final report

Project code:	V.LDL.1904
Prepared by:	Dr Alex Ball, Kimbal Curtis, Johan Boshoff and Dr Paul Blazkiewicz Rural Analytics, Livestock Dynamics and UNE
Date published:	22/10/2019

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

Investigation into Lean Meat Yield (LMY) algorithms in Livestock Data Link (LDL)

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

This publication is published by Meat & Livestock Australia Limited ABN 39 081 678 364 (MLA). Care is taken to ensure the accuracy of the information contained in this publication. However MLA cannot accept responsibility for the accuracy or completeness of the information or opinions contained in the publication. You should make your own enquiries before making decisions concerning your interests. Reproduction in whole or in part of this publication is prohibited without prior written consent of MLA.

Abstract

Executive summary

Meat and Livestock Australia through the Integrity Systems Company (ISC) has created a platform for accessing and benchmarking carcase data for the Australian red meat industry. This program called Livestock Data Link (LDL) is delivering carcase feedback to producers and processors that enables the prediction, understanding and manipulation of production and processing efficiency. One of the key indicators of efficiency is Lean Meat Yield (LMY) which is often calculated as the percentage of lean meat relative to carcase weight or LMY%.

Within the LDL program, prediction algorithms or equations for LMY% have been developed by Australian research providers using carcase data from controlled research herds and flocks such as the Sheep CRC information nucleus. These algorithms have relied on having accurate measurements of LMY that have been derived from CT-scanning of lamb and beef carcases and relating that information to linear measurements of carcase performance that are routinely collected in commercial processing facilities. Whilst these LMY% algorithms are statistically significant within the research datasets from which they have been derived, the application of these equations in large scale commercial datasets such as LDL has not been reviewed.

As a result, ISC have commissioned this review of the LMY% prediction algorithms in LDL for both sheep and cattle with the following objectives: -

- Prepare a register of existing LMY% prediction algorithms, and what attributes (independent variables) they use and what the expected range and accuracies of such algorithms are.
- Using industry data, test the accuracy and range of LMY% predictions from a range of cattle and sheep processors and different markets.
- Determine the correlations between the LMY% predictions and the industry collected independent variables and investigate potential for biases in both the scale and accuracy of LMY% prediction.
- Investigate the implications of variations in available data to LDL i.e. what attribute is not available and hence prevents implementing a specified prediction equation. Assess potential to capture the missing attributes.
- Undertake a comparison of the performance of some of the “available algorithms which have lesser accuracy” against actual LMY% data or predictions from currently implemented algorithms.
- Provide recommendations of the application of alternative LMY% algorithms, against the objective of industry credibility and standardisation
- Provide recommendations on internal monitoring procedures within LDL for LMY% prediction.

In August 2019, the project team received an extract from the LDL database of carcase data from four beef and 2 sheep processors. This data had been deidentified by ISC prior to receipt to ensure that processor confidentiality of the processor was maintained. The data supplied contained information on carcase traits collected as standard AUSMEAT assessment for sheep and cattle, being hot standard carcase weight (HSCW) and a measure of fat depth (Fat-score for sheep and P8 for cattle). In addition, there was carcase data collected during MSA grading that included Rib-Fat, ossification, eye muscle area and the MSA index. Also, within the data supplied were production characteristics including sex, dentition and feed type (grass vs grain). There was no data available for breed.

ISC supplied the current LDL prediction algorithms for both sheep and cattle as well as additional algorithms that had been calculated either prior to the current equations or using different carcase parameters. These algorithms are summarised within the report.

For cattle the current LDL algorithm for predicting percent lean meat yield (LMY%) uses hot standard carcase weight (HSCW) and Rib-fat depth cold (Rib-fat) as follows: -

LMY% =63.52120055 + (HSCW × 0.00635548) + (Rib-fat × -0.80109465)

There were 2.198 million cattle records available in the extracted dataset. These data sets were predominately from NSW producers, 57% male (castrate) and only 13% has Rib-fat available and therefore had estimates of LMY%. An initial investigation of the cattle data showed that for the traits used in the LMY% prediction, there were extreme values (measures) that sat outside the normal range of expectation (3 standard deviations) and as a result the LMY% calculated in these circumstances should be highlighted as potentially being inaccurate.

An analysis of the LMY% calculated from the current algorithm showed that the Rib-fat was having a much stronger influence on LMY% than HSCW, with a standard deviation in Rib-fat resulting in a 2.7% change in comparison to a 0.35% change for a standard deviation in HSCW. LMY% was also negatively correlated with both marbling and the MSA index. The implication of these results is that both producers and processors that are using LMY% to rank performance should be very conscious that those ranking will favour animals that are leaner and therefore may not suit current market scenarios. A systematic analysis of the correlations between LMY% and the component traits across different production criterion (sex, dentition, feed type) did not reveal any obvious biases or statistical irregularities, however when the data supplied did not contain an independent estimate of LMY then some caution needs to be stated in interpretation of this observation.

To improve the scope and scale of prediction of LMY in LDL, additional equations that predicted using P8 fat rather than Rib-fat were compared to the current LDL prediction of LMY%. The results showed a lower than expected correlation of 0.41 and a significant difference in the scale. The authors concluded that these P8 equations were predicting a LMY% that was not compatible with the current LMY%. However, it was recommended that further research be developed that looked for LMY algorithms that used P8 fat. Further historical equations that included additional factors such as sex and EMA were compared and again the correlations obtained were lower than expected. These equations also suffered from the reliance on data only collected during MSA grading and as a result also had limitations in application within LDL.

For sheep the current LDL equation as supplied by ISC was as follows: -

LMY% = 66.4616 + (-0. 113985 x HSCW) + (-1.92786 x GRFatScore)

The data extract contained carcase records on approximately 13 million sheep. Ninety-two percent of those records were coded as having 0 teeth and assumed to be lambs. Less than 0.05% of the records hadsex recorded. The carcase traits supplied were HSCW and a measure of GR fat score that ranged from 1-5. A key finding in this report was that in sheep the predicted LMY% is overwhelmingly dependent on the GR fat score. So much so, that different GR fat scores resulted in clear separation in LMY% irrespective of the HSCW. Importantly higher LMY% was associated with low GR fat scores which is of concern given that the highest LMY% was exclusively associated with a fat score that has no market suitability and is heavily discounted. This outcome should be carefully considered by both producers and processors that are ranking lambs on LMY%.

Whilst the data supplied did not have a measure of GR fat depth in mm, the project team constructed a synthetic dataset to examine the value of using a measurement of GR fat depth vs the GR fat score. That analysis showed that there were some advantages in using GR fat depth as it allowed for the LMY% to form more of a continuous range and provided extra discrimination across carcase weight ranges. Additional analysis using data from the Sheep CRC showed the potential impact of fitting breed in the prediction of LMY%. However, to implement these equations would require the collection of breed data at slaughter. A detailed series of analytical methodologies for calculating LMY% in sheep are included in appendix 1.

New and emerging technologies that provide alternative methods for predicting LMY% offer both challenges and opportunities for LDL. This report showed that when correlations were less than 0.5 there was significant reranking in animals and potentially in producers. However as these technologies are implemented within processing plants, they do provide LDL with the opportunity to collect more data on a greater demographic of animals than could be obtained within one R&D project and therefore they may enable better and more accurate LMY% predictions to be calculated. Regardless, until formal statistical analysis has been completed all new LMY% measurements should be individually coded and stored within the LDL database.

This report has identified several areas in which improvements could be made or additional activities could be included in the LDL program and specifically for the prediction of LMY% from algorithms based on carcase traits in cattle and sheep. Ten recommendations have been made and these are tabulated below.

No.	Recommendation
1	That LDL considers the use of a LMY% predictor based on carcase weight and P8 fat given that 87% of the available data did not have a LMY% prediction.
2	Review function used to calculate LMY% to make sure it does not use zero (or [null] values for HSCW, LeftHSCW or RightHSCW or Rib-fat.
3	Consider how to adjust the prediction equation to exclude, or allow for, independent values outside acceptable limits. This might be based on a review of the data used to derive the equation, or by setting the range at the 95% confidence interval for each variable.
4	That LMY% calculated from data that is outside 95% confidence limits is shaded or coded differently to reflect that one or both of input variables might be inaccurate.
5	That beef processors that want to produce LMY% using LDL need to make the investment to measure Rib-Fat or invest in alternative technologies that potentially measure LMY% directly such as DEXA.
6	Data verification or validation should be conducted to understand how LMY% is being calculated for carcases with no recorded fat score and in some cases no carcase weight.
7	Lean meat yield and aggregate statistics should only be calculated on carcases with a HSCW greater than or equal to defined weight thresholds for each fat score. The lower limit of the 95% confidence interval is proposed as a starting value for consideration. If LMY% is being calculated for values outside the 95% confidence interval (or some other agreed range), then those should be identified or shaded as based on inputs that may be inaccurate.
8	Whenever available, GR fat depth should be recorded in mm as this conveys additional data that is lost when converted to a fat score.

9	LDL should seek to capture breed data, even if only in broad categories (i.e. Merino, Maternal, Terminal). This would allow for more tailored prediction of LMY% that could remove some of the inherit breed biases that are within the data.
10	ISC should investigate whether the addition of a non-linear term for either carcase weight or fat-score would improve the prediction accuracy of LMY%
11	As alternate methods of estimating LMY are implemented by processors, LDL should upload and store both the measurements and the technology used. The data should be used to verify and refine LMY predictions, examine equivalence between methodologies, and potentially improve predictions for carcases that have only have carcase grading measurements available.

Table of contents

1 Background to the LMY% LDL project

Meat and Livestock Australia (MLA) established Livestock Data Link (LDL) as a platform for producers and processors to source and interrogate carcase data, compliance rates to market specifications in grids and the cost of non-compliance. LDL has implemented algorithms or equations for the prediction of Lean Meat Yield (LMY%) from carcase traits measured within processing plants for both cattle and sheep. The accuracy and range of values derived from such algorithms are inherently dependent on the accuracy of data inputs from the processor. The research data used to develop and calibrate the LDL beef and sheepmeat algorithms have been collected under strict protocols, in contrast the data used in LDL reflects industry commercial practice and as a result there is a potential for the equations to produce LMY% predictions with lower accuracy or potentially biases.

LMY% is reported as a percentage of carcase weight and is an important measurement as it impacts on processor efficiency and ultimately the volume of meat sold. Lower LMY% means more trim (lower value product) and increased labour costs. Improving LMY% should lead to improved efficiencies along the value chain – less trim and lower labour requirements for the processor, and for producers, less feed to produce the same amount of lean meat and ultimately better prices.

This project will establish a register and conduct an audit of available LMY% algorithms within LDL that are used for cattle and for sheep. The audit will determine whether the range and accuracy of LMY% predictions are consistent with both scientific and industry expectations and provide recommendations on the application of algorithms. The project will look at the correlations between components of the algorithms and the LMY% predictions to see if there are any inherent biases being developed. There is an important trade-off whereby higher LMY% is associated with a negative impact on eating quality. This relationship will be examined in the beef data where eating quality data is available.

The project will examine the impact of variable data collection on animals across different processing plants, production types, sexes and market categories and undertake a comparison of other available LMY% algorithms and their relative performance, accuracy and range. From this, recommendations on the current application of the LDL LMY% algorithms will be made. In addition, where appropriate an assessment of the impacts of implementation of different LMY% algorithms will be provided.

The project will provide recommendations on methodologies that can implemented within LDL to determine when data that is missing or potentially inaccurate is affecting the performance of LMY% predictions.

2 Project objectives

The following project objectives were defined in the request for quotation (RFQ) that the Integrity Systems Company (ISC) provided for this project. Each of the objectives were considered for both beef and sheepmeat.

2.1 Project Objectives

- Prepare a register of existing LMY% prediction algorithms, and what attributes (independent variables) they use and what the expected range and accuracies of such algorithms are.
- Using industry data, test the accuracy and range of LMY% predictions from a range of cattle and sheep processors and different markets.

- Determine the correlations between the LMY% predictions and the industry collected independent variables and investigate potential for biases in both the scale and accuracy of LMY% prediction.
- Investigate the implications of variations in available data to LDL i.e. what attribute is not available and hence prevents implementing a specified prediction equation. Assess potential to capture the missing attributes.
- Undertake a comparison of the performance of some of the “available algorithms which have lesser accuracy” against actual LMY% data or predictions from currently implemented algorithms.
- Provide recommendations of the application of alternative LMY% algorithms, against the objective of industry credibility and standardisation
- Provide recommendations on internal monitoring procedures within LDL for LMY% prediction.

3 Methodology

3.1 Data provided

The project team was provided with a database of records from LDL by the ISC in August 2019. The data came from four beef and two sheep processors. The data sets were deidentified prior to being made available to the project team to avoid any potential biases and issues with producer and processor confidentiality. The database, as supplied, contained separate tables for cattle and sheep carcases that identified the carcase traits and derived variables such as LMY% and the MSA index, provided some indication of production type, dentition and sex; and finally provided dates for both slaughter (sheep and cattle) and grading (cattle). An overview of this data is presented in the following sections.

The data was loaded into a SQL database and examined with standard database and statistical software.

3.1.1 Cattle

For cattle, there were carcases with kill dates from 6-Jan-2014 through to 25-Jul-2019.

The table below shows basic statistics for the primary measurements recorded for cattle carcases. Only 13% of the 2.198 million cattle carcases in the database had a full complement of these measurements. Some of the key features of the data received included: -

- Male cattle accounted for 57% of the 1.576 million carcases for which a sex was recorded.
- Over two thirds of the cattle were sourced from PICs located within NSW (Figure 1).

Table 1 Mean, standard deviation and 25th and 75th percentiles for measurements on cattle carcases (n=281,182)

Variable	Mean	Standard deviation	25th percentile	75th percentile
HSCW (kg)	328.7	55.3	291.5	366.0
Rib-fat (mm)	5.60	3.35	3.34	7.86
Fat depth P8 (mm)	16.3	7.2	11.9	21.1
Eye muscle area(cm)	76.5	12.3	68.2	84.9
AUS marbling (score)	1.56	1.12	0.80	2.31
MLA index (score)	58.2	4.4	55.2	61.1
LMY% (%)	61.1	2.6	59.4	62.9

Of those cattle with the measurements in the table above

- 28% were grain finished, as compared with the 72% finished on grass.
- 23% had received HGP

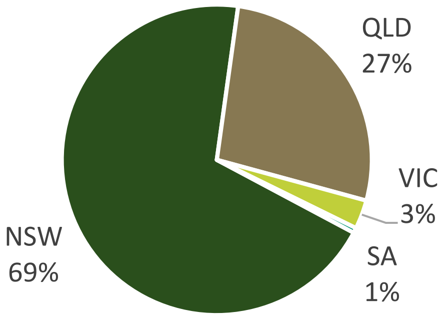


Figure 1 Origin of carcases based on producer PIC. (n= 1,501,644)

3.1.2 Sheep

For sheep, there were just under 13 million carcases with kill dates from 18-Dec-2013 to 24-Jul-2019.

The following table gives summary statistics for the measurements reported for the ovine carcases.

- Figure 2 shows the distribution of fat scores.
- Sex was reported for less than 0.05% of the carcases.
- Dentition was report as ‘0’ for 92% of the carcases (lambs).
- Only 34% of the carcases had information to identify a state of origin.

Table 2 Mean, standard deviation and 25th and 75th percentiles for HSCW and LMY(n=5,513,742)

Variable	Mean	Standard Deviation	25th percentile	75th percentile
HSCW (kg)	22.94	4.38	19.99	25.89
Fat score	3.09	0.94	3	4
LMY% (%)	57.43	2.14	55.99	58.87

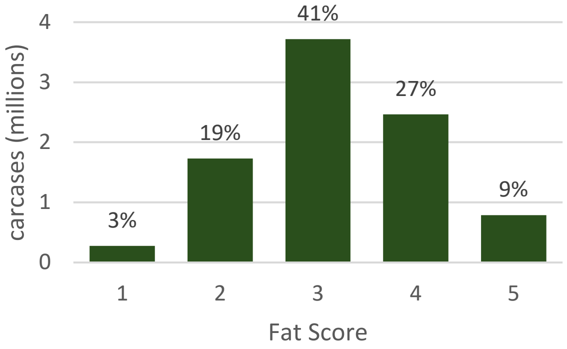


Figure 2 Distribution of lamb carcases between fat scores one through five. (n=8,976,632)

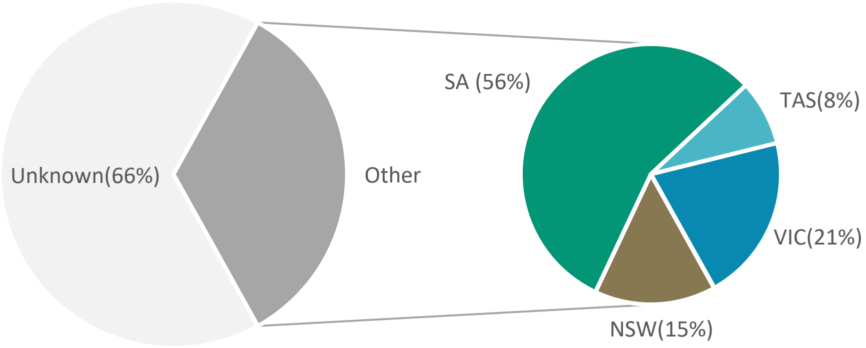


Figure 3 Distribution of carcases by state as identified from the producer PIC field. (n=4,389,824 with a known state).

Whilst the fat score distribution appears to be consistent with industry data, the distribution of carcases by state probably reflects the catchment areas of the contributing processors.

3.2 LMY% Algorithms

3.2.1 Cattle

The timeline for prediction equation development and implementation shown in Equation list 1 was provided to the project team by the ISC. The final equation (June 2018) is the LMY% prediction algorithm (equation) that is currently implemented in LDL.

Equation list 1 Lean meat yield prediction equations for Cattle

2016 – June

For Steers

$$\text{Predicted LMY\%} = 62.1109 + (\text{LeftsideHSCW} \times -0.09244) + (\text{EMA} \times 0.1645) + (\text{Rib-Fat} \times -0.4936)$$

For Heifers

$$\text{Predicted LMY\%} = 59.3974 + (\text{LeftsideHSCW} \times -0.09244) + (\text{EMA} \times 0.1645) + (\text{Rib-Fat} \times -0.4936)$$

2017 – September

$$\text{CTLean\%} = 63.4689 + (-0.604415 \times \text{Rib-fat}) + (-0.007643 \times \text{HSCW})$$

2018 – January

$$\text{CTLean\%} = 59.90525657 + (-0.54566852 \times \text{P8}) + (0.0157628 \times \text{HSCW})$$

2018 – June

$$\text{CTLean\%} = 63.52120055 + (\text{HSCW} \times 0.00635548) + (\text{Rib-fat} \times -0.80109465)$$

Table 3 Variables used in the prediction equations for cattle

Variable	Description
HSCW	Hot standard carcase weight (kg)
Leftside HSCW	HSCW of left half of carcase (kg)
EMA	Eye muscle area (cm^2)
Rib-Fat	Rib fat thickness (cold) (mm) at the 12/13th rib
P8	Fat thickness at the P8 (rump) site (mm)
Predicted LMY%	Predicted lean meat yield as a percent of HSCW (%)
CTLean%	Lean meat yield as a percent of HSCW estimated by Computed Tomography scanning (CT scan) (%)

3.2.2 Sheep

For sheep, the algorithms (equations) provided by the ISC to the project team (Equation list 2) and used in LDL included HSCW and GR fat score (1-5 scale). They are based on an analysis of the Sheep CRC Information Nucleus Flock (INF) carcase data. G.Gardner (Murdoch University; pers comm) provided two sets of equations (Equation list 2). The first equation (for lambs from Terminal sires) is currently the LMY% algorithm used in LDL.

Equation list 2 Lean meat yield prediction equations by breed of lamb

CTLean% (Terminal) = 66.4616 + (-0. 113985 x HSCW) + (-1.92786 x GRFatScore)

CTLean% (Merino) = 66.4616 + (-0.113985 x HSCW) + (-1.92786 x GRFatScore) + (-1.51103)

CTLean% (Maternal) = 66.4616 + (-0. 113985 x HSCW) + (-1.92786 x GRFatScore) + (-1.79238)

CTLean% (MerinoXB Merino) = 65.4128 + (-0. 079609 x HSCW) + (-1.99410 x GRFatScore) + (-1.03115)

CTLean% (MerinoXB XB) = 65.4128 + (-0. 079609 x HSCW) + (-1.99410 x GRFatScore)

Table 4 Variables used in the prediction equations for sheep

Variable	Description
HSCW	Hot standard carcase weight (kg)
GRFatScore	Fat score estimated from fat depth at the GR site. When the depth is measured using a GR knife, the depth in mm is converted to scores thus: 1-5mm is score 1; 6-10mm is score 2, 11-15 is 3, 16-20 is 4 and 21-25mm is score 5.
CTLean%	Lean meat yield as a percent of HSCW estimated by Computed Tomography scanning (CT scan) (%)

4 Results and Discussion

4.1 Cattle

4.1.1 Lean meat yield estimates

The analysis of lean meat yield percent (“leanmeatyieldperc”, LMY%) in this report is derived from the data in table “factcarcase” as supplied in the LDL database. There was a total of 2,198,377 carcase records for cattle in the database. Of these carcases, only 281,320 or 13% had an estimated LMY% mostly because Rib-fat (cold) was not recorded. It appears that only animals that had been presented for MSA grading had this measurement recorded.

***Recommendation 1** That LDL considers the use of a LMY% predictor based on carcase weight and P8 fat given that 87% of the available data did not have a LMY% prediction.*

The equation for LMY% (leanmeatyieldperc) being used by LDL should be

LMY% = 63.52120 + 0.00636 * HSCW - 0.80109 * Rib-fat (1)

The model in equation 1 was fitted to the LDL dataset resulting in equation (2) with an R-squared of 1.0.

LMY% = 63.521 + 0.0063555 * HSCW – 0.80111 * Rib-Fat (2)

This confirmed that equation 1 was the equation that was implemented in LDL to predict LMY.

4.1.2 Missing HSCW values

Two records have “HSCW” set to [null]. For these records, both “lefthscw” and “righthscw” were set to zero (rather than [null]). Each record has a calculated “leanmeatyieldperc” (see table).

Table 5 Records with a predicted LMY% but no valid carcase weight

HSCW	Left HSCW	Right HSCW	Rib fat thickness cold	LMY%
[null]	0.00	0.00	4	60.31
[null]	0.00	0.00	6	58.71

If a value of zero is entered for HSCW in the prediction equation, it gave the values in the table above. It would appear that the process for implementation of the LMY% algorithm checks to determine if there is a positive value in “HSCW”, if none is found i.e. it is [null], it tries “lefthscw” and/or “righthscw”. These were zero rather than [null], so the function doubled them and used zero for “hscw”.

***Recommendation 2** A check should be made of the function for calculating “leanmeatyieldperc” for cattle to make sure it does not use zero or [null] values for “hscw”, “lefthscw”, or “righthscw” and Rib-fat.*

4.1.3 Predicted LMY%

For the 281,000 cattle records with measurements for HSCW and Rib-fat, the reported “leanmeatyieldperc” is within rounding error of the value calculated using the prediction equation above (equation 1). All LMY% values in the provided table were between -0.010% and +0.005% of the value calculated using the prediction equation (1).

4.1.4 Unexpectedly low predicted LMY% values

The following table presents basic descriptive statistics for lean meat yield and the independent variables used in its prediction.

Table 6 Descriptive statistics for LMY%, HSCW and Rib fat

Variable	Mean	Standard Deviation	Minimum	Maximum	95% confidence interval
HSCW (kg)	329	55.3	111	822	220 – 437
Rib fat (mm)	5.60	3.4	0	60	3.3 – 7.9
LMY% (%)	61.1	2.6	16.7	65.8	56.0 – 66.2

Clearly there are some data measures of traits in the dataset that are well outside the 95% confidence interval e.g. a Rib-fat depth of 60mm when the upper limit of the 95% confidence interval is 7.9mm.

The table below (Table 7) shows the correlation between LMY% and the independent variables used for it’s (LMY%) prediction. This shows that Rib-fat is the overriding predictor and therefore primary driver of LMY% with a correlation coefficient of -0.99. In contrast HSCW only had a -0.17 correlation with LMY% which is significantly less than that for Rib-Fat.

Table 7 Correlation coefficient for LMY% and the independent variables in the LMY% prediction equation

	HSCW	Rib fat (cold)	LMY%
HSCW	1.00	0.30	-0.17
Rib-fat (cold)	0.30	1.00	-0.99
LMY%	-0.17	-0.99	1.00

The following figure shows the distribution of LMY% in the LDL dataset. A small number of values have very low predicted lean meat yields, as low as 17% with 2,800 carcases with LMY% of 49 or less. These values are associated with very high Rib-fat measures, up to 60 mm!

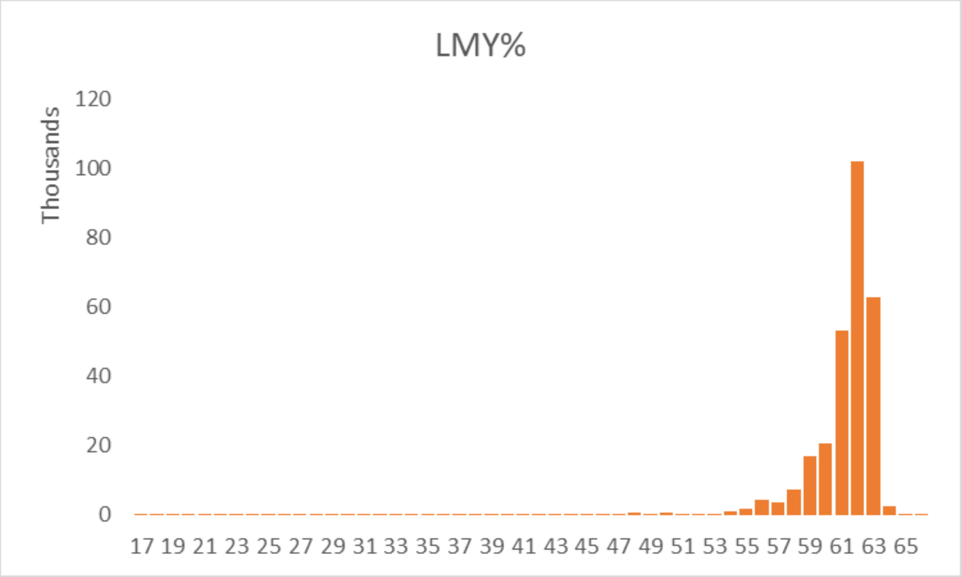


Figure 4 Distribution of predicted LMY% for cattle carcases where LMY% was predicted using equation (1) from HSCW and Rib-Fat (n>280,000).

From the coefficients in the prediction equation, LMY in cattle is much more sensitive to changes in Rib-fat depth than changes in HSCW. A one standard deviation change in HSCW (55 kg) gave a 0.35% change in LMY, while a one standard deviation change in Rib-fat depth (3.4mm) gave a 2.7% change in LMY. In real terms this indicates a 6.75:1 relative change in LMY% for Rib-fat in comparison to HSCW.

Thus, producers aiming to lift LMY% will respond by significantly reducing Rib-fat rather than increasing carcase weight. More importantly processors ranking producers on higher LMY% should be aware that they are preferencing producers that sell animals (cattle) that have less Rib-fat which could be to the detriment of eating quality. For example, in Table 15, the correlations between LMY% and two eating quality parameters, AUSMEAT marbling and the MSA index, are both negative (-0.29 and -0.2 respectively) indicating that a higher LMY% will be associated with lower AUSMEAT marbling and a lower index.

Furthermore, within the dataset provided, Rib-fat depths were reported as high as 60mm, or 16 standard deviations above the mean. This translates to a LMY% that is 43% below the mean (18.1%) which is not biologically possible. Another significant outcome is that given there is a minimum requirement of 3mm of Rib-Fat for cattle to be eligible for MSA grading, it is suspected that a proportion of cattle less than 3mm of Rib-fat were not recorded which may have skewed the dataset.

It is also suspected that the prediction equation (equation 1) is being used to extrapolate outside the range of data used for its development. The prediction equation should be reviewed with a view to setting lower and upper limits on the range of values for the independent variables. Ideally this would be done by considering the dataset used in formulating the prediction equation.

Alternatively (and in the short term), this might be done by limiting the independent variables to their 95% confidence interval with higher and lower values set to their respective limits. For Rib-fat depth, this would mean replacing "- 0.80109 * Rib-fat" with "- 0.80109 * MIN(Rib-fat,7.9)".

Recommendation 3 *Consider how to adjust the prediction equation to exclude, or allow for, independent values outside acceptable limits. This might be based on a review of the data used to derive the equation, or by setting the range at the 95% confidence interval for each variable.*

These acceptable limits may have to be different for different classes of cattle (sex, dentition, grain vs grass etc).

Recommendation 4 *That LMY% calculated from data that is outside 95% confidence limits is shaded or coded differently to reflect that one or both of input variables might be inaccurate.*

4.1.5 Relationship between predicted LMY% and other traits and measures

Carcase weight (HSCW) and Rib-fat depth cold (Rib-fat) are the two independent measures used in the current LMY% prediction equation for cattle. There are other factors that might be considered for their impact on lean meat yield (LMY%) including weight class, sex, age via dentition and feed type (grass versus grain). In this section, descriptive statistics are presented to indicate how HSCW, Rib-fat and predicted LMY% vary with these factors.

4.1.5.1 Weight class

Most carcases processed by the processing establishments providing data to LDL were within the 240-320 and 320-400 kg classes.

Table 8 Count, mean and standard deviation of hot standard carcase weight (HSCW) for carcases with a predicted LMY%

HSCW Class	Count	Mean	Standard deviation
'<240'	14,557	223	12.0
'240-320'	106,836	285	22.5
'320-400'	132,827	356	22.0
'>400'	26,962	425	26.5
'(all)'	281,182	329	55.3

Table 9 Variation in Rib-fat depth by weight class

HSCW Class	Count	Mean	Standard deviation	95% confidence interval	
'<240'	14,557	4.6	2.2	0.3	8.9
'240-320'	106,836	4.9	2.3	0.45	9.3
'320-400'	132,827	5.7	3.1	0*	11.8
'>400'	26,962	8.3	6.1	0*	20.3
'(all)'	281,182	5.6	3.4	3.3	7.9

** Indicates the statistical software estimated a negative Rib-fat due to the skewed distribution within a weight class.*

Unsurprisingly, Rib-fat is higher in heavier carcases suggesting an interaction term (HSCW*Rib-fat) may account for some of the variation in LMY%. Testing this would require an independent measure of LMY% (e.g. DEXA, CT or "bone-out").

Table 10 Variation in predicted LMY% by weight class

HSCW Class	Count	Mean	Standard deviation	95% confidence interval	
'<240'	14557	61.3	1.7	57.8	64.7
'240-320'	106836	61.4	1.8	57.9	65.0
'320-400'	132827	61.2	2.5	56.4	66.0
'>400'	26962	59.5	4.8	50.1	69.0
'(all)'	281,182	61.1	2.6	59.4	62.9

LMY% was relatively constant across the three lower weight class categories as the impact of higher carcase weight on predicted LMY% is partly offset by increasing Rib-fat. In the current LMY% prediction algorithm, a lift in LMY% due to a 100kg increase in HSCW would be offset by an increase in Rib-fat of 0.8mm. For the higher HSCW class (>400), predicted LMY% was lower reflecting the step up in average Rib-fat for this class.

4.1.5.2 Grain versus grass finished

Grain-fed cattle were heavier, had a higher Rib-fat depth and as a result a slightly lower predicted LMY% on average compared to grass-fed cattle. On average the Grain-fed cattle produced 31 kg more lean meat per head than grass-fed (223 kg versus 192 kg). This most likely reflects the intent of commercial feeders to optimise the carcase delivered as against the growing season limitations faced in a grass-fed system. There appears to be no bias in the estimate of LMY% between grain and grass finishing.

Table 11 Variation in HSCW, Rib-fat and predicted LMY% for grain versus grass fed cattle

Measure	Feed type	Count	Mean	Standard deviation	95% confidence interval	
HSCW						
	Grain	79,148	371	45.5	282	460
	Grass	202,034	312	49.6	215	409
Rib-fat						
	Grain	79,148	7.2	4.8	*	16.6
	Grass	202,034	5.0	2.3	0.5	9.5
LMY%						
	Grain	79,148	60.1	3.8	52.7	67.5
	Grass	202,034	61.5	1.8	57.9	65.1

* *Indicates the statistical software estimated a negative Rib-fat due to the skewed distribution.*

4.1.5.3 Sex

Male cattle (steers; castrates) were heavier (+66kg) and had a higher level of Rib-fat depth (+0.8mm) than female cattle, though their predicted LMY% were the same. There does not appear to be any systematic bias in the estimate of LMY% due to sex.

Table 12 Variation in HSCW, Rib-fat and predicted LMY% by sex of cattle processed

Measure	Sex	Count	Mean	Standard deviation	95% confidence interval	
HSCW						
	Female	69,900	279	41.5	197	360
	Male	211,282	345	48.9	249	441
Rib-fat						
	Female	69,900	5.0	2.7	*	10.4
	Male	211,282	5.8	3.5	*	12.7
LMY%						
	Female	69,900	61.2	2.2	57.0	65.5
	Male	211,282	61.1	2.7	55.7	66.4

4.1.5.4 Dentition

HSCW increased as dentition increased from 0 to 6 (307kg to 346kg), was lower (324kg) in cattle with 8 teeth which may be a reflection of an adult female slaughter. Rib-fat was highest in 2-tooth cattle, but then declined with age/teeth. As a result, predicted LMY% was lowest for 2-tooth cattle (60.9%) and highest in the older cattle (61.6%), however the range is less than 1%.

Table 13 Variation in HSCW, Rib-fat and predicted LMY% by dentition of cattle processed

Measure	Dentition	Count	Mean	Standard deviation	95% confidence interval	
HSCW						
	0	73,758	307	56.1	197	417
	2	80,724	336	56.1	226	446
	4	47,596	343	53.3	239	447
	6	29,018	346	43.9	260	432
	7	1,222	346	44.0	259	432
	8	48,864	324	49.8	227	422
Rib-fat						
	0	73,758	5.5	3.0	*	11.3
	2	80,724	6.0	3.8	*	13.4
	4	47,596	5.8	3.8	*	13.4
	6	29,018	5.3	3.0	*	11.3
	7	1,222	5.2	2.5	0.3	10.0
	8	48,864	5.0	2.6	*	10.1
LMY%						
	0	73,758	61.0	2.3	56.5	65.5
	2	80,724	60.9	3.0	55.1	66.7
	4	47,596	61.0	3.0	55.2	66.8
	6	29,018	61.5	2.4	56.8	66.1
	7	1,222	61.6	1.9	57.8	65.4
	8	48,864	61.6	2.1	57.5	65.6

* *Indicates the statistical software estimated a negative Rib-fat due to the skewed distribution.*

4.1.5.5 Hormonal growth promotant (HGP)

Carcases from cattle that had used HGPs were 37kg heavier, had a higher Rib-fat depth (+0.9mm), but had an estimated LMY% that was 0.5% lower than the average of cattle not treated with HGPs. There is significant overlap in the distributions of HGP and non-HGP cattle for yield reflecting the fact that the HGP effect is probably being accounted for in the changes in carcase weight and Rib-fat depth.

Table 14 Variation in HSCW, Rib-fat and predicted LMY% with use of HGP

Measure	HGP use	Count	Mean	Standard deviation	95% confidence interval	
HSCW						
	No	216,989	320	56.1	210	430
	Yes	64,193	357	41.2	277	438
Rib-fat						
	No	216,989	5.4	3.3	*	11.8
	Yes	64,193	6.3	3.5	*	13.2
LMY%						
	No	216,989	61.2	2.5	56.3	66.2
	Yes	64,193	60.7	2.8	55.3	66.2

* *Indicates the statistical software estimated a negative Rib-fat due to the skewed distribution.*

4.1.6 Assessment of P8 fat and Rib fat thickness cold in the LDL database

Table 15 shows the correlation coefficient matrix for all records in the LDL dataset with both Rib fat and P8 fat depth as well as other measures.

Table 15 Correlation coefficients for cattle LDL data (n=281,182)

	HSCW	Rib-fat cold	P8 fat depth	Left HSCW	EMA	AUS marbling	MSA index	LMY%
1 HSCW	1.00	0.30	0.52	1.00	0.50	0.42	0.22	-0.17
2 Rib-fat cold	0.30	1.00	0.47	0.30	0.19	0.34	0.23	-0.99
3 Fat depth	0.52	0.47	1.00	0.52	0.25	0.43	0.08	-0.41
4 Left HSCW	1.00	0.30	0.52	1.00	0.50	0.42	0.23	-0.17
5 EMA	0.50	0.19	0.25	0.50	1.00	0.31	0.14	-0.13
6 AUS marbling	0.42	0.34	0.43	0.42	0.31	1.00	0.52	-0.29
7 MSA index	0.22	2.23	0.08	0.23	0.14	0.52	1.00	-0.20
8 LMY	-0.17	-0.99	-0.41	-0.17	-0.13	-0.29	-0.20	1.00

This table shows that P8 fat depth is not strongly correlated with predicted LMY% (-0.41), Rib-fat (0.47) or HSCW (0.52).

In an attempt to find for LDL a LMY% solution for a greater proportion of cattle in the database (i.e. those cattle that are likely to have been graded for AUSMEAT but not for MSA), equation 1 was modified by replacing Rib-fat with the P8 fat depth measurement. This gave the following equation that had an R^2 of 0.17 with the current LDL LMY% algorithm (equation 1):

LMY% (fat depth) = 62.86 + 0.00268 * HSCW – 0.160 * P8 fat depth (3)

With an R^2 of only 0.17, this does not seem to be a useful alternative.

The P8 fat depth equation previously proposed for use in LDL (as supplied by ISC) was then applied to the available data:

LMY% (fat depth) = 59.9 + 0.0158 * HSCW – 0.546 * P8 fat depth (4)

Predicted LMY% from equation 4 was calculated and compared to the prediction from equation 1, the current LMY% algorithm in LDL. This was done for all records in the database that had both Rib-fat and P8 fat depth (13% of the data; 282,182 animals). The correlation coefficient between these predicted values for LMY% was quite poor at 0.41. The figure below shows their distributions within the dataset.

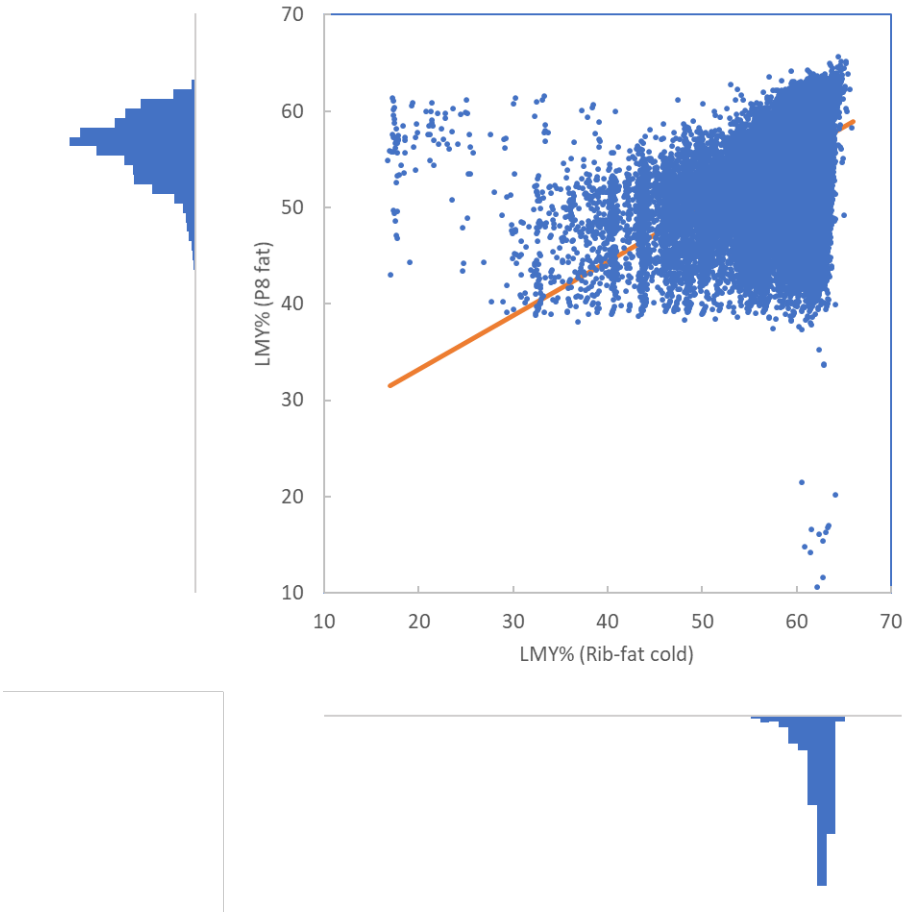


Figure 5 Correlation between LMY% predicted using Rib-fat (horizontal axis) and LMY% predicted using P8 fat depth (vertical axis). (n=281,182)

The least squares linear regression equation for Figure 5 is: -

$$LMY\% (P8) = 22.0 + 0.559 * LMY\% (Rib\text{-}fat) \tag{5}$$

In practical terms, regression equation (5) shows that where the LMY% from a Rib-fat measure is 50%, the LMY% predicted using P8 fat depth will also average 50%. For every one percent increase (or decrease) in the LMY% (Rib-fat), the LMY%(P8) changes in the same direction by a little over a half a percent (0.559%). Importantly not only is the relationship between the two predictions not one to one, but the distribution of the LMY% (P8) from equation 4 was significantly broader than that from the current LDL equation (1). The LMY% from equation 4 is clearly not the same 'trait' as that of equation 1. As an independent measurement of LMY% (either DEXA, CT or bone-out) was not available in the LDL data set, it is difficult to determine which equation has a greater accuracy, however it is assumed that equation 1 has been recommended to the ISC by other R&D providers for accuracy and transportability reasons. This however does not solve the issue of the limited volume of Rib-fat data in LDL and therefore potentially limits the application of LDL LMY% beyond those animals that have been MSA graded.

Including other independent variables in the regression – Dentition, Feed type (grass, grain), saleyard (yes, no), AUS marbling and Eye muscle area – only increased the R^2 to 0.20. {Note that the LMY% dependent variable was that calculated from equation 1, and not an independently measured value.}.

It would be easy to suggest that further analysis of data that includes an independent measure of LMY% is conducted to find an equation with higher accuracy that uses P8 fat. However, given the correlation between Rib-fat and P8 fat is less than 0.5, it is unlikely that any new derivation would improve the correlations.

Recommendation 5 *That beef processors that want to produce LMY% using LDL need to make the investment to measure Rib-Fat or invest in alternative technologies that potentially measure LMY% directly such as DEXA.*

4.1.7 Comparison of LMY% prediction equations for cattle

There are four sets of prediction equations provided for consideration.

The equation currently used by LDL (1), introduced in June 2018, uses carcase weight and Rib-fat thickness.

$$\text{CTLean\%} = 63.52120055 + 0.00635548 * \text{HSCW} - 0.80109465 * \text{Rib-fat} \tag{1}$$

The original prediction equations (June 2016) derived from Teys data (2m and 2f) use carcase weight and rib fat, but also eye muscle area. Separate intercepts were set for male and female cattle.

$$\text{PredLMY\% (Steers)} = 62.1109 - 0.09244 * \text{LeftHSCW} + 0.1645 * \text{EMA} - 0.4936 * \text{Rib-fat} \tag{2m}$$

$$\text{PredLMY\% (Heifers)} = 59.3974 - 0.09244 * \text{LeftHSCW} + 0.1645 * \text{EMA} - 0.4936 * \text{Rib-fat} \tag{2f}$$

Prior to the introduction to LDL of the current equation, equation 3 had been developed (September 2017) based on CT lean measurements.

$$\text{CTLean\%} = 63.4689 - 0.604415 * \text{Rib-fat} - 0.007643 * \text{HSCW} \tag{3}$$

In January 2018, a prediction equation based on P8 fat depth (4) rather than Rib-fat thickness was developed as Rib-fat was not being measured on all carcases.

$$\text{CTLean\%} = 59.90525657 + 0.0157628 * \text{HSCW} - 0.54566852 * \text{P8} \tag{4}$$

The following tables show a correlation matrices for LMY% predicted by the various equations using the carcase data in the LDL dataset, firstly for all data (Table 16), then for male cattle (Table 17) and finally for female cattle (Table 18).

Table 16 Correlation matrix for LMY% prediction, all cattle. Equations 2m and 2f were applied to male and female cattle respectively. (n=281,182)

	LMY% in LDL	Teys eqns	LMY% v2	LMY% P8
1 LMY% in LDL	1.0000	0.6061	0.9509	0.4119
2 Teys eqns	0.6061	1.0000	0.6955	0.4435
3 LMY% v2	0.9509	0.6955	1.0000	0.4708
4 LMY% P8	0.4119	0.4435	0.4708	1.0000

Table 17 Correlation matrix for LMY% prediction, male cattle only. (n=281,182)

	LMY% in LDL	Teys eqns	LMY% v2	LMY% P8
1 LMY% in LDL	1.0000	0.6355	0.9650	0.4378
2m Teys eqns	0.6355	1.0000	0.7478	0.4605
3 LMY% v2	0.9650	0.7478	1.0000	0.5088
4 LMY% P8	0.4378	0.4605	0.5088	1.0000

Table 18 Correlation matrix for LMY% prediction, female cattle only. (n=69,900)

	LMY% in LDL	Teys eqns	LMY% v2	LMY% P8
1 LMY% in LDL	1.0000	0.4971	0.9553	0.3348
2f Teys eqns	0.4971	1.0000	0.6338	0.3820
3 LMY% v2	0.9553	0.6338	1.0000	0.4275
4 LMY% P8	0.3348	0.3820	0.4275	1.0000

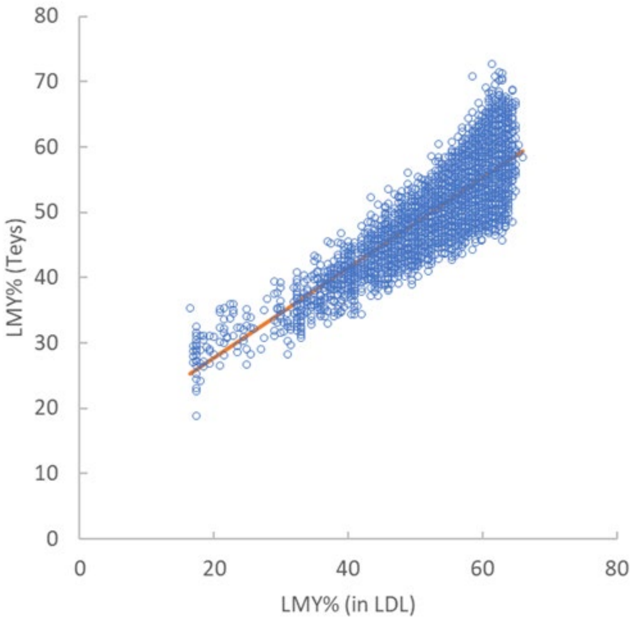


*Figure 6 Comparison of LMY% predicted with the Teys equations (2m and 2f) and LMY% predicted by equation 1 as currently used in LDL. LMY(Teys) = 13.9 + 0.688 * LMY(LDL). (r=0.606, n=281,182)*

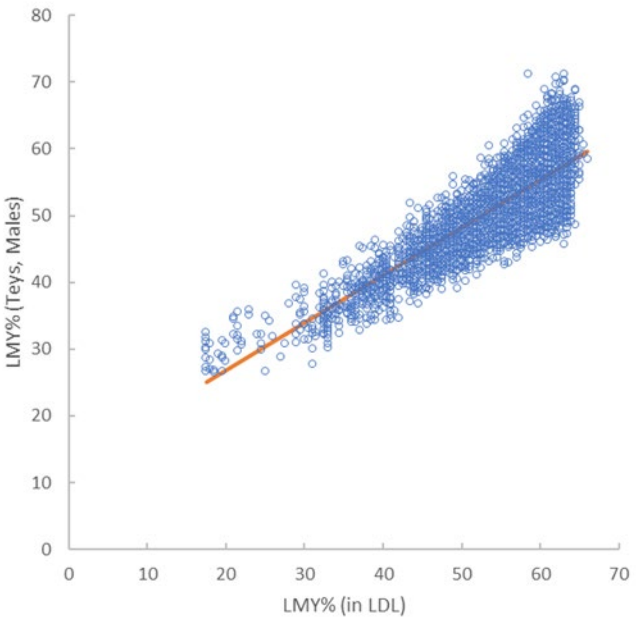


*Figure 7 Comparison of LMY% predicted with the Teys equations (2m) and LMY% predicted by equation 1 as currently used in LDL for male cattle only. LMY(Teys) = 12.5 + 0.714 * LMY(LDL). (r=0.635, n=211,282)*

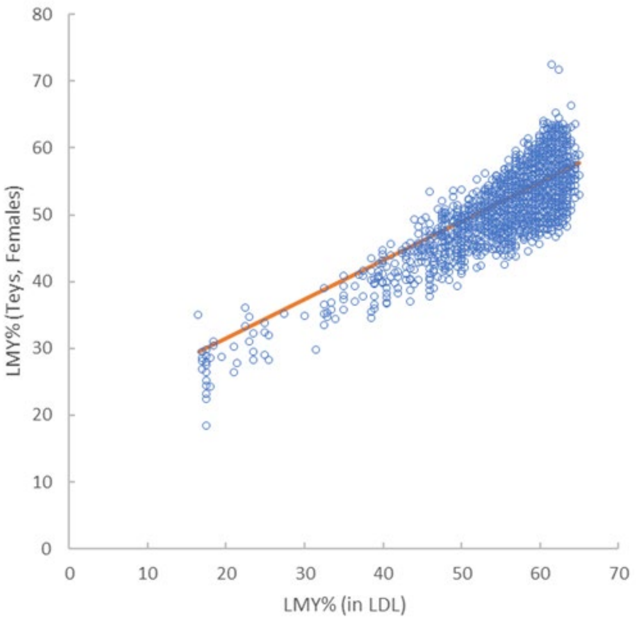


*Figure 8 Comparison of LMY% predicted with the Teys equations (2f) and LMY% predicted by equation 1 as currently used in LDL for female cattle only. LMY(Teys) = 19.9 + 0.582 * LMY(LDL). (r=0.497, n=69,900)*

As shown in the tables 16 – 18 and figures 6 – 8, the correlations between LDL LMY% and the Teys equations have a moderate correlation, with the correlation being higher in the males compared to females. Interestingly the Teys equations provide a greater spread of yield at the higher end of the yield spectrum, which may be useful in separating performance. However, as these equations are dependent on the collection of EMA in addition to Rib-fat they do not solve the issue of scale across the LDL dataset. A further complication is that it appears that the slope of the correlations are quite different to the current LMY% equation being 0.71 and 0.58 respectively for males and females which would have an impact on interpretation of results. There is also a clear intercept effect with the LMY% in Teys equations being between 12.5 and 19.9. This suggests that the Teys equations again are not a good replacement option for LDL.

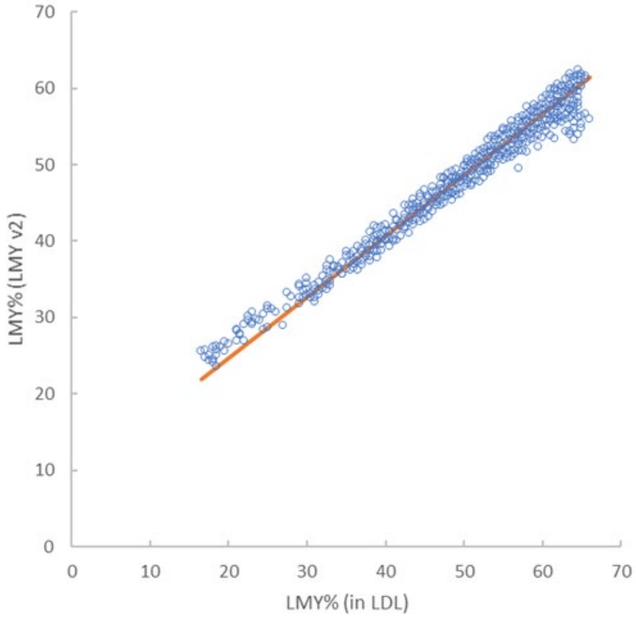


*Figure 9 Comparison of LMY% predicted with equation 3 and LMY% predicted by equation 1 as currently used in LDL. LMY(v2) = 8.72 + 0.799 * LMY(LDL). (r=0.951, n=182,281)*

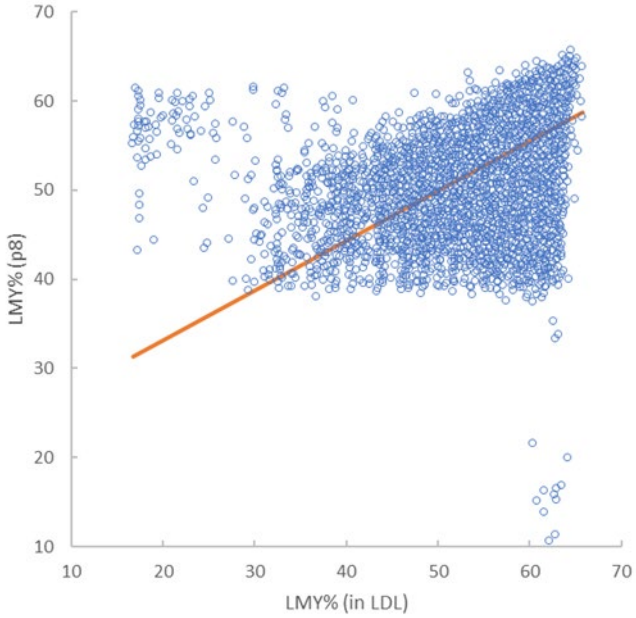


*Figure 10 Comparison of LMY% predicted using P8 fat depth (equation 4) and LMY% predicted from equation 1 as used in LDL. LMY(P8) = 22.0 + 0.559 * LMY(LDL). (r=0.412, n=281,182)*

The correlation between the current equation (LDL LMY) and the previous equation (3; v2) in Figure 9 is very high at 0.95 with a slope of 0.8 and an intercept effect of 8.72. It appears that given the range in the data, there would be little difference in these two equations. There is some increase in variance between these equations at the higher end of the LMY% spectrum.

4.2 Sheep

4.2.1 LMY% analysis – sheep

This analysis relates to the data provided in table “factcarcasesheep”. Of the almost 13 million rows in the table, 8.98 million contained a predicted “leanmeatyieldperc” (LMY%).

The following diagram (Figure 11) shows the reason for LMY% not being calculated as well as the cases where LMY% was calculated despite missing some of the necessary independent data. There were 25 carcases with an LMY% estimate but no fat score and 155 carcases with an LMY% but neither carcase weight nor fat score. All the carcases without a predicted LMY% (over 3.99 million) were missing a fat score including 174,721 that were also missing a carcase weight. Highlighted in the diagram are cases where LMY% is recorded yet fat score (FS) and/or carcase weight (CWT) is missing.

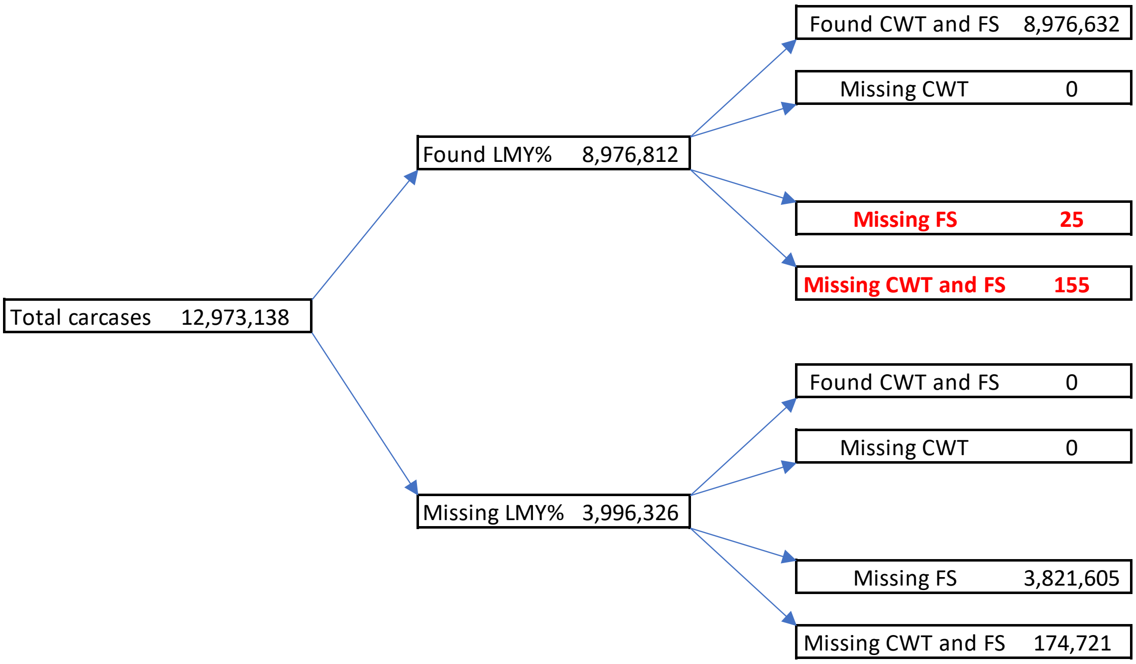


Figure 11 Number of sheep carcases with or without values for LMY%, carcase weight (CWT) and fat score (FS).

Recommendation 6 *Data verification or validation should be conducted to understand how LMY% is being calculated for carcases with no recorded fat score and in some cases no carcase weight.*

The prediction equation for LMY% provided by the ISC as being used by LDL (Equation list 2) included hot standard carcase weight (HSCW, kg) and fat score (related to the measurement of fat depth at the GR site on a 1-5 scale). The following table provides basic statistics on these data.

Table 19 Basic statistics (mean, standard deviation, minimum and maximum) for hot standard carcase weight (HSCW), fat score and predicted lean meat yield per cent (LMY%) for sheep carcases (n = 8.98 million)

Variable	Mean (std dev)	Minimum, Maximum
HSCW	23.75 (4.47)	2 – 81.2
Fat score	3.20 (0.95)	1 – 5
LMY%	57.2 (2.2)	19.0 – 63.2

A minimum HSCW of 2 kg suggests a data problem, particularly as all the 2kg carcases were reported with fat scores of higher than 1. Table 20 shows statistics on HSCW by fat score. The lower end of the 95% confidence interval for each fat score could be used as a cut-off value for HSCW. The same approach should be considered for the upper limit as there are some extreme values such as a carcase weight of 72 kg at a fat score of 1. This should be considered for application to calculations of lean meat yield and aggregate statistics.

Table 20 Basic statistics plus 95% confidence interval for hot standard carcase weight (HSCW)by fat score

Fat score	Count	Mean (std dev)	Minimum, Maximum	95% confidence interval
1	275,532	15.6 (3.4)	3.2, 72	8.9 – 22.3
2	1,729,461	19.8 (2.9)	2.0, 54.1	14.1 – 25.5
3	3,719,549	23.1 (2.7)	2.0, 67.7	17.9 – 28.3
4	2,463,997	26.2 (3.1)	2.0, 72.1	20.2 – 32.2
5	788,093	30.9 (3.5)	6.84, 81.2	24.0 – 37.8

Recommendation 7 *Lean meat yield and aggregate statistics should only be calculated on carcases with a HSCW greater than or equal to defined weight thresholds for each fat score. The lower limit of the 95% confidence interval is proposed as a starting value for consideration. If LMY% is being calculated for values outside the 95% confidence interval (or some other agreed range), then those should be identified or shaded as based on inputs that may be inaccurate.*

4.2.2 Distribution of LMY%

The following figure shows the frequency distribution in the LDL dataset of predicted LMY% for each fat score. It shows that lean meat yield percent is inversely related to fat score, and the spread within each fat score peak is being driven by carcase weight.

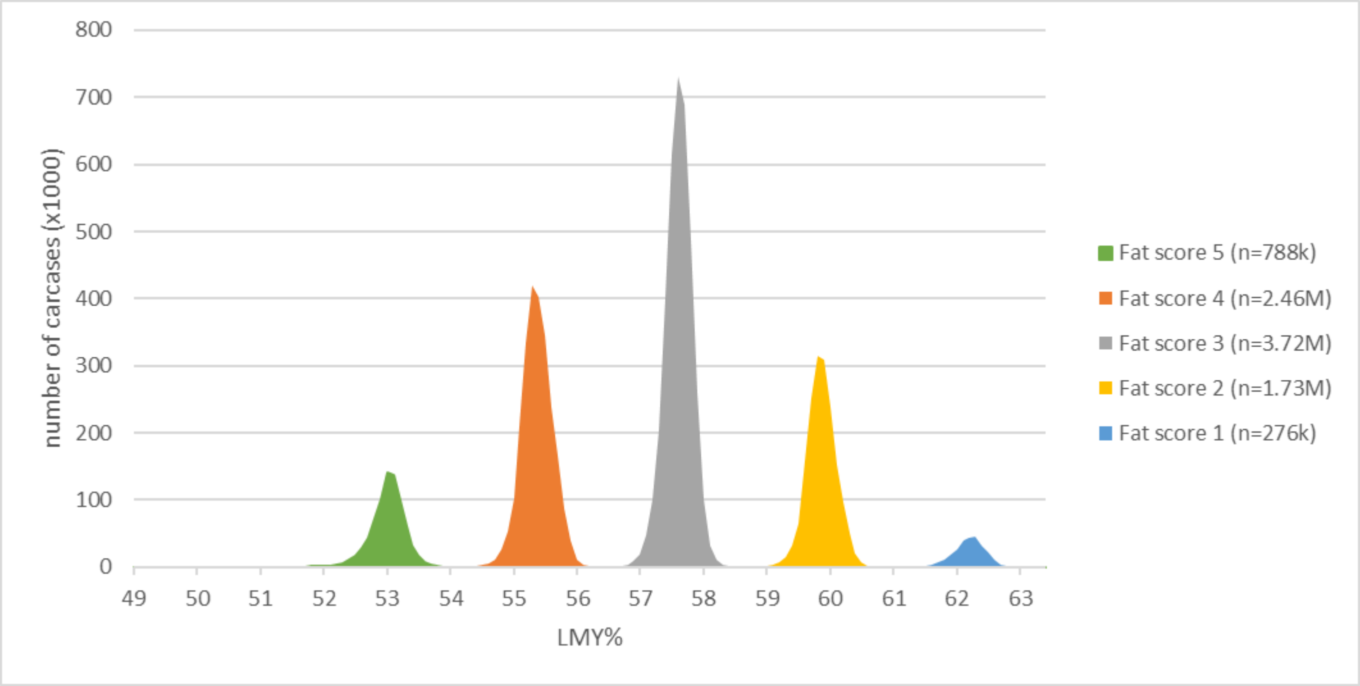


Figure 12 Distribution of carcases by predicted lean meat yield (%) and fat score.

The multimodal distribution reflects the strong relationship between fat score and LMY% in the prediction equation, but also the impact of rounding GR fat depth measures (mm) to fat scores[*]. Rounding fat depth measurements to fat scores removes any overlap in LMY% between fat scores. This makes LMY% overwhelming dependent on fat score in lambs and sheep. As a result, producers submitting sheep with a fat score of 1 (i.e. fat depth of less than five mm) will see the highest LMY%, despite those animals generally receiving a significant discount.

An important consideration for ISC is that the current equation essentially directs producers who seek to increase LMY% toward leaner animals which may not be suitable for domestic or international markets. It also means that producers that rank higher on LMY% are actually simply consigning animals that are lean rather than a combination of optimal carcase weight and leanness. For processors, the implications are more significant in that a processor that is using the current LMY% in LDL for sheep to rank producers will actually be preferentially targeting producers with very lean animals and within each fat score lighter animals.

Although the data is not available within LDL, previous research from the Sheep CRC has confirmed that selection for animals that are leaner at the GR site will effectively reduce eating quality potential through its direct negative correlation with intramuscular fat.

This latter effect can be demonstrated using a synthetic dataset generated to mimic actual data (next section).

---

[*] Fat score is the fat depth (measured in mm at the GR site) divided by five and rounded up to the next whole number. The maximum fat score allowed is five.

4.2.3 Demonstrating impact of converting fat depth to fat score

To demonstrate the adverse impact of rounding fat depth to fat scores, a synthetic dataset of 5000 records was generated using the means, standard deviations (Table 21) and covariance matrix (Table 22) derived from data on Terminal lambs in the INF dataset. The means of the synthetic dataset are slightly higher than those of the INF as a result of excluding samples with negative HGRFAT values.

Table 21 Mean and standard deviation for each measure both in the original dataset and in the synthetic dataset

	INFO Nucleus		**Synthetic dataset**	
	Mean	**Std. Dev**	**Mean**	**Std. Dev**
HCWT	23.83	3.60	23.86	3.53
HGRFAT	14.38	5.34	14.42	5.26

Table 22 Covariance matrix for terminal lambs

	HCWT	**HGRFAT**
HCWT	12.94	13.77
HGRFAT	13.77	28.54

The relationship between HCWT and HGRFAT in the synthetic dataset is shown in Figure 13.

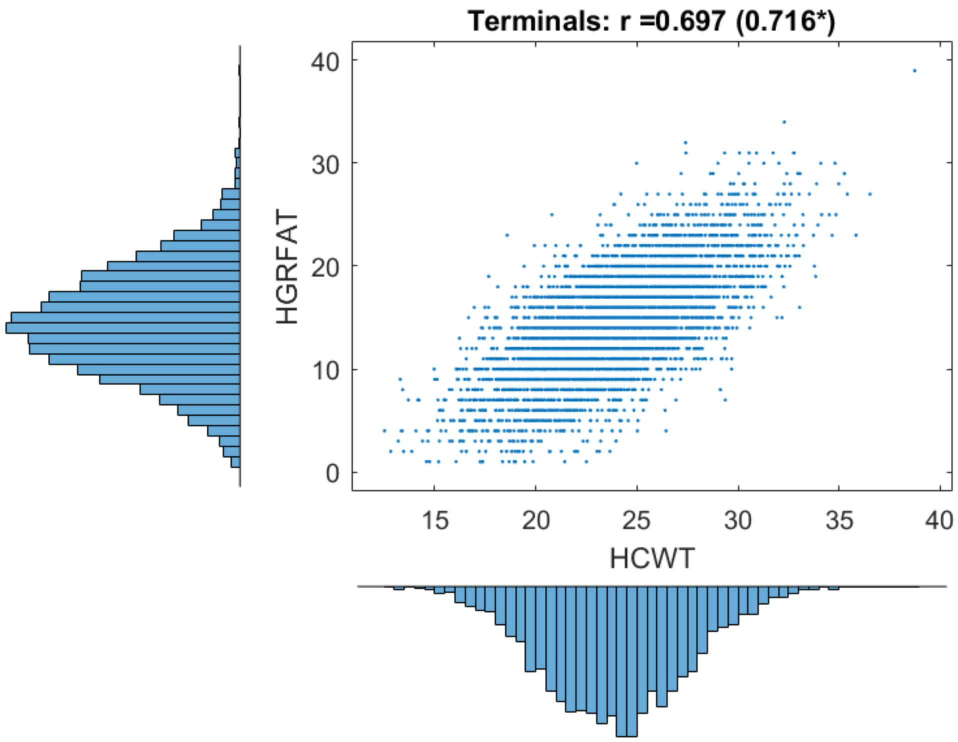


Figure 13 This chart shows the distribution of the HCWT and HGRFAT from the synthetic dataset. The correlation coefficient for this data is 0.697 compared to 0.716 for the original data.

The prediction equation for lean meat yield percent (LMY%) used to compare the use of GR fat score and GR fat depth is given below (G Gardner, pers comm).

CTLean = 66.462 – 0.113985 * HCWT – 1.92786 * GRFATSCORE (for sire type Terminal)

To use this equation, the synthetic HGRFAT data was converted to fat scores in two ways – either to whole number scores as is being currently recorded in LDL and as was used in fitting the prediction equation, or as numbers with a decimal part reflecting fat depth measured in mm.

Table 23 Correspondence between measured HGRFAT (mm) and integer or real Fat Scores

HGRFAT (mm)	Fat Score (integer)	Fat Score (real)*
1, 2, 3, 4, 5	1	0.6, 0.8, 1.0, 1.2, 1.4
6, 7, 8, 9, 10	2	1.6, 1.8, 2.0, 2.2, 2.4
11, 12, 13, 14 ,15	3	2.6, 2.8, 3.0, 3.2, 3.4
16, 17, 18, 19, 20	4	3.6, 3.8, 4.0, 4.2, 4.4
21, 22, 23, 24, 25	5	4.6, 4.8, 5.0, 5.2, 5.4
26 +	5	5.4

* Fat Score (real) was calculated using the formula:
Fat Score (real) = MIN (HGRFAT / 5 + 0.4, 5.4)

The following chart (Figure 14) shows the distribution of LMY% generated using whole number fat scores (1, 2, 3, 4, 5)(vertical axis) versus using fat scores as decimal numbers (horizontal axis). The correlation coefficient between these two predictions is 0.97.

Clearly, the use of scores is having a strong influence on the predicted LMY% with the distribution showing five separate peaks. At the same time, using continuous data (i.e. GR fat depth in mm) gives a normal distribution.

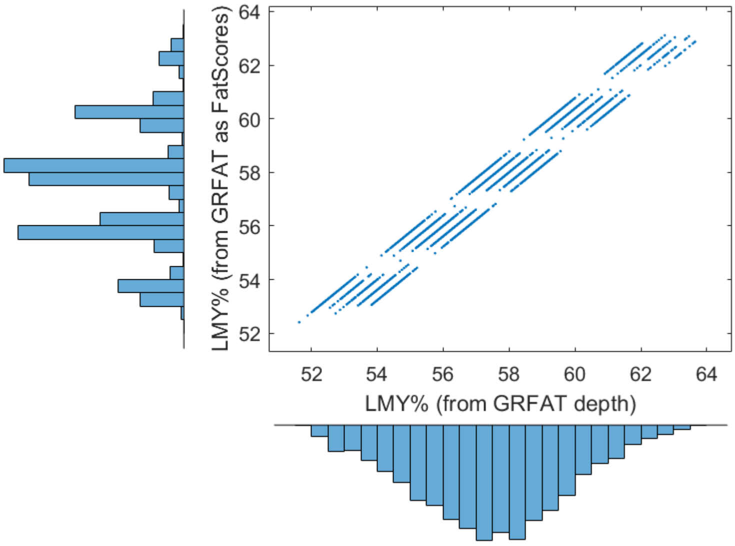


*Figure 14 Comparison of the distributions of LMY% predicted using fat score versus fat depth. LMY% (Fat Score) = 1.21 + 0.98 * LMY% (Fat Depth)* [†]

[†] Fat Depth scaled to the same range (1-5) as Fat Score.

As shown in Figure 14, the slope (regression coefficient) for GR fat score vs GR fat depth is very close to 1, however the actual LMY% is 1.21% higher for the fat score vs the fat depth equation. Another way of looking at this data is to plot predicted LMY% against HCWT with different colours for each fat score. Figure 15 shows this for LMY% predicted from fat depth (left-hand chart) and from fat score (right-hand chart).

The LMY% from fat score chart (right-hand side) suggests there is no overlap in LMY% between fat scores, whereas the LMY% predicted from fat depth (left hand side) shows there can be overlap which may be more useful in separating the actual performance of animals.

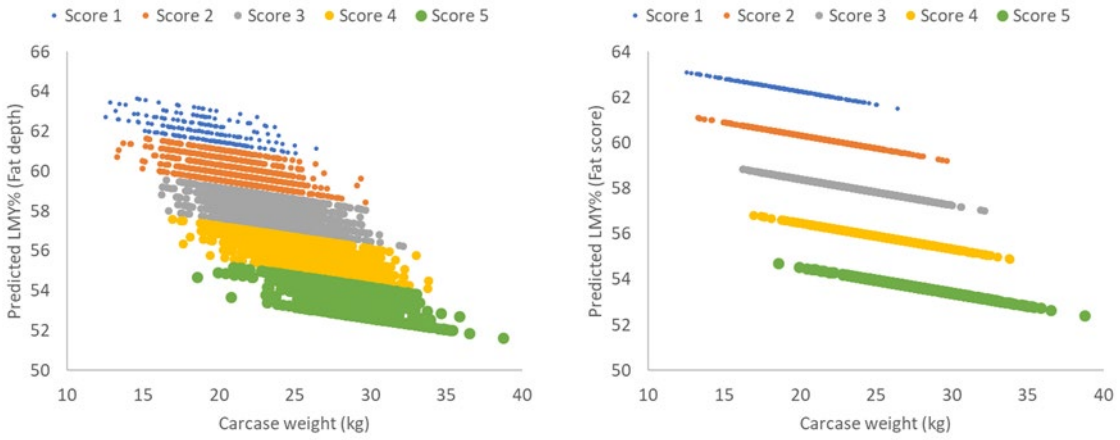


Figure 15 Comparison of the distribution of predicted lean meat yield (%) versus hot standard carcase weight using fat depth (mm) (left-hand chart) and fat score (right-hand chart). The legend shows the colours used for each fat score.

Recommendation 8 *Whenever available, GR fat depth should be recorded in mm as this conveys additional data that is lost when converted to a fat score.*

The following table shows the mean and standard deviation for HSCW, Fat score and predicted LMY% by standard carcase weight ranges and by dentition. The data presented refer to processing data where the same processor reported multiple dentition levels.

Table 24 Mean (and standard deviation) plus number of carcases reported for HSCW, fat score and predicted LMY% by dentition and weight class. Results are only presented for processors reporting more than one level of dentition

Weight class	Dentition	Count	HSCW	Fat score	LMY%
16-18 kg	0	236,006	17.0 (0.58)	2.04 (0.64)	60.0 (1.27)
	2	1,614	16.9 (0.57)	1.52 (0.55)	61.0 (1.09)
	4	14,235	17.0 (0.57)	1.46 (0.52)	61.2 (1.04)
	8	103	17.0 (0.52)	1.85 (0.72)	60.4 (1.34)
18-20 kg	0	353,424	19.0 (0.57)	2.53 (0.74)	58.9 (1.49)
	2	1,573	18.9 (0.58)	1.74 (0.61)	60.4 (1.21)
	4	16,150	19.0 (0.57)	1.65 (0.53)	60.6 (1.06)
	8	139	19.0 (0.55)	1.77 (0.67)	60.4 (1.34)
20-22 kg	0	521,122	21.0 (0.57)	2.99 (0.73)	57.8 (1.47)
	2	1,112	20.9 (0.58)	2.03 (0.64)	59.7 (1.28)
	4	15,014	20.9 (0.58)	1.83 (0.52)	60.1 (1.04)
	8	191	21.0 (0.58)	1.99 (0.72)	59.8 (1.43)
22-24 kg	0	503,410	22.9 (0.57)	3.28 (0.69)	57.1 (1.38)
	2	731	22.9 (0.57)	2.29 (0.70)	59.0 (1.40)
	4	12,237	22.9 (0.58)	2.00 (0.56)	59.6 (1.11)
	8	260	23.0 (0.58)	2.17 (0.75)	59.3 (1.50)

As would be expected, LMY% declined with increasing HSCW. Higher yields were observed in lighter carcases independent of dentition. An interesting finding was there was a trend for the mean fat score to be higher in zero tooth (lambs) animals compared to other dentition classes and as a result LMY% was lower for lambs when compared to hoggets and adult sheep. This may be a reflection that the lamb category had finished lambs whereas hoggets and adult sheep are more likely to be in store condition when slaughtered.

4.2.4 Comparison of LMY% prediction equations for lambs

There were three groups of equations for predicting percent lean meat yield (LMY%) of lambs available for evaluation.

1. Predictions from equations distributed by Graham Gardner (in email received from Demelsa Lollback). This model was fitted with sire breed as a dummy variable. In these equations, fat score rather than fat depth was used. The source data was the Sheep CRC Information Nucleus Flock (INF). LDL currently uses equation C to predict LMY% for sheep carcases.
   A) LMY%_Merino = 66.462 – 0.1140*HCWT – 1.9279* GR_fat_score – 1.5110 ($R^2 = 0.577$)
   B) LMY%_Maternal = 66.462 – 0.1140*HCWT – 1.9279* GR_fat_score – 1.7924 ($R^2 = 0.577$)
   C) LMY%_Terminal = 66.462 – 0.1140*HCWT – 1.9279* GR_fat_score ($R^2 = 0.577$)

   A separate regression model was fitted with sire groups “MerinoXB Merino” and “MerinoXB XB”.

D) LMY%_MerinoXB Merino = 65.413 – 0.0796*HCWT – 1.9941* GR_fat_score – 1.03115 ($R^2 = 0.540$)

E) LMY%_MerinoXB XB = 65.413 – 0.0796*HCWT – 1.9941* GR_fat_score ($R^2 = 0.540$)

2. Equations developed and fitted by Dr Paul Blazkiewicz as part of this project of the form LMY% = a + b * HSCW + c * GR fat depth.

   Separate equations were fitted for Merinos (n=309), Maternals (n=390) and Terminals (n=1033). The source data was the Sheep CRC Information Nucleus Flock (INF). The statistical approach used for these equations differs from that taken by Gardner in that the sire groups were fitted separately.

F) LMY%_Merino = 58.0317 + 0.2029*HCWT – 0.5066* GR_fat_depth ($R^2 = 0.377$)

G) LMY%_Maternal = 61.24137 – 0.03507*HCWT – 0.39118* GR_fat_depth ($R^2 = 0.443$)

H) LMY%_Terminal = 63.40199 – 0.04903*HCWT – 0.39836* GR_fat_depth ($R^2 = 0.491$)

3. Predictions based on equations that predict lean meat yield weight which can then be divided by carcase weight to give LMY%. The equations for Merino and for Maternals include a carcase weight squared term.

I) LMYwt_Merino = –4.527 + 1.0638*HCWT – 0.0106*$HCWT^2$ – 0.1025*GR fat depth

J) LMYwt_Maternal = –2.289 + 0.866*HCWT – 0.00697*$HCWT^2$ – 0.09107*GR_fat_depth

K) LMYwt_Terminal = 1.872932 + 0.540481* HCWT – 0.09604* GR_fat_depth

L) Then, LMY% = LMYwt / HCWT

For the Merino, Maternal and Terminal equations, LMY% was predicted for a matrix of all combinations of HCWT and GR fat depth. HCWT was varied over the range 10 to 30 kg, and GR fat depth from 1 to 25 mm.

The following groups of charts arranged by breed illustrate how predicted lean meat yield percent (LMY%) varies by HCWT and GR fat depth for each of the prediction equations listed above.

It is acknowledged that the response surface would not be equally represented in the population of carcases processed. Both the top left corner (high GR fat depth, low HCWT) and the lower right corner (low GR fat depth, high HCWT) of each chart represent less populated combinations. See appendix 1 and the end of this section for the boundary capturing 99% of lamb carcases.

4.2.4.1 LMY% for Merinos

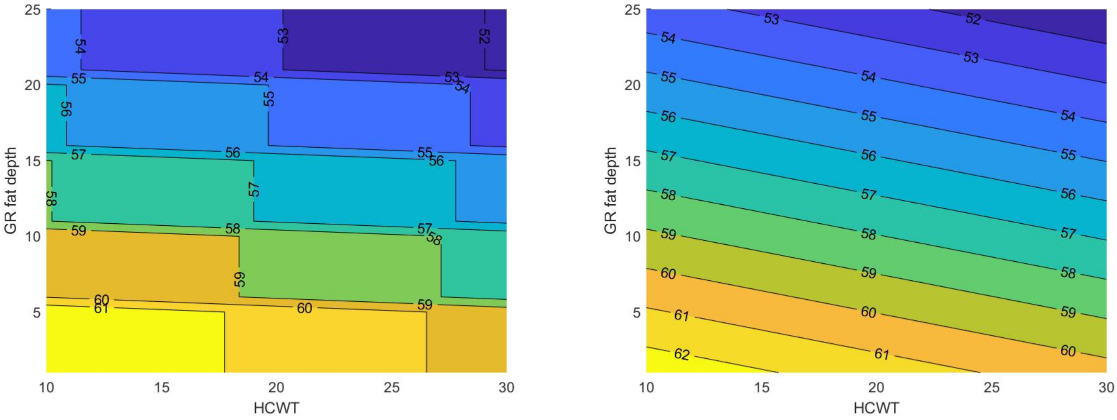


Figure 16 LMY% contours for Merino lambs over a GR fat depth by carcase weight matrix. Calculated using equation (A). Left-hand chart uses GR fat depth converted to GR fat scores. Right-hand chart uses GR fat depth.

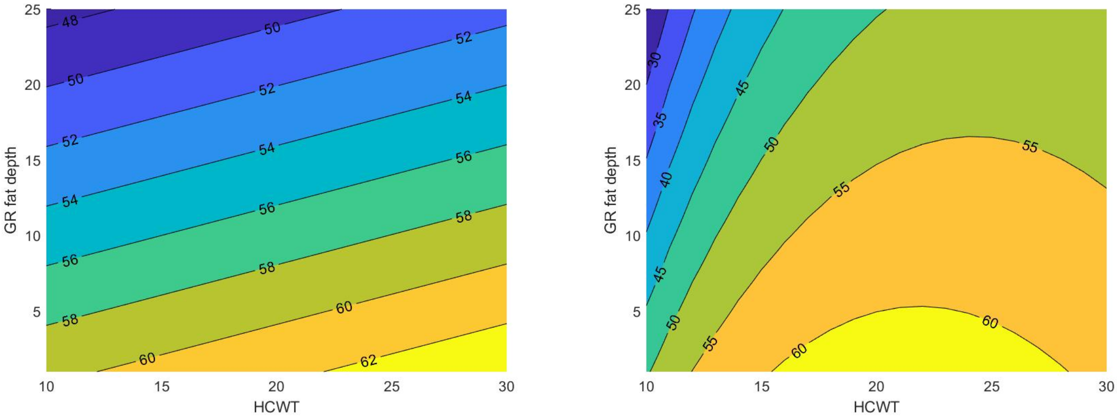


Figure 17 LMY% contours for Merino lambs over a GR fat depth by carcase weight matrix. Calculated using equation (F) (left-hand chart) and equations (I) and (L) (RHS). Both charts use GR fat depth (right-hand chart).

4.2.4.2 LMY% for Maternals

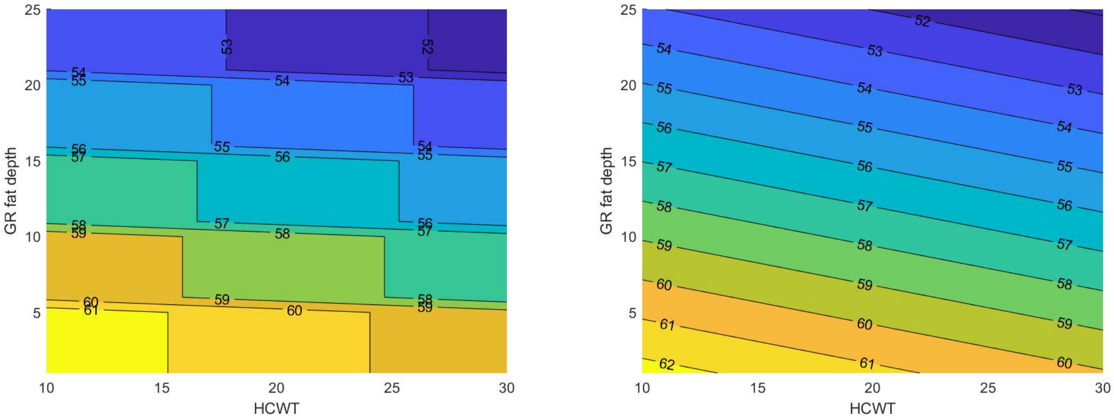


Figure 18 LMY% contours for Maternal lambs over a GR fat depth by carcase weight matrix. Calculated using equation (B). Left-hand chart uses GR fat depth converted to GR fat scores. Right-hand chart uses GR fat depth.

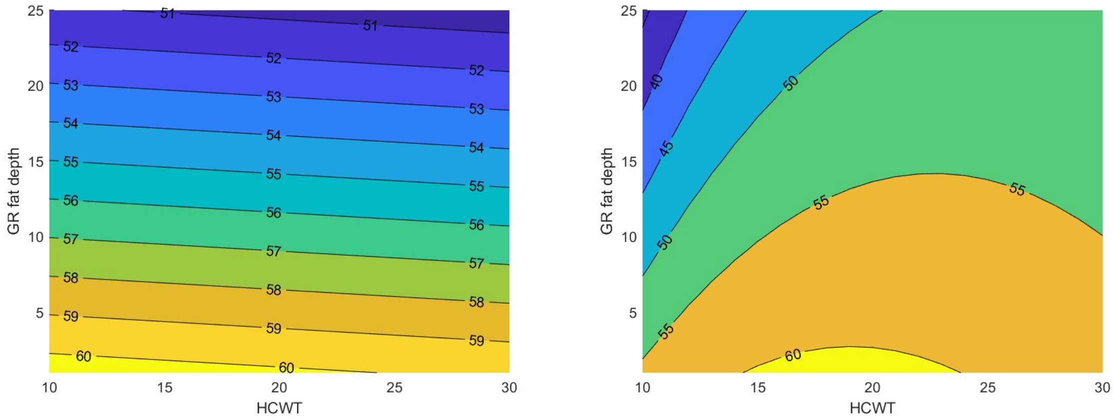


Figure 19 LMY% contours for Maternal lambs over a GR fat depth by carcase weight matrix. Calculated using equation (G) (left-hand chart) and equations (J) and (L) (RHS). Both charts use GR fat depth (right-hand chart).

4.2.4.3 LMY% and Terminals

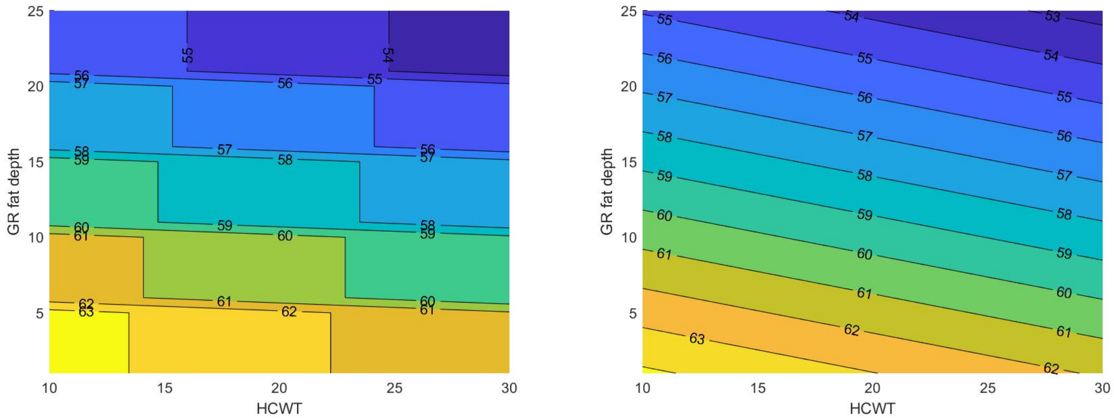


Figure 20 LMY% contours for Terminal lambs over a GR fat depth by carcase weight matrix. Calculated using equation (C). Left-hand chart uses GR fat depth converted to GR fat scores. Right-hand chart uses GR fat depth.

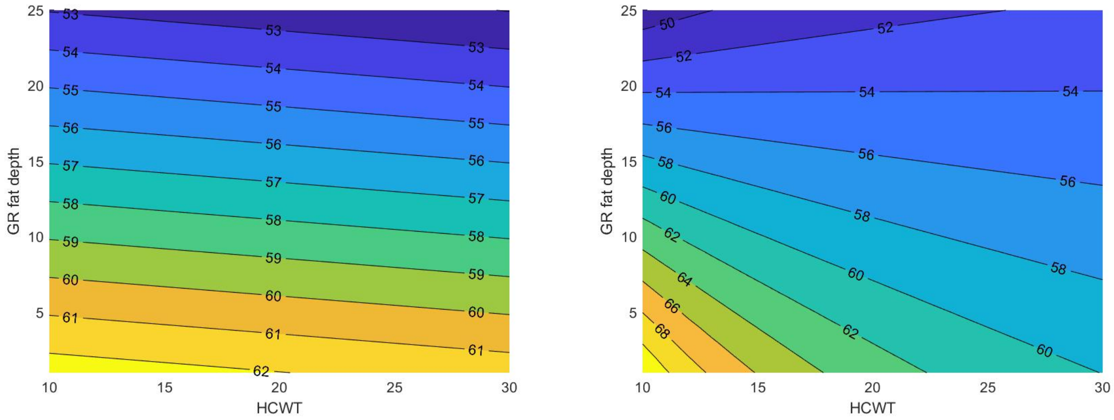


Figure 21 LMY% contours for Terminal lambs over a GR fat depth by carcase weight matrix. Calculated using equation (H) (left-hand chart) and equations (K) and (L) (RHS). Both charts use GR fat depth (right-hand chart).

4.2.4.4 Observations on the use of breed in an LMY% LDL equation

- Clearly using GR fat score (left-hand chart in Figure 16, Figure 18 and Figure 20) rather than GR fat depth (right-hand chart in the same figures) results in a discontinuous surface. This complicates understanding of the interaction.
- Unsurprisingly, the right-hand charts in Figure 16, Figure 18 and Figure 20 are almost the same. The equations behind these charts (A, B, C) are identical apart from a breed constant that adjusts the intercept.
- The right-hand charts in Figure 16, Figure 18 and Figure 20 should be similar to the left-hand charts in Figure 17, Figure 19, and Figure 21 respectively. Their regression models use the same independent variables. The former is based on an analysis with data pooled across breeds, while the prediction equations for the latter charts were fitted for each breed separately.
    - For Maternals and Terminals, the comparison holds well – each chart pair appearing similar.
    - This is not the case for Merinos. Using equation 'A', LMY% increases as either GR fat depth or HCWT decrease (Figure 1, right-hand chart). But using equation F (Figure 2, left-hand chart), for a constant GR fat depth, LMY% increases as HCWT increases; while for a constant HCWT, LMY% increases as GR fat depth decreases.
- The right-hand charts in Figure 17, Figure 19, and Figure 21 were calculated using a function that first predicted lean meat weight before this was converted to LMY%. For Merinos and Maternals, the regression fit was improved by adding a HCWT squared term resulting in the curved contours.
    - For both Merinos and Maternals, these charts show that for a given GR fat depth, there is a HCWT that maximises LMY%.

      The result is slightly different for Terminals. At low GR fat depth levels, LMY% increases as HCWT decreases. But at higher GR fat depth levels, LMY% is largely unchanged by HCWT, though the likelihood of having low HCWT at high GR fat depths is low. These results show that if breed was fitted to available LDL equations then there would be a significant shift in the prediction for Merinos relative to the other breeds. Currently, there is no breed data captured in LDL. As a result, only a generic prediction equation can be used, although it does appear that fitting the equation without breed adjustments is biasing results. This has implications for producers that have more than one breed of sheep and particularly for processors that are trying to rank lamb producers for yield.

Recommendation 9 *LDL should seek to capture breed data, even if only in broad categories (i.e. Merino, Maternal, Terminal). This would allow for more tailored prediction of LMY% that could remove some of the inherit breed biases that are within the data.*

Recommendation 10 *ISC should investigate whether the addition of a non-linear term for either carcase weight or fat-score would improve the prediction accuracy of LMY%*

4.2.4.5 Extent of distribution of carcases on a fat score by carcase weight grid.

The following figure shows on a fat score by carcase weight grid the boundary that captures 99% of the lamb carcases (i.e. dentition of "0") in the supplied LDL data.

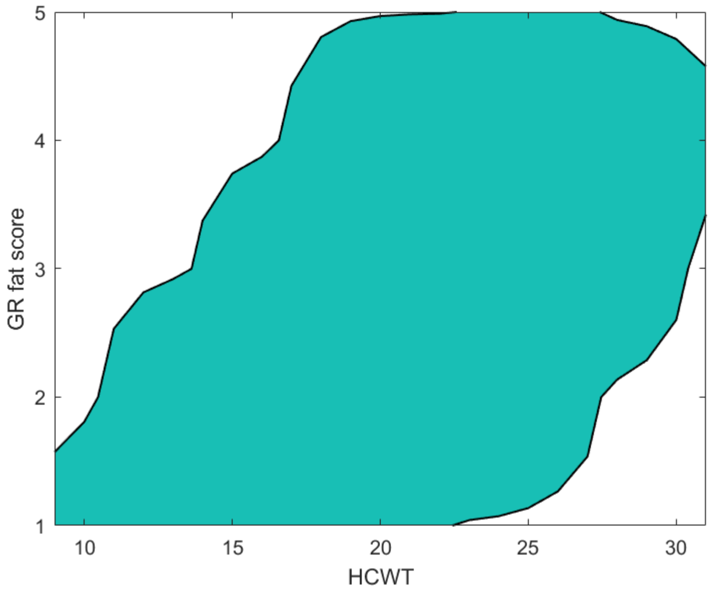


Figure 22 Boundary for distribution of 99% of carcases with '0' recorded for dentition. Note this analysis used fat score rather the preferred GR fat depth.

It may be possible to use the contour map in figure 22 as a guide for the acceptance of data ranges in the LDL dataset. This contour map could also be used to plot animals that fall outside of the accepted range for producers to understand why LMY% predictions may be inaccurate.

4.3 Implications for implementation of alternative methodologies for measuring LMY%

In the future, there will be alternative methods for estimating LMY% in sheep and cattle using technologies such as DEXA, microwave and video image analysis. This presents both a challenge and an opportunity for the LDL management team. Although those technologies will likely be calibrated against a gold standard such as CT scanning, it is highly likely that the correlations between estimates from those technologies and the current LDL prediction algorithms will be variable and, in some cases, significantly different. For cattle as shown in section 4.1, the alternative prediction algorithms that use Rib-Fat and P8 fat respectively have a correlation that suggests that the LMY% predicted will show significant re-ranking and may in fact be a different trait as correlations with other traits are different. It was suggested that if LDL implemented the P8 algorithm that it should be notated with a different script to indicate that it is a different trait.

The opportunity for LDL management is that as processors implement these technologies, they will capture LMY estimates for a far wider range of carcase and animal attributes than can be collected in research studies. This data might then be used to improve the prediction equations used for untested carcases, notwithstanding the complication of how close the alternative measurement is to the CT scanned gold standard.

Recommendation 11: As alternate methods of estimating LMY are implemented by processors, LDL should upload and store both the measurements and the technology used. The data should be used to verify and refine LMY predictions, examine equivalence between methodologies, and potentially improve predictions for carcases that have only have carcase grading measurements available.

In Appendix A, the analysis of the prediction algorithm for LMY% in sheep shows that the current LDL algorithm has a moderate correlation with the gold standard for LMY% (CT scanning) of 0.5 and a residual standard deviation of 2.7% , this compares with unpublished estimates from DEXA of greater than 0.8 and less than 2% respectively. It is highly likely that the correlation between DEXA LMY% and the LMY% within LDL will result in some reranking. LDL management will need to complete a formal analysis of this data when it is available to understand what the impact of those reranking are likely to be in terms of ranking animals and potentially producers for LMY%. In sheep as DEXA LMY% is unlikely to be as dependent on fat score as is the current LDL predicted LMY%, use of DEXA LMY% may reduce the current issues with LMY% being overwhelmingly dependent on fat measurements. This should allow for a greater discrimination on the value of LMY% across sheep value chains.

Overall, the introduction of alternative methods for estimating LMY% will be both challenging and complex, even if a “gold standard” validation procedure such as CT scanning is in place. Regardless moving from LMY% that is dependent on one linear measurement of fat (in both cattle and sheep) to more robust methods that measure whole carcase characteristics is likely to result in key differences. While the estimates recorded by the processors may not have a perfect correlation with the gold standard CT scans, or even be measuring the same trait, it must be recognised and accepted that each processor in implementing a technology will use that estimate in their decision making.

The key issue will be for producers that are comparing LMY% across the different methods and attempting to reconcile why there are differences. LDL management will have to conduct formal statistical analysis to determine the correlations and reranking’s between those alternative methods and to see whether there is a possibility for harmonisation through application of fixed correction factors or if the various LMY% predictions need to be notated to reflect that they are different traits. Perhaps an intermediate position is to group or classify the different LMY% methods into groups which have high correlations and then report those as the one LMY% prediction. To ensure confidence is maintained in LDL, it is probably more pragmatic to report them as different traits as is done for marbling (AUSMEAT vs MSA) so that producers don’t become distracted with method differences and stay focussed on using the LMY% in management and selection decisions.

5 Conclusions/recommendations

This report provides an assessment of the objectives defined in the RFQ that broadly investigated the implementation of LMY% algorithms for both cattle and sheep in LDL. The project received data from the LDL database from 4 beef and 2 sheep processors over extended period of time. The report has summarised a number of potential LMY% equations, as supplied by ISC and derived by the project team specifically for this project, that exist for the prediction of lean meat yield percent based on carcase data that is collected either on the slaughter floor or during carcase grading. The report found that there are some issues with the data in that there are examples of extreme data for those measurements that are used in the prediction of LMY%. Furthermore, for several of the key production and market segments, data that was expected to be included was not available in the LDL database. This included breed, sex, dentition and feeding type (cattle) for some and often most of the records.

For cattle a significant issue for the implementation of LMY% in LDL occurs as a result of the absence of Rib-fat measurements for many carcases in the LDL database. It is assumed that this absence is because only cattle that are eligible for MSA grading would be measured for Rib-fat as it is not a current AUSMEAT requirement.

In both sheep and cattle, the report found that LMY% is strongly dependent on the fat measurement that is used. For beef, Rib-fat has an overwhelming influence on LMY% and this is consistent across all production and market segments. For sheep, LMY% is very strongly dependent on fat score, which totally discriminates LMY% independent of HSCW. Again, this is consistent across dentition and carcase weight ranges. These dependencies are being driven by the prediction equations where the coefficient on the fat term (Rib-fat or GR fat score) is having a much greater impact than the coefficient on the carcase weight term, despite carcase weight and the fat measurement being highly correlated.

The implications of these findings are significant for producers and processors using LDL. For producers, using LMY% as a management tool will direct them towards producing animals that are leaner independent of all other factors and regardless of species. This is potentially of concern for the Australian red meat industry where maintaining or improving eating quality is a key objective. The findings are significant for the processors in that if they are using or intent on using LMY% as a criterion in the ranking of producers for carcase performance and compliance, then they need to be aware that they are preferencing producers that supply lean animals that might not be suitable for both market and consumer requirements.

When the LMY% based on the current LDL equation was compared to other documented equations in cattle in order to find possible alternatives, it was clear that all other options predicted an LMY% that was significantly different to that which is currently used. The most important finding was the comparison between LMY% from Rib-fat vs that derived from P8 fat which had a moderate correlation, a different intercept and a regression coefficient of 0.56. This showed that the implementation of an equation based on P8 fat which could improve the breadth of LMY% prediction in the industry, would be problematic. The implication for the Australian beef industry and LDL is that it is unlikely that single point measures of carcase traits are going to provide robust estimates LMY%.

For sheep, two important comparisons of LMY% equation variations are contained within this report. Firstly, using a derived synthetic dataset based on that statistical proprieties of a research database (Sheep CRC) for HSCW and fat depth, it was shown that the conversion of fat depth measured in mm to a fat score resulted in a loss of discriminatory information for LMY% and a small change in the actual LMY% calculated. This has implications for the sheep industry that generally relies on manual palpation to estimate fat score. Moving towards a direct measurement of GR fat depth would improve the LMY% prediction.

This report has identified several areas in which improvements could be made or additional activities could be included in the LDL program and specifically for the prediction of LMY% from algorithms based on carcase traits in cattle and sheep both now and into the future. These recommendations are based on the project team's industry experiences, analysis and review of the data provided from LDL and the subsequent comparisons of LMY% predictions from different equations and across a range of different production and market segments. Whilst there is no priority ranking given to the recommendations, they should be considered by LDL and ISC based on merit and difficulty of implementation.

Eleven recommendations have been made and these are tabulated below.

Table 25 Table of recommendations

No.	Recommendation	Page
1	That LDL considers the use of a LMY% predictor based on carcase weight and P8 fat given that 87% of the available data did not have a LMY% prediction.	13
2	Review function used to calculate LMY% to make sure it does not use zero (or [null] values for HSCW, LeftHSCW or RightHSCW or Rib-fat.	14
3	Consider how to adjust the prediction equation to exclude, or allow for, independent values outside acceptable limits. This might be based on a review of the data used to derive the equation, or by setting the range at the 95% confidence interval for each variable.	16
4	That LMY% calculated from data that is outside 95% confidence limits is shaded or coded differently to reflect that one or both of input variables might be inaccurate.	16
5	That beef processors that want to produce LMY% using LDL need to make the investment to measure Rib-Fat or invest in alternative technologies that potentially measure LMY% directly such as DEXA.	22
6	Data verification or validation should be conducted to understand how LMY% is being calculated for carcases with no recorded fat score and in some cases no carcase weight.	26
7	Lean meat yield and aggregate statistics should only be calculated on carcases with a HSCW greater than or equal to defined weight thresholds for each fat score. The lower limit of the 95% confidence interval is proposed as a starting value for consideration. If LMY% is being calculated for values outside the 95% confidence interval (or some other agreed range), then those should be identified or shaded as based on inputs that may be inaccurate.	27
8	Whenever available, GR fat depth should be recorded in mm as this conveys additional data that is lost when converted to a fat score.	33
9	LDL should seek to capture breed data, even if only in broad categories (i.e. Merino, Maternal, Terminal). This would allow for more tailored prediction of LMY% that could remove some of the inherit breed biases that are within the data.	36
10	ISC should investigate whether the addition of a non-linear term for either carcase weight or fat-score would improve the prediction accuracy of LMY%	36
11	As alternate methods of estimating LMY are implemented by processors, LDL should upload and store both the measurements and the technology used. The data should be used to verify and refine LMY predictions, examine equivalence between methodologies, and potentially improve predictions for carcases that have only have carcase grading measurements available.	37

6 Appendix 1. Technical Note: Different statistical approaches to predicting lean meat yield (LMY) and LMY percent (LMY%).

6.1 Introduction

It has been observed that the R^2 of typical lean meat yield percent (LMY%) regressions are quite low. A new method of fitting equations to the data has been tested in this paper as it has been observed that the R^2 for regressions of Computation Tomography estimated lean meat yield (ctleankg) against the typical fitting parameters are much higher. This paper tests the new methodology against the old methodology for sheep using three categories of sire type, Merino, maternal and terminal using data collected by the Sheep CRC. At the end of this paper additional examination is performed to see if segmenting the hot GR fat depth (hgrfat) values into fat scores improves the prediction results.

6.2 Merino Results

6.2.1 Ctleankg

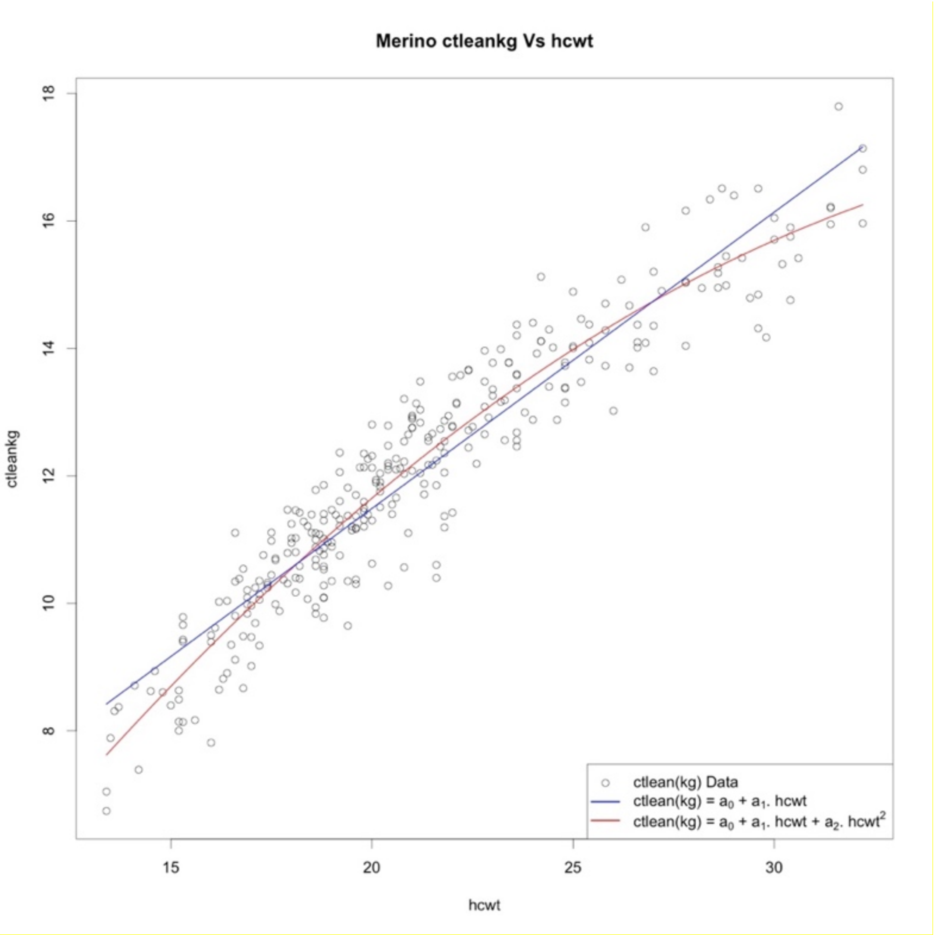


Figure A1: Scatter plot of ctleankg verses hcwt for Merinos, with polynomial fits.

In figure A1, it can be seen through visual inspection that there is some negative curvature in the ctleankg and hot carcase weight (hcwt) scatter plot. It is common to approximate more complicated functions using simple power series expansions as seen in equation 1.

$$ctleankg = a_o + a_1x + a_2x^2 + a_3x^3 + \cdots \qquad \text{(A1)}$$

In figure A1, two truncated power series are fitted to the raw data, the blue curve being a power series truncated at the first order term (giving a linear fit), and the second curve which is a power series truncated at the second order term (giving a quadratic fit).

Coefficients	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	-3.887894	0.854768	-4.548	7.80e-06
$\boldsymbol{hcwt}$	1.025140	0.077368	13.250	< 2e-16
$\boldsymbol{hcwt^2}$	-0.012412	0.001704	-7.286	2.73e-12

Table A1: Results of quadratic fit for ctleankg verses hcwt for Merinos.

The residual standard error of the quadratic fit in table A1 was 0.6387 kg on 306 degrees of freedom. The multiple R^2 was 0.906, the F-statistic was 1474 on 2 and 306 degrees of freedom, and the p-value of the fit was < 2.2e-16. From figure A1 and table A1, we can see that the quadratic fit is statistically significant. Although, this is a standard method of modelling deviations from linearity by adding additional power series terms, this methodology can give strange results if the fitted equation is used outside of the fit range. This is why extrapolation should be used with great care if at all.

For example, the quadratic equation described in Table A1 has ctleankg increasing as hcwt increases reaching a maximum when hcwt is 41.3 kg before declining rapidly. Biologically, it does not make sense for ctleankg to decline as hcwt continues to increase and from figure A1, we can see the range of hcwt values used to fit the quadratic is from about 13 to 32 kg. In the LDL dataset, 1.5% of the lamb carcases (dentition = 0) were heavier than 32 kg including 0.05% that were heavier than 41.3 kg, though it is not known how many of these are Merino.

For this reason, we probably need a different functional form to fit to the convexity of our data while avoiding the curve turning down.

$$ctleankg = a + be^{c.hcwt} \qquad \text{(A2)}$$

Instead of the truncated power series fits, an exponential function was tested as shown in equation A2, here we would expect the values of b and c to be negative.

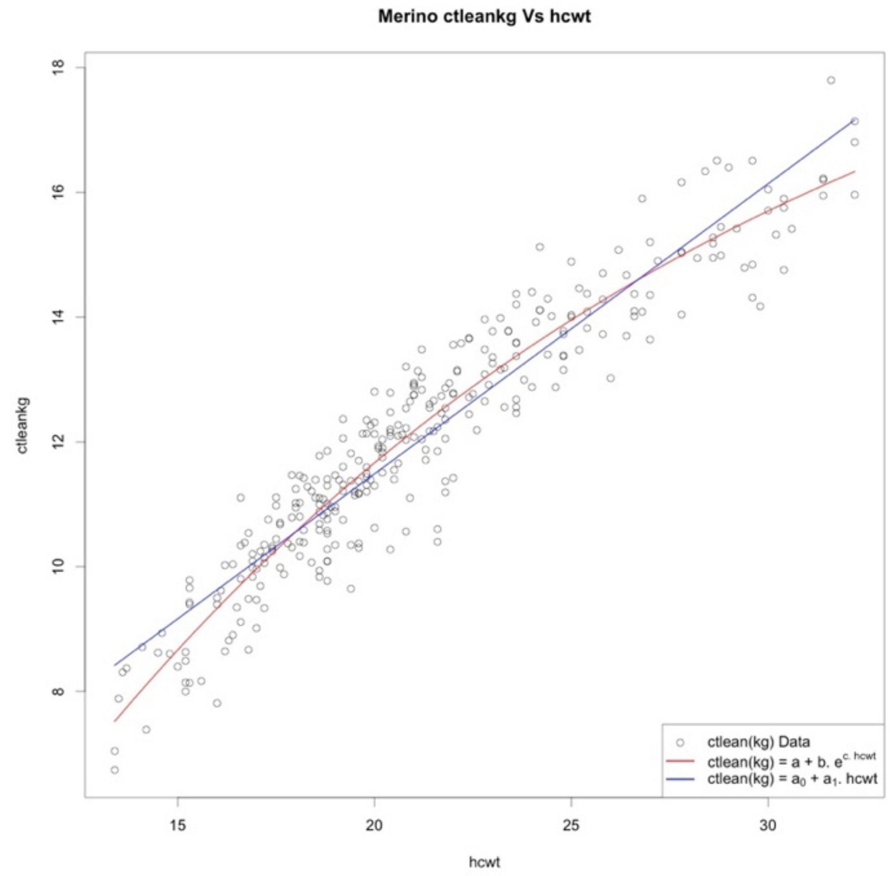


Figure 2: Scatter plot of ctleankg verses hcwt for Merinos, with exponential and linear fits.

Coefficients	**Estimate**	**Std. Error**	**t value**	**Pr(>|t|)**
$\boldsymbol{a}$	21.331197	1.201173	17.759	< 2e-16
$\boldsymbol{b}$	-28.502488	0.807908	-35.279	< 2e-16
$\boldsymbol{c}$	-0.054070	0.007363	-7.343	1.9e-12

Table A2: Results of exponential fit for ctleankg verses hcwt for Merinos.

The residual standard error of the fit was 0.6377 kg on 306 degrees of freedom which is 10 grams smaller than the RSE of the quadratic fit. The multiple R^2 of the exponential fit was 0.9062 which is slightly better than the quadratic fit which was 0.906. So now using the exponential function in equation A2, we can get a statistically significant fit on par with the previous quadratic fit. But now, we do not have the problem of a localised peak occurring, and the predicted ctleankg then rolling over and approaching zero for large carcase weights. In fact, with this exponential model the ctleankg increases toward its asymptotic value of a at infinity, which in this case is 21.3 kg.

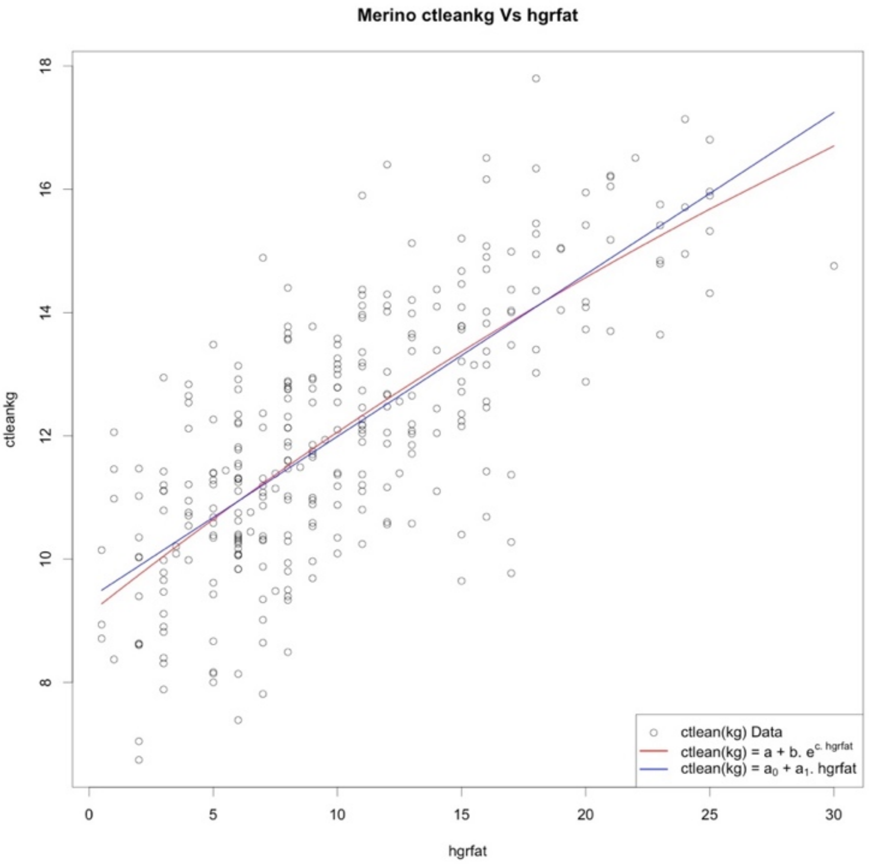


Figure A3: Scatter plot of ctleankg verses hgrfat for Merinos, using exponential and linear fits.

Coefficients	Estimate	Std. Error	t value	Pr(>|t|)
a	28.93736	15.32864	1.888	0.060
b	-19.82226	15.10743	-1.312	0.190
c	-0.01609	0.01503	-1.070	0.285

Table A3: Results of exponential fit for ctleankg verses hgrfat for Merinos.

The residual standard error of the fit was 1.408 mm on 306 degrees of freedom, and the multiple R^2 was 0.5432. In table A3, we can see that none of the fit parameters are significant at the 5% level. Also, from visual inspection there does not appear to be any significant convexity in the raw data. So, it is probably best to keep the model as simple as possible by using a linear fit.

Coefficients	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	9.36245	0.16199	57.80	<2e-16
$\mathbf{hgrfat}$	0.26282	0.01381	19.04	<2e-16

Table A4: Results of linear fit for ctleankg verses hgrfat for Merinos.

The residual standard error of the linear fit was 1.408mm on 307 degrees of freedom. The multiple R^2 was 0.5414, and the F-statistic was 362.4 on 1 and 307 degrees of freedom with a p-value of < 2.2e-16. So, from table A3, we can achieve the same RSE with our linear hgrfat fit, without resorting to a more complicated model. So now, we can make a model that combines the hcwt and hgrfat fits using equation A3.

$$ctleankg = a + b.e^{c.hcwt} + d.hgrfat \qquad \text{(A3)}$$

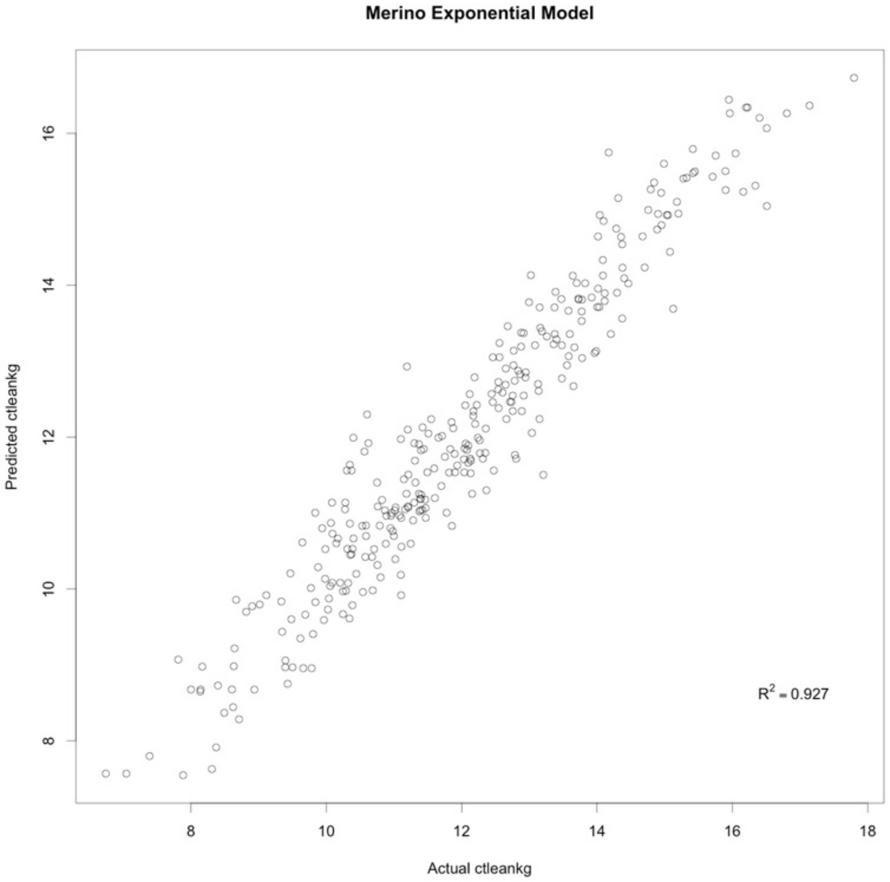


Figure A4: Scatter plot of predicted ctleankg verses actual ctleankg for Merinos, using equation A3.

Coefficients	**Estimate**	**Std. Error**	**t value**	**Pr(>|t|)**
a	30.069459	2.587802	11.620	< 2e-16
b	-36.305086	1.616760	-22.455	< 2e-16
c	-0.036384	0.005369	-6.776	6.37e-11
d	-0.102251	0.011130	-9.187	< 2e-16

Table A5: Results of equation 3 fit for ctleankg verses hgrfat and hcwt for Merinos.

The residual standard error of the fit was 0.5654 kg on 305 degrees of freedom and the multiple R^2 was 0.9265. In table A5, we can see that all of the selected terms are statistically significant well below the 1% level. Next, we can compare the result from the exponential fit with a simple linear fit on hcwt and hgrfat, as shown in equation A4.

$$ctleankg = a + b.hcwt + c.hgrfat \tag{A4}$$

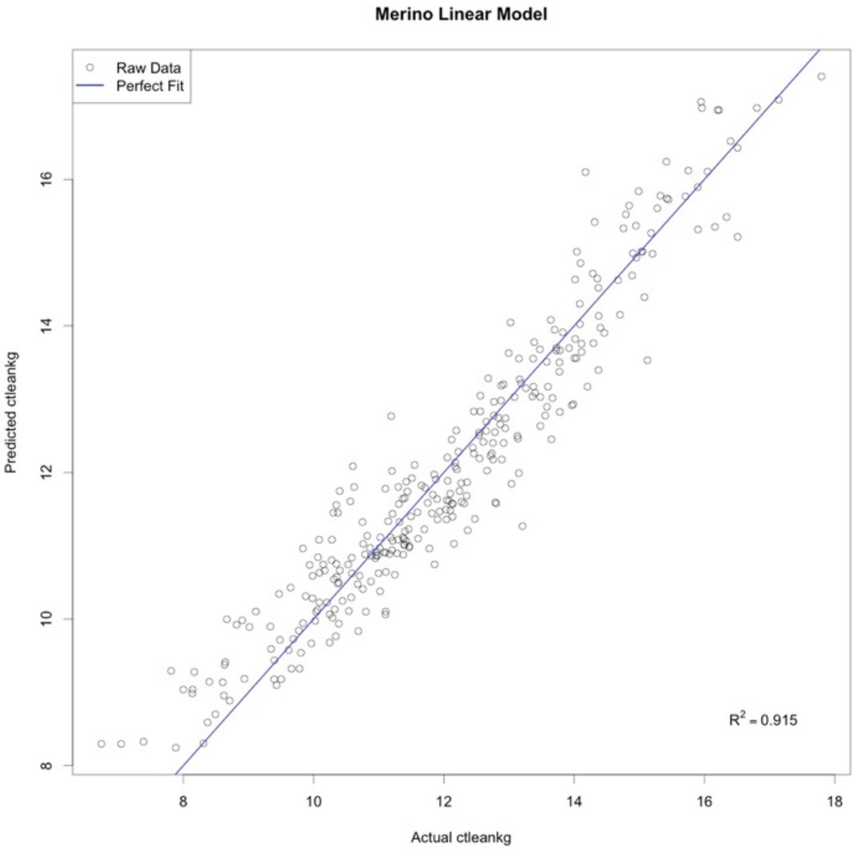


Figure A5: Scatter plot of predicted ctleankg verses actual ctleankg for Merinos, using equation A4.

Coefficients	**Estimate**	**Std. Error**	**t value**	**Pr(>|t|)**
Intercept	0.48879	0.25216	1.938	0.0535
hcwt	0.59949	0.01637	36.628	<2e-16
hgrfat	-0.11272	0.01186	-9.505	<2e-16

Table A6: Results of equation A4 fit for ctleankg verses hgrfat and hcwt for Merinos.

The residual standard error was 0.6079 kg on 306 degrees of freedom a little higher than the 0.5654 kg achieved with the hcwt exponential fit. The multiple R^2 for the linear fit was 0.9148 which is slightly less than the 0.9265 achieved with the exponential fit. The F-statistic was 1643 on 2 and 306 degrees of freedom with a p-value < 2.2e-16. Visual inspection of figure A5 indicates that there is a small amount of positive (upward) curvature in the prediction response. This is presumably because the linear fits do not correct for the negative curvature of the ctleankg versus hcwt relationship.

6.2.2 Lean Meat Yield%

Now we have formulas which are accurate at predicting ctleankg, we can now modify it to calculate LMY%, as shown in equation A3.

$$LMY\% = \frac{ctleankg}{hcwt} \qquad \text{(A5)}$$

Using the exponential function in equation A3 we get,

$$LMY\% = \frac{a + b.e^{c.hcwt} + d.hgrfat}{hcwt} \qquad \text{(A6)}$$

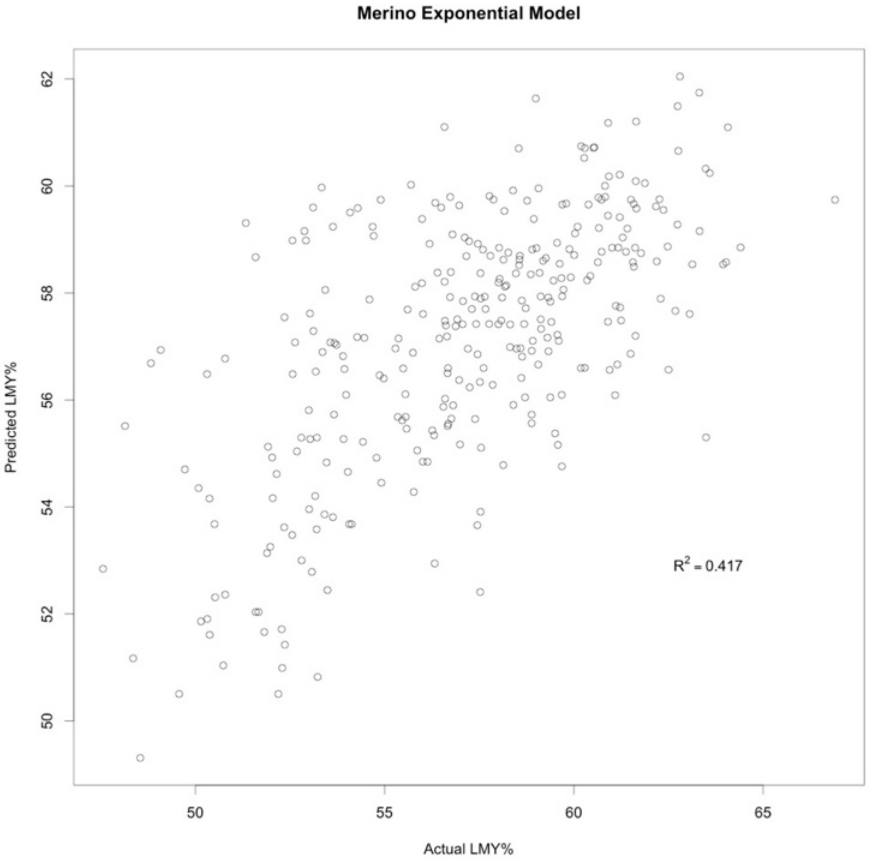


Figure A6: Scatter plot of new exponential model predicted LMY% verses actual LMY% for Merinos.

A scatterplot using the result of equation A6 is shown in figure A6, the R^2 is measured to be 0.417 and the residual standard error was 2.802%. Next, the old methodology was examined to compare the R^2 from the new and old methods. For the old methodology a linear model of the form shown in equation A7 was used.

$$LMY\% = a_0 + a_1 hcwt + a_2 hgrfat \tag{A7}$$

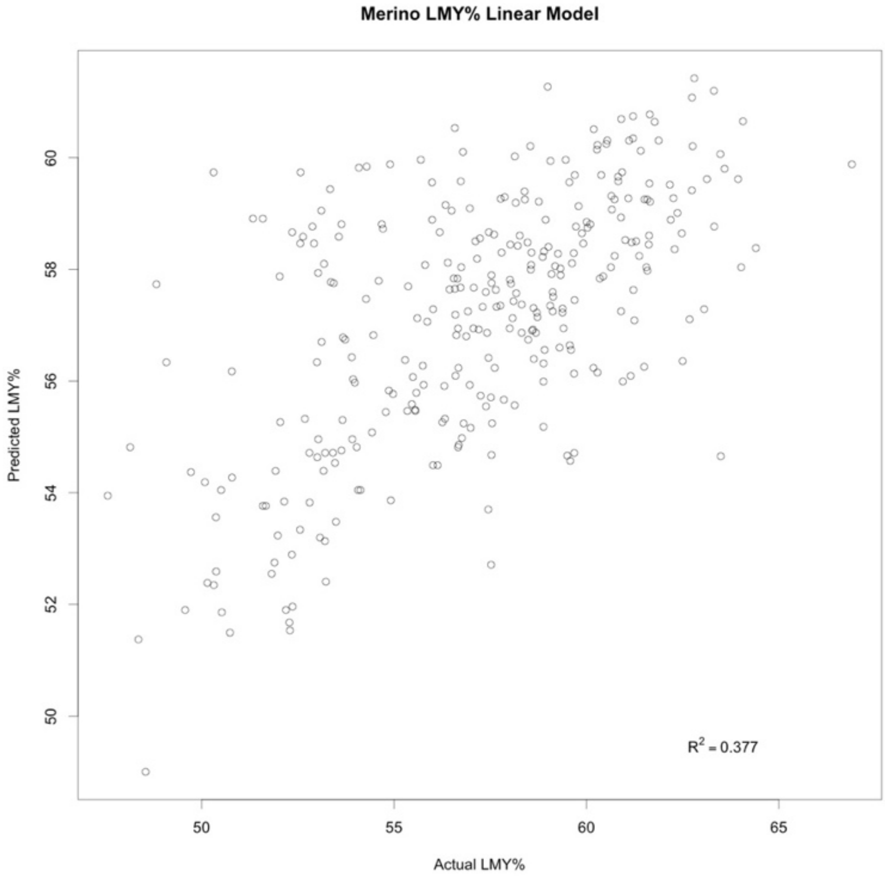


Figure A7: Scatter plot of old model predicted LMY% verses actual LMY% for Merinos.

Coefficients	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	58.03166	1.20189	48.284	< 2e-16
$hcwt$	0.20293	0.07801	2.601	0.00974
$hgrfat$	-0.50662	0.05653	-8.962	< 2e-16

Table A7: Results of polynomial fit for LMY% verses hgrfat and hcwt for Merinos.

The residual standard error of the fit was 2.898% on 306 degrees of freedom, compared to 2.802% for the exponential model. The multiple R^2 was 0.3771 compared to 0.417 for the exponential model, and the F-statistic was 92.61 with a p-value of < 2.2e-16. In table A7, we can see that all of the selected terms are statistically significant at the 1% level.

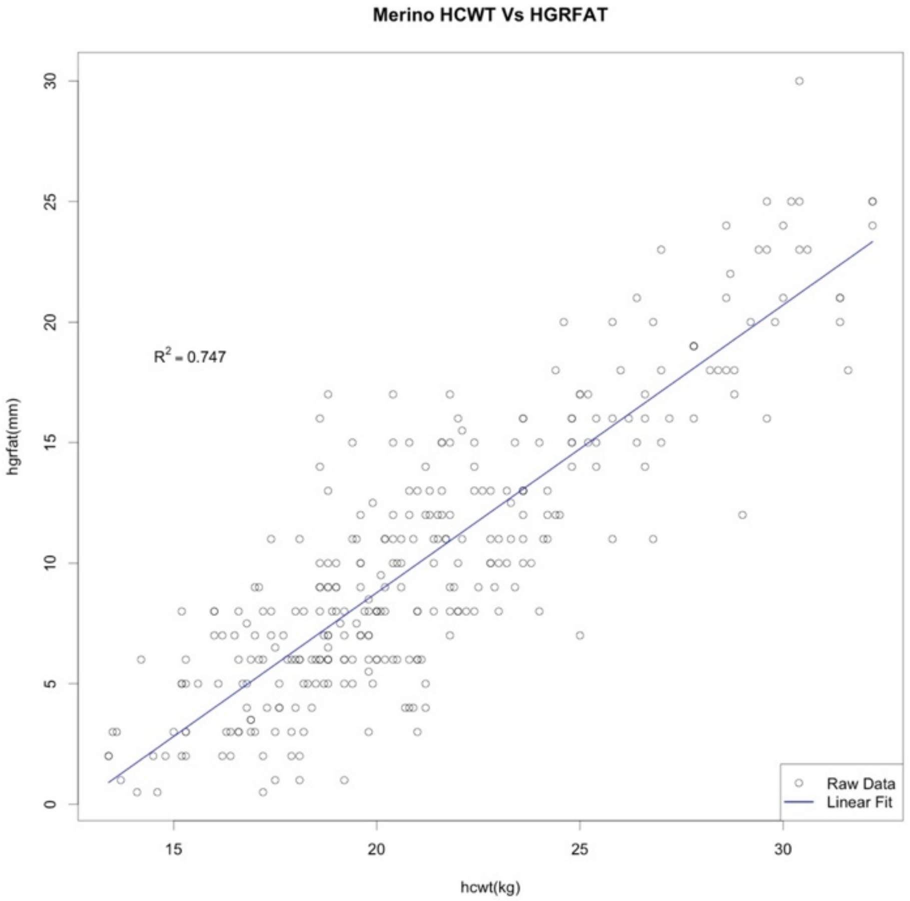


Figure A8: Scatter plot of hgrfat verses hcwt for Merinos.

Now, it was also observed that there is a linear relationship between hgrfat and hcwt as seen in figure A8. The results of the linear fit are shown in table A8.

Coefficients	**Estimate**	**Std. Error**	**t value**	**Pr(>|t|)**
Intercept	-15.08597	0.85512	-17.64	<2e-16
$\boldsymbol{hcwt}$	1.19316	0.03958	30.14	<2e-16

Table A8: Results of linear fit for hgrfat verses hcwt for Merinos.

The residual standard error of the linear fit was 2.926mm on 307 degrees of freedom with a multiple R^2 of 0.7475. The F-statistic was 908.6 on 1 and 307 degrees of freedom with a p-value of < 2.2e-16. Now, if we were using a linear model like equation A4 for ctleankg, and we applied equation A5 on it we would get equation A8.

$$LMY\% = \frac{ctleankg}{hcwt} = \frac{a + b.hcwt + c.hgrfat}{hcwt}$$

$$LMY\% = \frac{a}{hcwt} + b + \frac{c.hgrfat}{hcwt} \qquad \text{(A8)}$$

Now, from table A8, hgrfat has the linear relationship shown in equation A8.

$$hgrfat = k.hcwt + l \qquad \text{(A9)}$$

Substituting equation A9 into equation A8 we get equation A10.

$$LMY\% = \frac{a}{hcwt} + b + \frac{c.(k.hcwt + l)}{hcwt}$$

$$LMY\% = b + c.k + \frac{a + c.l}{hcwt}$$

Let $q = b + c.k$ and $p = a + c.l$ then we get

$$LMY\% = q + \frac{p}{hcwt} \qquad \text{(A10)}$$

Coefficients	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	49.701	1.027	48.375	< 2e-16
$\mathbf{1/hcwt}$	152.379	20.602	7.396	1.35e-12

Table A9: Results of linear fit for LMY% verses 1/hcwt for Merinos.

We find that this fit is not as good as the previous attempts, it has a residual standard error of 3.377% on 307 degrees of freedom. It has a multiple R^2 of only 0.1512, with an F-statistic of 54.71 on 1 and 307 degrees of freedom and a p-value: 1.346e-12. We extend this model by changing equation A10 to equation A11, by including a linear hgrfat term.

$$LMY\% = q + \frac{p}{hcwt} + r.hgrfat \qquad \text{(A11)}$$

Coefficients	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	69.26637	1.94279	35.653	< 2e-16
$\mathbf{1/hcwt}$	-131.70966	30.64390	-4.298	2.32e-05

$hgrfat$	-0.55391	0.04924	-11.249	< 2e-16

Table A10: Results of linear hgrfat and reciprocal hcwt terms for Merino LMY%.

Repeating the fitting exercise, we find that the addition of the linear hgrfat term gives a residual standard error of 2.845% on 306 degrees of freedom (similar to previous models). A multiple R^2 of 0.3995 which is much better than the 0.1512. The F-statistic was 101.8 on 2 and 306 degrees of freedom with a p-value of < 2.2e-16. It appears that using the reciprocal of hcwt instead of hcwt in the regression we can improve the R^2 from 0.3771 to 0.3995, presumably because the reciprocal function has some curvature which can be used to compensate for the curved relationship between ctleankg and hcwt.

ctleankg	RSE	RSQ	DF
$a + b.e^{c.hcwt} + d.hgrfat$	0.5654	0.9265	305
$a + b.hcwt + c.hgrfat$	0.6079	0.9148	306
$a + b.e^{c.hcwt}$	0.6377	0.9062	306
$a + b.hcwt + c.hcwt^2$	0.6387	0.906	306
$a + b.e^{c.hgrfat}$	1.408	0.5432	306
$a + b.hgrfat$	1.408	0.5414	307
LMY%	**RSE**	**RSQ**	**DF**
$\frac{a + b.e^{c.hcwt} + d.hgrfat}{hcwt}$	2.802	0.417	305
$a + \frac{b}{hcwt} + c.hgrfat$	2.845	0.3995	306
$a + b.hcwt + c.hgrfat$	2.898	0.3771	306
$a + \frac{b}{hcwt}$	3.377	0.1512	307
HGRFAT Vs HCWT	**RSE**	**RSQ**	**DF**
$a + b.hcwt$	2.926	0.7475	307

Table A11: Summary of Merino ctleankg and LMY% models and their RSE, R^2 and DF values.

Table A11 shows all the RSE and R^2 values for each of the Merino models discussed. The equations are ordered from highest R^2 to lowest R^2 within their respective categories. We can see that the models where hgrfat is linear, and hcwt is exponential give the best results with the smallest R^2 and RSE. It is also found that using the reciprocal of hcwt together with a linear hgrfat term leads to a great improvement over the typical linear hcwt and linear hgrfat configuration. It was observed that the Merino data sets showed the greatest amount of curvature in ctleankg versus hcwt. In the next sections we will examine the response of the INF maternal and terminal sheep data.

6.3 Maternal Results

6.3.1 Ctleankg

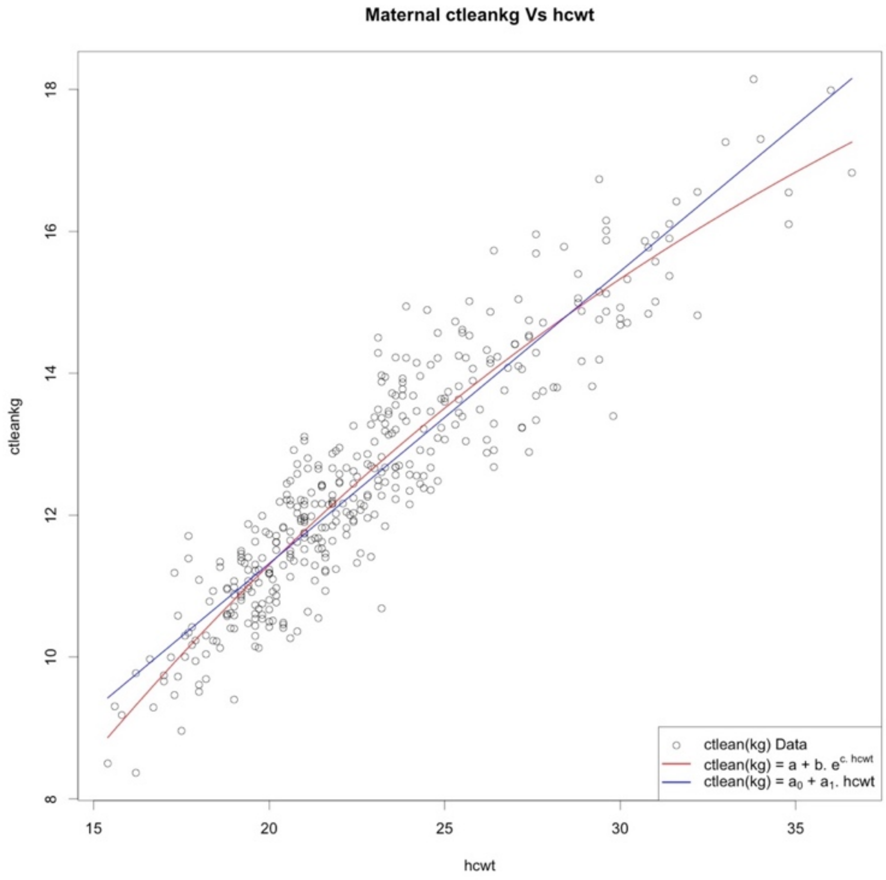


Figure A9: Scatter plot of ctleankg verses hcwt for maternals, with linear and exponential fits.

In figure A9, it can be seen that there is some slight negative curvature in the ctleankg and hcwt scatter plot. The exponential function shown in equation A2 was fitted (see red curve) as well as a linear fit (blue curve) and plotted together on the scatter plot.

Coefficients	Estimate	Std. Error	t value	Pr(>|t|)
$\boldsymbol{a}$	23.97737	2.42753	9.877	< 2e-16
$\boldsymbol{b}$	-27.23983	0.79963	-34.065	< 2e-16
$\boldsymbol{c}$	-0.03825	0.00851	-4.495	9.2e-06

Table A12: Results of exponential fit for ctleankg verses hcwt for maternals.

The residual standard error of the fit was 0.6646 kg on 387 degrees of freedom. The multiple R^2 of the exponential fit was 0.8525 which is worse than the Merino fit with 0.9062.

Coefficients	**Estimate**	**Std. Error**	**t value**	**Pr(>|t|)**
Intercept	3.081209	0.207709	14.83	<2e-16
hcwt	0.411877	0.008969	45.92	<2e-16

Table A13: Results of exponential fit for ctleankg verses hcwt for maternals.

The linear fit results are shown in table A13, the residual standard error was 0.6813 kg (total lean) on 388 degrees of freedom. The multiple R^2 was 0.8446 which is smaller than for the exponential model with 0.8525, the F-statistic was 2109 on 1 and 388 degrees of freedom with a p-value of < 2.2e-16.

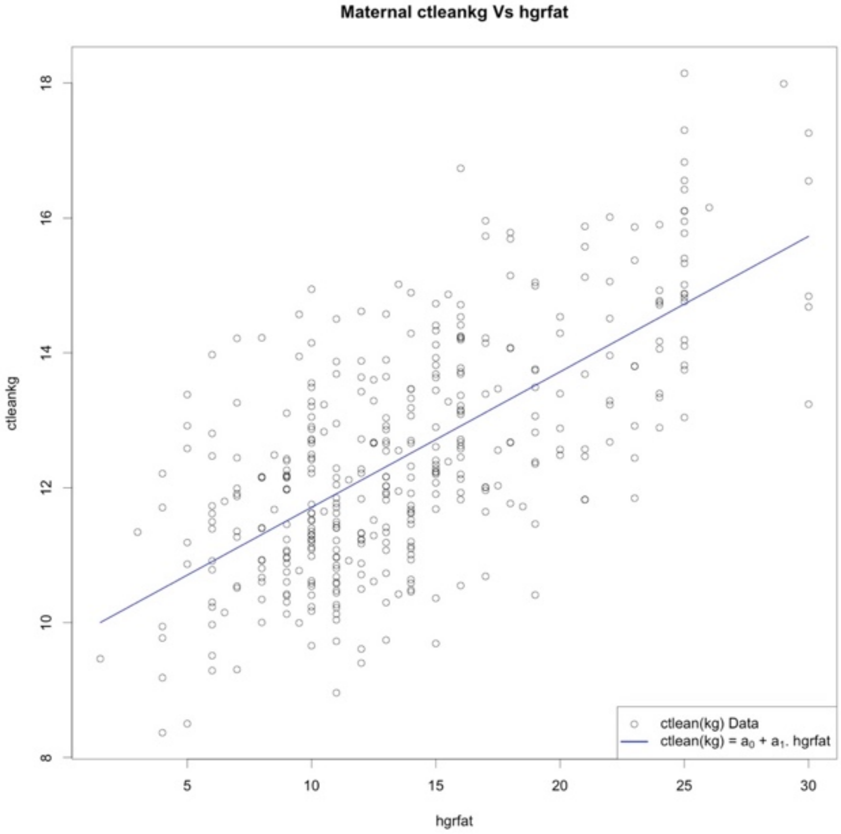


Figure A10: Scatter plot of ctleankg verses hgrfat for maternals and linear fits.

The linear and exponential models were tested against hgrfat, and it was found that the nonlinear regression solver was not able to converge to a result for the hgrfat relationship. Since, the ctleankg versus hgrfat relationship appears linear and a linear relationship was used for the Merino's then the same approach was used here for the maternals.

Coefficients	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	9.69962	0.17195	56.41	<2e-16
$hgrfat$	0.20094	0.01146	17.53	<2e-16

Table A14: Results of linear fit for ctleankg verses hgrfat for maternals.

The residual standard error of the linear fit was 1.291mm on 388 degrees of freedom. The multiple R^2 was 0.4419, and the F-statistic was 307.2 on 1 and 388 degrees of freedom with a p-value of < 2.2e-16. So, from the results of table A12 and table A14, we can make a model that combines the hcwt and hgrfat fits using equation A3.

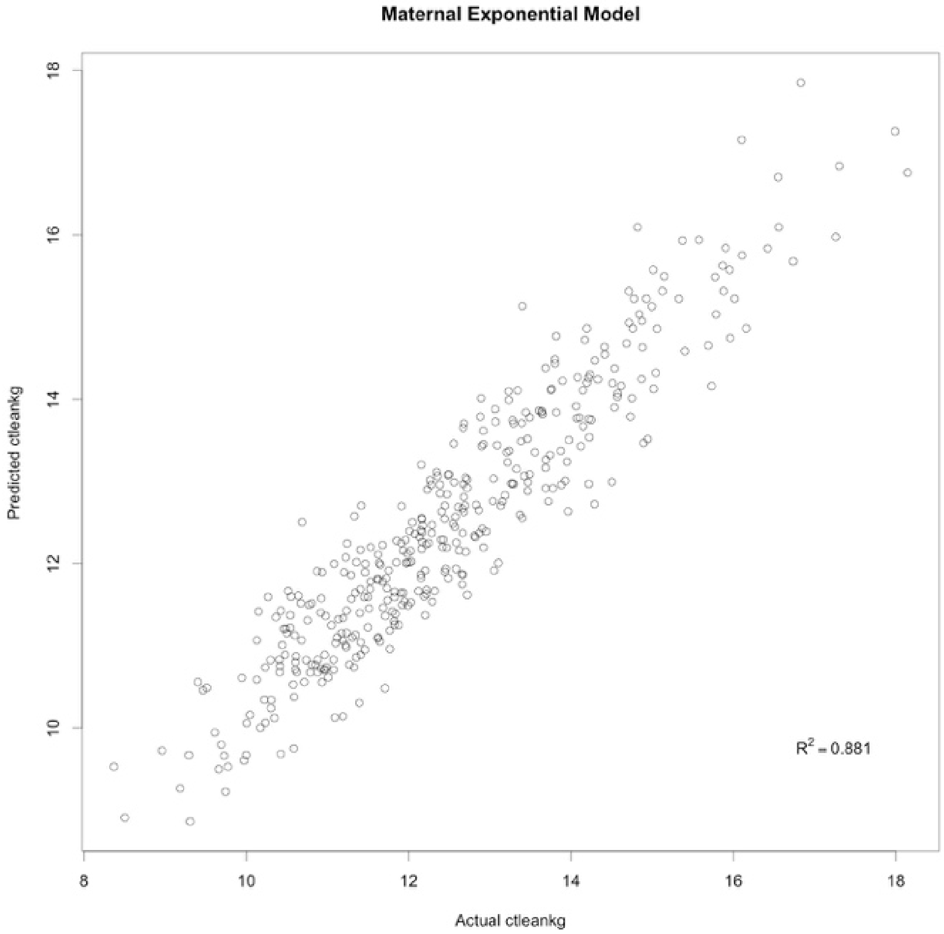


Figure A11: Scatter plot of predicted ctleankg verses actual ctleankg for maternals, using equation A3.

Coefficients	Estimate	Std. Error	t value	Pr(>|t|)
a	34.107116	4.503546	7.573	2.71e-13
b	-37.449191	3.252098	-11.515	< 2e-16
c	-0.026912	0.006020	-4.470	1.03e-05
d	-0.090791	0.009499	-9.558	< 2e-16

Table A15: Results of equation A3 fit for ctleankg verses hgrfat and hcwt for maternals.

The residual standard error of the fit was 0.5985 kg on 386 degrees of freedom and the multiple R^2 was 0.8807. In table A15, we can see that all of the selected terms are statistically significant well below the 1% level. Next, we can compare the result from the exponential fit with a simple multiple linear fit on hcwt and hgrfat, as shown in equation A4.

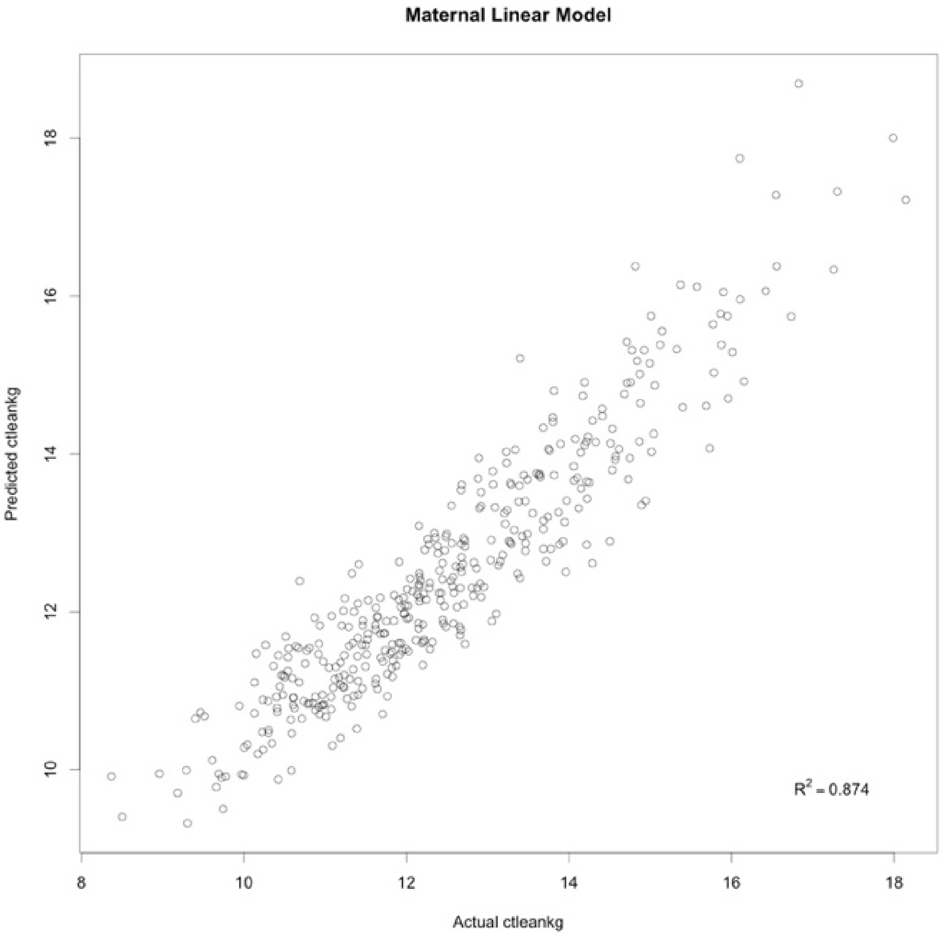


Figure A12: Scatter plot of predicted ctleankg verses actual ctleankg for maternals, using equation A4.

Coefficients	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	1.768021	0.232420	7.607	2.15e-13
hcwt	0.525684	0.014419	36.457	< 2e-16
hgrfat	-0.092686	0.009725	-9.530	< 2e-16

Table A16: Results of equation A4 fit for ctleankg verses hgrfat and hcwt for maternals.

The residual standard error was 0.6139 kg on 387 degrees of freedom which was higher than the 0.5985 kg achieved with the hcwt exponential fit. The multiple R^2 for the linear fit was 0.8741which is slightly less than the 0.8807achieved with the exponential fit. The F-statistic was 1344 on 2 and 387 degrees of freedom with a p-value < 2.2e-16. Visual inspection of figure A12 indicates that there is a small amount of positive (upward) curvature in the prediction response. This is presumably because the linear fits do not correct for the negative curvature of the ctlean versus hcwt relationship.

6.3.2 Lean Meat Yield%

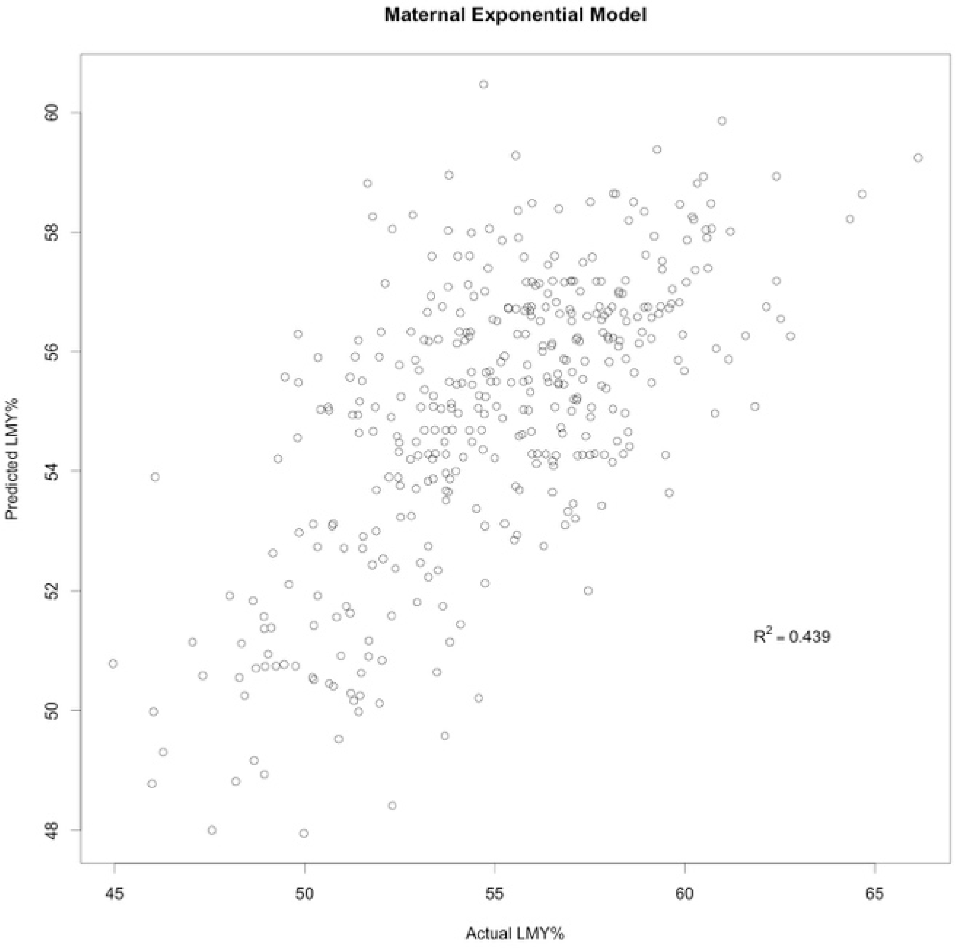


Figure A13: Scatter plot of new exponential model predicted LMY% verses actual LMY% for maternals.

A scatterplot using the result of equation A6 is shown in figure A13, the R^2 is measured to be 0.439 and the residual standard error was 2.647%. Next, the old methodology was examined to compare the R^2 from the new and old methods. For the old methodology a linear model of the form shown in equation A7 was used.

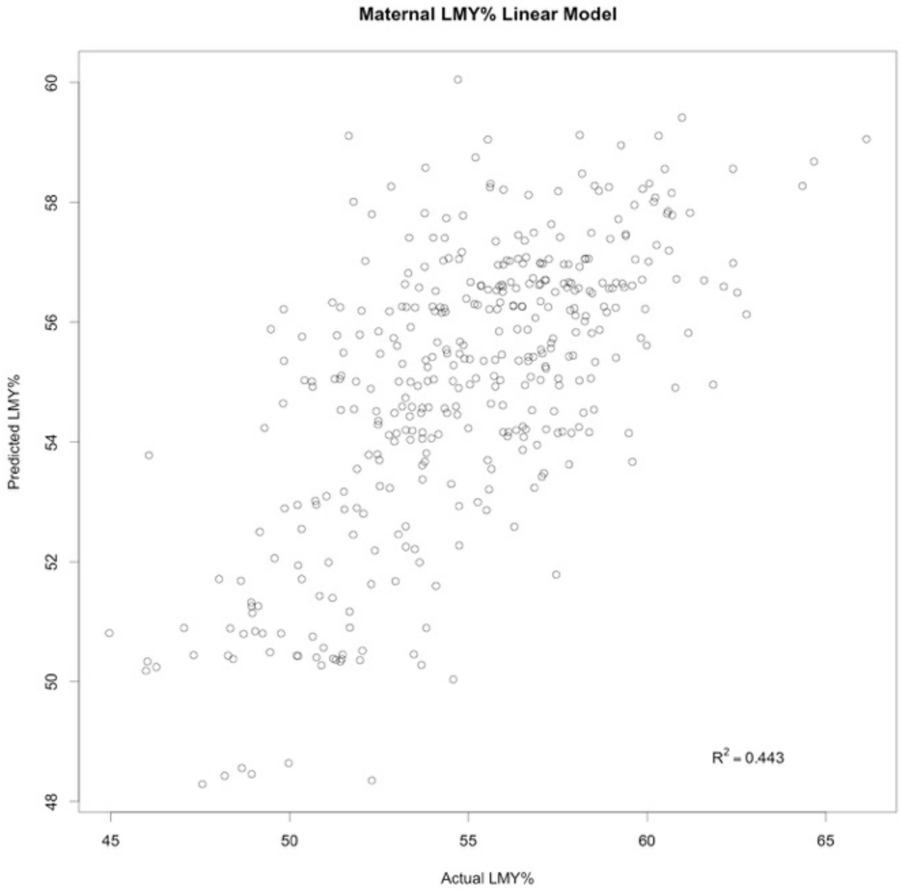


Figure A14: Scatter plot of old model predicted LMY% verses actual LMY% for maternal's.

Coefficients	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	61.24137	0.99782	61.375	<2e-16
$hcwt$	-0.03507	0.06191	-0.566	0.571
$hgrfat$	-0.39118	0.04175	-9.369	<2e-16

Table A17: Results of linear fit for LMY% verses hgrfat and hcwt for maternal's.

The residual standard error of the fit was 2.636 % on 387 degrees of freedom, compared to 2.647% for the exponential model. The multiple R^2 was 0.4435 compared to 0.439 for the exponential model, and the F-statistic was 154.2 with a p-value of < 2.2e-16. In table A17, we can see that hcwt is not statistically significant at any reasonable level for this linear model.

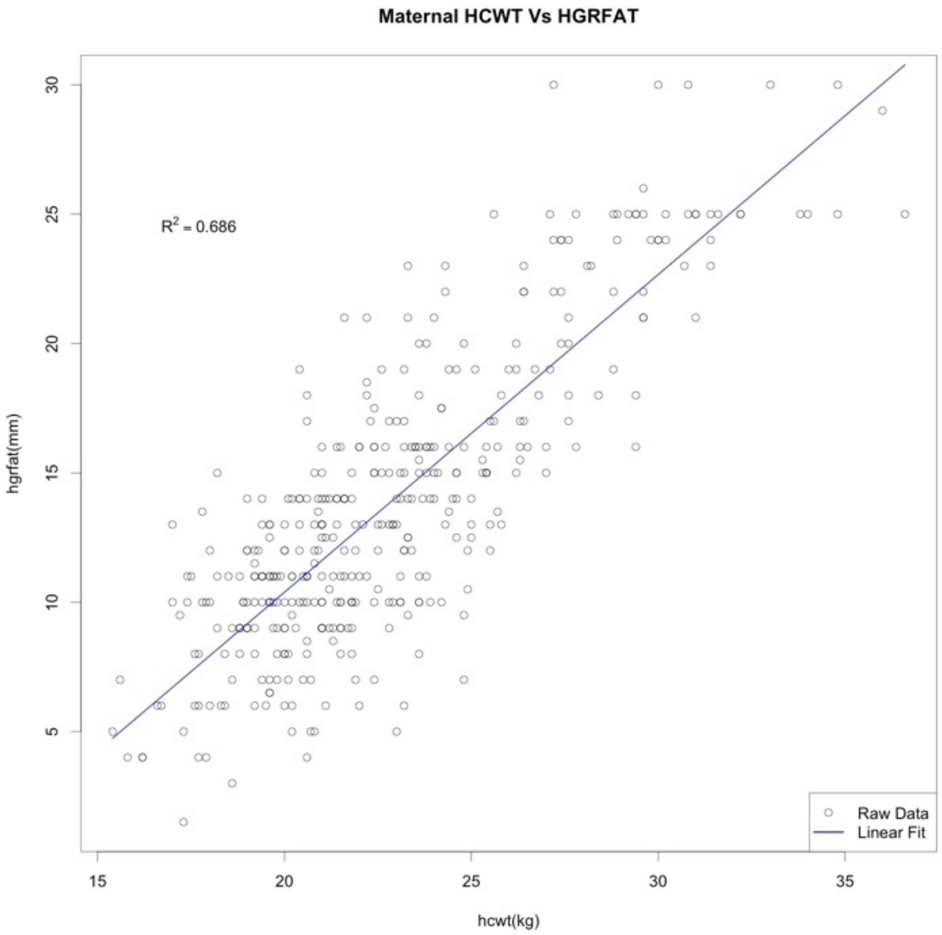


Figure A15: Scatter plot of hgrfat verses hcwt for maternals.

Now, it was also observed that that there is a linear relationship between hgrfat and hcwt as see in figure A15, the results of the linear fit are shown in table A18.

Coefficients	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	-14.16812	0.97704	-14.50	<2e-16
$hcwt$	1.22787	0.04219	29.11	<2e-16

Table A18: Results of linear fit for hgrfat verses hcwt for maternals.

The residual standard error of the linear fit was 3.205 mm on 388 degrees of freedom with a multiple R^2 of 0.6858. The F-statistic was 847.1 on 1 and 388 degrees of freedom with a p-value of < 2.2e-16. We can see from figure A15, that in some sense the hgrfat values and the hcwt carry the same information, as there is a strong positive correlation between hgrfat and hcwt. Because they supply redundant information, removing one of the correlated factors shouldn't drastically reduce the R^2.

Next, as was shown in the Merino section, we will test equation A10 which indicates that the lean meat yield percentage should be proportional to a constant plus the reciprocal of hcwt.

Coefficients	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	42.9307	0.9765	43.96	<2e-16
$\boldsymbol{1/hcwt}$	268.8372	21.4669	12.52	<2e-16

Table A19: Results of linear fit for LMY% verses 1/hcwt for maternals.

We find that this fit is not as good as the previous attempts, it has a residual standard error of 2.978% on 388 degrees of freedom. It has a multiple R^2 of only 0.2879, with an F-statistic of 156.8 on 1 and 388 degrees of freedom and a p-value < 2.2e-16. However, if we extend this model by changing equation A10 to equation A11, by including a linear hgrfat term.

Coefficients	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	60.52222	1.90200	31.820	<2e-16
$\boldsymbol{1/hcwt}$	3.23949	31.86733	0.102	0.919
$\boldsymbol{hgrfat}$	-0.40756	0.03925	-10.384	<2e-16

Table A20: Results of linear hgrfat and reciprocal hcwt terms for maternal LMY%.

Repeating the fitting exercise, we find that the addition of the linear hgrfat term gives a residual standard error of 2.637% on 387 degrees of freedom. A multiple R^2 of 0.443, which is much better than the 0.2879. The F-statistic was 153.9 on 2 and 387 degrees of freedom with a p-value of < 2.2e-16. It appears that using the reciprocal of hcwt instead of hcwt in the regression we can improve the R^2 from 0.2879 to 0.443, presumably because the reciprocal function has some curvature, which can be used to compensate for the curved relationship between ctleankg and hcwt. However, again the hcwt term is not statistically significant due to the strong linear relationship between hcwt and hgrfat.

$$LMY\% = a + b.hgrfat \qquad \text{(A11)}$$

Next, we can test just a linear fit between LMY% and hgrfat shown in equation A11, to see if any improvement can be achieved compared to the fits which use a combination of hcwt and hgrfat.

Coefficients	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	60.71224	0.35072	173.11	<2e-16
$\boldsymbol{hgrfat}$	-0.41076	0.02338	-17.57	<2e-16

Table A21: Results of linear hgrfat fit for LMY% in maternals.

The residual standard error is 2.633% on 388 degrees of freedom with a multiple R^2 of 0.443, which is the same as the model with the additional linear hgrfat term. The F-statistic was 308.6 on 1 and 388 degrees of freedom with a p-value of < 2.2e-16.

ctleankg	**RSE**	**RSQ**	**DF**
$\boldsymbol{a + b.e^{c.hcwt} + d.hgrfat}$	0.5985	0.8807	386
$\boldsymbol{a + b.hcwt + c.hgrfat}$	0.6139	0.8741	387
$\boldsymbol{a + b.e^{c.hcwt}}$	0.6646	0.8525	387
$\boldsymbol{a + b.hcwt}$	0.6813	0.8446	388
$\boldsymbol{a + b.hgrfat}$	1.291	0.4419	388
LMY%	**RSE**	**RSQ**	**DF**
$\boldsymbol{a + b.hcwt + c.hgrfat}$	2.636	0.4435	387
$\boldsymbol{a + \frac{b}{hcwt} + c.hgrfat}$	2.637	0.443	387
$\boldsymbol{a + b.hgrfat}$	2.633	0.443	388
$\boldsymbol{\frac{a + b.e^{c.hcwt} + d.hgrfat}{hcwt}}$	2.647	0.439	386
$\boldsymbol{a + \frac{b}{hcwt}}$	2.978	0.2879	388
HGRFAT Vs HCWT	**RSE**	**RSQ**	**DF**
$\boldsymbol{a + b.hgrfat}$	3.205	0.6858	388

Table A22: Summary of maternal ctleankg and LMY% models and their RSE, R^2 and DF values.

Table 22 shows all of the RSE and R^2 values for each of the maternal models discussed. The equations are ordered from highest R^2 to lowest R^2 within their respective categories. We can see that the models where hgrfat is linear, and hcwt is exponential give the best results with the smallest RSE and largest R^2 for predicting ctleankg. In addition, it is possible to eliminate complicated models for maternal LMY% prediction as you can get almost the same results by just using a linear fit against hgrfat. This is because hcwt and hgrfat have a strong relationship causing these two parameters to carry the same information. Thus, we can eliminate the hcwt term and achieve pretty much the same results.

6.4 Terminal Results

6.4.1 Ctleankg

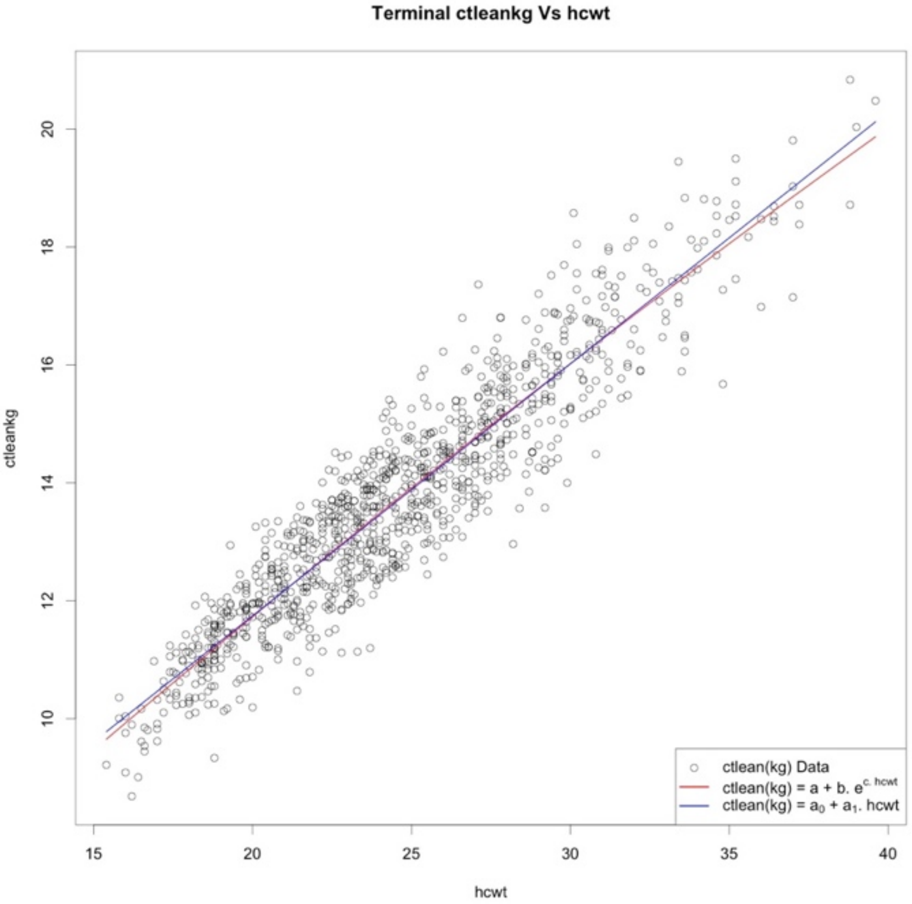


Figure A16: Scatter plot of ctleankg verses hcwt for terminals, with linear and exponential fits.

In figure A16, it can be seen through visual inspection that there is almost no curvature in the ctleankg and hcwt scatter plot compared to Merino's and maternals. So, for terminals we may not need to bother with the complicated exponential function model. The exponential function shown in equation A2 was fitted (see red curve) as well as a linear fit (blue line) and plotted together on the scatter plot.

Coefficients	Estimate	Std. Error	t value	Pr(>|t|)
a	75.271630	36.815172	2.045	0.0411
b	-73.081826	36.178444	-2.020	0.0436
c	-0.006994	0.004211	-1.661	0.0971

Table A23: Results of exponential fit for ctleankg verses hcwt for terminals.

The residual standard error of the fit was 0.7377 kg on 1030 degrees of freedom. The multiple R^2 of the exponential fit was 0.8684 which is worse than the Merino fit with 0.9062 but better than the maternal fit. Here, we can see that none of the parameters satisfy statistical significance at the 1% level. At best we can claim that a, b and c are statistically significant at the 10% level.

Coefficients	Estimate	Std. Error	t value	Pr(>|t|)
$Intercept$	3.19859	0.12959	24.68	<2e-16
$hcwt$	0.42736	0.00519	82.34	<2e-16

Table A24: Results of exponential fit for ctleankg verses hcwt for terminals.

The linear fit results are shown in table A24, the residual standard error was 0.7383 kg on 1031 degrees of freedom. The multiple R^2 was 0.868 which is slightly smaller than for the exponential model with 0.8684, the F-statistic was 6780 on 1 and 1031 degrees of freedom with a p-value of < 2.2e-16.

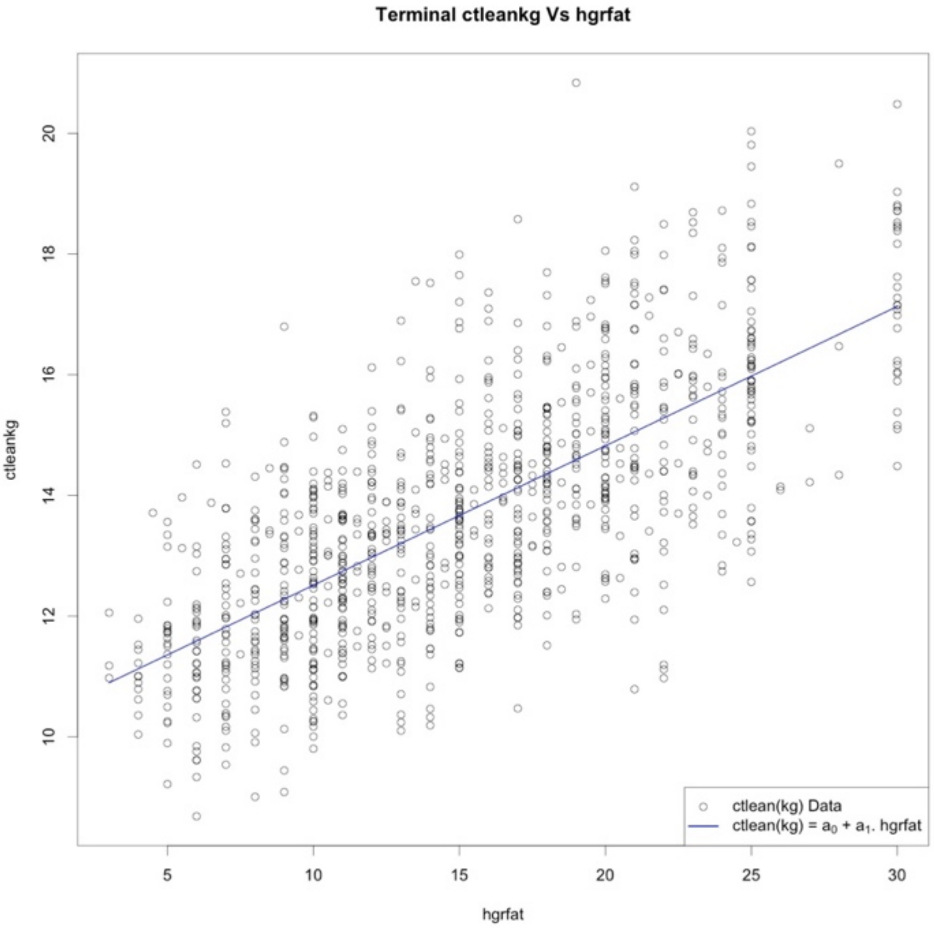


Figure A17: Scatter plot of ctleankg verses hgrfat for terminals and linear fits.

The linear and exponential models were tested against hgrfat, and it was found that the nonlinear regression solver was not able to converge to a result for the hgrfat relationship. Since, the ctleankg versus hgrfat relationship appears linear and a linear relationship was used for the Merinos then the same approach was used for the terminals.

Coefficients	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	10.204014	0.119155	85.64	<2e-16
$hgrfat$	0.230903	0.007286	31.69	<2e-16

Table A25: Results of linear fit for ctleankg verses hgrfat for terminals.

The residual standard error of the linear fit was 1.446mm on 1031 degrees of freedom. The multiple R^2 was 0.4934, and the F-statistic was 1004 on 1 and 1031 degrees of freedom with a p-value of < 2.2e-16. So, from the results of table A23 and table A25, we can make a model that combines the hcwt and hgrfat fits using equation A3.

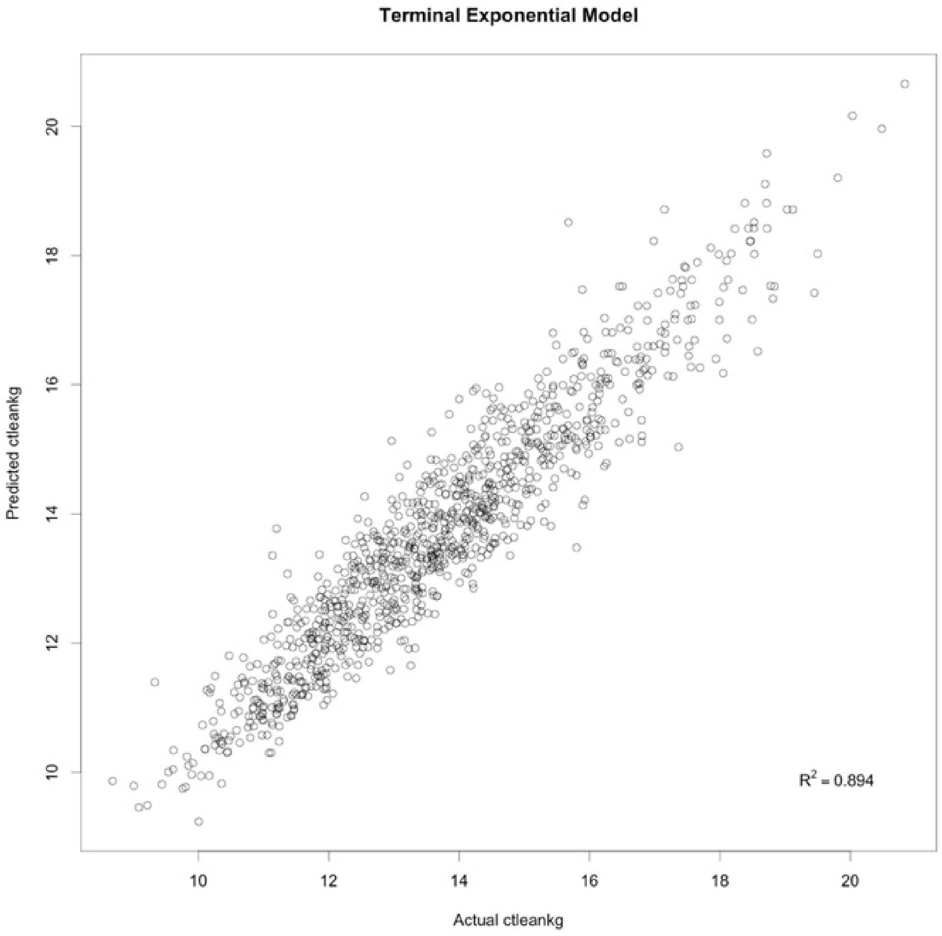


Figure 18: Scatter plot of predicted ctleankg verses actual ctleankg for terminals, using equation A3.

Coefficients	Estimate	Std. Error	t value	Pr(>|t|)
a	72.015043	17.476620	4.121	4.08e-05
b	-71.974232	16.879809	-4.264	2.19e-05
c	-0.009648	0.002998	-3.218	0.00133
d	-0.097864	0.006258	-15.639	< 2e-16

Table A26: Results of equation A3 fit for ctleankg verses hgrfat and hcwt for terminals.

The residual standard error of the fit was 0.6634 kg on 1029 degrees of freedom and the multiple R^2 was 0.8936. In table A26, we can see that all of the selected terms are statistically significant well below the 1% level, in contrast to the results from table A23. Next, we can compare the result from the exponential fit with a simple linear fit on hcwt and hgrfat, as shown in equation A4.

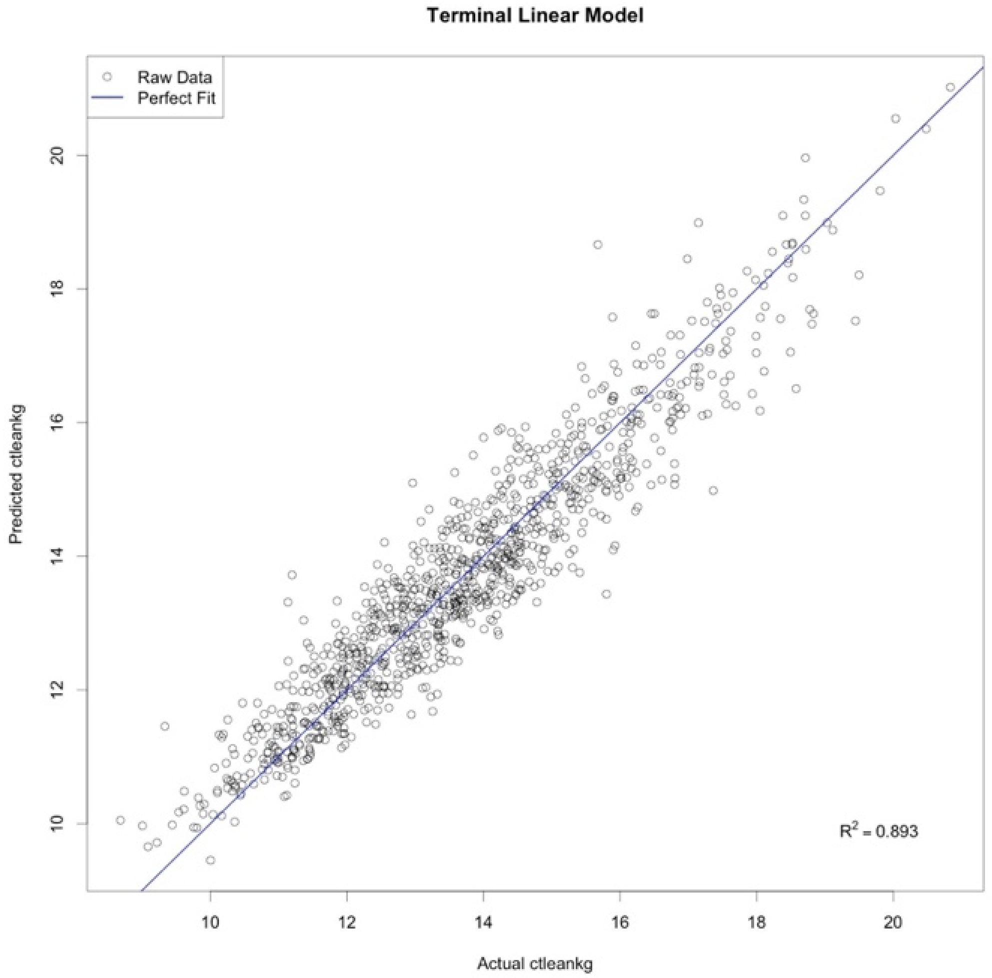


Figure A19: Scatter plot of predicted ctleankg verses actual ctleankg for terminals, using equation A4.

Coefficients	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	1.872932	0.145447	12.88	<2e-16
hcwt	0.540481	0.008738	61.85	<2e-16
hgrfat	-0.096040	0.006262	-15.34	<2e-16

Table A27: Results of equation 4 fit for ctleankg verses hgrfat and hcwt for terminals.

The residual standard error was 0.6665 kg on 1030 degrees of freedom which was slightly higher than the 0.6634 kg achieved with the hcwt exponential fit. The multiple R^2 for the linear fit was 0.8926 which is slightly less than the 0.8936 achieved with the exponential fit. The F-statistic was 4278 on 2 and 1030 degrees of freedom with a p-value < 2.2e-16.

6.4.2 Lean Meat Yield%

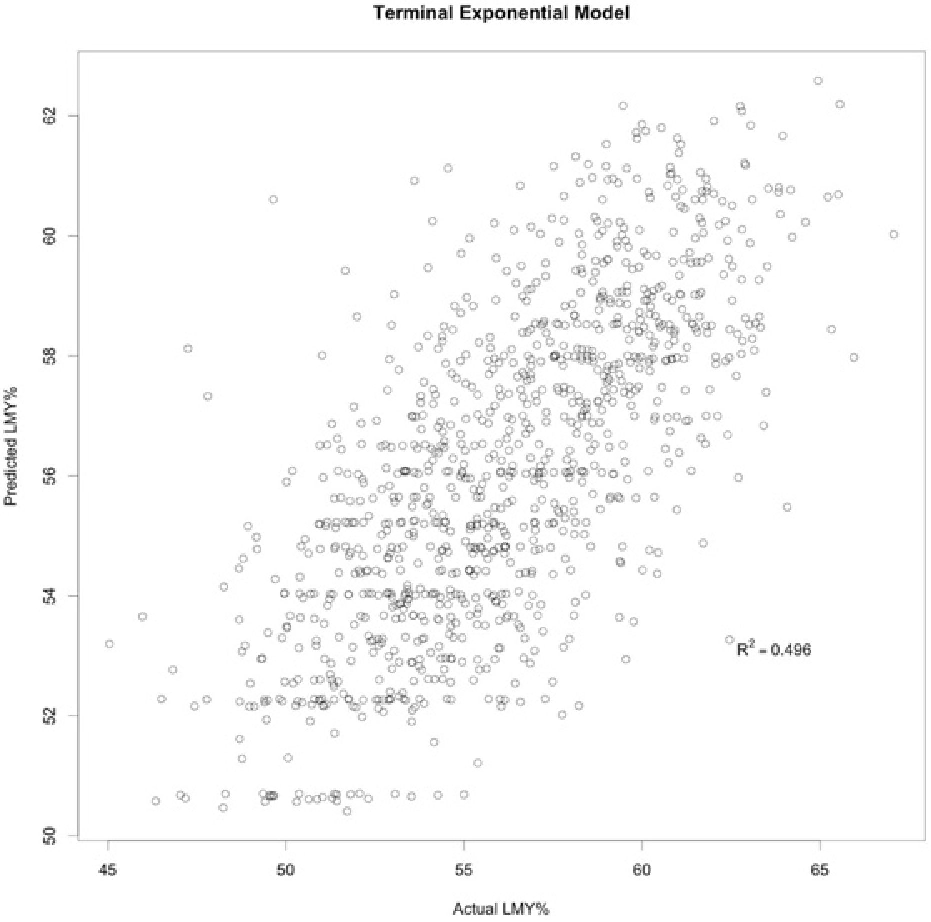


Figure A20: Scatter plot of new exponential model predicted LMY% verses actual LMY% for terminals.

A scatterplot using the result of equation A6 is shown in figure A20, the R^2 is measured to be 0.4958 and the residual standard error was 2.685%. Next, the old methodology was examined to compare the R^2 from the new and old methods. For the old methodology a linear model of the form shown in equation A7 was used.

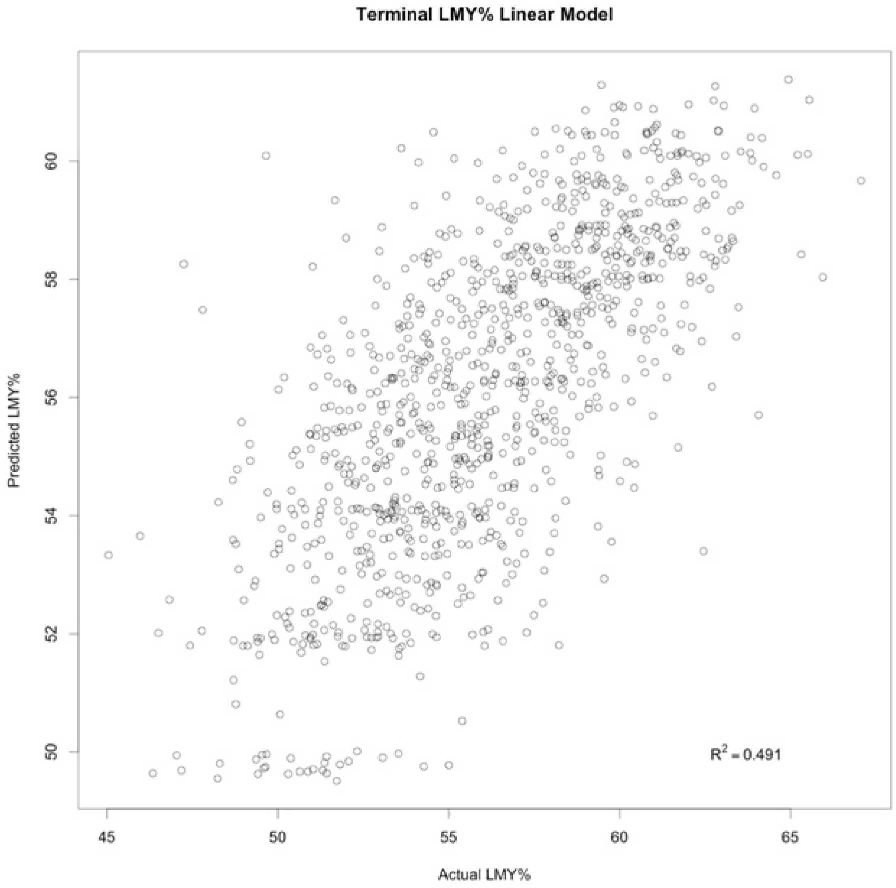


Figure A21: Scatter plot of old model predicted LMY% verses actual LMY% for terminals.

Coefficients	**Estimate**	**Std. Error**	**t value**	**Pr(>|t|)**
Intercept	63.40199	0.58868	107.701	<2e-16
$hcwt$	-0.04903	0.03537	-1.386	0.166
$hgrfat$	-0.39836	0.02534	-15.718	<2e-16

Table A28: Results of linear fit for LMY% verses hgrfat and hcwt for terminals.

The residual standard error of the fit was 2.697% on 1030 degrees of freedom, compared to 2.685% for the exponential model. The multiple R^2 was 0.4911 compared to 0.4958 for the exponential model, and the F-statistic was 497 with a p-value of < 2.2e-16. In table A28, we can see that hcwt is not statistically significant at the 10% level for this linear model. This is similar to how hcwt was not statistically significant for the maternal linear LMY% model.

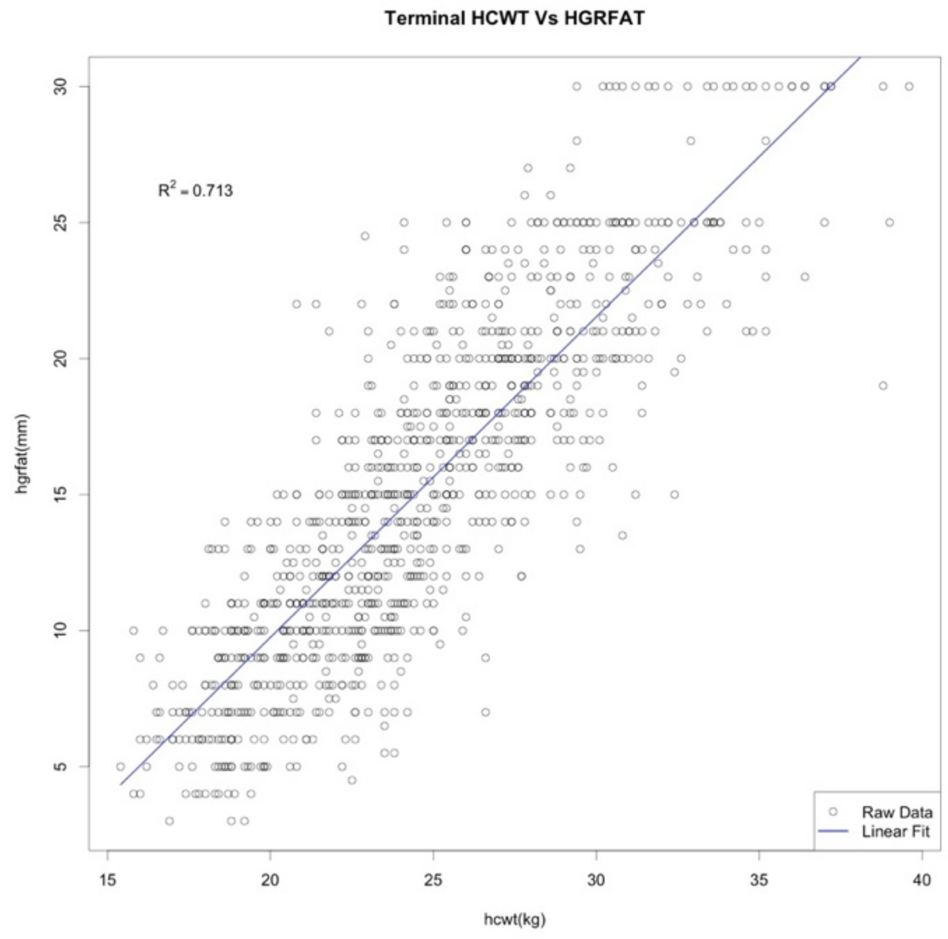


Figure A22: Scatter plot of hgrfat verses hcwt for terminals.

Now, it was also observed that that there is a linear relationship between hgrfat and hcwt as see in figure A22, the results of the linear fit are shown in table A29.

Coefficients	**Estimate**	**Std. Error**	**t value**	**Pr(>|t|)**
Intercept	-13.8032	0.5818	-23.72	<2e-16
$hcwt$	1.1779	0.0233	50.55	<2e-16

Table A29: Results of linear fit for hgrfat verses hcwt for terminals.

The residual standard error of the linear fit was 3.315 mm on 1031 degrees of freedom with a multiple R^2 of 0.7125. The F-statistic was 2555 on 1 and 1031 degrees of freedom with a p-value of < 2.2e-16. We can see from figure A22, that in some sense the hgrfat and hcwt values carry the same information, as there is a strong positive correlation between hgrfat and hcwt. Because they supply redundant information, removing one of the correlated factors shouldn't drastically reduce the R^2.

Now, if we were using a linear model like equation A4 for ctleankg, and we applied equation A5 on it we would get equation A8. Substituting equation A9 into equation A8 we get equation A10 which indicates that the lean meat yield percentage should be proportional to a constant plus the reciprocal of hcwt. Next, we test this hypothesis by performing a linear regression on the reciprocal of hcwt.

Coefficients	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	43.2624	0.5315	81.39	<2e-16
$\mathbf{1/hcwt}$	307.1388	12.4560	24.66	<2e-16

Table A30: Results of linear fit for LMY% verses 1/hcwt for terminals.

We find that this fit is not as good as the previous attempts, it has a residual standard error of 2.998% on 1031 degrees of freedom. It has a multiple R^2 of only 0.371, with an F-statistic of 608on 1 and 1031 degrees of freedom and a p-value < 2.2e-16. However, if we extend this model by changing equation A10 to equation A11, by including a linear hgrfat term.

Coefficients	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	60.06088	1.17052	51.311	<2e-16
$\mathbf{1/hcwt}$	45.15874	20.07537	2.249	0.0247
$\mathbf{hgrfat}$	-0.38257	0.02434	-15.719	<2e-16

Table A31: Results of linear hgrfat and reciprocal hcwt terms for terminal LMY%.

Repeating the fitting exercise, we find that the addition of the linear hgrfat term gives a residual standard error of 2.693% on 1030 degrees of freedom. A multiple R^2 of 0.4927, which is much better than the 0.371 achieved with the reciprocal of hcwt. The F-statistic was 500.1 on 2 and 1030 degrees of freedom with a p-value of < 2.2e-16. It appears that by using the reciprocal of hcwt instead of hcwt in the regression there is only a marginal improvement of the R^2 from 0.4911 to 0.4927. However, again the hcwt term is not statistically significant at the 1% level due to the linear relationship between hcwt and hgrfat.

Next, we can test just a linear fit between LMY% and hgrfat shown in equation A11, to see if any improvement can be achieved compared to the fits which use a combination of hcwt and hgrfat.

Coefficients	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	62.64618	0.22232	281.79	<2e-16
$\mathbf{hgrfat}$	-0.42803	0.01359	-31.48	<2e-16

Table 32: Results of linear hgrfat fit for terminal LMY%.

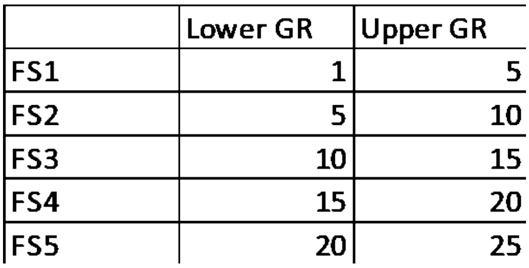
The residual standard error of 2.699 % on 1031degrees of freedom. A multiple R^2 of 0.4902, which is slightly lower than the model with the reciprocal hcwt and linear hgrfat term which was 0.4927. The F-statistic was 991.3 on 1 and 1031 degrees of freedom with a p-value of < 2.2e-16.

ctleankg	**RSE**	**RSQ**	**DF**
$a + b.e^{c.hcwt} + d.hgrfat$	0.6634	0.8936	1029
$a + b.hcwt + c.hgrfat$	0.6665	0.8926	1030
$a + b.e^{c.hcwt}$	0.7377	0.8684	1030
$a + b.hcwt$	0.7383	0.868	1031
$a + b.hgrfat$	1.446	0.4934	1031
LMY%	**RSE**	**RSQ**	**DF**
$\frac{a + b.e^{c.hcwt} + d.hgrfat}{hcwt}$	2.685	0.4958	1029
$a + \frac{b}{hcwt} + c.hgrfat$	2.693	0.4927	1030
$a + b.hcwt + c.hgrfat$	2.697	0.4911	1030
$a + b.hgrfat$	2.699	0.4902	1031
$a + \frac{b}{hcwt}$	2.998	0.371	1031
HGRFAT Vs HCWT	**RSE**	**RSQ**	**DF**
$a + b.hcwt$	3.315	0.7125	1031

Table A33: Summary of ctleankg and LMY% models and their RSE, R^2 and DF values for terminal lambs.

Table A33 shows the all of the RSE and R^2 values for each of the terminal models discussed. The equations are ordered from highest R^2 to lowest R^2 within their respective categories. We can see that the models where hgrfat is linear, and hcwt is exponential give the best results with the smallest RSE and largest R^2. This is the same result as we got for the Merino regression fits even though the ctleankg versus hcwt did not seem to have as much curvature as the Merino data had. However, if we used only a linear hgrfat model for LMY% prediction we would only sacrifice 0.4958 - 0.4911 = 0.0047 in R^2, making the LMY% much simpler without giving up a significant amount of accuracy.

6.5 Segmenting by Fat score



	Lower GR	Upper GR
FS1	1	5
FS2	5	10
FS3	10	15
FS4	15	20
FS5	20	25

Table A34: Fat score segmentation.

$$LMY\% = a_0 + a_1 hcwt + a_2 FS_1 + a_3 FS_2 + a_4 FS_3 + a_5 FS_4 + a_6 FS_5 \tag{A12}$$

The hgrfat values were segmented using the range values given in table A34, then fat score dummy variables were created as FS_{1-5}. A linear regression was performed using these dummy variables using equation A12. A non-linear fit of the segmented using the exponential function for hcwt and the dummy variables with fatscore, as seen in equation A13.

$$LMY\% = \frac{a+b.e^{c.hcwt}+e.FS_1+f.FS_2+g.FS_3+h.FS_4+i.FS_5}{hcwt} \tag{A13}$$

6.5.1 Merino Fat Segmenting Results

Coefficients	**Estimate**	**Std. Error**	**t value**	**Pr(>|t|)**
Intercept	57.16634	1.51210	37.806	< 2e-16
$hcwt$	-0.02487	0.06792	-0.366	0.71449
FS_1	3.15199	0.67975	4.637	5.27e-06
FS_2	1.49252	0.53102	2.811	0.00527
FS_3	0.73353	0.57840	1.268	0.20570
FS_4	-2.31224	0.72595	-3.185	0.00160
FS_5	-4.82061	0.98032	-4.917	1.44e-06

Table A35: Results of equation 12 fit for Merinos.

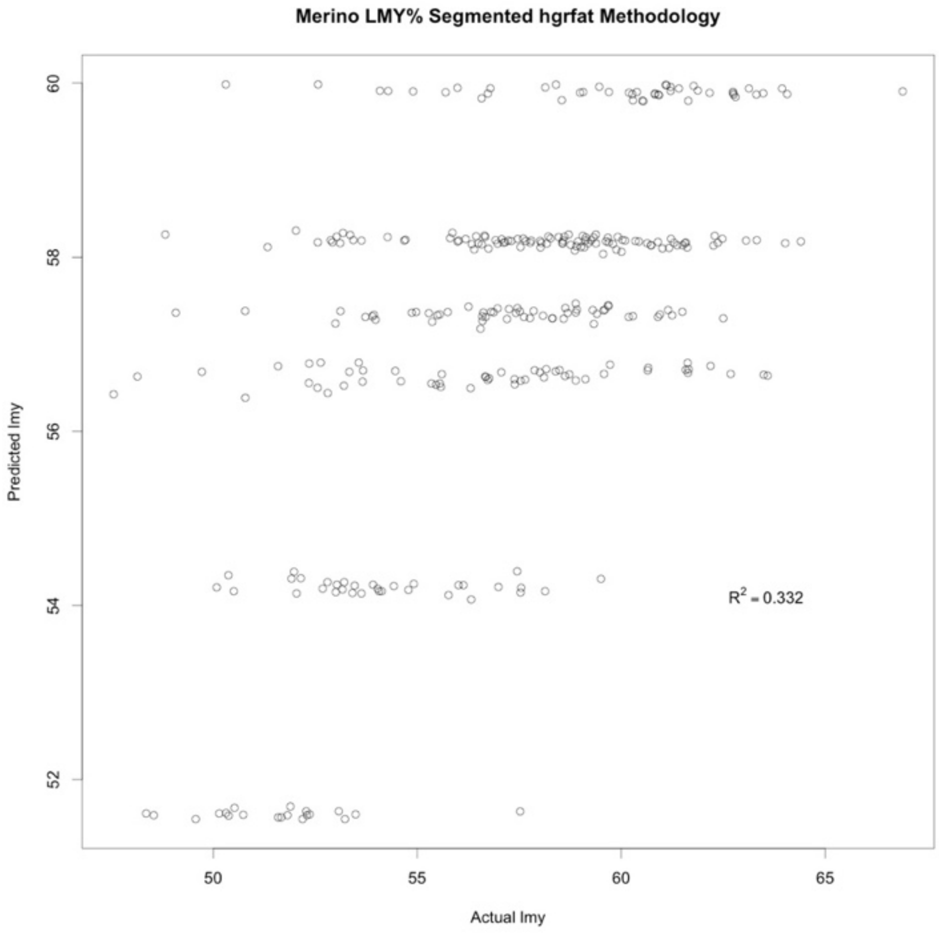


Figure A23: Scatter plot of model predicted LMY% verses actual LMY% for Merinos

The residual standard error of the fit was 3.019% on 302 degrees of freedom. The R^2 was 0.3325, and the F-statistic was 25.07 and the p-value was < 2.2e-16. In table A16, we can see that hcwt and FS_3 are not statistically different from zero at the 1% level. From figure A23, we can see that because the coefficient for hcwt is close to zero then the fat score dummy variables dominate the response and produce an undesirable segmented predicted response. This response is similar to what we have seen with the LMY% given in the factcarcasesheep table.

If we remove the linear hcwt term from equation A12, then we get a residual standard error of 3.015 on 303 degrees of freedom with multiple R^2 of 0.3322, which is very similar to the above result. However, if we don't segment the hgrfat data into five fat score values we can get a residual standard error of 2.925 and a multiple R^2 of 0.3633.

Coefficients	Estimate	Std. Error	t value	Pr(>|t|)
a	25.165830	2.520005	9.986	< 2e-16
b	-32.233053	1.044585	-30.857	< 2e-16
c	-0.042661	0.008131	-5.246	2.93e-07

FS_1	0.712144	0.138589	5.139	4.99e-07
FS_2	0.281340	0.106014	2.654	0.00838
FS_3	0.117506	0.116215	1.011	0.31277
FS_4	-0.370260	0.146160	-2.533	0.01181
FS_5	-0.710653	0.221921	-3.202	0.00151

Table A36: Results of equation A12 fit for Merinos.

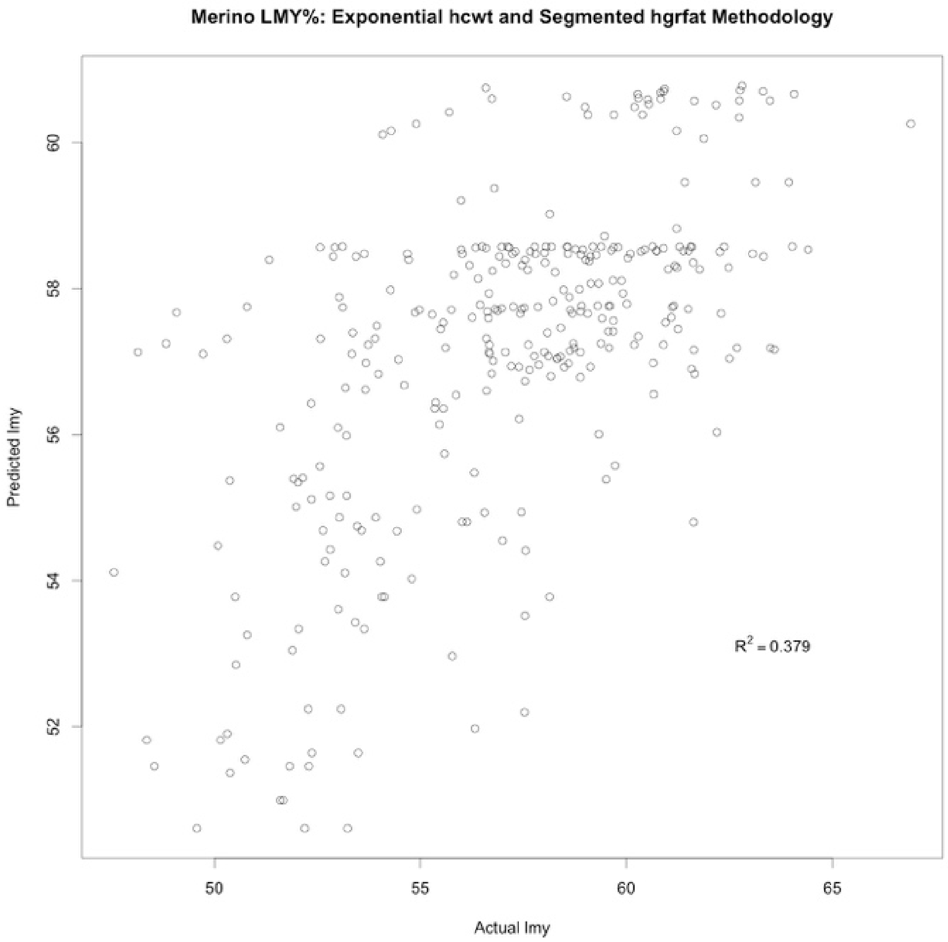


Figure A24: Scatter plot of exponential hcwt model with fatscores predicted LMY% verses actual LMY% for Merinos

The residual standard error for predicting LMY% using equation A12 was 2.916% on 301 degrees of freedom, with an R^2 of 0.3795. Removing the segmentation and using a continuous hgrfat we get an R^2 of 0.4958 and an RSE of 2.685%.

Maternal Fat Segmenting Results

Coefficients	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	60.74914	1.33783	45.409	< 2e-16
$hcwt$	-0.20847	0.05939	-3.510	0.000501
FS_1	1.24754	1.05471	1.183	0.237608

FS_2	0.74685	0.47231	1.581	0.114646
FS_3	-0.91475	0.41058	-2.228	0.026466
FS_4	-1.33466	0.49875	-2.676	0.007771
FS_5	-4.20912	0.64464	-6.529	2.1e-10

Table A37: Results of equation A12 fit for maternals.

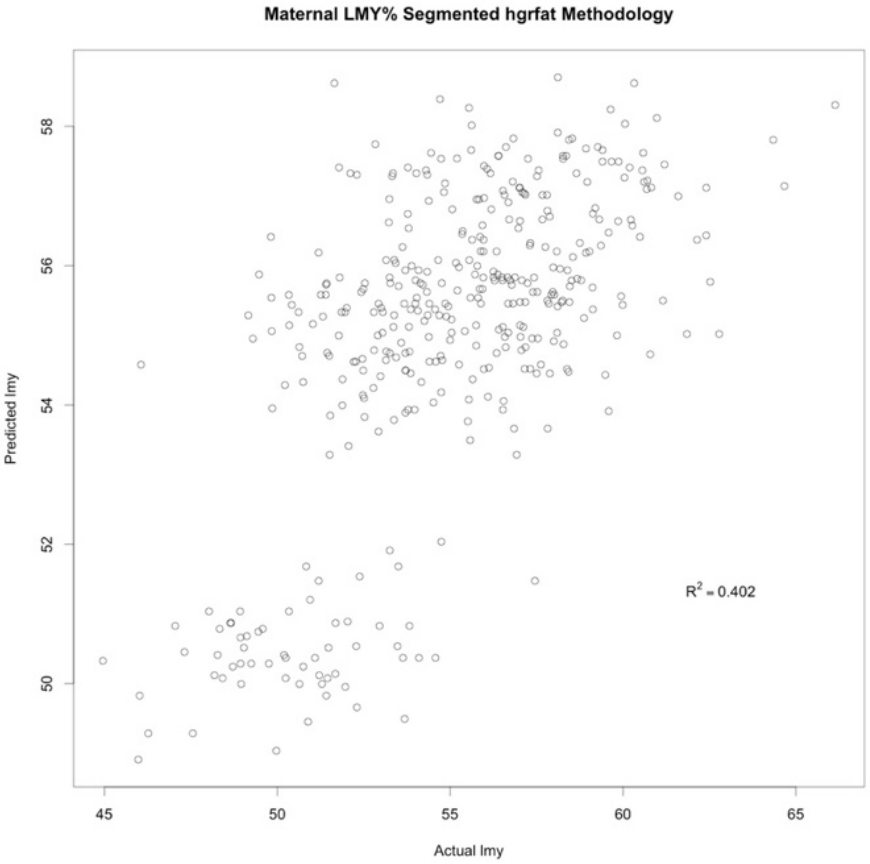


Figure A25: Scatter plot of model predicted LMY% verses actual LMY% for maternals

The residual standard error of the fit was 2.747% on 383 degrees of freedom, and the multiple R^2 was 0.4019. The F-statistic was 42.9 and the p-value was < 2.2e-16. In table A37, we can see that FS_1 and FS_2 are not statistically different from zero at the 1% level. From figure A25, we can see that there is still some segmenting of the LMY% values at least into a two moded distribution due to the hgrfat segmenting into five fat scores.

If we remove the linear hcwt term from equation 12 then we get a residual standard error of 2.787 on 384 degrees of freedom with multiple R^2 of 0.3827. However, if we don't segment the hgrfat data into five fat score values we can get a residual standard error of 2.623 and a multiple R^2 of 0.4546.

Coefficients	Estimate	Std. Error	t value	Pr(>|t|)
a	36.644098	8.673510	4.225	2.99e-05
b	-38.617410	7.262556	-5.317	1.80e-07

c	-0.021083	0.007756	-2.718	0.00686
FS_1	0.298091	0.247741	1.203	0.22963
FS_2	0.169207	0.108107	1.565	0.11837
FS_3	-0.179061	0.093236	-1.921	0.05554
FS_4	-0.277203	0.113753	-2.437	0.01527
FS_5	-1.028818	0.153633	-6.697	7.65e-11

Table A38: Results of equation A12 fit for maternals.

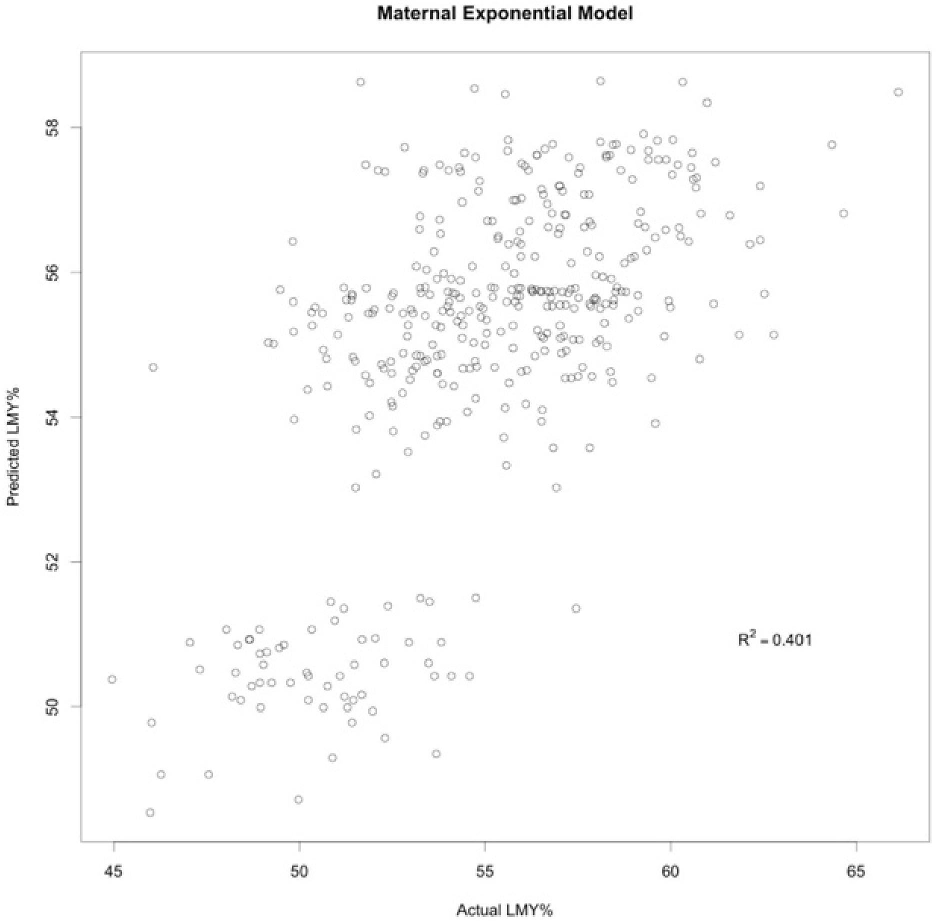


Figure A26: Scatter plot of exponential hcwt model with fat scores predicted LMY% verses actual LMY% for Merinos

The residual standard error for predicting LMY% using equation A12 was 2.753% on 382 degrees of freedom, with an R^2 of 0.4006. Removing the segmentation and using a continuous hgrfat we get an R^2 of 0.4435 and an RSE of 2.636%.

Terminal Fat Segmenting Results

Coefficients	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	63.26760	0.73945	85.560	< 2e-16
$hcwt$	-0.27490	0.03023	-9.094	< 2e-16
FS_1	3.47894	0.77721	4.476	8.45e-06

$\boldsymbol{FS_2}$	1.83005	0.31329	5.841	6.94e-09
$\boldsymbol{FS_3}$	0.24391	0.27421	0.889	0.374
$\boldsymbol{FS_4}$	-1.30102	0.29447	-4.418	1.10e-05
$\boldsymbol{FS_5}$	-2.37812	0.33602	-7.077	2.72e-12

Table A39: Results of equation 12 fit for terminals.

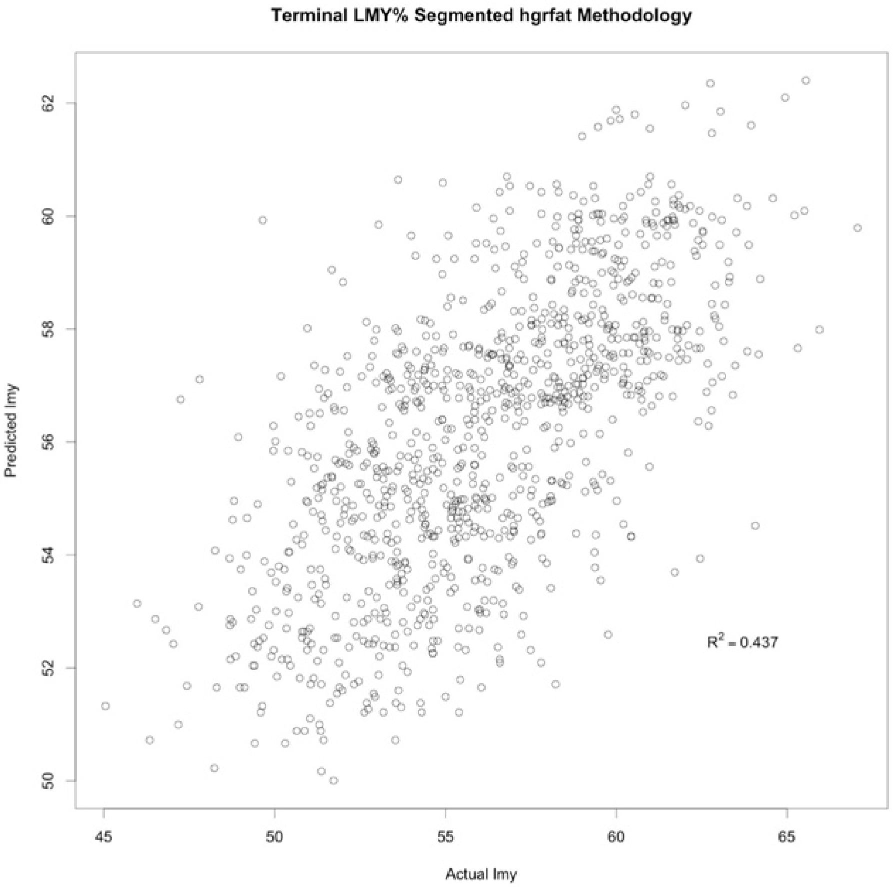


Figure A27: Scatter plot of model predicted LMY% verses actual LMY% for terminals

The residual standard error of the fit was 2.843% on 1026 degrees of freedom. The multiple R^2 was 0.437, the F-statistic was 132.8 and the p-value was < 2.2e-16. In table A39, we can see that FS_3 is not statistically different from zero at the 1% level. However, if we don't segment the hgrfat data into five fat score values we can get a residual standard error of 2.697 and a multiple R^2 of 0.4911.

Coefficients	**Estimate**	**Std. Error**	**t value**	**Pr(>|t|)**

a	66.631386	22.087627	3.017	0.002619
b	-66.372953	21.381167	-3.104	0.001960
c	-0.009316	0.003934	-2.368	0.018048
FS_1	0.620536	0.195697	3.171	0.001565
FS_2	0.390271	0.078995	4.940	9.10e-07
FS_3	0.093492	0.067878	1.377	0.168701
FS_4	-0.275166	0.073915	-3.723	0.000208
FS_5	-0.643609	0.083382	-7.719	2.78e-14

Table A40: Results of equation A12 fit for terminals.

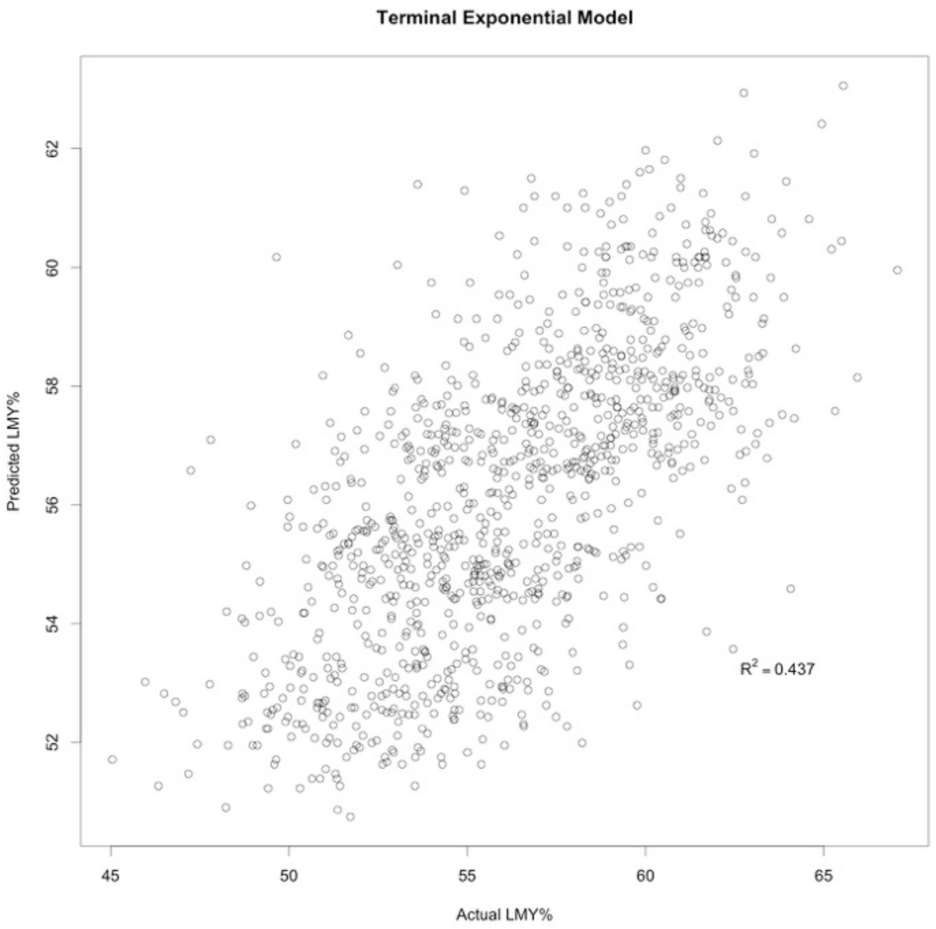


Figure 28: A Scatter plot of exponential hcwt model with fatscores predicted LMY% verses actual LMY% for Terminals

The residual standard error for predicting LMY% using equation 12 was 2.8438% on 1025 degrees of freedom, with an R^2 of 0.4372. Removing the segmentation and using a continuous hgrfat we get an R^2 of 0.4958 and an RSE of 2.685%.

Conclusion

From the investigation in this paper we can see that for Merino sheep the best lean meat yield percentage equation is:

$$LMY\%_{Merino} = \frac{a + b.e^{c.hcwt} + d.hgrfat}{hcwt}$$

where a = 30.069, b = -36.305, c = -0.0364 and d = -0.1023.

For maternal lean meat yield percentage prediction, the simplest and best equation is:

$$LMY\%_{Maternal} = a + b.hgrfat$$

where a = 62.65 and b = -0.4280.

Whereas for terminals, the best LMY% prediction equation is of the same form as for Merinos where the coefficients are a = 72.015, b = -71.974, c = -0.0096 and d = -0.0979.

It is possible to use the equation:

$$LMY\%_{Maternal} = a + b.hgrfat$$

for terminals without much reduction in R^2, in this case for terminals a = 62.646 and b = -0.4280. The segmentation of hgrfat into fat scores typically degrades the R^2 on average by 16.32%.

Overall, despite extensive statistical modelling using research data from the Sheep CRC, there appeared to be only minor improvements in prediction of LMY% by moving to a more complex prediction algorithm in Merino and no improvements in maternal or terminal over the base use of a linear model. However, a key point is that currently LDL does not collect breed information and therefore the application of models that discriminate due to breed is limited. Furthermore, these analyses reveal that converting GR fat measured in mm to a GR fat score significantly reduces the amount of variation in LMY% that is accounted for by the model as well as prediction accuracy.