final report

Project code: B.FLT.7009

Prepared by: Dr Stuart McCarthy[1], Daniel Mcleod[1], Dr Joseph McMeniman[2]
[1]Manabotix Pty Ltd, [2]Meat & Livestock Australia

Date published: 24 August 2018

PUBLISHED BY
Meat and Livestock Australia Limited
Locked Bag 1961
NORTH SYDNEY NSW 2059

Feedlot adoption – bunk scanner field trials

Meat & Livestock Australia acknowledges the matching funds provided by the Australian Government to support the research and development detailed in this publication.

Executive summary

Feed bunk management is the process of determining feed allocation for pens of feedlot cattle for a 24-hour feeding cycle. Objectives of bunk management focus on consistently maximising feed intake, whilst minimising feed wastage and digestive disorders (bloat and acidosis). Bunk calling is a critical input for this process, and the human callers' actions directly determine feed intake and carcase weight gain of pens of feedlot cattle.

With advancements in mapping, sensors, and robotics technologies, it is now feasible to automate feed bunk management. Enabled by levy-funds, our recent MLA project B.FLT.0166 achieved world-first determinations of the precision and accuracy of humans when predicting feed remaining in feed bunks versus a bespoke prototype automatic bunk calling solution based on appropriate technologies. The prototype's performance was assessed through a rigorous and varied experimental process, under non-operational conditions, providing highly repeatable and accurate predictions for all feed remaining quantities assessed. Moreover, the prototype system was consistently more accurate and precise than human predictions.

Feedlot trials were subsequently undertaken to assess the prototype system against humans in an operational environment under normal conditions. Due to logistics of mapping empty feed bunks in the commercial feedlot, feed remaining predictions were determined through post-processing of lidar point-cloud data to complete project B.FLT.0166.

To improve likelihoods of acceptance and adoption by all stakeholders for the prototype automatic bunk calling system, further in-field experiments were required across three diverse feedlot operations with different locations, bunk designs, and road conditions to validate the prototype system's technology robustness. This report presents the outcomes of these experiments and demonstrates that under normal operating conditions the prototype system predicted feed remaining in bunks accurately and repeatedly, outperforming human callers in both criteria, thereby representing a high-value outcome for the Australian red meat industry.

Table of contents

1 Introduction

This final report describes the results of three experiments to evaluate a prototype automatic system to estimate feed remaining in bunks of feedlot cattle.

1.1 Project intent

Feed bunk management is the process of determining feed allocation for pens of feedlot cattle for a 24-hour feeding cycle. Bunk calling is a critically important job, and the callers' actions directly determine feed intake and carcase weight gain of pens of feedlot cattle.

As a world-first, MLA project B.FLT.0166 determined that feed remaining in bunks could be automatically predicted more accurately and precisely than humans, and this was demonstrated through a rigorous series of experiments. However, due to logistical limitations of mapping empty feed bunks in the commercial feedlot during project B.FLT.0166, feed remaining predictions were determined from post-processing of collected lidar point-cloud data.

Further in-field observations are required across diverse sites with different bunk designs and road conditions to determine the technology robustness, and to drive industry confidence for adoption. This project's intention is to advance the prototype automatic bunk calling system through a series of additional validation experiments at three different feedlots.

2 Project objectives

The overall project objective that was agreed in the agreement is as follows,

1. Determine precision and accuracy of the bunk scanner versus human bunk callers to determine feed remaining across 3 diverse feedlot sites.

3 Methodology

The purpose of the research activities described here was to progress the validation of the prototype system for estimating feed remaining accurately and precisely. The experiment was serialised across three feedlot sites, and targeted finisher diet with 100 observations per site; appropriate environmental characteristics for each site are included in Table 1. The predictions provided by the prototype system and human bunk callers were generated in-field and in 'real-time' and were later assessed against manually weighed-back mass observations. Progression through the experiment was confirmed by MLA after successful completion of each site's experiment.

Table 1: Characteristics of experiment host sites

ID	Site (experimental order)	Standard cattle units (SCUs)	Pens	Overall pen extents	Environmental conditions (weather conditions are included later)
1	Feedlot A	¬30,000	142	0.5km x 1.2km	Roads were compacted fine base aggregate in good condition, with concrete aprons adjacent to bunks in rows 4 to 6. Bunks in rows 1 to 3 ('old section') were precast and uniform, though exhibit significant wear and tear in places. Bunks in rows 4 to 6 were new slip-form concrete and are very uniform.

ID	Site (experimental order)	Standard cattle units (SCUs)	Pens	Overall pen extents	Environmental conditions (weather conditions are included later)
2	Feedlot B	47,952	258	0.7km x 1.6km	Roads were bitumen in good condition. Occasional pot-holes and ground softness were experienced at edges of roads (bases of bunks) after rain event. Bunks in rows 1 to 12 were new slip-form concrete and were very uniform.
3	Feedlot C	9,800	55	0.3km x 1.3km	Roads were in good condition: rows 1 to 3 were bitumen, and rows 4 to 12 were compacted fine base aggregate. Bunks in rows 1 to 3 were new slip-form concrete and were uniform. Bunks in rows 4 to 10 had mixed geometries and were very non-uniform, even within a single pen's extents.

Project management activities performed through the project included provision of regular verbal and email updates to MLA and the host site so that all stakeholders were aware of project statuses, milestones, and interface and resourcing requirements.

Experiment protocols employed are provided in the following sub-sections.

3.1 Preliminaries

The GNSS RTK reference (base) station receiver (contained in weatherproof enclosure) supplied for project B.FLT.0166 was transferred, installed, and reconfigured at each host site, with bespoke ad hoc antenna and support infrastructure arrangements provided for Feedlots A and C.

It was established early at Feedlot A that the prototype's existing arrangement requiring RTK corrections provided over wireless broadband connection would not be appropriate due to insufficient internet coverage within the feedlot rows; the preferred network was Telstra. Instead, a temporary ad hoc UHF transceiver arrangement was also provided to enable RTK correction communications, and these were achieved with 100% reliability.

The prototype automatic bunk calling rover was installed on a light vehicle provided by the host sites: at Feedlot A this was their Toyota Hilux 2WD single-cab utility, and for the other sites it was Feedlot B's Mitsubishi Triton 4WD dual-cab utility.

The week before the experiment commenced, the start and end positions of all pens was georeferenced, i.e. located in global coordinates based on GNSS measurements from the rover, to enable automatic localisation within the feedlot during the experiment. The baseline geometry for all feedlot bunks was then measured, and required the light vehicle complete with prototype system to follow front-end loaders with bunk sweeper attachment, and scan and database all feed bunks to be targeted during the experiment.

The platform scale (CAS BW-L60, Brisbane, Q; ±0.01 kg readability), paddock vacuum (Greystone Maxi Vac), tarpaulins, and collection/weigh-back buckets provided during the previous experimental activities (B.FLT.0166) were also transferred in preparation for the experiment's commencement at that site.

3.2 Experiment protocol

The following experimental protocol was exercised for this project,

1. A daily dedicated bunk caller and prototype bunk scanner predicted feed remaining independently for a random selection of pens prior to feeding; this usually started at 0600h each day.
    a. The light vehicles for both the human and bunk scanner were operated at a maximum speed of 10 km/hr.
    b. Human predictions were recorded directly into a provided cabin HMI touchscreen, and bunk scanner volume predictions were automatically stored to file onboard its processor. All predictions were also automatically pushed to a cloud server as backup (and when internet coverage was available).
    c. Livestock were pushed back from the feed so no more was consumed before collection.
2. Feed was collected in buckets and vacuumed from each bunk, and the recovered masses were dumped onto segregated tarpaulins in a protected location in the feedlot away from normal operations. In some cases, to reduce double-handling, the recovered masses were not transferred to tarpaulins, and instead were measured immediately on the platform scale in situ (i.e. within the pen roads) and returned to the appropriate pen bunks. Humans were kept blind from feed recovery and weigh-back activities.
3. Steps 1 through 2 were repeated until the feed trucks commenced the morning deliveries.
4. At Feedlots B and C, after the morning feed delivery had enough time for some consumption, an additional round of measurements was undertaken; this usually commenced at 1030h, with a duration of a few hours prior to the afternoon feed delivery.
5. Feed density for the experiment was determined from a quotient calculated from significant known feed mass delivered (usually Rotomix 920-18; Dodge City, KS, USA) and its scale-head (Digistar EZ indicator; Fort Atkinson, WI, USA; ±5 kg readability) divided by the predicted volume, averaged over three bunks. The quotient was multiplied by scanned volume to determine predicted feed remaining.

The following section summarises the result and analyses outputs for the experiment by host site.

4 Results

The following section describes the statistical analyses employed to assess the outputs of the adoption experiment, followed by presentation of the results by site.

4.1 Statistical analyses

Several statistical analyses have been undertaken with a view to objectively assessing the performances of the prototype system and human bunk callers.

Observed feed remaining has been regressed on predicted feed remaining for both the prototype system and human bunk callers. The coefficient of determination (r^2) has been calculated on the line of regression as a measure of the strength of the relationship between observed and predicted feed remaining.

Evaluation of the model's precision has been enabled through employment of several commonly used measures of deviance, including mean absolute error (MAE), mean square prediction error (MSPE), and root mean square error (RMSPE). Shah and Murphy (2006) defined MSPE as: $\Sigma (O_i - P_i)^2/n$, where n = number of paired observed (O) and predicted (P) feed remaining values being compared. The MAE is defined as: $(\Sigma|O_i - P_i|)/n$.

Furthermore, the MSPE can be decomposed to assess sources of variation, viz, (1) variation in central tendency (mean bias), (2) variation resulting from regression (systematic bias or line bias), and (3) random variation.

Variation resulting from mean bias has been calculated by squaring the mean bias of the prediction, whereas variation resulting from systematic bias was calculated as the product of the variance of the predicted feed remaining and the square of the deviation from 1 of the slope of the regression of observed on predicted data. Random variation was calculated as the product of the variance of observed data and the deviation from 1 of the coefficient of determination of the regression of observed on predicted data. Shah and Murphy (2006) noted that mean bias is useful to test the robustness of the model, whereas line bias can be used to test inadequacy in model structure. Mean proportional bias has been calculated as the slope of the regression of the predicted data on observed data with an intercept of 0 (Shah and Murphy, 2006). Over the range of observed values, a value of mean proportional bias less than one (< 1) denotes underprediction, whereas a value more than one (> 1) denotes overprediction.

In addition, mean and linear biases were calculated by regression of residuals (observed minus predicted feed remaining) on mean-centred predicted feed remaining (St-Pierre, 2003). St-Pierre (2003) noted that by centering predicted feed remaining to the mean value, the intercept of the linear model is estimated at the mean value of the independent variable rather than a value of zero.

The intercept term at the mean value is a measure of the mean prediction bias, and a t-test on the estimate of the intercept has been used to determine the statistical significance of this bias. The slope of this mean-centred regression is an estimate of the linear prediction bias, and a t-test has been used again to test significance. When the linear prediction bias has been found to be significant ($P < 0.10$), the magnitude of the bias within the range of predicted values was determined by calculating the bias at the minimum and maximum data points of the predicted values (St-Pierre, 2003).

4.2 Results

The following section contains a summary of data analyses, employed across the scope of the experiments for the prototype system and human callers. Plots of the results for the total datasets have also been provided.

4.2.1 Feedlot A summary of regressions

Six human bunk callers and one diet (steam-flaked white cereal grain finisher ration) were used during the Feedlot A experiment. The average forward travel speed of the light vehicle and prototype automatic bunk calling system for the 100 observations was 7.4kmh^{-1}. Fig. 1 demonstrates the composition of feed remaining masses observed during the feedlot experiment, and then as a function of pen head (/hd) count respectively; this scale is also defined in the second histogram.

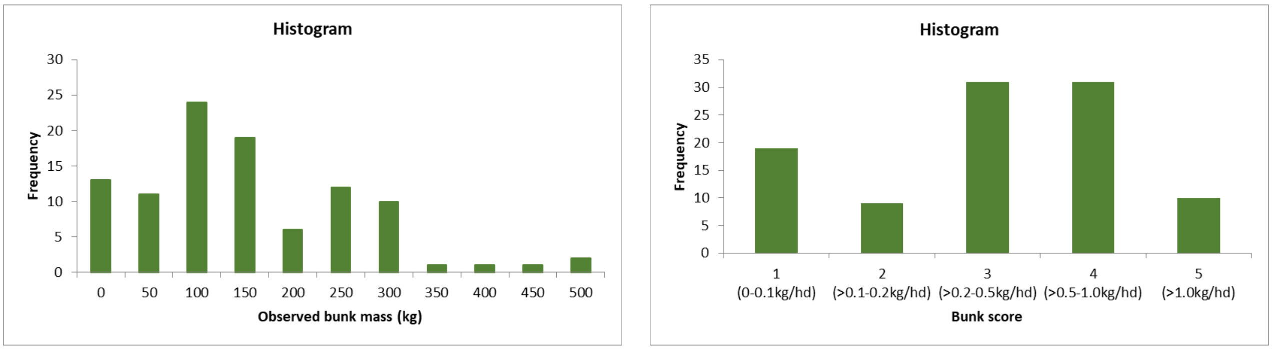


Fig. 1: Histogram of observed feed remaining masses (left), and these masses normalised over head count for nominal bunk scores for Feedlot A experiment.

Weather conditions for the Feedlot A experiment's duration are included in Table 2 below, sourced from the Bureau of Meteorology website. Data was also provided by the site operator from their Environdata Weather Station (Warwick, Q), however this system was offline for some time during the experiment period.

Table 2: Significant weather conditions at Feedlot A during experiment in June 2018, with no measurements made on 'greyed-out' days.

Day	11/06	12/06	13/06	14/06	15/06	16/06	17/06	18/06	19/06	20/06	22/06
Maximum temp. (°C)	22.7	25.8	19.7	20.4	20.7	20.9	17.2	15.8	16.9	19.1	21.6
Minimum temp. (°C)	3.9	2.0	7.7	10.4	7.1	(0.5)	1.0	(2.5)	2.6	1.5	0.5
Rainfall (mm)	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Results of the regression of observed on predicted feed remaining are shown in Table 3. Mean and linear biases are also reported, determined from the regression of residuals on mean-centred predicted feed remaining.

Table 3: Evaluation statistics of feed remaining predictions for humans and prototype bunk scanner system at Feedlot A, 100 observations from slick to 473kg. Mean and linear biases calculated using St-Pierre (2003) techniques.

Item	Humans	Prototype system
Mean bias, kg	14.52	8.37
P-value	0.01	< 0.01
Linear bias	-0.13	0.00
P-value	0.03	0.91
r^2, regression of observed on predicted feed remaining	0.70	0.99
RMSPE, kg	61.26	13.76
MSPE, kg^2	3,752.78	189.21
MAE, kg	34.61	11.22
Mean proportional bias	0.85	0.96
Decomposition of MSPE		
Mean bias, %	6%	37%
Systematic bias, %	5%	0%
Random bias, %	90%	63%
Bias at minimum predicted value, kg	29.51	-
Bias at maximum predicted value, kg	-22.43	-

It can be seen from Table 3 that the prototype system accurately and precisely predicted feed remaining in bunks, outperforming human callers in both criteria. The prototype system had a small amount of mean bias (P-value < 0.01), underpredicting feed remaining by 8.3kg. The prototype system had no linear bias (P-value = 0.91); that is, bias was consistent over the full range of feed remaining in the bunk, and so it follows that minimum and maximum biases have not been calculated. Precision of the prototype was excellent during the experiment (r^2 = 0.99). Mean absolute error for the prototype system was 11.2kg, and the RMSPE was 13.8kg.

Human performance was variable and less accurate. A significant mean bias (P-value = 0.01) was reported, with bunk callers underestimating (14.5kg) feed remaining, albeit moderate precision (r^2 = 0.70). Significant linear bias (P-value = 0.03) was reported, with magnitude much larger than the prototype (29.5 to (-22.4kg)). Mean absolute error for human callers was 34kg, and the RMSPE was 61kg.

4.2.2 Feedlot A chart results

The total experiment observed on predicted feed remaining for humans and prototype system (machine) are provided graphically in Fig. 2. These plots demonstrate the prototype system's results are very close to the ideal unity gradient with insignificant offset (based on y-intercept).

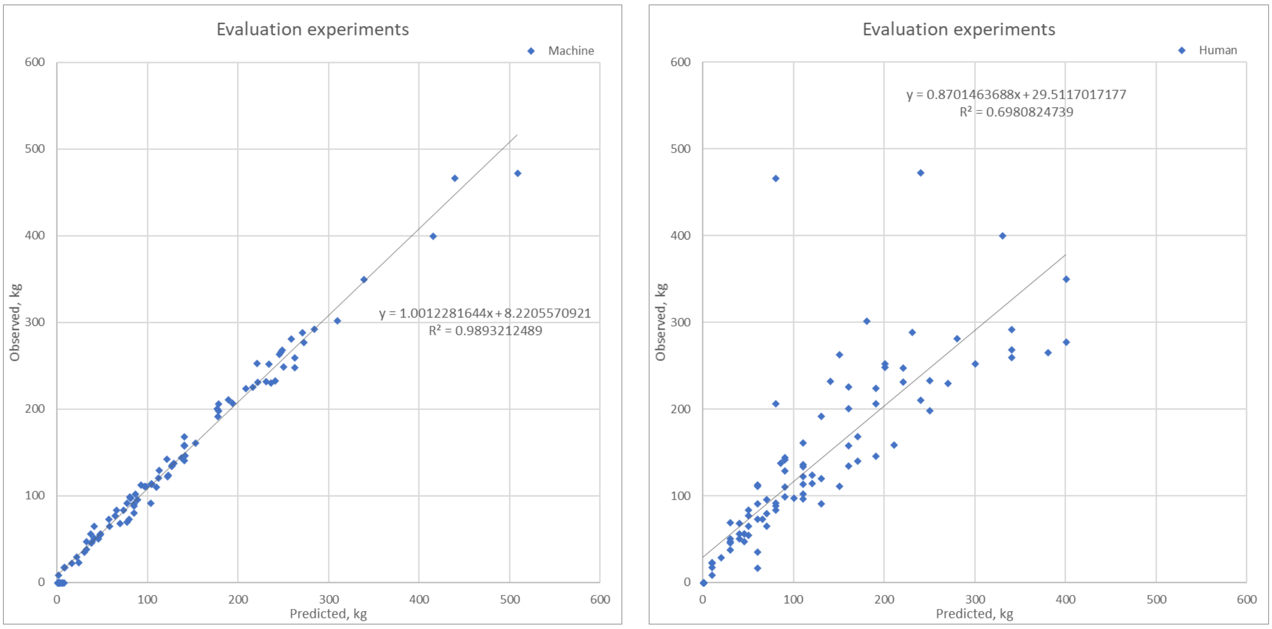


Fig. 2: Observed on predicted feed remaining masses for prototype system (machine, left), and human callers, for Feedlot A experiment.

Sources of errors for the total experiment for humans and the prototype system (machine) are represented graphically below in Fig. 3, i.e. the residuals (errors, observations minus predictions) over mean-centred predictions. It can be seen again from the prototype system's error decomposition that a small mean bias exists (y-offset = 8.3kg), with very limited linear bias (gradient near zero), and a very low level of variability (precision). Conversely, the human plot further reinforces the significance of both the mean bias (underpredicting by 14.5), linear bias (significant negative gradient as masses increase), as well as overall lack of precision.

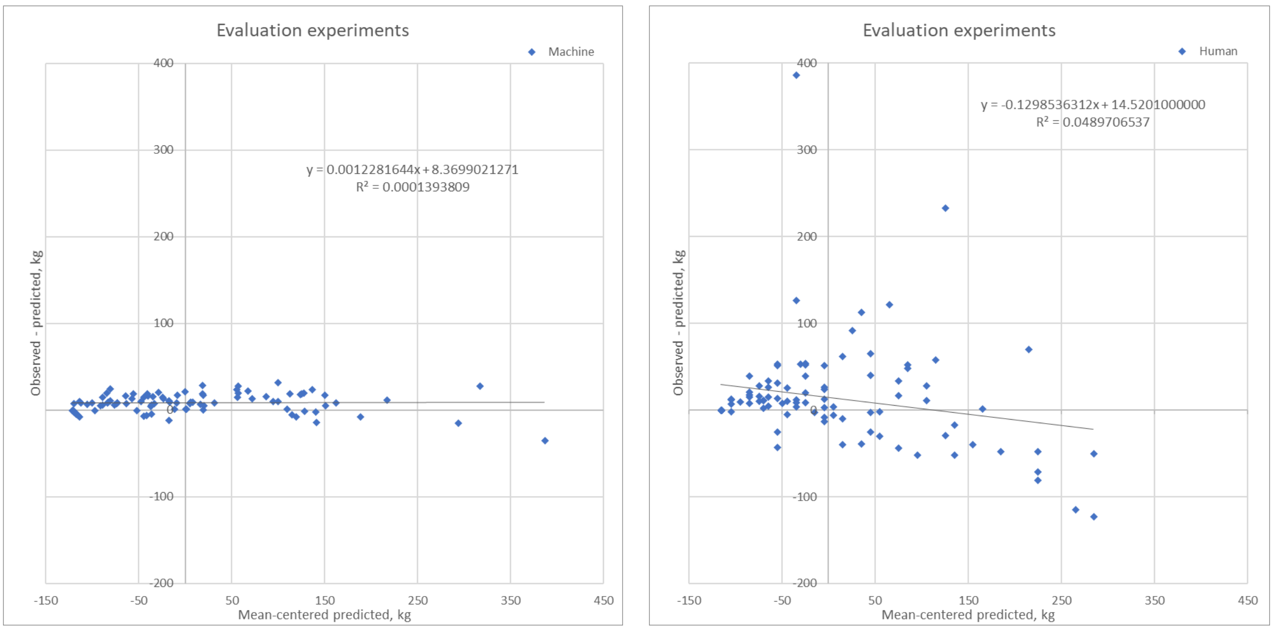


Fig. 3: Experiment residuals over mean-centred predictions for assessing sources of error in prediction models for Feedlot A experiment.

4.2.3 Impact of prediction errors on bunk scores for Feedlot A during experiment

Table 4 summarises the potential impact of incorrect human bunk calls for the pens observed through the course of the experiment at Feedlot A.

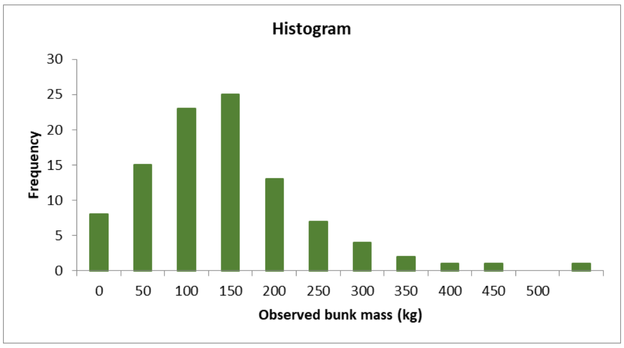
Table 4: Success table when applying bunk scores for humans and prototype at Feedlot A.

Bunk scores	1 (0-0.1kg)	2 (>0.1-0.2kg/hd)	3 (>0.2-0.5kg/hd)	4 (>0.5-1.0kg/hd)	5 (>1.0kg/hd)
Observations	19	9	31	31	10
Machine	19	8	28	28	9
	100.0%	88.9%	90.3%	90.3%	90.0%
Humans	18	6	24	21	7
	94.7%	66.7%	77.4%	67.7%	70.0%

From this table, human callers had a higher number of prediction errors, especially as the feed remaining masses increased, and this is consistent with the previous results. The prototype automatic system had less errors, and on review these were all on the cusps of the nominal correct scores. This also indicates that very accurate masses (and therefore scores) could be returned by the prototype automatic bunk calling system with some minor adjustment to the prediction model, consistent with the outputs provided earlier.

4.2.4 Feedlot B summary of regressions

Four human bunk callers and one diet (steam-flaked white cereal grain finisher ration) were used during the Feedlot B experiment. The average forward travel speed of the light vehicle and prototype automatic bunk calling system for the 100 observations was 6.2kmh^{-1}. Fig. 4 demonstrates the composition of feed remaining masses observed during the feedlot experiment, and then as a function of pen head count respectively.



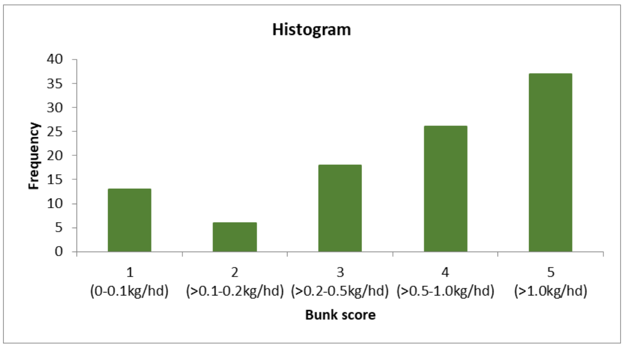


Fig. 4: Histogram of observed feed remaining masses (left), and these masses normalised over head count for nominal bunk scores for Feedlot B experiment.

Weather conditions for the Feedlot B experiment's duration are included in Table 5 below, sourced from the on-site digital weather station (Environdata Weather Station, Warwick, Q).

Table 5: Significant weather conditions at Feedlot B during experiment in July 2018, with no measurements made on 'greyed-out' days.

Day	03/07	04/07	05/07	06/07	07/07	08/07	09/07	10/07	11/07	12/07
Maximum temp. (°C)	19.0	20.4	22.8	25.4	20.8	21.2	15.5	17.0	20.4	21.6
Minimum temp. (°C)	11.2	10.6	8.6	12.8	16.5	4.7	2.0	1.0	5.5	6.7
Rainfall (mm)	0.0	0.0	0.0	0.0	1.8	0.0	0.0	0.0	0.0	0.0

Results of the regression of observed on predicted feed remaining are shown in Table 6. Mean and linear biases are also reported, determined from the regression of residuals on mean-centred predicted feed remaining.

Table 6: Evaluation statistics of feed remaining predictions for humans and prototype bunk scanner system at Feedlot B, 100 observations from slick to 510kg. Mean and linear biases calculated using St-Pierre (2003) techniques.

Item	Humans	Prototype system
Mean bias, kg	32.89	11.93
P-value	< 0.01	< 0.01
Linear bias	0.28	0.01
P-value	< 0.01	0.46
r^2, regression of observed on predicted feed remaining	0.71	0.99
RMSPE, kg	62.79	15.17
MSPE, kg^2	3,943.16	230.08
MAE, kg	41.20	12.65
Mean proportional bias	0.67	0.93
Decomposition of MSPE		
Mean bias, %	28%	62%
Systematic bias, %	8%	0%
Random bias, %	64%	38%
Bias at minimum predicted value, kg	7.19	-
Bias at maximum predicted value, kg	87.10	-

It can be seen from Table 6 that the prototype system accurately and precisely predicted feed remaining in bunks, outperforming human callers in both criteria. The prototype system had a small amount of mean bias (P-value < 0.01), underpredicting feed remaining by 11.93kg. The prototype system had no linear bias (P-value = 0.46); that is, bias was consistent over the full range of feed remaining in the bunk, and so it follows that minimum and maximum biases have not been calculated. Precision of the prototype was excellent during the experiment ($r^2 = 0.99$). Mean absolute error for the prototype system was 12.65kg, and the RMSPE was 15.17kg.

Human performance was variable and less accurate. A significant mean bias (P-value < 0.01) was reported, with bunk callers underestimating (32.89kg) feed remaining, albeit with moderate precision ($r^2 = 0.71$). Significant linear bias (P-value = 0.01) was reported, with magnitude much larger than the prototype (7.19 to 87.10kg). Mean absolute error for human callers was 41.2kg, and the RMSPE was 62.79kg. These results are very consistent with the Feedlot A experiment outcomes.

4.2.5 Feedlot B chart results

The total experiment observed on predicted feed remaining for humans and the prototype system (machine) are represented graphically below in Fig. 5.

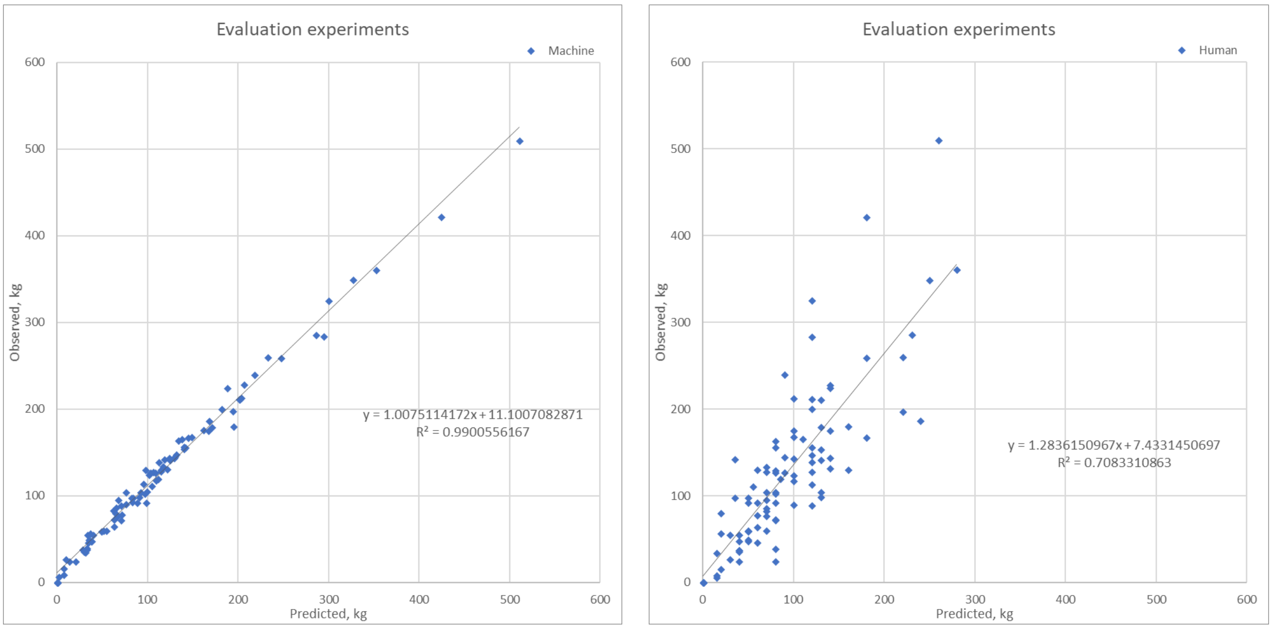


Fig. 5: Observed on predicted feed remaining masses for prototype system (machine, left), and human callers, for Feedlot B experiment.

Sources of errors for the total experiment for humans and the prototype system (machine) are represented graphically below in Fig. 6, i.e. the residuals (errors, observations minus predictions) over mean-centred predictions. It can be seen again from the prototype system's error decomposition that a small mean bias exists (y-offset = 11.93kg), with very limited linear bias (gradient near zero), and a very low level of variability (high precision). Conversely, the human plot further reinforces the significance of both the mean bias (underpredicting an average 32.88kg), linear bias (significant positive gradient as masses increase), as well as overall lack of precision.

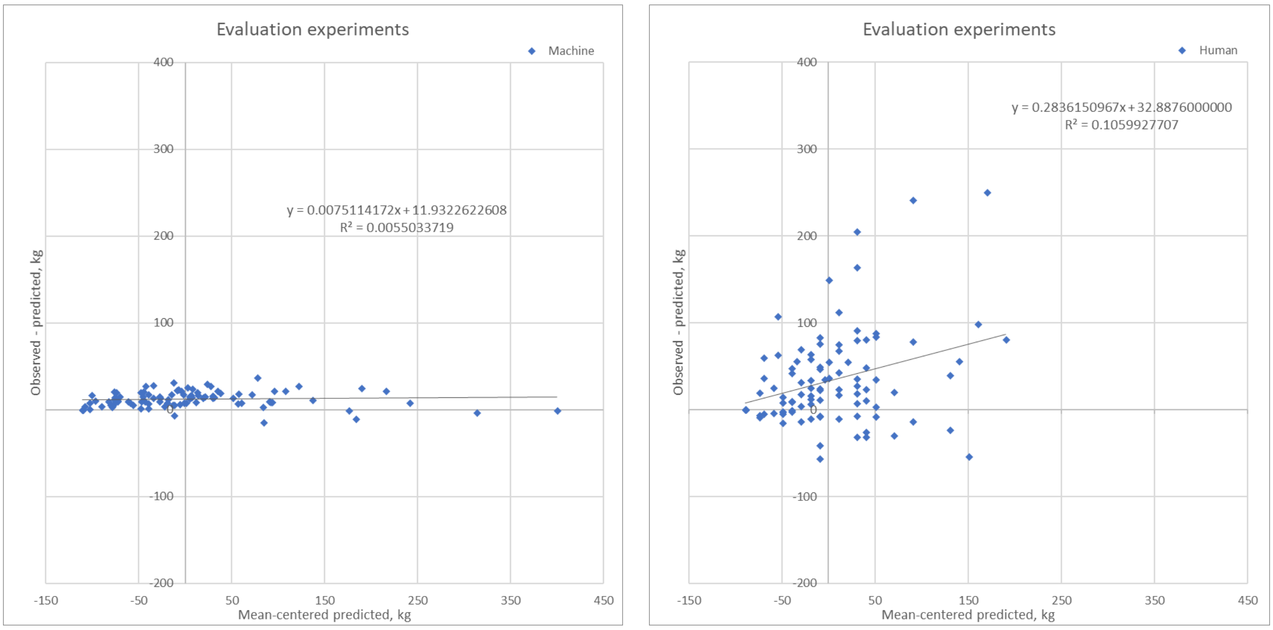


Fig. 6: Experiment residuals over mean-centred predictions for assessing sources of error in prediction models for Feedlot B experiment.

4.2.6 Impact of prediction errors on bunk scores for Feedlot B during experiment

Table 7 summarises the potential impact of incorrect human bunk calls for the pens observed through the course of the experiment at Feedlot B.

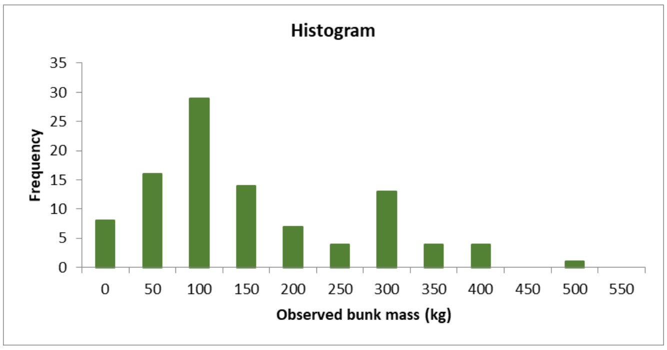
Table 7: Success table when applying bunk scores for humans and prototype at Feedlot B.

Bunk scores	**1 (0-0.1kg)**	**2 (>0.1-0.2kg/hd)**	**3 (>0.2-0.5kg/hd)**	**4 (>0.5-1.0kg/hd)**	**5 (>1.0kg/hd)**
Observations	13	6	18	26	37
Machine	13	4	14	16	32
	100.0%	66.7%	77.8%	61.5%	86.5%
Humans	9	3	13	10	26
	69.2%	50.0%	72.2%	38.5%	70.3%

From this table, human callers had a higher number of prediction errors, especially as the feed remaining masses increased, and this is consistent with the previous results. The prototype automatic system had less errors, and on review these were all on the cusps of the nominal correct scores. This also indicates that very accurate masses (and therefore scores) could be returned by the prototype automatic bunk calling system with some minor adjustment to the prediction model, consistent with the outputs provided earlier.

4.2.7 Feedlot C summary of regressions

One human bunk caller and one diet (steam-flaked white cereal grain finisher ration) were used during the Feedlot C experiment. The average forward travel speed of the light vehicle and prototype automatic bunk calling system for the 100 observations was 7.6kmh^{-1}. Fig. 7 demonstrates the composition of feed remaining masses observed during the feedlot experiment, and then as a function of pen head count respectively.



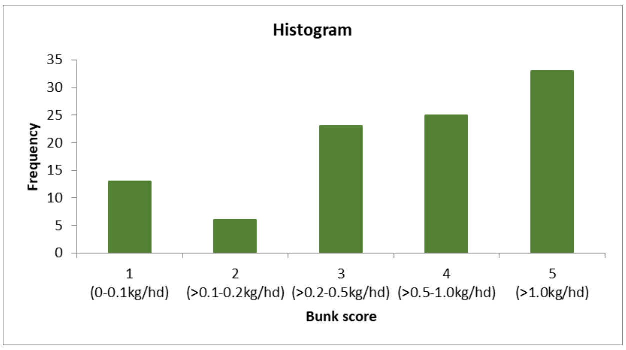


Fig. 7: Histogram of observed feed remaining masses (left), and these masses normalised over head count for nominal bunk scores for Feedlot C experiment.

Weather conditions for the Feedlot C experiment's duration are included in Table 8 below, sourced from the on-site digital weather station (MEA Feedlot Weather Station, Magill, SA).

Table 8: Significant weather conditions at Feedlot C during experiment in July and August 2018, with no measurements made on 'greyed-out' days.

Day	31/07	01/08	02/08	03/08	04/08	05/08	06/08	07/08	08/08	09/08	10/08
Maximum temp. (°C)	17.7	20.9	22.4	23.0	22.7	22.0	27.0	24.7	17.2	18.6	19.9
Minimum temp. (°C)	(1.2)	(0.6)	0.3	5.2	7.1	2.4	6.2	4.4	0.0	(0.7)	0.6
Rainfall (mm)	0.0	0.0	3.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Results of the regression of observed on predicted feed remaining are shown in Table 9. Mean and linear biases are also reported, determined from the regression of residuals on mean-centred predicted feed remaining.

Table 9: Evaluation statistics of feed remaining predictions for human and prototype bunk scanner system at Feedlot C, 100 observations from slick to 466kg. Mean and linear biases calculated using St-Pierre (2003) techniques.

Item	Human	Prototype system
Mean bias, kg	2.96	6.43
P-value	0.56	< 0.01
Linear bias	-0.17	0.02
P-value	< 0.01	0.09
r^2, regression of observed on predicted feed remaining	0.80	0.99
RMSPE, kg	53.79	13.94
MSPE, kg^2	2,893.67	194.44
MAE, kg	36.78	10.17
Mean proportional bias	0.96	0.96
Decomposition of MSPE		
Mean bias, %	0%	21%
Systematic bias, %	14%	2%
Random bias, %	86%	77%
Bias at minimum predicted value, kg	24.44	-
Bias at maximum predicted value, kg	-75.84	-

It can be seen from Table 9 that the prototype system accurately and precisely predicted feed remaining in bunks, outperforming human callers in both criteria. The prototype system had a small amount of mean bias (P-value < 0.01), underpredicting feed remaining by 6.43kg. The prototype system had no linear bias (P-value = 0.09); that is, bias was consistent over the full range of feed remaining in the bunk, and so it follows that minimum and maximum biases have not been calculated. Precision of the prototype was excellent during the experiment ($r^2 = 0.99$). Mean absolute error for the prototype system was 10.17kg, and the RMSPE was 13.94kg.

Human performance was variable and less accurate. The reported mean bias was not significant (P-value = 0.56); that is, predictions were equally over and under observed masses, as reflected by the high random bias contribution (86%) to the MSPE. Humans had moderate precision ($r^2 = 0.80$). Significant linear bias (P-value = 0.01) was reported, with magnitude much larger than the prototype (24.44 to -75.84kg). Mean absolute error for the human caller was 36.78kg, and the RMSPE was 53.79kg. These results are very consistent with the previous experiment outcomes, and the improved human result may be a function of only using one caller compared to the other sites, and this caller demonstrating reasonable performances against absolute metrics.

4.2.8 Feedlot C chart results

The total experiment observed on predicted feed remaining for humans and the prototype system (machine) are represented graphically below in Fig. 8.

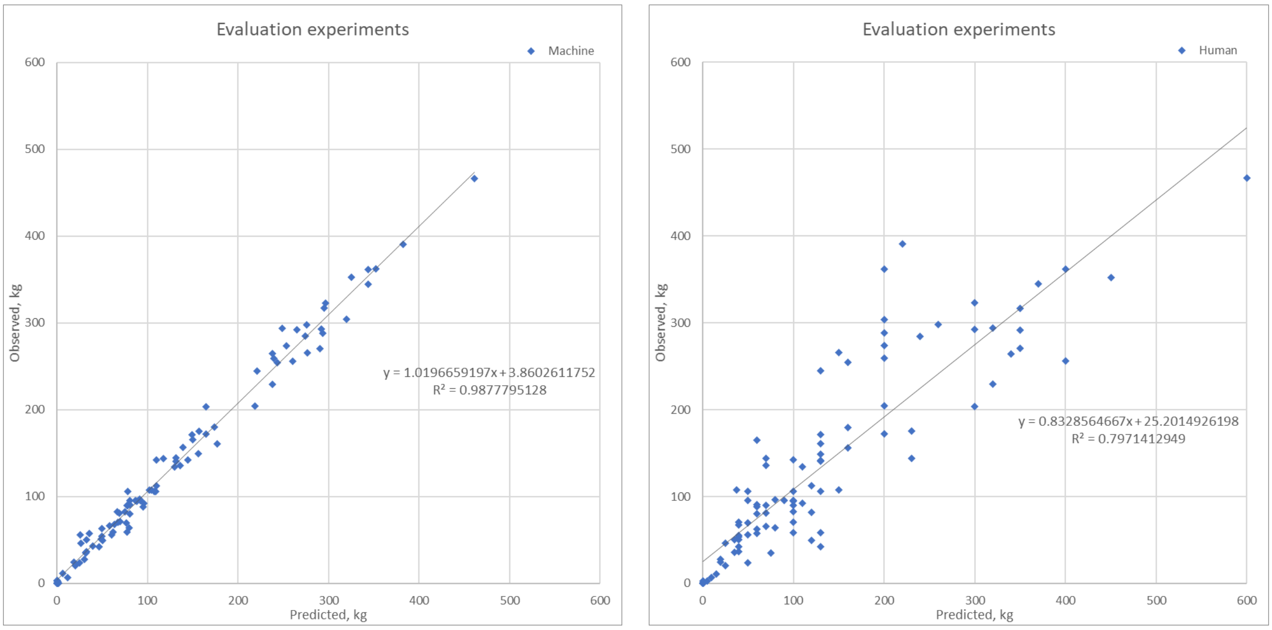


Fig. 8: Observed on predicted feed remaining masses for prototype system (machine, left), and human caller, for Feedlot C experiment.

Sources of errors for the total experiment for human and the prototype system (machine) are represented graphically below in Fig. 9, i.e. the residuals (errors, observations minus predictions) over mean-centred predictions. It can be seen again from the prototype system's error decomposition that a small mean bias exists (y-offset = 6.43kg), with very limited linear bias (gradient near zero), and a very low level of variability (precision). Conversely, the human plot further reinforces the insignificance of the very low mean bias (imprecision around y-intercept at 2.96kg), linear bias (significant negative gradient as masses increase), as well as overall lack of precision. In this case the caller tended to overpredict as the feed remaining masses increased, and this is behaviour is not consistent with the significant underpredictions with increasing masses observed at the previous sites.

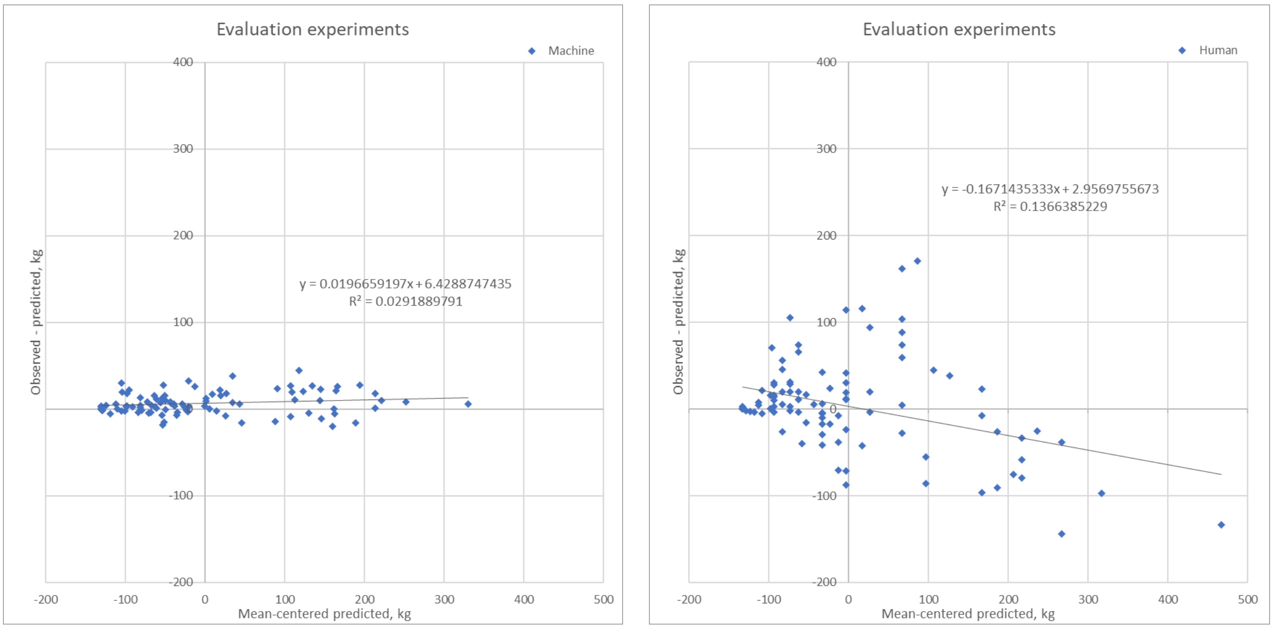


Fig. 9: Experiment residuals over mean-centred predictions for assessing sources of error in prediction models for Feedlot C experiment.

4.2.9 Impact of prediction errors on bunk scores for Feedlot C during experiment

Table 10 summarises the potential impact of incorrect human bunk calls for the pens observed through the course of the experiment at Feedlot C.

Table 10: Success table when applying bunk scores for humans and prototype at Feedlot C.

Bunk scores	**1**	**2**	**3**	**4**	**5**
	(0-0.1kg)	**(>0.1-0.2kg/hd)**	**(>0.2-0.5kg/hd)**	**(>0.5-1.0kg/hd)**	**(>1.0kg/hd)**
Observations	13	6	23	25	33
Machine	13	5	17	20	30
	100.0%	83.3%	73.9%	80.0%	90.9%
Humans	13	3	12	13	25
	100.0%	50.0%	52.2%	52.0%	75.8%

From this table, human callers had a higher number of prediction errors, especially as the feed remaining masses increased, and this is consistent with the previous results. The prototype automatic system had less errors, and on review these were all on the cusps of the nominal correct scores. This also indicates that very accurate masses (and therefore scores) could be returned by the prototype automatic bunk calling system with some minor adjustment to the prediction model, consistent with the outputs provided earlier.

5 Discussion

Against the results presented in this report, the prototype bunk calling system has been demonstrated to be highly repeatable (precise) and accurate for predicting feed remaining in bunks across a diverse range of commercial feedlots.

In contrast across all the experiments, human callers predict feed remaining in bunks with significantly less accuracy and precision, and these errors are particularly evident at higher masses. Their performances with lesser feed remaining masses are probably acceptable for normal operating requirements.

The potential impact of incorrect human bunk calls for the pens observed throughout the whole experiment (three feedlots) are shown in Table 11.

Table 11: Success table when applying bunk scores for humans and prototype at all three experiment feedlots.

Bunk scores	1	2	3	4	5
	(0-0.1kg)	(>0.1-0.2kg/hd)	(>0.2-0.5kg/hd)	(>0.5-1.0kg/hd)	(>1.0kg/hd)
Observations	45	21	72	82	80
Machine	45	17	59	64	71
	100.0%	81.0%	81.9%	78.0%	88.8%
Humans	40	12	49	44	58
	88.9%	57.1%	68.1%	53.7%	72.5%

It should be noted that the overall results achieved at each host site during this experiment are very consistent with the feedlot validation experiment completed during the prototype system's development at Feedlot B (55 observations completed in January 2018, results published in final report of project B.FLT.0166).

A statistical methodology has been exercised with a view to assess feed remaining predictions provided by the prototype system and human callers. The methodology has provided very clear and objective support for the prototype system's precision and accuracy. The feedlot experiment results also suggest the significance of feed density knowledge, and especially how it may change after-delivery through a 24-hour cycle. It is believed that this is the major contributing factor for the feedlot experiment's minor underpredictions, though the influence of this effect could be simply mitigated programmatically: simple simulations of the prototype system's predictions indicate a small increase in density (less than 10%) removes almost all underpredictions. We suggest that possible sources of changes in feed remaining density across the 24-hour cycle may be attributable through settling and compaction, probably because of the feed's duration in the bunk, as well as the livestock's interactions with the feed.

6 Conclusions/recommendations

Against the results presented in this report, the prototype bunk calling system provided highly repeatable (precise) and accurate feed remaining predictions across three disparate commercial feedlots. In contrast, the human callers provide significantly less accurate and precise feed remaining predictions across all the experiments, especially at higher masses. Their performances with lesser feed remaining masses are probably acceptable for normal operating requirements.

A statistical methodology has been exercised with a view to assess feed remaining predictions provided by the prototype system and human callers. The methodology has provided very clear and objective support for the prototype system's precision and accuracy. The feedlot experiment results also suggest the significance of feed density knowledge, and especially how it may change after-delivery through a 24-hour cycle. It is believed that this is the major contributing factor for the prototype system's minor underpredictions, and its influence will in practice be simply mitigated programmatically.

7 Key messages

We are very excited about the potential benefits that the red meat industry should garner through our technology solutions partnership. We have been very pleased with the establishing relationship between ourselves and MLA, and our ability to respond effectively to this opportunity through the prompt supply of the working prototype system.

With MLA's support, we have delivered a prototype that has demonstrated significant potential benefit to the red meat industry. We suggest that a commercially-viable product should be available in the near term based on the successful outcomes of this project.

8 Bibliography

The following literature was cited in this report.

Shah, M. A., and M. R. Murphy. 2006. Development and evaluation of models to predict the feed intake of dairy cows in early lactation. J. Dairy Sci. 89:294–306.

St-Pierre, N. R. 2003. Reassessment of biases in predicted nitrogen flows to the duodenum by NRC 2001. J. Dairy Sci. 86:344–350.