
THE PROCEEDINGS OF THE
NEW ZEALAND SOCIETY
OF
ANIMAL PRODUCTION
VOLUME 62 2002

Practical implications of theoretical modelling of in-line milk meters

O Dowson[a]*, G Anderson[b] and K Hempstalk[b]

[s]*Department of Engineering Science, University of Auckland, Private Bag 92019, Auckland Mail Centre, Auckland 1142 New Zealand;* [b]*Livestock Improvement Corporation, Private Bag 3016, Hamilton 3240, New Zealand.*
**Corresponding author. Email: odow003@aucklanduni.ac.nz*

Abstract

In-line milk meters are devices which measure the milk yield of dairy cows at every milking; however, they are also susceptible to error. This results in frequent, but inaccurate, measurements. This paper attempts to model a lactation curve and accumulated milk yields from these devices. Wood's model of a lactation curve was used to predict milk yield. This was modified by adding both systemic and random daily variation. The meters were modelled by taking the modified Wood's lactation curve for each cow and altering it with various error terms. Along with precision, each sensor also had a rate of drift, and a rate at which its variance grew. It was found that the precision of a sensor can be low, however by aggregating multiple records it is possible to obtain an accurate value for total yield. One strategy that improved the total yield accuracy was to scale the measurements obtained for a particular sensor in a given milking. Due to the uncertainty surrounding the errors involved, and the large assumptions that were made, this model is a good starting point for future work. It is able to identify the relative effect of errors on meter accuracy, and test strategies to improve accuracy of daily and total milk yield.

Keywords: in-line milk meter; frequent measurement; low precision; theoretical

Introduction

In-line milk meters (ILMMs) are devices which measure the milk yield of dairy cows at every milking. The current International Committee for Animal Recording (ICAR) standard for recording devices of milk yield is 2.5% coefficient of variation (CV) (ICAR, 2012). This paper will consider ILMMs that are low-precision (>10% CV). In 2012, ILMMs were only installed in 2% of New Zealand milking sheds (De LaRue et al. 2012). It is expected that the usage of ILMMs will grow as their accuracy improves, and their cost decreases.

In anticipation of their wider use, this paper presents a theoretical model of an ILMM, and then integrates it into a simulation of a herd of 200 cows. It identifies the relative effect of various errors on, and presents one method for improving the accuracy of, ILMMs so that practical steps can be taken to improve the accuracy and usefulness of ILMMs once they are installed.

Materials and Methods

A model was designed in Octave (Eaton, 2013) in which cows were modelled individually, then grouped together to form a theoretical herd. All data were simulated and kept generalised so that it may apply to any herd and any brand of ILMM.

Steenson (Unpublished) created a simple model of an ILMM with random measurement error to show how aggregation improved yield measurement accuracy. For simplicity, this work uses Wood's model of a lactation curve (Wood, 1967) as a basis for the theoretical yield of each cow (Row 1 in Table 1), and uses the same constants as Steenson.

To model the difference between morning and afternoon milkings, AM yield was defined as and PM yield as $, t)$. These values were also obtained from Steenson (Unpublished). Using different constants or a more complicated model of cow lactation is unlikely to change the general findings of this paper.

The error function $\varepsilon(cow, t)$ (Row 2 in Table 2) was constructed from a combination of production variation and measurement error that varied between cows, and over time. The production variation consisted of systemic variation between cows (S), random daily variation for each cow (R), systemic herd variation (S), and random daily variation of the herd (R). The measurement error was modelled by a systemic term ε_S, and a random term .

Production variation

Systemic variation between cows. Due to environmental factors, such as a cow's genetic makeup, there will be a systemic variation within the herd with regards to average yield of each animal. This variation was modelled by assigning each cow a normally distributed, random number S_{cow} with mean 0 and standard deviation $\sigma_{S_{cow}}$ at the start of each run. To simplify the model it was assumed that the systemic variation of each cow in relation to the herd remains constant over the lactation cycle.

Random daily variation of each cow. The daily yield for a cow can vary significantly due to biological factors. Mein (2000) estimates this value as 6.8% CV. It was assumed that this variation remained constant between cows and over the lactation cycle to simplify the model. Thus R_{cow} was modelled by a normally distributed random number with mean 0 and standard deviation 0.068.

Systemic herd variation. On a given farm, each paddock will have a unique level of potential milk production. This is caused by a number of factors, including pasture quality, herbage availability per cow, terrain and distance to the milking shed. This variation was modelled by generating a normally distributed, random number S_{herd} with mean 0 and standard deviation $\sigma_{S_{herd}}$ for each paddock at the start of each run. There were N_{pad} paddocks and the herd visited them consecutively, beginning at paddock one, moving to a new paddock after each milking. The value for S_{herd} was the random number that corresponded to the paddock that the cows were in before the milking. It was assumed that the variation between paddocks remains constant over the lactation cycle. This is a reasonable first approximation, however in practice some paddocks may perform better in different weather conditions, so seasonal variation could be incorporated in future.

Random daily herd variation. Entire herds will experience daily variation due to environmental factors that all cows on the same farm are exposed to. This term was modelled by generating a normally distributed, random number R_{herd} with mean 0 and standard deviation $\sigma_{R_{herd}}$ at the start of each milking. This remained constant for every cow in that milking. It assumes that all cows respond identically to environmental factors. This was done to simplify the model, however in practice some cows may respond better to dry spells than others, others to cold days, and so on.

Measurement error. ILMMs are prone to at least three types of error. They can be imprecise, the mean of their readings can drift over time, and they may experience sudden faults which cause either a missed reading, or implausible results. Only the first two were considered since sudden faults are usually easy to identify and remove from the dataset. In the model there were N_{bails} in the milking shed, and prior to each milking, a milking order was generated. From this order, cows were allocated to bails consecutively, beginning at bail one.

Systemic. ILMMs can 'drift' over time. This causes the mean of their measurements to deviate from the true mean until they are recalibrated. Such calibration is often done via the wash water, or against bulk milk data. It was assumed that this calibration was perfect and instantaneous. These two assumptions are unlikely in practice but were made to simplify the model. Future work could model actual calibration methods, or leave a residual error after calibration. In this model, recalibration happened every c_{int} days.

The systemic measurement error ε_S (Row 3 in Table 1) was modelled by generating a normally distributed, random "drift rate" δ with mean 0 and standard deviation σ_{Drift} for each sensor at the beginning of each run. This rate was multiplied by the number of months since the ILMM was last calibrated.

It was assumed that the drift rate of each meter was a fixed property, such that it continued to drift at the same rate after each calibration. It was also assumed that the meters drift linearly in one direction. This assumption was made to simplify the model and because it represents a more pessimistic approach than meters that can drift up and down. Linear drift was chosen as a first approximation, although other modes are possible.

Random. There are two parts to random measurement error: the base precision of the meter, and the rate at which it decreases. The random measurement error ε_R (Row 4 in Table 1) was modelled by generating a normally distributed, random variance growth rate β with mean 0 and standard deviation σ_{growth} for each sensor at the start of each run. This caused the base sensor variance α (initially a normally distributed random number with mean 0 and standard deviation σ_{sensor}) to increase linearly with time since calibration.

Table 1 List of formulae. t is the number of days in lactation, $\bar{t}$ is the number of days since the in-line milk meter was last calibrated.

1	$Y(cow,t) = 12\left(\frac{t}{7}\right)^{0.5} e^{-0.06\left(\frac{t}{7}\right)} \times \varepsilon(cow,t)$
2	$\varepsilon(cow,t) = \underbrace{1 + S_{cow} + R_{cow} + S_{herd} + R_{herd}}_{Production\,Variation} \times \underbrace{1 + \varepsilon_S + \varepsilon_R}_{Measurement\,Error}$
3	$\varepsilon_S = \delta \times \bar{t}/30$
4	$\varepsilon_R = \alpha \times (1 + \beta \times \bar{t}/30)$
5	$MPD = \sum_{d=T}^{N\_Lac} \sum_{c=1}^{N\_Cows} \left\lvert \frac{1}{N_{Cows} \times (N_{Lac} - T + 1)} \left(\frac{\sum_{t=d-T+1}^{d} Y_{measured}(c,t)}{\sum_{t=d-T+1}^{d} Y_{actual}(c,t)} - 1 \right) \right\rvert$

It was assumed that each meter has a fixed linear rate at which the precision decreased, and that this remained constant over time, and after recalibration. This assumption was made to simplify the model.

Milking order. In order to model an entire herd, animals need to be assigned to milking bails each milking. Various authors (Adamczyk et al. 2011, Berry et al. 2012, Grasso et al. 2007, Rathore 1982) have reported the milking order of a single herd to be fairly consistent. To model a semi-random milking order, the following algorithm was developed:

When initialising the model, assign each cow a uniformly distributed random number between 0 and 1. Sort these numbers to obtain an initial "favourite" milking order. Before each milking assign the first ρ positions (where ρ is the number of 'pushy' cows in the herd who always enter the shed first) to the cows who favour those positions. Then take the cow whose favourite position is $\rho + 1$ and assign it to the nearest empty position to where)) (i.e. some normally distributed random integer with mean 1). Increment ρ and repeat until all cows have been assigned a position in the milking order.

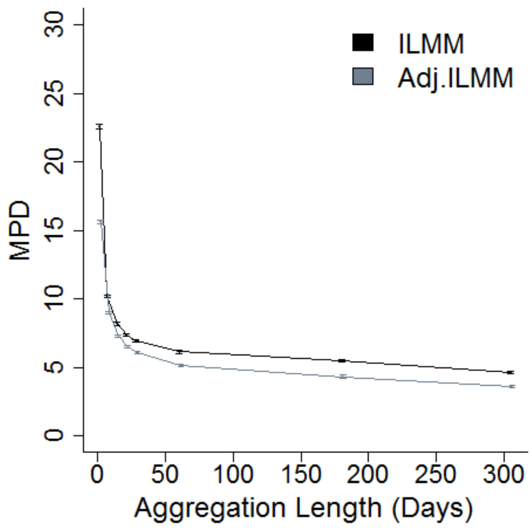
In this approach, cows at the beginning of the herd will tend to get positions closer to their favourite position, while cows in the middle are more jumbled. As the algorithm moves towards the back, there are fewer open slots available, and cows may be forced into a slot further away than desired.

Implementation

The mean percentage deviation (MPD) was chosen as a metric by which to quantify the accuracy of ILMMs. The MPD is a measure of absolute error, which is useful because over and under estimating milk yield can lead to undesirable consequences such as culling good cows that appear to be performing poorly. The MPD is defined in row 5 of Table 1 where *N_Lac* is the lactation length (in days), *N_Cows* is the number of cows in the herd, T is the number of days worth of ILMM data to aggregate, $Y_{actual}(c,t)$ is value obtained from the augmented Woods Model without measurement error, and $Y_{measured}(c,t)$ is value obtained from the augmented Woods Model with measurement error. If, for example the measured values were 20% above the actual values for half the time, and 10% below the actual values for the remaining time, then the MPD would be 15%.

Given the uncertainty in the values for various parameters, a Monte-Carlo simulation was conducted. The simulation was coded in Octave 3.6.4 and 2620 experiments were run using 4 cores of an Intel i5-3570 CPU, with 4GB of RAM, running Windows 7 (64bit). Each experiment consisted of a herd of 200 cows, with a lactation length of 305 days, and a random combination of parameters, each drawn from a uniform distribution. Each set of parameters was run 20 times and the mean MPD of the replications was recorded as the value for that set of parameters.

Figure 1 Mean ± 1 SEM mean percentage deviation (MPD) of all 2620 Monte Carlo experiments against aggregation length for adjusted and unadjusted in-line milk meter (ILMM) data. Points are at aggregation lengths of 1, 7, 14, 21, 28, 60, 180 and 305 days. Lines of best fit plotted using 'smooth.spline' in R3.0.2. SEM bars are small due to the large number of data points in each population. Restricting the sample space to realistic parameters shifts both curves downwards and retains vertical separation.

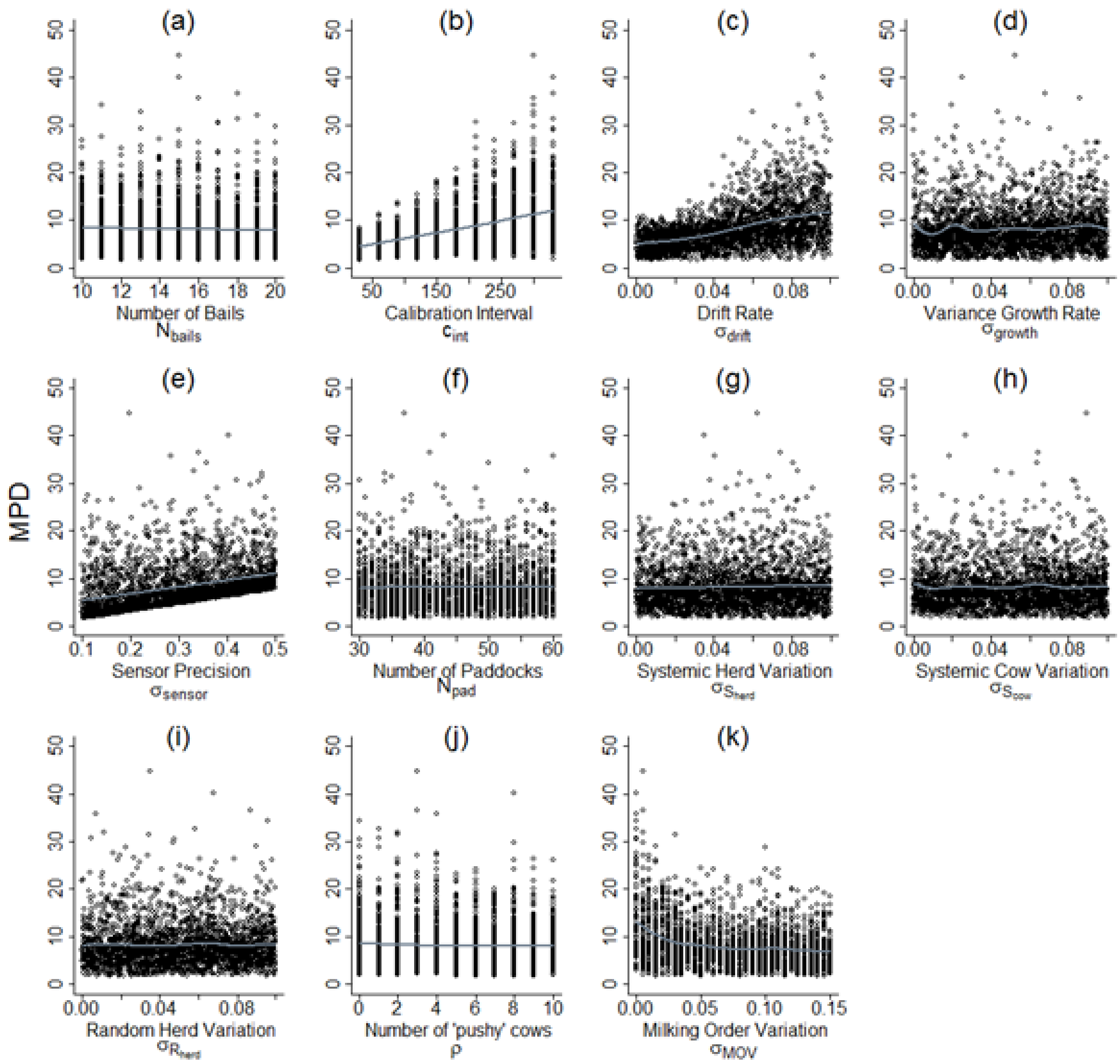


To simplify the model, all cows started lactating on day one and followed the same unadjusted Woods curve. A different lactation length is unlikely to change the general findings of this paper, but smaller lengths offer less data to aggregate. The range of in the Monte-Carlo simulation (10% – 50% CV) was intentionally made pessimistic to explore the effect of sensors that are far more imprecise than those currently available that meet ICAR standards (ICAR 2013, DeLaval).

Method of Adjusting the ILMM Data

At the end of each milking, a scale factor was calculated for each bail that was defined as the average yield of cows that were milked in that bail, divided by the average yield of all cows in the milking. Immediately before recalibration, for each bail, simple linear regression was applied to the scale factor time series since the last calibration. The slope of the line was used as an approximation of the rate at which the bail had been drifting. The yield of each cow that had been milked in that bail was scaled using the reciprocal of this value to obtain an adjusted value. Data modified by this method are herein referred to as 'adjusted ILMM' data.

Figure 2 Scatter plots of mean percentage deviation (MPD) for unadjusted in-line milk meters against the value of each Monte Carlo parameter with an aggregation length of 14 days. Lines of best fit plotted using 'smooth.spline' in R3.0.2.



Results and discussion

Mean percentage deviation

It was found that the MPD decays exponentially as the aggregation length is increased for both adjusted and unadjusted ILMMs (Fig. 1). Due to the large sample size (n = 2620), the standard error of each population is small, even though their standard deviations are large. The MPD of the adjusted data is consistently below the MPD of the unadjusted data (p-value < 1e-10 at each point). In addition, if the Monte-Carlo sample space is restricted to realistic, rather than pessimistic values, the standard deviation and vertical position of each point decreases, but the general trends remain.

Scatter plots of the unadjusted ILMM MPD against each parameter for an Aggregation Length of 14 days are shown in Fig. 2. Similar results are observed if the aggregation length is shorter or longer. The clearly defined lower bound on MPD with respect to sensor precision in Figure 2e is noticeable. It suggests that, assuming that the true values for the parameters in the model lie within the range explored in the Monte-Carlo simulation, it is possible to place a lower bound on the MPD of the unadjusted ILMM data that is a function of sensor precision $\alpha \in N(0, \sigma_{sensor}^2)$ and aggregation length. Using a realistic value for ILMM yield sensor precision CV of 10% and the curve contained in Figure 3 aggregating seven days of data can never, on average, produce a value for total milk yield closer than 2.2% of the true value. This places a fundamental limit on the accuracy of ILMMs. The implication of this is that regardless of the amount of data aggregated, ILMMs are limited by their underlying precision.

It was found that, on average, increasing the amount of time between recalibrations increased the

MPD (Fig. 2b). This was expected as an increased amount of time between recalibrations allows for more error to accumulate and the sensor to drift further from the true value. In addition, on average, increasing the rate at which meters drift also increases the MPD (Fig. 2c). Again, this is expected as the sensors will diverge faster from the true value, leading to the sensor either overestimating, or underestimating the cow's true yield. The greatest MPDs are therefore explained as being generated by experiments with both a long calibration interval, and high rate of drift. Therefore, ILMMs should be calibrated as often as possible to prevent sensors from drifting far from the true value.

A low milking-order variation results in a cow visiting a bail with a higher than expected probability (Fig. 2k). Should a herd have this feature, it could lead to biased results, particularly if the bail has a large rate of drift. However it is demonstrated in Figure 4k that applying the adjustment outlined earlier in this paper can almost entirely remove this effect.

Figure 3 Chart of θ against aggregation length where θ is the slope of the line MPD_LB = θ × α and α is the precision of the sensor and MPD_LB is the lower bound on the mean percentage deviation for unadjusted in-line milk meters. Points are at aggregation lengths of 1, 7, 14, 21, 28, 60, 180 and 305 days. Line of best fit plotted using 'smooth.spline' in R3.0.2.

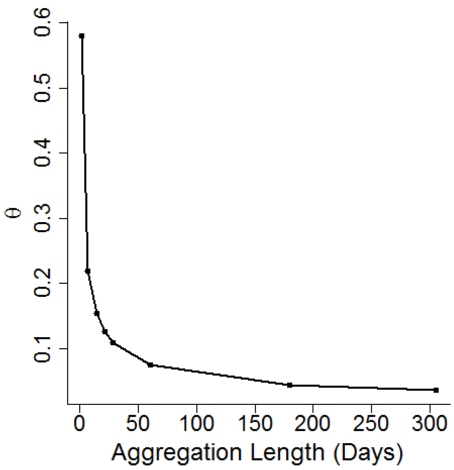
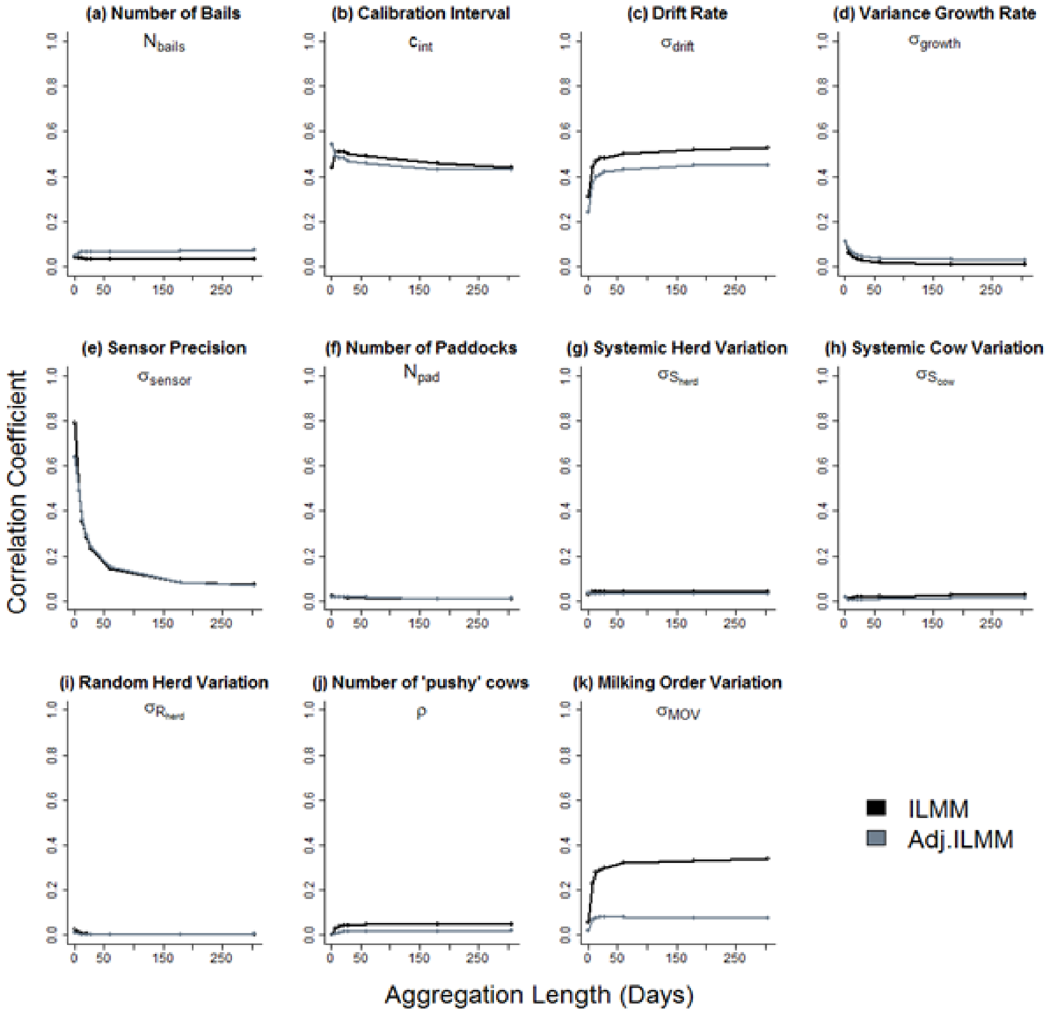


Figure 4 Correlation coefficient of each parameter and mean percentage deviation against aggregation length for unadjusted (black) and adjusted (grey) in-line milk meter (ILMM) data. Points are at aggregation lengths of 1, 7, 14, 21, 28, 60, 180 and 305 days. Lines of best fit plotted using 'smooth.spline' in R3.0.2.



Relative effect of parameters in the model

The correlation coefficient between each parameter and MPD for each aggregation length was used as a measure of the effect each the parameter had on the MPD of ILMMs. A low correlation coefficient was taken as an indication that the parameter had little effect on the value of the MPD. In contrast, a parameter with a high correlation coefficient had a large effect on the MPD. The Calibration Interval (Fig. 4b), rate of Sensor drift (Fig. 4c), and Milking Order Variation (Fig. 4k) all had a large effect on the MPD of the unadjusted ILMMs. Adjusting the data slightly reduced the effect of the rate of drift, and reduced the effect of Milking Order Variation. When the aggregation length was small (i.e., less than 50 days), sensor precision (Fig. 4e) had a large correlation coefficient, however, as the aggregation length increased, the correlation coefficient decreased exponentially.

It was found that aggregating data improves the accuracy of ILMMs. The potential accuracy of ILMMs is reduced if there is a high rate of drift, low milking-order variation, or long periods (e.g., six months) between calibrations. Adjusting the ILMM data in the way outlined in this paper increases the accuracy of ILMMs. The precision of ILMMs places a limit on their accuracy. Future work could include investigating some of the assumptions in the model and extending it to multiple lactation seasons.

References

Adamczyk K, Slania A, Zygmunt G, Felenczak A, Bulla J. 2011. Relationships between milk performance and behaviour of cows under loose housing conditions. Annals of Animal Science 11(2): 283-293.

Berry DP, McCarthy J 2012. Genetic and non-genetic factors associated with milking order in lactating dairy cows Applied Animal Behaviour Science 136(1): 15-19.

De LaRue B, Kamphuis C, Jago J 2012. Smart ways to detect mastitis: What do we know about automated in-line detection systems? Proceedings of the New Zealand Milk Quality Conference 7(12): 1-4.

DeLaval. Milk Meter MM27BC. http://www.delaval.co.nz/en/-/Product-Information1/Milking/Products/Milking-point/Milk-recording/DeLaval-milk-meter-MM27BC/ [accessed 9 Jan 2014].

Eaton J, Octave 2013. https://www.gnu.org/software/octave/ [accessed 28 March 2014]

Grasso F, Rosa GD, Napolitano F, Di Francia A, Bordi A 2007. Entrance order and side preference of dairy cows in the milking parlour. Italian Journal of Animal Science, 6:187–194.

ICAR 2012. International Agreement of Recording Practices, Pg. 316. http://www.icar.org/Documents/Rules%20and%20regulations/Guidelines/Guidelines_2012.pdf [accessed 9 Jan 2014].

ICAR 2013. Approved Milk Meters for Cattle. http://www.icar.org/pages/sub_committees/sc_recording_devices_approved_milkmeters.htm [accessed 9 Jan 2014, last updated 12 July 2013].

Mein, Hannah, Clarke 2000. Limits of error for permanently installed milk meters used for herd recording or for daily herd management purposes. Proceedings of the 32nd Biennial Session of ICAR, Bled, Slovenia, May 2000. Performance Recording of Animals: State of the Art, Pg. 159-162.

Rathore AK 1982. Order of cow entry at milking and its relationships with milk yield and consistency of the order. Applied Animal Ethology 8: 45-52.

Wood 1967. Algebraic Model of the Lactation Curve in Cattle. Nature 216:164-165.