Detection of Estrus and Mastitis: Field Performance of a Model

R. M. de Mol, W. Ouweltjes, G. H. Kroeze, M. M. W. B. Hendriks

Abstract. *A new detection model ('IMAG model') for estrus and mastitis in dairy cows was tested on four farms during several years. Such a test is necessary because information is lacking about the performance of detection models under field conditions. The test gave insight into the field performance of the IMAG model and the results were compared with the results of older models and with the results predicted by experts. Sensor data of milk yield, milk temperature, electrical conductivity of milk and animal activity were the inputs for the IMAG model. The IMAG model is based on time series analysis combined with a Kalman filter. This structure yields cow–dependent model parameters and combines data of different sensors. Results were compared with the manufacturer's model (supplied with the sensors), based only on exponential smoothing on data from one sensor. The sensor equipment differed between farms. The sensitivity (percentage of estruses detected) for estrus varied from 63 to 80%, depending on the threshold used. Specificity (non–estruses not detected as estrus) varied from 94 to 98%. The sensitivity for clinical mastitis varied from 55 to 80%, depending on the threshold used. The specificity for mastitis varied from 94 to 99%. Significant differences existed between farms, in sensitivity for estrus and mastitis. The applied equipment could only partially explain the differences in estrus and mastitis detection results between farms. No relation between stage of lactation and activity level was found, although a lower activity level in the first period of lactation might be expected. The main conclusion is that a detection model can give good results, but only if the equipment is working properly. The new model outperformed the manufacturer's model.*

Keywords. *Detection, Estrus, Mastitis, Models.*

Sensors can be used for livestock monitoring (Frost et al., 1997; Geers, 1994; Schlünsen et al., 1987). The sensor data are input for detection models for estrus and mastitis. Sensor systems (sensors and detection model) for the detection of estrus and mastitis are commercially available. The sensors measure activity (for estrus), electrical conductivity (for mastitis), or yield and temperature of milk (for both). Activity is measured by a pedometer attached to a collar around the neck or to the leg of the cow. Electrical conductivity is measured in the milk flow during milking; either in total milk or in quarter milk. Milk meters record the yield. Milk temperature is related to body temperature and is measured in the milk flow. Some data from experiments with sensor systems are available now. Gil et al. (1997) detected 79% out of 38 estrus cases by temperature increase. Koelsch et al. (1994) detected ca 70% out of 29 estrus cases by three activity–comparison procedures. In a literature review, Lehrer et al. (1992) reported a success rate of estrus detection in the range 60 to 100%. Mastitis in 22 cows in first lactation was clearly identifiable by conductivity and yield changes (Graupner and Barth, 1994). Maatje et al. (1992) detected 100% of 25 clinical mastitis cases by increased conductivity. Milner et al. (1996) detected 100% of 12 clinical mastitis cases after experimental infection with *Streptococcus uberis* by changes in the conductivity, and likewise 95% of 19 cases after infection with *Staphylococcus aureus*. Hamann and Zecconi (1998), concluded in their evaluation that the published information on electrical conductivity in milk, as a mastitis indicator, comprises quite variable results; the published information is too varied to justify the claim that mastitis can be detected, under field conditions, by electrical conductivity measurements.

Detection results under field conditions are not available. There might be a great discrepancy between experimental detection results and results in the field. A lower sensitivity was found in experiments with low clinical mastitis prevalence, as in the field, than in experiments with a high prevalence, as in experimental circumstances (Hamann and Zecconi, 1998). Experts were consulted to evaluate the implication of research results for field use, the effect of combinations of sensors on the sensitivity and specificity of the detection of estrus and mastitis was elicited (Van Asseldonk et al., 1998). The results of this consultation can be compared with the performance in the field, on a farm equipped with activity meters, milk–yield sensors, conductivity sensors, and milk–temperature sensors. In the opinion of the experts (Van Asseldonk et al., 1998), the sensitivity (percentage of all cases detected) on a farm equipped with these sensors would be 81% for estrus detection, the specificity (percentage of non–estrus detected

Article was submitted for review in June 1999; approved for publication by the Information & Electrical Technologies Division of ASAE in January 2001.

The authors are **Rudi M. de Mol,** Researcher,Institute of Agricultural and Environmental Engineering (IMAG), 6700 AA Wageningen, The Netherlands, **Wijbrand Ouweltjes,** Researcher,Research Institute for Animal Husbandry (PV), Lelystad, The Netherlands, **Gerrit H. Kroeze,** Researcher,Institute of Agricultural and Environmental Engineering (IMAG), Wageningen, The Netherlands, and **Margriet M. W. B. Hendriks,** Researcher, Institute of Agricultural and Environmental Engineering (IMAG), and Centre for Biometry Wageningen, Wageningen, The Netherlands. **Corresponding author:** R. M. de Mol, Institute of Agricultural and Environmental Engineering (IMAG), P.O. Box 43, Wageningen, The Netherlands; phone: +31317476459; fax: +31317425670; e–mail: R.M.deMol@imag.wag–ur.nl.

as non–estrus) 90%. The sensitivity for mastitis would be 71% and the specificity 86%.

The expected performance in the field of existing sensor systems may hamper a rapid introduction as a management tool. A new detection model for estrus and diseases in dairy cattle was developed by IMAG (De Mol et al., 1999) to process the measured variables in a combined way. Other models use only single variables for the detection: mostly activity for estrus and conductivity for mastitis. The IMAG model is based on time series models for yield, temperature and electrical conductivity of milk, and for the cow's activity. The parameters of the time series models are fitted on–line by a Kalman filter for each cow after each milking. The IMAG model combines the variables and generates cow–specific alerts. The model calculates the deviation between the expected values and measured values of the sensor measurements, for each cow and each milking. An alert is given when a combination of deviations falls outside a certain confidence interval, which can be 95, 99, or 99.9%. In a previous research the IMAG model was tested on two experimental farms (De Mol et al., 1997; overall results in table 1). For example, with the 99% confidence interval, 79% of the estrus cases are detected (one or more alerts are given) and there is no alert for 96.4% of the milkings outside estrus periods.

The detection methods and the equipment used may influence the results. For example, activity can be measured by pedometers attached to the leg or to the neck, and conductivity can be measured for the milk of each quarter of the udder or for total milk. The activity of cows, which is used for estrus detection, may be influenced by the lactation phase. A test in the field may indicate the prospects of the application of sensors and detection models in dairy husbandry. Therefore testing was done at four farms for which sensor data as well as reference data were available. These data have been used as input for the detection model. The results will be compared with the results for the same data of the manufacturer's model (supplied with the sensors) and with the results of the IMAG model for different data in the previous research (De Mol et al., 1997).

The objectives of this research were to get insight into the performance of the IMAG model in the field and to compare the results with those of the manufacturer's model and with the prediction of the experts (Van Asseldonk et al., 1998).

Table 1. Sensitivity and specificity for estrus and mastitis by the IMAG model, on two experimental farms and with three confidence intervals.[a]

Confidence Interval (%)	Estrus Sensitivity (%)[b] (537 cases)	Estrus Specificity (%)[c] (60,665 milkings)	Clinical Mastitis Sensitivity (%)[b] (53 cases)	Mastitis Specificity (%)[c] (6495 milk-ings)
95	87	93.7	76	95.2
99	79	96.4	59	98.1
99.9	74	97.8	36	99.4

[a] The figures are based on data by De Mol et al. (1997) using different methods for sensitivity and specificity (see text).
[b] Proportion of positive cases with one or more alerts.
[c] Proportion of negative milkings without an alert.

Materials and Methods

Data collection

Data were collected on four farms equipped with commercially available sensors. These farms, unlike the farms used by De Mol et al. (1997), are managed like commercial farms but they are also used for extension purposes and practical research. All farms had loose housing systems, and the cows were milked twice a day. The farm names are based on the equipment (table 2). On farm ALCQ, Activity was measured by Leg transponders and Conductivity was measured in Quarter milk. On farm ANCQ1, as well as ANCQ2, Activity was measured by Neck transponders and Conductivity was measured in Quarter milk. On farm ALCT, Activity was measured by Leg transponders and Conductivity was measured in Total milk. The periods of data collection and the average number of cows milked are given for each farm in table 3. The dates, but not the time of the day, of observed cases of estrus and of clinical mastitis were recorded. Milk samples for cell counts were taken once in three weeks on ANCQ1, ANCQ2, and ALCT; and once a week on ALCQ.

The farms differed not only in data collection period (table 3) and equipment (table 2), but also in housing system, management practice, geographical location and breed. The cows went to pasture during the summer period on ANCQ1, ANCQ2, and ALCT; cows were kept indoors the whole year round on ALCQ. The frequency of estrus and clinical mastitis in the data collection period is given in table 4 for all farms. Estrus and clinical mastitis cases were based on visual detection. Clinical mastitis was treated after observation. The observed estrus frequency depends on the quality of estrus detection, the insemination success rate and possibly on other farm–dependent factors. The observed mastitis frequency is also farm–dependent; the mean mastitis frequency in practice is approximately one case in 2000 to 2400 milkings (Brand et al., 1996). On ALCQ and ANCQ1, the mastitis frequency was above the normal level. Other farm characteristics, such as the attitude toward sensor usage and management goals, may also be important but these are difficult to quantify.

Table 2. Sensor equipment used on the four farms: manufacturers (X, Y, Z) and sensor types used to assess the cow's activity and the electrical conductivity of milk.

Farm	Milk Yield	Milk Temperature	Cow's Activity	Conductivity of Milk
ALCQ	X	X	X: leg transponder	X: quarter milk
ANCQ1	X	X	X: neck transponder	X: quarter milk
ANCQ2	X	X	X: neck transponder	X: quarter milk
ALCT	Y	not present	Z: leg transponder	Z: total milk

Table 3. Description of the size of the data sets of the four farms.

Farm	From	Until	No. of Milkings	No. of Cows Milked	Avg. No. of Cows per Milking
ALCQ	16 Nov. '95	6 Apr. '97	1015	41 949	41
ANCQ1	4 Oct. '94	30 Mar. '97	1816	124 919	69
ANCQ2	1 Feb. '96	27 Mar. '97	839	55 914	67
ALCT	4 July '95	20 May '97	1350	157 755	117

Table 4. Observed frequency of cases of estrus and clinical mastitis on the four farms.

Farm	Estrus Frequency (1/milkings)	Clinical Mastitis Frequency (1/milkings)
ALCQ	1/236	1/999
ANCQ1	1/234	1/1288
ANCQ2	1/206	1/2542
ALCT	1/336	1/3093

DETECTION MODEL

Sensor measurements and other cow information (such as calving dates) of all four farms were stored in the database of the IMAG model, using the same format as in De Mol et al. (1999). These data were inputs for the detection model. Alerts for estrus and mastitis are the main outputs of the model. The structure of the model is depicted in figure 1.

After each milking an input file is built with for each cow:

- sensor measurements (yield, temperature, conductivity, activity, and concentrate intake) for the actual milking and some (depending on the variable, 2–4) previous milkings;
- cow–dependent parameters of the time series model (TSM) for each variable, and the cow–dependent variance–covariance of these parameters;
- cow information: days in lactation, the cow's status (calved, in estrus, inseminated, or in calf), days since last estrus observation.

For each variable (yield, temperature, conductivity, and activity) the parameter values and some previous sensor measurements are used to calculate the expected sensor measurement value for a specific cow and a specific milking. These results are compared with the actual sensor measurements. An alert is given when a combination of deviations falls outside a confidence interval, using the variance–covariance of the TSM parameters. Thresholds for alerts correspond with the chosen confidence interval: 95, 99, or 99.9%. The following combinations are used to generate alerts:

- for estrus: a combination of increased activity, decreased yield, and increased temperature;
- for mastitis: a combination of increased conductivity, decreased yield, and increased temperature;

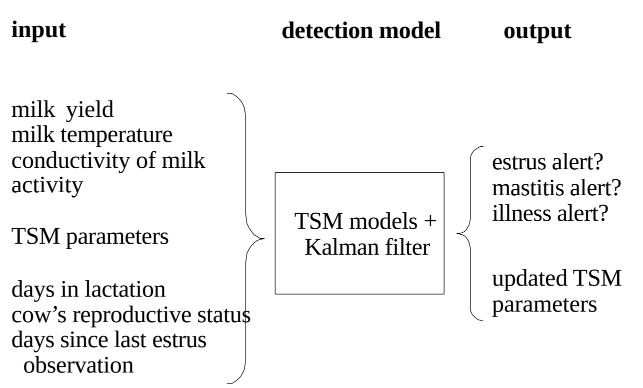

Figure 1. The structure of the detection model, based on time series models (TSM) and a Kalman filter with a description of input and output (see text for explanation).

- for illness: a combination of decreased yield, increased temperature, and decreased activity.

The Kalman filter is a statistical technique to estimate the state of a system on–line. In the IMAG model, the filter is used to update the parameter values and variance–covariance matrix, which are used for the next milking. For the first milking in a lactation of a cow, average values for the parameters and variance–covariance are used. The model output consists of the alerts (if any) and the updated parameters and variance–covariance values. These alerts, parameters, and variance–covariance values are contained in an output file, which is transferred back to the database of the farm. More details, including formulas, are given by De Mol et al. (1999).

The IMAG model is flexible in the number of variables used and in the input settings, therefore only minor adaptations were needed for application in the present research. Milk temperature was not recorded on ALCT. Thus estrus alerts for ALCT were only based on activity and yield, whereas mastitis alerts were only based on conductivity and yield. When conductivity of total milk was recorded, only one variable for conductivity was used in the model (instead of four). Parameter values were updated per cow and per milking by the Kalman filter, so there were no farm–dependent parameter settings needed. Starting values for the parameter values were needed for the first milking in a lactation of a cow. The same starting parameter values were used for all farms. These parameter values were adapted to a cow–specific level after the first few milkings in lactation.

The comparison of the detection performance on the four experimental farms was based on:

- sensitivity to estrus: the proportion of all estrus cases observed, with one or more estrus alerts;
- specificity to estrus: the proportion of all milkings of cows that are not in estrus, without an estrus alert;
- sensitivity to clinical mastitis: the proportion of all cases of clinical mastitis observed, with one or more mastitis alerts;
- specificity to mastitis: the proportion of all milkings of cows that never showed mastitis in the experimental period, without mastitis alert.

For three farms the results based on estrus and mastitis alerts of the IMAG model were compared with the results of the manufacturer's model that is supplied with the equipment. The manufacturer's model is based on exponential smoothing on single variables: for estrus on activity, for mastitis on conductivity. Details of the manufacturer's model are not available for commercial reasons.

Comparison of the results on the four farms, with the previous results on two experimental farms was not straightforward because in the previous test more reference data were available; samples of progesterone levels in milk, and bacteriological examinations of milk, were taken. Moreover, milk samples for cell counts were taken more frequently in the previous test.

Alerts are based on the deviation between actual and expected values of a variable. A variable is considered indeterminable when:

- the actual value is missing due to measurement errors (no measurement data available or outlying values considered as measurement errors; see table 5), or

- the expected value of the yield, temperature, or activity is undefined due to start–up effects (first two milkings in a new lactation or after a sequence of measurement errors); the expected value of conductivity is based on average values if data of one of the two previous milkings are missing.

Indeterminable variables of the manufacturer's model are indicated by the absence of output of the model. The detection results were influenced by the measurement errors indicated as indeterminable variables. Estrus and mastitis cases associated with measurement errors were difficult to classify. Therefore, sensitivity and specificity for estrus were based only on cases without indeterminable activity; sensitivity and specificity for mastitis were based only on cases without indeterminable conductivity. Cases with indeterminable yield or temperature were still used in the tests.

Estrus Detection

Results were based on dates of recorded cases of estrus (dates of observed estrus or insemination) on the farms with some restrictions:

- dates were used only if sensor measurements of the cow were available for the specific period;
- insemination dates were discarded if estrus was observed the previous day;
- estrus and insemination dates were discarded if they did not comply with the estrus cycle of 3 weeks. Dates were sometimes discarded in retrospect by assessing the estrus cycle of a cow. E.g. an estrus case within one week before a next estrus date, was not used in the tests.

Only the dates of estrus cases were recorded and not the hour. Therefore, an estrus case was considered true positive (TP) if one or more estrus alerts were given in a period of five milkings: two milkings on the day of estrus, two milkings on the previous day and the morning milking on the next day. Alerts on the estrus day were TP, alerts on the previous day were TP because estrus signs might already be present, and alerts on the morning milking of the next day were TP because the estrus may have started (and been observed) during the evening on the estrus date. An estrus case without estrus alerts in this five–milkings period was classified as False Negative (FN). Milkings outside estrus periods were False Positive (FP) if an estrus alert was given, otherwise they were True Negative (TN).

Estrus detection results may depend on the stage of lactation. Cows in early lactation often have a negative energy balance and this may influence the cow's activity as well as the estrus detection results (Brand et al., 1996). The hypothesis of a relation between stage of lactation and activity level was tested by dividing the lactation into three periods. Period 1: from the calving day up to day 30, period 2: day 31 up to day 75, and period 3: day 76 and further. A generalized linear mixed model, fitted by iterative re–weighted residual maximum likelihood algorithm (Engel and Keen, 1994), was imposed to analyze the influence of the lactation period and other factors on the activity level (change in activity/hour):

$$y_{dglpcn} = \exp\left(m + d_d + g_g + i_{dg} + l_l + p_p + j_{lp} + c_{lc}\right) + \varepsilon_{dglpcn} \quad (1)$$

with

- $\exp(x)$ = e^x
- y_{dglpcn} = the *n*–th observation of the activity level of cow *c* with part of day *(d)*, with grazing system *g*, in lactation *l* and lactation period *p*
- m = the overall mean
- d_d = fixed effect of part of day *(d)*, $d = 1$ (nighttime) or 2 (daytime)
- g_g = fixed effect of grazing system *(g)*, $g = 0$ (in stall) or 1 (in pasture)
- i_{dg} = interaction effect between part of day *d* and grazing system *g*
- l_l = fixed effect of lactation number *(l)*, $l = 1$ (first lactation), 2 (second), or 3 (third or higher)
- p_p = fixed effect of lactation period *(p)*, $p = 1$, 2, or 3
- j_{lp} = interaction effect between lactation number *(l)* and lactation period *(p)*
- c_{lc} = random effect of cow *c* in lactation *(l)*
- ε_{dglpcn} = random error

These factors were chosen to clarify the effect of the lactation period separate from other factors that may influence the cow's activity.

Table 5. Acceptable ranges for the milk variables.

Variable	Lower Limit	Upper Limit
Yield (L)	0	99.99
Temperature (°C)	30	45
Conductivity (mS/cm)	3	12

Mastitis Detection

Automated detection of mastitis has two applications: detection of subclinical mastitis and early detection of clinical cases. In our study, the first application was not examined because appropriate reference data were not available.

Each observed case of clinical mastitis was TP or FN. In a TP case, one or more mastitis alerts were given on one of the eight milkings up to the day mastitis was observed (the two milkings on the mastitis day and all milkings on the three preceding days). Three preceding days were included because mastitis signs might already be present. Alerts given earlier (more than three days before the observation) might not be related with the actual case. Cases without any mastitis alert were classified as FN.

The IMAG model also generates illness alerts. These illness alerts were based on yield, temperature, and activity, and indicate that the cow might suffer from illness (not necessarily mastitis). In addition, the sensitivity for mastitis based on these illness alerts was calculated, taking a mastitis case as TP when one or more illness alerts were given in the mastitis period, otherwise FN. The illness alerts in this case were used as mastitis alerts, however they were not based on conductivity.

Calculating the specificity for mastitis was complicated because the real mastitis status of a cow was not always known. A cow without clinical mastitis might suffer from subclinical mastitis. The specificity was determined only with non–mastitis cows, i.e.:

- no case of clinical mastitis in the data collection period
- cell counts always below 500,000 cells/cm^3
- cow milked at least 300 times

- number of samples of cell counts at least 10

Mastitis alerts for non–mastitis cows were considered FP and were used to calculate the specificity.

RESULTS

INDETERMINABLE VARIABLES

Figure 2 gives the variables that were classified as indeterminable by the IMAG model, as a percentage of the number of cows milked, for each farm and each variable. The indeterminable variables were mostly caused by the absence of measurement values. A minor part (less than 5% of the indeterminable variables) was due to values observed outside the ranges defined in table 5. Exceptions to this rule were the milk temperature for ALCQ and ANCQ1, and conductivity (all quarters) for ANCQ1.

Indeterminable variables were found over the whole data collection period. The percentage of milkings without any indeterminable variable differed between farms: 82 for ALCQ, 72 for ANCQ1, 68 for ANCQ2, and 67 for ALCT. For milkings with indeterminable variables, the number of indeterminable variables varied (fig. 3). This number varied between 1 and 7 on ALCQ, ANCQ1, and ANCQ2, but 3 was the maximum on ALCT because only three variables were recorded (yield, activity, and conductivity of total milk). For milkings with one or more indeterminable variables, the occurrence of one indeterminable variable was predominant, with ANCQ2 as an exception. In indeterminable variables of ANCQ2, there were mostly four conductivity measurements missing (one for each quarter).

ESTRUS

Estrus detection results depended on the model used and the farm (tables 6 and 7). Choosing a higher threshold in the IMAG model (a higher confidence interval) gave a lower sensitivity but a higher specificity, and vice versa.

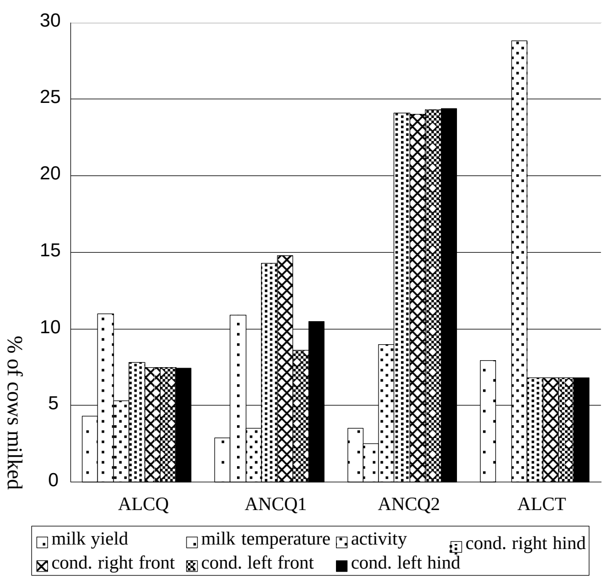

Figure 2. Indeterminable variables as a percentage of the number of cows milked, for each farm and each variable (cond. = conductivity).

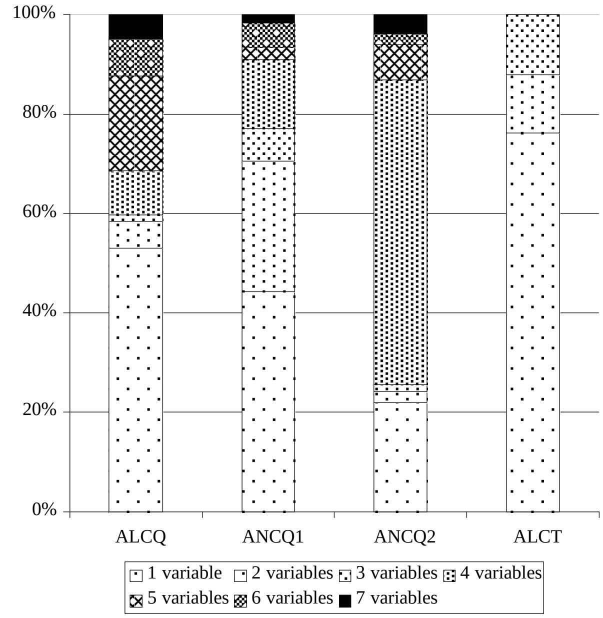

Figure 3. The distribution of the milkings with one or more indeterminable variables over the number of indeterminable variables per milking, for each farm.

The difference between farms for the detection of estrus was tested with a logistic regression model (Genstat, 1993). There were differences (P < 0.05) between farms in sensitivity and specificity of the IMAG model. The difference in sensitivity was also significant for the

Table 6. Number of cases found with the IMAG model (with three confidence intervals, % in brackets) and with the manufacturer's model (not available for ALCT) for all farms together and each farm separately.

Farm (No. Cases)	Model	TP[a]	FN[b]	?/TP[c]	?/FN[d]	Sensitivity (%)[e]
All farms	IMAG (95)	919	236	98	199	80[f]
(1452)	IMAG (99)	820	335	83	214	71[f]
	IMAG (99.9)	726	429	67	230	63[f]
(982)	Manufacturer	567	326	26	63	63[f]
ALCQ	IMAG (95)	138	11	25	4	93
(178)	IMAG (99)	134	15	25	4	90
	IMAG (99.9)	129	20	25	4	87
	Manufacturer	132	21	16	9	86
ANCQ1	IMAG (95)	395	110	11	17	78
(533)	IMAG (99)	350	155	8	20	69
	IMAG (99.9)	313	192	5	23	62
	Manufacturer	324	187	3	19	63
ANCQ2	IMAG (95)	152	66	25	28	70
(271)	IMAG (99)	127	91	21	32	58
	IMAG (99.9)	101	117	15	38	46
	Manufacturer	111	118	7	35	48
ALCT	IMAG (95)	234	49	37	150	83
(470)	IMAG (99)	209	74	29	158	74
	IMAG (99.9)	183	100	22	165	65

[a] True positive estrus cases.
[b] Number of false negative cases.
[c] Number of true positive cases with indeterminable activity.
[d] Number of false negative cases with indeterminable activity.
[e] Sensitivity (TP/TP+FN).
[f] Farm effect significant at P < 0.05.

Table 7. Number of milkings found with the IMAG model (with three confidence intervals, % in brackets) and with the manufacturer's model (not available for ALCT) for all farms together and each farm separately.

Farm (No. Milkings)	Model	TN[a]	FP[b]	?[c]	Specificity (%)[d]
All farms	IMAG (95)	280,704	19,138	54,832	93.6[e]
(354,674)	IMAG (99)	289,866	9976	54,832	96.7[e]
	IMAG (99.9)	294,292	5550	54,832	98.1[e]
(206,907)	Manufacturer	196,626	4576	5705	97.7[f]
ALCQ	IMAG (95)	33,230	2129	1911	94.0
(37,270)	IMAG (99)	34,227	1132	1911	96.8
	IMAG (99.9)	34,756	603	1911	98.3
	Manufacturer	35,591	424	1255	98.8
ANCQ1	IMAG (95)	107,998	6620	4144	94.2
(118,762)	IMAG (99)	111,632	2986	4144	97.4
	IMAG (99.9)	113,211	1407	4144	98.8
	Manufacturer	113,771	2977	2014	97.5
ANCQ2	IMAG (95)	43,622	2696	4557	94.2
(50,875)	IMAG (99)	45,001	1317	4557	97.2
	IMAG (99.9)	45,661	657	4557	98.6
	Manufacturer	47,264	1175	2436	97.6
ALCT	IMAG (95)	95,854	7693	44,220	92.6
(147,767)	IMAG (99)	99,006	4541	44,220	95.6
	IMAG (99.9)	100,664	2883	44,220	97.2

[a] True negative milkings for estrus.
[b] Number of false positive milkings.
[c] Number of milkings with indeterminable activity.
[d] Specificity (TN/TN+FP).
[e] Farm effect significant at $P < 0.05$.
[f] Significance of farm effect not determined.

manufacturer's model. It was not possible to test the differences in specificity for the manufacturer's model with the logistic regression model, because the major part of the deviance was caused by a cow effect (a few cows had a high number of FP alerts). For farms ANCQ1 and ANCQ2, having the same sensor equipment, the pair wise differences in sensitivity and specificity were all significant, except specificity in case of the 95% confidence interval. The sensitivity on ALCQ was higher ($P < 0.05$) than the sensitivity on the other three farms; and the sensitivity on ANCQ2 was lower ($P < 0.05$).

The estrus sensitivity, as given in table 6, is based only on cases without indeterminable activity. Sensitivity based on all cases, with or without indeterminable activity, can also be calculated from the figures in table 6: [(TP+?/TP)/ (number of cases)] × 100. The sensitivity based on all cases on all farms varied between 70% for the IMAG model with a 95% confidence interval (919 + 98 out of 1452 cases) and 55% for a 99.9% confidence interval. These lower percentages express the estrus sensitivity presented to the farmer by the sensor system. The percentages in table 6 express the sensitivity found by the IMAG model.

Fitting the generalized linear mixed model for the activity level (eq. 1), showed that the activity level is not always lower in the first lactation period than in the next periods (table 8), as might be expected due to a negative energy balance in early lactation. The predicted mean activity level was lower for the first lactation period on farms ANCQ1, ANCQ2, and ALCT, but the differences were relatively small. The predicted mean activity level was even higher in the first lactation period for farm ALCQ. The predicted mean activity level generally decreased as the lactation number increased (data not shown).

Mastitis

Results for clinical mastitis detection were calculated based on mastitis alerts (table 9) and based on illness alerts (table 10). Mastitis specificity results were only based on mastitis alerts (table 11).

A logistic regression model was used to test the differences between farms (significance level 0.05). The differences in sensitivity (table 9) of mastitis alerts produced by the IMAG model were all significant. The differences in sensitivity between farms of the manufacturer's alerts (table 9) and illness alerts by the IMAG model (table 10) were not significant. The differences in specificity between farms (table 11) were significant for the mastitis alerts by the IMAG model. It was not possible to test the difference in specificity of the manufacturer's model with the logistic regression model because the major part of the deviance was caused by

Table 8. Predicted mean activity level (change in activity/hour) for each lactation period and for each farm

Farm	Days in Lactation ≤ 30	> 30 and ≤75	> 75
ALCQ	4.03	3.81	3.62
ANCQ1	1.44	1.63	1.68
ANCQ2	1.99	2.21	2.30
ALCT	1.16	1.24	1.45

Table 9. Number of mastitis cases found with the IMAG model (with three confidence intervals, % in brackets) and with the manufacturer's model (not available for ALCT) for all farms together and each farm separately.

Farm (No. Cases)	Model	TP[a]	FN[b]	?/TP[c]	?/FN[d]	Sensitivity (%)[e]
All farms	IMAG (95)	75	20	77	40	79[f]
(212)	IMAG (99)	64	31	66	51	67[f]
	IMAG (99.9)	51	44	56	61	54[f]
(161)	Manufacturer	47	97	5	12	33
ALCQ	IMAG (95)	21	4	12	5	84
(42)	IMAG (99)	20	5	10	7	80
	IMAG (99.9)	18	7	10	7	72
	Manufacturer	11	23	2	6	32
ANCQ1	IMAG (95)	48	6	37	6	89
(97)	IMAG (99)	40	14	35	8	74
	IMAG (99.9)	31	23	31	12	57
	Manufacturer	32	57	3	5	36
ANCQ2	IMAG (95)	2	1	12	7	67
(22)	IMAG (99)	2	1	8	11	67
	IMAG (99.9)	1	2	5	14	33
	Manufacturer	4	17	0	1	19
ALCT	IMAG (95)	4	9	16	22	31
(51)	IMAG (99)	2	11	13	25	15
	IMAG (99.9)	1	12	10	28	8

[a] True positive clinical mastitis cases.
[b] False negative mastitis cases.
[c] Number of true positive cases with indeterminable conductivity.
[d] Number of false negative cases with indeterminable conductivity.
[e] Sensitivity (TP/TP+FN).
[f] Farm effect significant at $P < 0.05$.

Table 10. Mastitis cases found with illness alerts of the IMAG model (with three confidence intervals, % in brackets) for all farms together and each farm separately.

Farm (No. Cases)	Model	TP[a]	FN[b]	Sensitivity (%)[c]
All farms	IMAG (95)	174	38	82
(212)	IMAG (99)	159	53	75
	IMAG (99.9)	132	80	62
ALCQ	IMAG (95)	37	5	88
(42)	IMAG (99)	33	9	79
	IMAG (99.9)	25	17	60
ANCQ1	IMAG (95)	80	17	82
(97)	IMAG (99)	76	21	78
	IMAG (99.9)	60	37	62
ANCQ2	IMAG (95)	17	5	77
(22)	IMAG (99)	14	8	64
	IMAG (99.9)	14	8	64
ALCT	IMAG (95)	40	11	78
(51)	IMAG (99)	36	15	71
	IMAG (99.9)	33	18	65

[a] True positive clinical mastitis cases.
[b] Number of false negative cases (FN).
[c] Sensitivity (TP/TP+FN).

Table 11. Number of milkings found with mastitis alerts of the IMAG model (with three confidence intervals, % in brackets) for all farms together and each farm separately.

Farm (No. Cows; No. Milkings)	Model	TN[a]	FP[b]	?[c]	Specificity (%)[d]
All farms	IMAG (95)	119,576	8011	12,682	93.7[e]
(164; 140,269)	IMAG (99)	124,847	2740	12,682	97.9[e]
	IMAG (99.9)	126,696	891	12,682	99.3[e]
(105; 85,983)	Manufacturer	82,364	1189	2430	98.6[f]
ALCQ	IMAG (95)	12,669	1342	738	90.4
(20; 14,749)	IMAG (99)	13,543	468	738	96.6
	IMAG (99.9)	13,893	118	738	99.2
	Manufacturer	14,160	212	377	98.5
ANCQ1	IMAG (95)	39,137	2688	2784	93.5
(47; 44,609)	IMAG (99)	40,833	992	2784	97.6
	IMAG (99.9)	41,388	437	2784	98.9
	Manufacturer	42,204	826	1579	98.1
ANCQ2	IMAG (95)	19,388	1433	5804	93.0
(38; 26,625)	IMAG (99)	20,338	483	5804	97.7
	IMAG (99.9)	20,678	143	5804	99.3
	Manufacturer	26,000	151	474	99.3
ALCT	IMAG (95)	48,382	2548	3356	94.8
(59; 54,286)	IMAG (99)	50,133	797	3356	98.4
	IMAG (99.9)	50,737	193	3356	99.6

[a] True negative milkings for mastitis.
[b] Number of false positive milkings.
[c] Number of milkings with indeterminable conductivity.
[d] Specificity (TN/TN+FP).
[e] Farm effect significant at $P < 0.05$.
[f] Significance of farm effect not determined.

a cow effect (a few cows had a high number of FP alerts). Three farms (ALCQ, ANCQ1, and ANCQ2) used the same equipment for mastitis detection. The pair wise differences in sensitivity between these farms were not significant, while the significance of the pair wise differences in specificity depended on the confidence interval chosen.

The sensitivity (table 9) is based only on cases without indeterminable conductivity. Sensitivity based on all cases, with or without indeterminable conductivity, on all farms was lower and varied between 73% (IMAG 95%) and 50% (IMAG 99.9%). These lower percentages express the mastitis sensitivity presented to the farmer by the sensor system, the percentages in table 9 express the sensitivity of the detection model.

Discussion

Indeterminable Variables

There were many indeterminable variables, mostly caused by measurement errors, with great differences between farms (fig. 2). A level of 5% indeterminable variables of the number of cows milked appeared to be normal. A high number of indeterminable variables was found for conductivity on ANCQ2 (almost 25%), but also on ANCQ1 (10–15%) and for activity on ALCT (30%). These high values indicate hardware problems. On ALCT the pedometer tightening strips often went loose, and many cows lost their transponders. The problems causing the conductivity measurement errors were more difficult to explain because there were great deviations between the three farms using the same equipment (ALCQ, ANCQ1, and ANCQ2). Measurement errors seem inevitable, but the high percentages of milkings with measurement errors should have called earlier for attention on the farm. A monitoring system to check the functioning of the sensor equipment, and immediate servicing at problems are recommended.

Sensitivity for estrus and mastitis based on all cases is lower than sensitivity based only on cases without indeterminable variables. The occurrence of indeterminable variables (e.g. due to measurement errors) thus devaluates the practical applicability of the detection model.

Estrus

The estrus detection results of the previous research (table 1) differ from the results reported by De Mol et al. (1997), although they are based on the same data set. The differences are caused by:

- The length of the estrus period: Estrus cases reported by De Mol et al. (1997) were based on progesterone samples. Visual observations were not available for all estrus cases. Therefore, exact dates were not always known and a period of four days (eight milkings) was used to calculate the sensitivity. In the present report, a period of five milkings was used, which might be too short in cases based only on progesterone without visual observation. The influence of the use of estrus cases without visual observations is illustrated by the higher estrus frequency measured in the former study (1 case in 172 milkings, and 1 in 129) compared with the present results (table 4). In the former study, many cases were based only on progesterone with an assessed estrus date.
- Estrus cases with indeterminable activity: De Mol et al. (1997) reported that all TP cases with or without indeterminable activity were used for calculating the sensitivity. In this article, only TP cases without any indeterminable activity in the estrus period were used.

The length of estrus periods can be shorter than five milkings if not only the date of an observed case is recorded, but also the hour. Including the previous day in the estrus period is only useful when the estrus is observed in the morning. Including the next day is useful when the estrus is observed in the evening. The hour of estrus observation (and detection) is important for timing the insemination (Maatje et al., 1997).

The sensitivity for estrus calculated on all farms in the present research (table 6) was lower than the sensitivity calculated in the previous research (table 1). There were significant farm effects. Farm ALCQ had the same equipment as the farms of the experiments reported in De Mol et al. (1997), but outperformed the results presented in table 1. This difference in results may be because only observed estrus cases were taken into account in the present study. The changes in activity might be greater in observed cases compared with cases based on progesterone samples only. The poor results of ANCQ1 and ANCQ2 might partly be caused by the use of neck transponders, which give worse results then leg transponders (Koelsch et al., 1994). However, the equipment used was not the only cause of farm differences, as indicated by differences between ANCQ1 and ANCQ2 (same equipment, significant differences in results).

Furthermore, the results of ALCQ might be influenced by the housing system: the cows of ALCQ were kept inside all year, while the cows of the other farms were out in the pasture during the summer period. The results of ALCT were influenced by the absence of milk temperature sensors. Alerts for estrus on this farm were based on activity and yield and not on a combination of activity, yield, and milk temperature, as on the other farms.

The sensitivity results obtained with the manufacturer's model were comparable with the results of the IMAG model with a confidence interval of 99.9% (table 6), as witnessed by the number of TP cases (without or with indeterminable activity). However, the number of FP milkings from the manufacturer's model (table 7) was much higher on some farms (doubled for ANCQ1 and ANCQ2). Therefore, the performance of the IMAG model was better than that of the manufacturer's model: less FP milkings and the same sensitivity.

Our results are in accordance with estrus detection results from experimental farms found in literature. Comparing these results with the performance predicted in Van Asseldonk et al. (1998) makes clear that only the IMAG model with the 95% confidence interval met the expectations (81% sensitivity and 90% specificity). In practice the 99.9% confidence interval might be preferred because of the lower number of FP alerts. The number of FP alerts should not be much higher than the number of TP alerts, otherwise only a minority of the alerts has a practical value for the farmer.

Some FP milkings were probably due to true estrus cases that were detected by the model but not observed on the farm. This imperfect estrus observation means that the actual specificity was probably higher than reported here. It was difficult to quantify the effect of inadequate farm observations. Some of these FP estrus alerts might be classified as TP looking at the estrus cycle, but such a classification is subjective. However, the specificity of all farms together (table 7) was already higher than the specificity on the farms used in De Mol et al. (1997) as given in table 1.

Although the lactation period was a significant factor in the generalized mixed linear model for the activity level, the predicted means did not always show a decreased activity level in early lactation (table 8). The activity of cows in the first period of lactation was mostly at a lower level than the activity of cows in the second or third period of lactation. However, the differences were small and farm ALCQ showed the opposite trend. If the cows had a negative energy balance in the first period of lactation (as suggested in Brand et al., 1996), than it had no effect on the cow's activity.

Mastitis

The mastitis detection results in table 1 differ from the results reported by De Mol et al. (1997), although they are based on the same data set. These differences are caused by:

- The length of the mastitis period: De Mol et al. (1997) reported that a mastitis period of 18 days (36 milkings) was used (from 10 days before until 7 days after the mastitis date). In the present study, a shorter period of only eight milkings was used.
- De Mol et al. (1997) reported that all TP cases with or without indeterminable conductivity were used to calculate sensitivity. In the present article, only TP cases without indeterminable conductivity were used.

There were significant differences in sensitivity and specificity of mastitis alerts between the farms (tables 9 and 11). The sensitivity on ALCT was lower ($P < 0.05$) than on the other three farms. This lower sensitivity is an indication that conductivity of quarter milk gives better detection results than when the conductivity of total milk is used as a variable. The increase in conductivity in case of mastitis may be lower and more difficult to detect when the milk of four quarters is mixed.

The sensitivity on ALCQ and ANCQ1 was higher than the sensitivity on the farms used in De Mol et al. (1997) and presented in table 1. There were differences in specificity between farms; the specificity on ALCT was significantly higher than on other farms. There were no differences in specificity between the other farms.

The sensitivity of the IMAG model was much higher than the sensitivity of the manufacturer's model (table 9); the differences in specificity were less (table 11). Many FP alerts given by the manufacturer's model were caused by a small number of cows. This observation made it useless to test for the significance of farm effects with a logistic regression model because the cow effects were greater than the farm effects.

The mastitis sensitivity of all farms based on illness alerts (table 10) was higher than the sensitivity based on mastitis alerts (table 9). This difference was highest on ALCT. Illness alerts were based on deviating yield, temperature, and activity. Mastitis alerts were based on conductivity, yield, and temperature. The higher sensitivity for illness alerts indicates that deviations in yield and temperature were clearer than deviations in conductivity. Deviations in conductivity were not detectable or their detection might be too late. When a milker notified mastitis, the milk was separated (so no conductivity was recorded) and the cow was treated for mastitis. It was not possible to calculate the specificity of illness alerts with the information available, the occurrence of diseases other than mastitis was not known.

Mastitis detection in our experiments was lower than those found in literature (Graupner and Barth, 1994; Maatje et al., 1992; Milner et al., 1996). The differences might be explained by the circumstances. Literature data usually refer to more controlled conditions, while this research was a field trial utilizing four commercial locations. Hamann and Zecconi (1998) also indicate that sensitivity is much lower in experiments with a low prevalence (as in the present situation). However the results, 71% sensitivity with 86% specificity for the IMAG model with the 95% confidence interval, are within the range expected by experts (Van Asseldonk et al., 1998). The specificity should be high for practical application, taken into account the low prevalence of mastitis in practice. The number of FP alerts is much higher than the number of TP alerts even when the specificity is 99% (IMAG model with 99.9% confidence interval).

CONCLUSIONS

- The detection results predicted by an expert panel (Van Asseldonk et al., 1998), were achieved in this study. The results may attain the same level as found under experimental conditions by De Mol et al. (1997), which implies that estrus detection has been developed far enough for practical usage. Mastitis detection results show that practical usage is difficult with the available sensors. Both sensitivity and specificity are not high enough, and better detection results are attained by using only yield, temperature, and activity sensors (and not taking an increased conductivity as a prerequisite). The applicability for mastitis detection may be improved by a further development of sensors.
- Good detection results are only possible when the data collection equipment is functioning well. The farmer should monitor the equipment at regular intervals, otherwise detection based on sensor measurements will not yield acceptable results. Thus, implementation of a detection model will only add value to a farm, when accompanied by good management. Data collection might be improved by the use of auto calibration software.
- The IMAG model performed better than the manufacturer's model. Combined processing of the variables based on a more complex algorithm appears to be worthwhile. The advanced software used in the IMAG model gives promising results compared with currently available software.
- The sensor equipment used might explain some differences found between the farms. The results indicate that activity measured by neck transponders may result in lower estrus sensitivity and that conductivity data of total milk may give lower mastitis sensitivity than data of quarter milk.

Further research was directed towards reducing the number of false positive alerts by taking into account other influences, like group influences or the reproductive status of the cow (De Mol and Woldt, 2000). A manual for practical usage of sensors and a detection model, describing and explaining what the farmer should do in case of alerts, may be needed to make these systems ready for introduction in practice.

ACKNOWLEDGEMENTS

We wish to thank the farm managers and researchers of the four experimental farms of the Research Institute for Animal Husbandry (PV): Bosma Zathe, Cranendonck, De Marke and Waiboerhoeve, for their co–operation in this research.

REFERENCES

Brand, A., J. P. T. M. Noordhuizen, and Y. H. Schukken. 1996. *Herd health and production management in dairy practice*. Wageningen, The Netherlands: Wageningen Pers.

De Mol, R. M., G. H. Kroeze, J. M. F. H. Achten, K. Maatje, and W. Rossing. 1997. Results of a multivariate approach to automated oestrus and mastitis detection. *Livestock Prod. Sci.* 48: 219–227.

De Mol, R. M., A. Keen, G. H. Kroeze, and J. M. F. H. Achten. 1999. Description of a detection model for oestrus and diseases in dairy cattle based on time series analysis combined with a Kalman filter. *Comp. and Electronics in Agric.* 22: 171–185.

De Mol, R. M., and W. E. Woldt. 2001. Application of fuzzy logic in automated cow status monitoring. *J. of Dairy Sci.* 84: 400–410.

Engel, B., and A. Keen, 1994. A simple approach for the analysis of generalized linear mixed models. *Statistica Neerlandica* 48(1): 1–22.

Frost, A. R., C. P. Schofield, S. A. Beaulah, T. T. Mottram, J. A. Lines, and C. M. Wathes. 1997. A review of livestock monitoring and the need for integrated systems. *Computers and Electronics in Agriculture* 17: 139–159.

Geers, R. 1994. Electronic monitoring of farm animals: a review of research and development requirements and expected benefits. *Computers and Electronics in Agriculture* 10: 1–9.

Genstat 5 Release 3 Reference Manual, 1993. Oxford, United Kingdom: Oxford University Press.

Gil, Z., J. Szarek, and J. Kural. 1997. Detection of silent oestrus in dairy cows by milk temperature measurement. *Animal Sci.* 65: 25–29.

Graupner, M., and K. Barth. 1994. Udder health and milk quality control by using electrical conductivity and quarter milk yield. In: *Proc. of the Intl. Symp. Prospects for Future Dairying: A Challenge for Science and Industry*, eds. O. Lind and K. Svennersten. Alfa Laval Agri AB, Tumba, Sweden and Swedish University of Agricultural Sciences, Uppsala, Sweden.

Hamann, J., and A. Zecconi. 1998. Evaluation of the electrical conductivity of milk as a mastitis indicator. *Bulletin of the IDF* 334: 4–27.

Koelsch, R. K., D. J. Aneshansley, and W. R. Butler. 1994. Analysis of activity measurement for accurate oestrus detection in dairy cattle. *J. of Agricultural Engineering Research* 58: 107–114.

Lehrer, A. R., G. S. Lewis, and E. Aizinbud. 1992. Oestrus detection in cattle: recent developments. *Animal Reprod. Sci.* 28: 355–361.

Maatje, K., P. J. M. Huijsmans, W. Rossing, and P. H. Hogewerf. 1992. The efficacy of in–line measurement of quarter milk electrical conductivity, milk yield and milk temperature for the detection of clinical and subclinical mastitis. *Livestock Production Science* 30: 239–249.

Maatje, K., S. H. Loeffler, and B. Engel. 1997. Predicting optimal time of insemination in cows that show visual signs of estrus by estimating onset of estrus with pedometers. *J. of Dairy Sci.* 80: 1098–1105.

Milner, P., L. Page, A. W. Walton, and J. E. Hillerton. 1996. Detection of clinical mastitis by changes in electrical conductivity of foremilk before visible changes in milk. *J. of Dairy Sci.* 79: 83–86.

Schlünsen, D., H. Roth, H. Schön, W. Paul, and H. Speckmann. 1987. Automatic health and oestrus control in dairy husbandry through computer aided systems. *J. of Agric. Eng. Res.* 38: 263–279.

Van Asseldonk, M. A. P. M., R. B. M. Huirne, and A. A. Dijkhuizen. 1998. Quantifying characteristics of information technology applications based on expert knowledge for detection of oestrus and mastitis in dairy cows. *Preventive Veterinary Medicine* 66: 273–286.