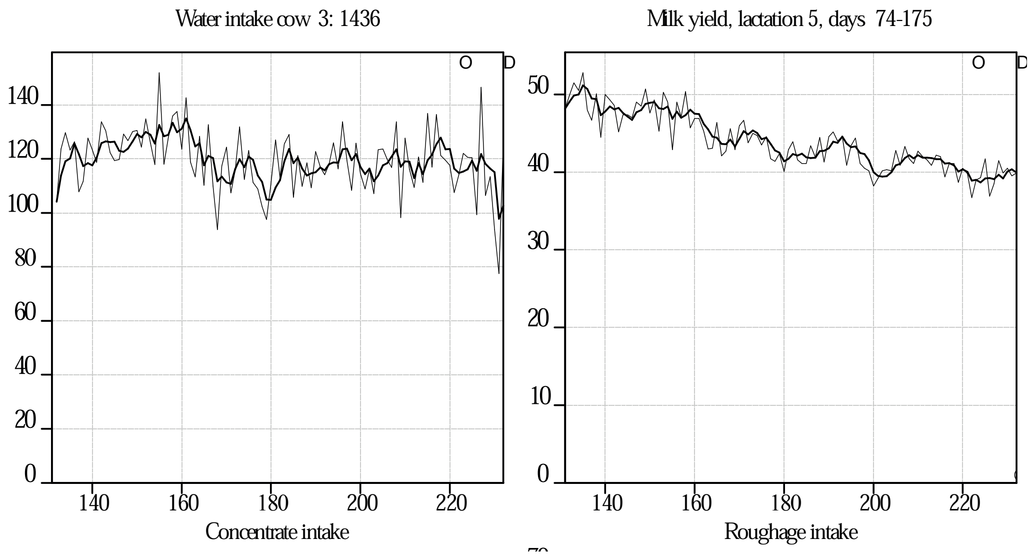
Data were available from an experiment with 40 Holstein/Friesian cows (12 first parity cows, 8 second parity cows and 20 older cows) during 102 days at the Dairy research farm "De Waiboerhoeve", of Wageningen UR Livestock Research in Lelystad, the Netherlands. The cows were housed without grazing in a free-stall barn with individual cubicles and a concrete slatted floor. The cows were fed ad libitum roughage and additional concentrate. To measure feed intake of the individual cows 32 electronic feed weighing troughs with transponder controlled access gates were used and two transponder controlled concentrate dispensers were used to determine concentrate intake. Water intake was measured by water troughs equipped with flow sensors and transponder controlled access gates. The cows were milked in a ten stands open tandem milking parlour with electronic cow identification and milk flow recording and cluster removal.

Sensor data were available for the detection model:

- daily water intake (kg);
- daily roughage intake, fresh and dry matter (kg);
- daily concentrate intake (kg);
- milk yield per milking (kg).

These variables were correlated but principal component analysis showed that all variables did have an added value.
Observations from the farm works related to health and fertility recorded in the management system were available as reference data.

<u>Data analysis</u>
Sensor data should be converted into alerts to make it applicable as management information. This was achieved by pre-processing the data and analyse this data by a detection model.
In the first step, data was visualised and corrected for errors caused by missing measurements and incorrect measurements. An example is given in Figure 1. Some of the first measurements of water and concentrate intake were obvious too low or too high. This was corrected by replacing the first values with missing values if the value was higher or lower than two times the standard deviation over the full experimental period. This method was also used for some specific dates and measurements when there had been technical problems with the sensor equipment.



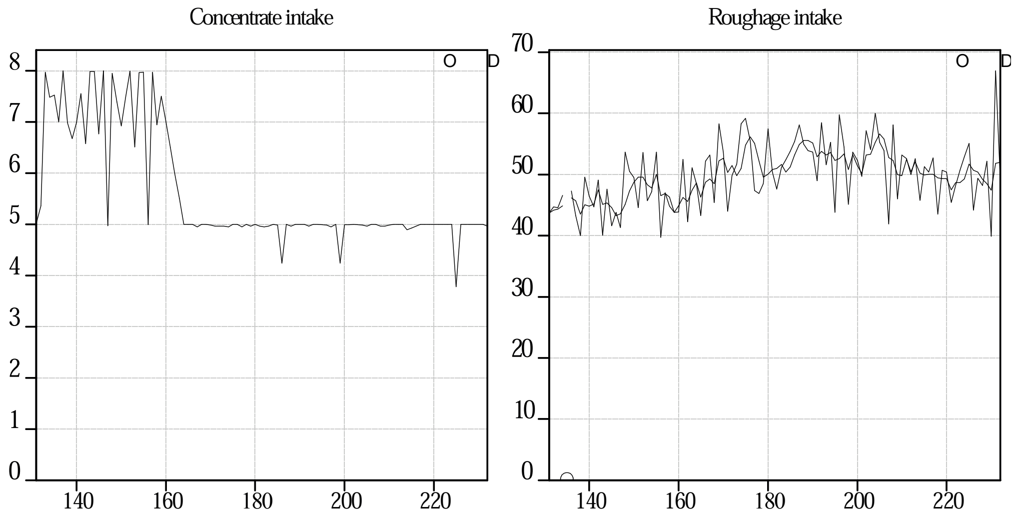


Figure 1: Example of the data visualisation for the third cow (1436) for water intake per day (top left), milk yield per day (top right), concentrate intake per day (bottom left) and roughage intake (bottom right); this cow was in oestrus at day 222 and a disease was recorded on day 232

The detection model gave cow-specific alerts for the univariate variables if the deviation between running average and actual value of the variable was larger than the standard deviation (calculated over the previous seven days) times a factor. A combined alert was given if the number of alerts for a cow in a certain period exceeded a set threshold. The running average of variable x was calculated according to:

$$\text{Running average} = (x_{(i-n)} + .. + x_i)/n \qquad (1)$$

with:

$x_{(i-n)}$, x_i = value of variable x at time (i-n), i;
i = delay time (days);
n = time span of calculation (days); n is equal to 4 in the initial model.

The deviations for water and roughage intake and milk yield were based on the difference between actual value and running average. The deviation of concentrate intakes was based on leftovers, the difference between actual intake and maximum available concentrate. An alert was given for a variable when the deviation was outside a pre-set confidence interval.
Data fusion was applied to integrate data from different variables to make more confident disease detection decisions than is possible with univariate variables (Sohn *et al.*, 2004). The used method for data fusion was a combined alert based on the sum

of univariate variables. A combined alert was given for a specific day if the total alerts exceeded the set threshold over a defined period. This number could be higher than four, because multiple alerts of the same variable on different days in the accumulation period were counted separately.

The performance of the model was determined with block sensitivity and specificity. A disease case was true positive (TP) if there were alerts within a block of six days before and up to and including the day of recorded disease, it was false negative (FN) if there where none alerts in this period. The sensitivity was defined as the percentage of detected cases TP/(TP+FN). Days with an alert outside these blocks (and more than two days after a disease case) were considered false positive (FP), otherwise such a day was true negative (TN). The specificity was the percentage of healthy cows that were classified correctly: TN/(TN+FP).

Results and discussion

During the experimental period eight disease cases were recorded: three mastitis cases, three severe lameness cases and two other diseases (damaged udder and respiratory disease) in seven cows (one cow suffered from both lameness and mastitis).

The performance of the model, expressed in sensitivity and specificity, to detect these diseases with the initial settings is summarized in Table 1. Detection based on water intake resulted in the highest sensitivity, all cases were detected. Sensitivity was lower when detection was based on other variables or on combined alerts. Also specificity was highest for detection based on water intake, the specificity was at the same level or lower when detection was based on other variables or combined alerts.

Table 1: Sensitivity (sens.) and specificity (spec.) for the univariate alerts and combined alerts with the initial settings of the model

	Water intake		Milk yield		Concentrate intake		Roughage intake		Combined alert	
	Sens. (%)	Spec. (%)	Sens. (%)	Spec. (%)	Sens. (%)	Spec. (%)	Sens. (%)	Spec. (%)	Sens. (%)	Spec. (%)
Initial model	100	86	62	86	62	76	75	86	62	85

The effects of the settings of the model on the performance have been examined by varying these settings. First the performance with varying confidence intervals was studied; results are included in Table 2. A broader confidence interval resulted in a higher sensitivity and lower specificity; a smaller confidence had a reverse effect.

Table 2: Sensitivity (sens.) and specificity (spec.) with varying confidence intervals (initial setting in bold) for the univariate alerts and combined alerts

Confidence interval (%)	Water intake		Milk yield		Concentrate intake		Roughage intake		Combined alert	
	Sens. (%)	Spec. (%)	Sens. (%)	Spec. (%)	Sens. (%)	Spec. (%)	Sens. (%)	Spec. (%)	Sens. (%)	Spec. (%)
60	100	77	88	77	62	68	88	78	75	73
70	100	82	75	82	62	72	88	82	75	80
80	100	86	62	86	62	76	75	86	62	85
90	88	91	50	91	62	80	62	91	50	92
95	88	94	38	94	62	84	62	93	50	95
99	62	97	25	97	62	89	25	96	25	98

The influence of time span of the standard deviations was determined by varying the time span between 3 and 13 (Table 3). A time span of n days corresponded with a period from n days before till one day before the current day. The detection performance was not strongly influenced by this setting.

Table 3: Sensitivity (sens.) and specificity (spec.) with varying time span for the standard deviations (initial setting in bold) for the univariate alerts and combined alerts

Time span std. dev. (day)	Period (day)	Water intake		Milk yield		Concentrate intake		Roughage intake		Combined alert	
		Sens. (%)	Spec. (%)	Sens. (%)	Spec. (%)	Sens. (%)	Spec. (%)	Sens. (%)	Spec. (%)	Sens. (%)	Spec. (%)
3	-3/-1	100	83	62	83	75	71	75	83	50	80
5	-5/-1	88	84	62	84	75	74	75	85	50	83
7	-7/-1	100	86	62	86	62	76	75	86	62	85
9	-9/-1	88	86	50	86	62	77	75	87	50	87
11	-11/-1	88	87	50	87	62	78	62	88	50	88
13	-13/-1	88	87	50	87	62	79	75	88	62	88

The influences of the time span for the running average on the performance were determined by varying the time span from 1 to 5 days (Table 4). The running average was used to establish a standard value where the current value is to be compared with. Also this setting did only have minor effects on the detection results.

Table 4: Sensitivity (sens.) and specificity (spec.) with varying time span for the running average (initial setting in bold) for the univariate alerts and combined alerts

Time span (day)	Period (day)	Water intake		Milk yield		Concentrate intake		Roughage intake		Combined alert	
		Sens.	Spec.	Sens.	Spec.	Sens.	Spec.	Sens.	Spec.	Sens.	Spec.
1	-1/-1	88	85	75	85	62	76	88	85	75	85
2	-2/-1	88	86	75	86	62	76	75	86	62	85
3	-3/-1	88	86	62	86	62	76	75	86	62	85
4	-4/-1	100	86	62	86	62	76	75	86	62	85
5	-5/-1	88	86	75	86	62	76	75	87	62	86

The influence of delay time of running average on the performance of the system was determined by varying the delay time from zero until four days, while the time span was fixed on four days (Table 5). The setting influences also the standard where a new value was compared with. Also this setting had only minor effects.

Table 5: Sensitivity (sens.) and specificity (spec.) with varying delay time for the running average (initial setting in bold) for the univariate alerts and combined alerts

Delay time (day)	Period (day)	Water intake		Milk yield		Concentrate intake		Roughage intake		Combined alert	
		Sens.	Spec.	Sens.	Spec.	Sens.	Spec.	Sens.	Spec.	Sens.	Spec.
0	-3/0	75	91	50	91	62	76	62	91	62	91
1	-4/-1	100	86	62	86	62	76	75	86	62	85
2	-5/-2	75	86	50	86	62	76	75	87	62	85
3	-6/-3	75	86	62	86	62	76	62	87	50	85
4	-7/-4	75	86	62	86	62	76	75	87	62	84

The influence of time period and threshold of accumulated alerts on the performance of the combined alert was also studied. The time period was varied between 0 and 2 days and threshold was varied from 1 to 4 cumulative alerts. The results are included in Table 6. Both settings had a great impact on the results. A lower threshold gave a higher sensitivity and lower specificity; a higher threshold had reverse effects. A broader period also gave a higher sensitivity and lower specificity.

Table 6: Sensitivity (sens.) and specificity (spec.) of combined alerts with varying thresholds and periods (initial setting in bold).

	Period: n		Period: n-1...n		Period: n-2...n	
Threshold	Sens. (%)	Spec. (%)	Sens. (%)	Spec. (%)	Sens. (%)	Spec. (%)
1	100	50	100	28	100	17
2	62	85	88	60	100	41
3	25	96	62	83	88	66
4	0	100	38	95	50	83
5	-	-	25	99	38	93
6	-	-	-	-	25	98

The timeliness of alerts for all eight disease cases was examined. In Table 7 for each case the number of days between first alert and observation is given. Automated detection, especially when based on water intake, resulted in early warnings for diseases.

Table 7: Days of earlier detection by model with initial settings per univariate variables and combined alerts (a minus sign is given when no alert was given)

Cow nr	Type of disease	Number of days between first alert and observation				
		Water intake	Milk yield	Concentrate intake	Roughage intake	Combined alert
2382	Mastitis	1	4	-	2	-
3492	Mastitis	4	-	5	-	-
3667	Mastitis	2	4	-	-	-
3667	Lameness	2	1	1	5	1
3674	Lameness	1	1	5	2	1
9490	Lameness	1	-	2	1	1
1436	Other disease	2	-	-	2	2
3468	Other disease	5	5	3	5	5

It was decided to use quite simple data processing techniques for the detection model. The sensitivity can be at a reasonable level, the specificity might be too low for practical application of the detection model. Results might be improved by applying more advanced data processing techniques, but that was outside the scope of this research.

It is a known fact that the settings influence the performance of the model: lower thresholds give an increased sensitivity and a decreased specificity; higher thresholds have the reverse effect. It is also a choice of the end-user which settings are preferred.

Only eight disease cases were recorded in the experiment. Therefore the results should be taken with reservation and further testing on larger data sets is advised.

Conclusions

Water intake can be very useful for automated detection of diseases. Water intake measurements are easier to realize than roughage intake measurements and are therefor a realistic option for practical application. For a data with 40 cows during 102 days detection based on water intake gave better results than detection based on milk yield, concentrate intake and roughage intake. All eight disease cases were detected, the specificity was 86%.

Acknowledgements

This research was supported by the Dutch research program Smart Dairy Farming, which is financed by Friesland Campina (Amersfoort, the Netherlands), CRV (Arnhem, the Netherlands), Agrifirm (Apeldoorn, the Netherlands), Dairy Valley (Leeuwarden, the Netherlands), Investment and Development Agency for the Northern Netherlands (Groningen, the Netherlands), the Dutch Dairy Board (Zoetermeer, the Netherlands) and the ministry of Economic Affairs, Agriculture and Innovation, Pieken in de Delta (Den Haag, the Netherlands).

References

Cardot, V., Le Roux, Y. & Jurjanz, S. (2008). Drinking behavior of lactating dairy cows and prediction of their water intake. Journal of Dairy Science 91(6): 2257-2264.

Lukas, J. M., J. K. Reneau and J. G. Linn (2008). Water Intake and Dry Matter Intake Changes as a Feeding Management Tool and Indicator of Health and Estrus Status in Dairy Cows. Journal of Dairy Science 91(9): 3385-3394.

Meyer, U., Everinghoff, M., Gädeken, D. & Flachowsky, G. (2004). Investigations on the water intake of lactating dairy cows. Livestock Production Science 90(2–3): 117-121.

Sohn, H., C. R. Farrar, F. M. Hemez, D. D. Shunk, D. W. Stinemates, B. R. Nadler and Others (2004). A review of structural health monitoring literature: 1996--2001. Los Alamos National Laboratory, Los Alamos, NM.

Winchester, C. F. & Morris, M. J. (1956). Water Intake Rates of Cattle. Journal of Animal Science 15(3): 722-740.