Kansas State University Libraries

New Prairie Press

DEVELOPING PREDICTION EQUATIONS FOR CARCASS LEAN MASS IN THE PRESCENCE OF PROPORTIONAL MEASUREMENT ERROR

Zachary J. Hass
Purdue University, zhass@purdue.edu

Ziqi Zhou
George Washington University

Bruce A. Craig
Purdue University

Follow this and additional works at: https://newprairiepress.org/agstatconference

🟠 Part of the Agricultural Economics Commons, Applied Statistics Commons, and the Meat Science Commons

Recommended Citation

DEVELOPING PREDICTION EQUATIONS FOR CARCASS LEAN MASS IN THE PRESCENCE OF PROPORTIONAL MEASUREMENT ERROR

Zachary J Hass[1], Ziqi Zhou[2], and Bruce A Craig[1]

[1] Department of Statistics
Purdue University
250 N. University Street
West Lafayette, IN 47907

[2] Department of Statistics
George Washington University
801 22nd Street NW
Washington, DC 20052

Abstract: Published prediction equations for carcass lean mass are widely used by commercial pork producers for carcass valuation. These regression equations have been derived under the assumption that the predictors, such as back fat depth, are measured without error. In practice, however, it is known that these measurements are imperfect, with a variance that is proportional to the mean. In this paper, we consider both a linear and quadratic true relationship and compare regression fits among two methods that account for this error versus simply ignoring the additional error. We show that biased estimates of the relationship result if measurement error is ignored. Between our version of regression calibration and a Bayesian model approach, the Bayesian inference approach produced the least biased predictions. The benefits of our Bayesian approach also increased with an increase in the measurement error.

INTRODUCTION

Fat free lean content is often used to value pork carcasses in order to incentivize production of leaner pigs.[1,5,8] Due to the impracticality of measuring fat free lean directly, indirect measures such as back fat depth and muscle depth are taken and used in published equations to predict fat free lean.[2,5,8] These prediction equations, however, are known to be biased, such that fat free lean is often under predicted for lean pigs and over predicted for fatter pigs.[2]

One possible reason for this bias is due to the lack of any adjustment for measurement error. Various methods exist for taking measurements of back fat depth including using a ruler, calipers, ultrasound, and optical probe.[5] Operator error associated with the use of the optical probe can lead to measurement error of the back fat depth value that is proportional to the true value.[1,4,8] Previous research has shown that the presence of measurement error with variance proportional to the mean greatly increases the probability of a false positive quadratic term.[8] In addition, the residual standard deviation was inflated. This is expected as measurement error increases the variability in the data.[10]

Because the published prediction equations based on optical probe data have included a quadratic relationship between back fat depth and fat free lean content, a logical question is whether this quadratic term is truly a false positive. To help answer this, we use a simulation study to compare two different methods to adjust for measurement error, relative to doing nothing, under both a true linear and quadratic setting and see how often the true model is selected. We also see how well these methods reduce the estimated residual standard error. We consider a range of simulation parameters that span the values commonly found in the literature.

A natural approach to eliminate the effects of measurement error would be to eliminate, or at least minimize, the degree of measurement error through replicate measurements of the same pig (i.e., using an average back fat depth measurement). While this would certainly remedy the problem, we feel the additional measurements would be overly burdensome and costly. The two adjustment methods we consider do not require any additional measurements.

The remainder of the paper is organized as follows. First we give a very brief overview of the measurement error problem and common methods used to adjust for it. We then describe our simulation study followed by a detailed description of our two adjustment methods. This is followed by the results of the study, which are summarized in both figures and tables. We conclude with a brief discussion.

MEASUREMENT ERROR

Measurement error refers to the situation when a predictor X is a quantity measured with random error. If this measurement error is a relatively large component of the total variability of X, it can bias regression parameters. Although adjustment for measurement error in regression has been widely studied, the primary focus has been on errors with a constant variance over the range of X. Our focus here is on multiplicative measurement error, or error variance that is proportional to the mean.

For review, consider the most basic measurement error case. Consider $X \sim N(\mu, \sigma_X^2)$ and $Y = \beta_0 + \beta_1 X + \varepsilon$ where $\varepsilon \sim N(0, \sigma^2)$. Instead of observing X, one observes $X^* \sim N(X, \delta^2)$. If the measurement error is ignored and least squares is performed using the observed X^* and Y, then

$$E(b_1) = \frac{Cov(X^*, Y)}{Var(X^*)} = \beta \frac{\sigma_X^2}{\sigma_X^2 + \delta^2}$$

Thus the slope is shrunk towards 0 based on the relative amount of variance in X versus the variance in X^*. This ratio $R = \frac{\sigma_X^2}{\sigma_X^2 + \delta^2}$ is called the reliability ratio.

The techniques to remedy such ill effects require some knowledge of the amount of measurement error present. For example, methods such as SIMEX or regression calibration, require an estimate of the measurement error variance.[10]

For regression calibration, one adjusts X^* prior to performing the least squares fit. It is common to use the $E(X \mid X^*)$ as the predictor. Because both $X^* \mid X$ and X are normally distributed in the case above, we can show that

$$E(X \mid X^*) = \mu_x + R(X^* - \mu_X)$$

Thus, if the reliability ratio is high, the new predicted value is close to the observed X^*. If the reliability ratio is low, the new predicted value is shrunk towards the mean. One can use the estimate of measurement error variance and the observed X^* to obtain estimates of μ_X and σ_X^2.

When there is measurement error proportional to the mean of the explanatory variable, $Var(X^* \mid X = KX^2)$, this calibration method is not directly applicable. While one might anticipate knowing the proportionality constant K, the reliability ratio R is still unknown because one does not know the true X.

Throughout this paper we will assume that K, the proportionality constant, is known. Knowledge of K could reasonably come from prior studies with multiple measurements on pigs with different values of actual back fat depth (ABFD). We utilize knowledge of K to augment the regression calibration and also consider a Bayesian approach to this problem. We investigate both methods via a simulation study and compare results to those obtained using a least squares fit ignoring the measurement error.

METHODS

One thousand data sets were simulated for each of 18 (2 x 3 x 3) scenarios given in Table 1. Previous studies have used between 50-627 pigs for single studies and have used as many as 1024 pigs when estimating prediction equations across multiple studies. The 250 pigs represents a medium single study, 500 a large single study, and 1000 pigs represents combining data from multiple studies.

Table 1: Simulation Scenarios

Factor	Levels
Relationship of FFL and ABFD	Linear, Quadratic
Constant of Proportionality (K)	0.01, 0.03, 0.06
Number of Pigs	250, 500, 1000

ABFD values were sampled from a Normal distribution, $\mu = 28$, $\sigma = 8$, left truncated at zero to avoid unrealistic values.[8] Measurement error was simulated by drawing a BFD for each ABFD from a Gamma distribution with $\alpha = \frac{1}{K}$ and $\beta = K * ABFD$ so that values would be unbiased for ABFD, non-negative, with variance proportional to ABFD. Proportionality constants K = 0.01, 0.03, and 0.06 result in correlations of approximately 0.94, 0.85, and 0.75 respectively. These correlations are contained within the range of those found in previous research and can be thought of as very well-tuned, moderately well-tuned, and poorly-tuned measurement processes.[8]

FFL was generated using either:

$$FFL = 56.2 - 0.4 * ABFD + \epsilon; \quad \epsilon \sim N(0, 2.8467) \qquad (1)$$

$$FFL = 54.46 - 0.543 * ABFD + 0.006 * ABFD^2 + \epsilon \qquad (2)$$

These coefficients were derived from previous research.[8] For the quadratic model, coefficients were derived by substituting the average values for other model variables and collapsing their effects into the intercept.[2]

Each data set was analyzed using each of our three approaches. For the ordinary least squares estimation (i.e., ignoring the measurement error), a t-test was performed for the quadratic term. This was done to determine the percent of simulations that correctly rejected the quadratic when the true relationship was linear and what percent of simulations correctly kept the quadratic term when the true relationship was quadratic. Additionally, the ratio of the mean squared error to the true regression variance (σ_ϵ^2) was calculated as a measure of variance inflation due to the measurement error. In order to investigate coefficient bias and variability, we calculated the model residual for each estimated equation at the 2.5th, 5th, 10th, 50th, 90th, 95th, and 97.5th

percentiles of the prior distribution on ABFD. This was done across all 18 scenarios and summarized using side-by-side box plots.

As noted previously, regression calibration refers to a collection of techniques that replaces the predictor value measured with error by an estimate of the true value of the predictor. Once the substitution is made, analysis proceeds on the new data set as it would on data without measurement error.[10] The goal here is to remove some of the variability caused by measurement error by replacing each value of BFD with the conditional expectation of ABFD|BFD. Given the way the data were generated, the conditional distribution of ABFD|BFD is proportional to the product of a Gamma and a Normal distribution. We use method of moment estimators (3) and (4) for the variance and mean of ABFD respectively. Numerical integration was then used to determine the estimated mean for the conditional distribution given in (5).

$$\hat{\sigma}^2_{ABFD} = \sqrt{\frac{\sigma^2_{BFD} - K\mu^2_{BFD}}{K+1}} \tag{3}$$

$$\hat{\mu}_{ABFD} = \overline{BFD} \tag{4}$$

$$\pi(ABFD|BFD) \propto Gamma\left(\frac{1}{K}, BFD * K\right) N(\hat{\mu}_{ABFD}, \hat{\sigma}^2_{ABFD}) \tag{5}$$

Since least-squares regression was used to fit these calibrated data, the same statistics (i.e., t-test for the quadratic term, the measure of variance inflation, and prediction box plots) were recorded for each data set.

Lastly, we considered a Bayesian inference approach assuming the distributions used in the data generation. A Gibbs Sampler with a Metropolis-Hastings step was used to sample from the posterior distribution given in (6). We placed an improper prior on the regression parameters $\pi(\beta, \sigma) \propto \frac{1}{\sigma}$ in order to "mimic" least squares regression.[12] Placing a Normal prior on ABFD yields the conditional distributions for the three random quantities given in (7), (8), and (9).

$$\pi(ABFD, \beta, \sigma^2 | BFD, FFL, K) \tag{6}$$

$$\pi(\beta | ABFD_t, \sigma^2, FFL) \sim MVN_p\left((X^TX)^{-1}X^TY, \ \sigma (X^TX)^{-1}\right) \tag{7}$$

$$\pi(\sigma^2 | ABFD_t, \beta, FFL) \sim RSS * \frac{1}{\chi^2(NPig)} \tag{8}$$

$$\pi(ABFD_{t+1} | \beta, \sigma^2, \text{BFD}, FFL) \propto N(\mu_{ABFD}, \sigma_{ABFD})$$
$$* N(\beta_0 + \beta_1 ABFD_t, \sigma_\epsilon) * Gamma(\frac{1}{K}, K * ABFD_t) \tag{9}$$

Here X is the design matrix containing the current estimate of ABFD, Y is a vector containing FFL, RSS is the residual sum of squares, and NPig is the number of pigs in the data set. The first

two full conditional distributions, the multivariate normal and the scaled inverse chi-square can be sampled from directly making a Gibbs update possible. The last distribution is not straightforward and so it is updated using a Metropolis-Hastings update using a Normal random walk proposal. For all datasets 80,000 iterations of the Markov Chain were run with a 30,000 iteration burn-in. In both relationship settings, a quadratic model was considered. The posterior means of the regression coefficients and regression variance were considered the parameter estimates.

The ratio of the estimated variance to the true regression variance was used to measure variance inflation as before. In order to assess model selection, we calculated the 95% credible interval for the coefficient on the quadratic term and determined whether 0 fell in the interval. Prediction bias and variability was assessed with box plots as with both regression fits.

RESULTS

Tables 2, 3, and 4 summarize the least squares fit, regression calibration, and Bayesian approach, respectively, for the linear data. Tables 5, 6, and 7summarize these approaches for the quadratic data. The first two columns of these tables give the constant of proportionality and the number of pigs in the simulated data set.

For the approaches that used least squares, columns 3-5 give the average coefficients across all simulations where the quadratic term was found to be significant and columns 7-8 give the same information for models where it was not significant. The percent of simulations where the quadratic term was retained in the model is given in column 6. The ratio of the selected model's mean squared error to the true variance is averaged across datasets and is given in column 9.

For the Bayesian approach, only the quadratic model was fit so columns 3-5 give the average of estimated coefficients over all data sets. Column 6 gives the number of simulations failing to reject the quadratic term. The last column gives the ratio of the posterior mean of the model variance and the true variance.

The least squares results agree with previous research, higher acceptance of the spurious quadratic term and variance inflation (running from approximately 14-52%).[8] When the relationship is truly quadratic and correlation between BFD and ABFD is high (~94%), the correct model is chosen most of the time and variance inflation is only roughly 3%. This is the best case for ignoring measurement error, but we still see bias in the estimated coefficients.

For the regression calibration approach, the number of simulations finding a spurious quadratic drops much closer to the desired Type I error of 5%, but the same level of variance inflation is still present. In the quadratic case, the regression calibration adjusted data performs worse in correctly identifying the true relationship, particularly in the smaller sample size of 250

pigs. For both relationships, when the correct model is chosen, the coefficients are very nearly unbiased.

The Bayesian approach resulted in estimates that are very nearly unbiased for the true coefficients even though a quadratic model was fit to the linear data sets. The method also chooses the correct model when the relationship is linear in approximately 95% of simulations regardless of the amount of correlation between BFD and ABFD or the number of pigs. Additionally, the variance inflation has been removed. In the quadratic case, we still see unbiased coefficients and removal of variance inflation, but the ability to correctly infer about the quadratic term is not quite as strong as the least squares fit when the number of pigs is small (250).

Figure 1 graphically displays the comparison between least squares fit, regression calibration fit, and the Bayesian fit in terms of ability to reject the spurious quadratic term. Figure 2 compares the three methods for removing the variance inflation for the linear relationship.

Prediction bias and variability across the simulations were evaluated visually using box plots at each of seven percentiles from the Normal distribution assumption on ABFD. Three of these are given in Figure 3, one for each of the model estimation methods when K is 0.01, number of pigs is 250, and the data is linear. The least squares fit shows what is expected when a quadratic relationship is fit to linear data, the under prediction of FFL for lean pigs and over prediction for fat pigs. The regression calibration adjusted least squares fit shows a slight bias for under prediction which grows worse as K increases (correlation between ABFD and BFD declines). The estimates from the Bayesian approach were very similar to those obtained from the regression calibration, but bias did not increase as K increased. Across scenarios, the Bayesian approach gave the least biased predictions, but at the cost of higher variability, particularly at lower number of pigs (250).

In summary, ignoring measurement error and simply using least squares regression will result in biased coefficients and predictions as well as being more likely to select a quadratic model whether the true relationship supports it or not. Applying a regression calibration adjustment first will discover the true linear model as often as type I error dictates, but requires larger sample sizes to find a true quadratic. When the right relationship is found, coefficient estimates are very nearly unbiased. Bayesian inference produces nearly unbiased coefficients even when fitting the quadratic relationship to linear data. Estimates are slightly less biased than those from regression calibration, Figure 4 gives a graphical representation of the typical behavior of the three fits on linear data. Notably, Bayesian inference is the only method of the three to recover the true model variance.

DISCUSSION

Although the results are very encouraging in terms of the possibility of correcting the prediction bias caused by measurement error, the methods employed here represent an

oversimplification of reality. While correlation level between BFD and ABFD may be a realistically obtainable quantity, how well the methods hold up if K is only roughly known was not studied. Future work will verify how robust these adjustment approaches are to an unknown K.

Additionally, the majority of prediction equations for FFL contain other predictors besides BFD, such as muscle depth and carcass weight. Since muscle depth is also measured with an optical probe, it too is likely to have measurement error. Future work will extend our approaches to including additional covariates as well as handling two predictors measured with error. We also think further work should be done to determine exactly how many pigs are needed for the method to perform well.

Although it appears the Bayesian approach is the better of the two adjustment methods, it did well in part because the data were generated using the model considered for inference. In practice, this will rarely be the case. As a result, we need to investigate its robustness to data generated under alternative models. Similarly the calculation of the E(ABFD | BFD) in regression calibration also depends on the correctness of the distributional assumptions.

In this simplified version of the problem, using MCMC sampling techniques produced prediction equations that gave very nearly unbiased predictions whose variability was comparable to the least squares fit, especially as sample size increased. The posterior mean for the regression model residual variance was also unbiased for the true variance. These positive results are not unexpected, but do indicate that the extension of these ideas to the more realistic scenario is a worthwhile venture in order to produce unbiased prediction equations for FFL in the presence of measurement error.

TABLES AND FIGURES

Table 2: Least Squares Fit on Linear Data: $FFLM = 56.2 - 0.4 * ABFT$

Simulation Parameters		Average Coefficients across all simulations where Quadratic model was chosen.			Percent of Sims Accepting Quadratic	Average Coefficients across all simulations were Linear model was chosen.		Average ratio of model MSE over true model variance
K	NPig	β_0	β_1	β_2	$\% \beta_2 \neq 0$	β_0	β_1	$\dfrac{MSE}{\sigma_\epsilon^2}$
0.01	250	58.46	-0.6241	0.0047	24.4%	54.93	-0.3546	1.139
0.01	500	57.58	-0.5592	0.0036	44.5%	54.92	-0.3547	1.141
0.01	1000	56.95	-0.5107	0.0027	72.6%	54.90	-0.3537	1.140
0.03	250	56.68	-0.5594	0.0046	61.5%	53.06	-0.2885	1.326
0.03	500	56.15	-0.5183	0.0039	87.0%	53.06	-0.2878	1.337
0.03	1000	55.91	-0.5010	0.0036	99.6%	53.02	-0.2854	1.331
0.06	250	54.67	-0.4729	0.0040	70.5%	51.27	-0.2237	1.515
0.06	500	54.24	-0.4402	0.0034	95.1%	51.26	-0.2231	1.517
0.06	1000	54.10	-0.4304	0.0033	99.7%	51.09	-0.2175	1.516

Table 3: Regression Calibration Adjusted Least Squares Fit on Linear Data: $FFLM = 56.2 - 0.4 * ABFT$

Simulation Parameters		Average Coefficients across all simulations where Quadratic model was chosen.			Percent of Sims Accepting Quadratic	Average Coefficients across all simulations were Linear model was chosen.		Average ratio of model MSE over true model variance
K	NPig	β_0	β_1	β_2	$\% \beta_2 \neq 0$	β_0	β_1	$\dfrac{MSE}{\sigma_\epsilon^2}$
0.01	250	56.59	-0.4344	0.0007	5.2%	56.25	-0.4014	1.137
0.01	500	57.05	-0.4654	0.0012	5.5%	56.21	-0.4004	1.139
0.01	1000	56.30	-0.4085	0.0002	4.5%	56.20	-0.4001	1.139
0.03	250	57.66	-0.5160	0.0020	3.5%	56.26	-0.4020	1.322
0.03	500	55.99	-0.3816	-0.0004	5.9%	56.25	-0.4015	1.334
0.03	1000	56.29	-0.4116	0.0003	4.2%	56.22	-0.4010	1.329
0.06	250	54.57	-0.2576	-0.0029	5.0%	56.33	-0.4046	1.506
0.06	500	56.44	-0.4151	0.0002	4.5%	56.29	-0.4030	1.510
0.06	1000	55.60	-0.3554	-0.0008	3.9%	56.23	-0.4010	1.509

Table 4: Reg Calibration Adjusted Least Squares Fit on Quadratic Data: $FFLM = 54.46 - 0.543 * ABFT + 0.006 * ABFT^2$

Simulation Parameters		Average Coefficients across all simulations where Quadratic model was chosen.			Percent of Sims Accepting Quadratic	Average Coefficients across all simulations were Linear model was chosen.		Average ratio of model MSE over true model variance
K	NPig	β_0	β_1	β_2	% $\beta_2 \neq 0$	β_0	β_1	$\dfrac{MSE}{\sigma_\epsilon^2}$
0.01	250	55.17	-0.5975	0.0070	70.9%	50.32	-0.2131	1.026
0.01	500	54.59	-0.5523	0.0062	94.2%	50.31	-0.2141	1.031
0.01	1000	54.44	-0.5402	0.0060	100.0%	0	0	1.031
0.03	250	55.78	-0.6437	0.0079	54.9%	50.46	-0.2186	1.073
0.03	500	54.80	-0.5641	0.0064	86.0%	50.40	-0.2177	1.075
0.03	1000	54.43	-0.5365	0.0059	98.7%	50.23	-0.2113	1.076
0.06	250	56.55	-0.7008	0.0090	38.6%	50.54	-0.2222	1.130
0.06	500	55.32	-0.6024	0.0072	66.8%	50.47	-0.2186	1.127
0.06	1000	54.64	-0.5459	0.0061	92.4%	50.46	-0.2186	1.130

Table 5: Least Squares Fit on Quadratic Data: $FFLM = 54.46 - 0.543 * ABFT + 0.006 * ABFT^2$

Simulation Parameters		Average Coefficients across all simulations where Quadratic model was chosen.			Percent of Sims Accepting Quadratic	Average Coefficients across all simulations were Linear model was chosen.		Average ratio of model MSE over true model variance
K	NPig	β_0	β_1	β_2	% $\beta_2 \neq 0$	β_0	β_1	$\dfrac{MSE}{\sigma_\epsilon^2}$
0.01	250	54.06	-0.5322	0.0061	89.2%	49.43	-0.1824	1.026
0.01	500	53.78	-0.5120	0.0057	99.6%	49.02	-0.1679	1.032
0.01	1000	53.75	-0.5102	0.0057	100.0%	0	0	1.031
0.03	250	52.58	-0.4527	0.0051	89.9%	48.40	-0.1449	1.077
0.03	500	52.34	-0.4335	0.0047	99.8%	48.79	-0.1569	1.079
0.03	1000	52.26	-0.4286	0.0047	100.0%	0	0	1.081
0.06	250	50.81	-0.3526	0.0038	83.7%	47.42	-0.1111	1.139
0.06	500	50.52	-0.3303	0.0034	98.8%	47.46	-0.1099	1.137
0.06	1000	50.49	-0.3275	0.0034	100.0%	0	0	1.142

Table 6: MCMC Quadratic Fit on Linear Data: $FFLM = 56.4 - 0.4 * ABFT$

Simulation Parameters		Average Coefficients: Posterior means averaged across all simulations.			Average Posterior Probability	Percentage of Simulations Accepting β_2	Average ratio of model MSE over true model variance
K	NPig	β_0	β_1	β_2	$P(\beta_2 = 0)$	Accepting β_2	$\dfrac{\widehat{\sigma}^2}{\sigma_\epsilon^2}$
0.01	250	56.21	-0.40	-0.0001	24.9%	4.9%	1.007
0.01	500	56.28	-0.41	0.0001	24.7%	5.8%	1.005
0.01	1000	56.19	-0.40	0.0000	25.8%	4.8%	1.002
0.03	250	56.27	-0.40	0.0000	25.0%	4.4%	0.997
0.03	500	56.22	-0.40	-0.0001	24.8%	6.6%	1.005
0.03	1000	56.16	-0.40	-0.0001	25.1%	5.2%	0.999
0.06	250	56.37	-0.40	-0.0001	24.7%	6.3%	0.989
0.06	500	56.30	-0.40	0.0000	25.0%	4.7%	0.994
0.06	1000	56.24	-0.40	0.0000	25.2%	4.8%	0.998

Table 7: MCMC Quadratic Fit on Quadratic Data: $FFLM = 54.46 - 0.543 * ABFT + 0.006 * ABFT^2$

Simulation Parameters		Average Coefficients: Posterior means averaged across all simulations.			Average Posterior Probability	Percentage of Simulations Accepting β_2	Average ratio of model MSE over true model variance
K	NPig	β_0	β_1	β_2	$P(\beta_2 = 0)$	Accepting β_2	$\dfrac{\widehat{\sigma}^2}{\sigma_\epsilon^2}$
0.01	250	54.46	-0.54	0.0059	3.6%	71.7%	1.002
0.01	500	54.48	-0.54	0.0060	0.5%	94.9%	1.004
0.01	1000	54.45	-0.54	0.0060	0.0%	100.0%	1.002
0.03	250	54.59	-0.55	0.0061	6.2%	57.6%	0.994
0.03	500	54.55	-0.55	0.0061	1.3%	87.7%	0.995
0.03	1000	54.45	-0.54	0.0060	0.1%	98.7%	0.998
0.06	250	54.65	-0.56	0.0062	8.6%	47.5%	0.989
0.06	500	54.56	-0.55	0.0061	2.8%	77.2%	0.992
0.06	1000	54.52	-0.55	0.0060	0.3%	97.3%	0.996

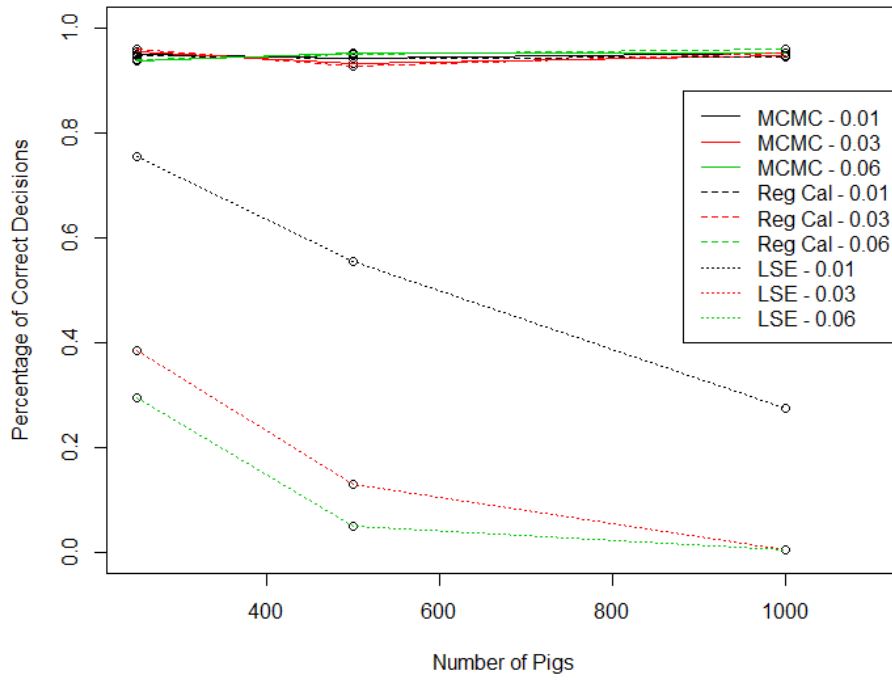**Figure 1: Correctly Removing Spurious Quadratic Terms (MCMC vs Reg Cal vs LSE)**

Figure 2: Removing Variance Inflation for Linear Data (MCMC vs Reg Cal vs LSE)
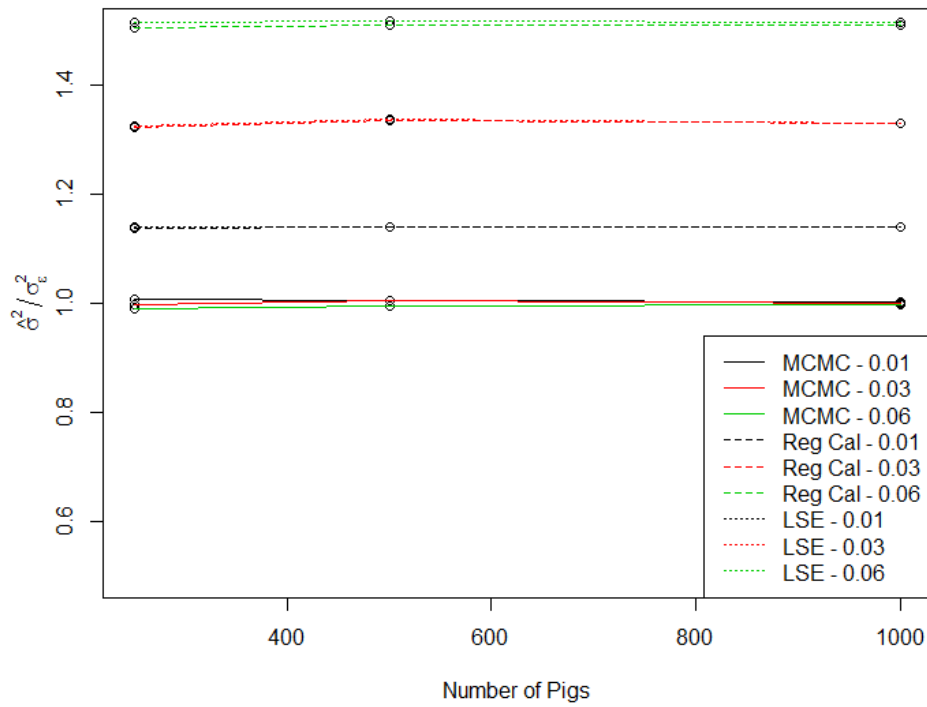
Figure 3: Boxplots of FFL Prediction Bias (K = 0.01, Number of Pigs = 250)

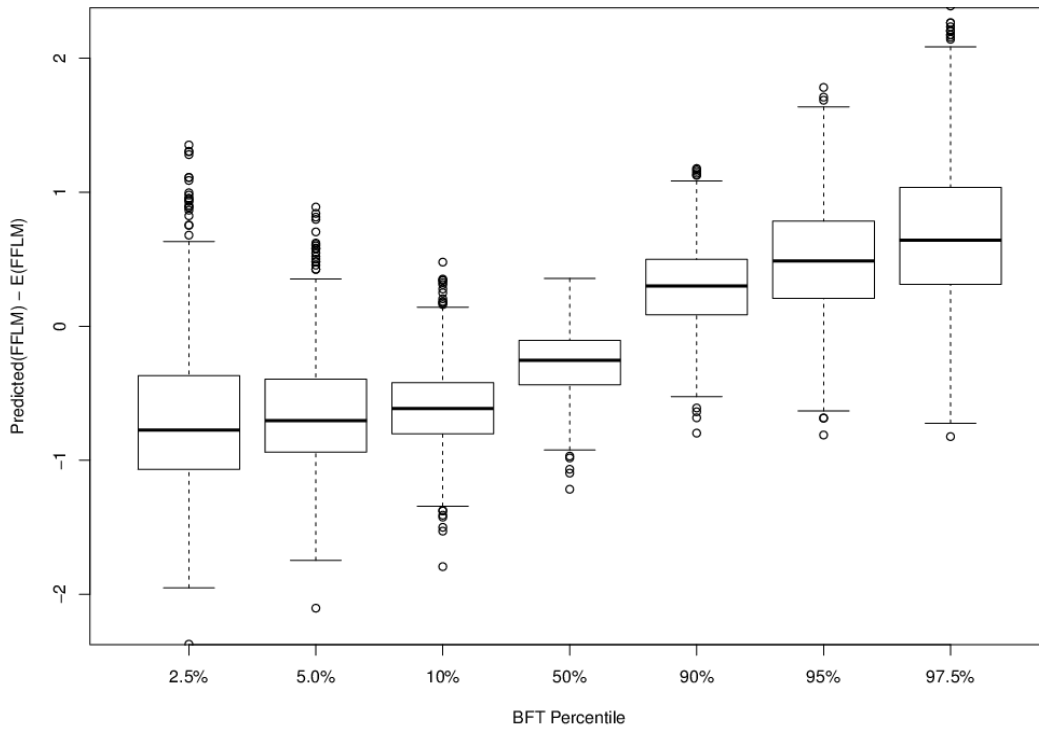**Figure 3 A-Least Squares Fit**

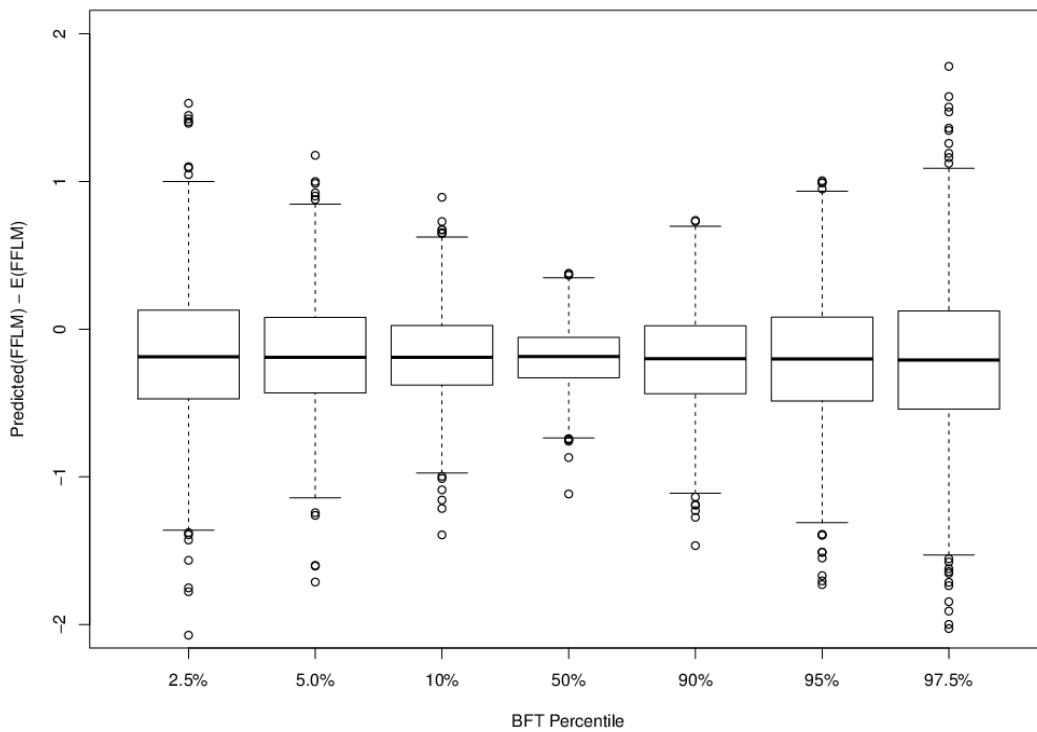**Figure 3 B – Regression Calibration Adjusted Least Squares Fit**
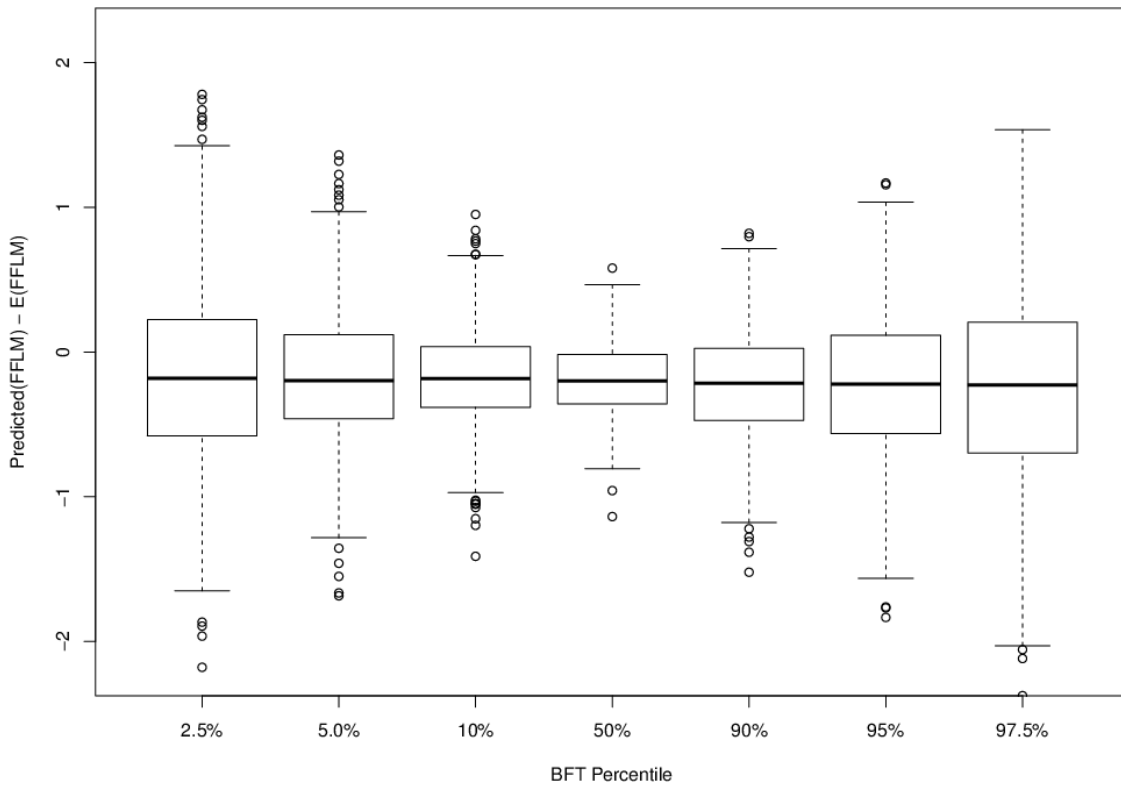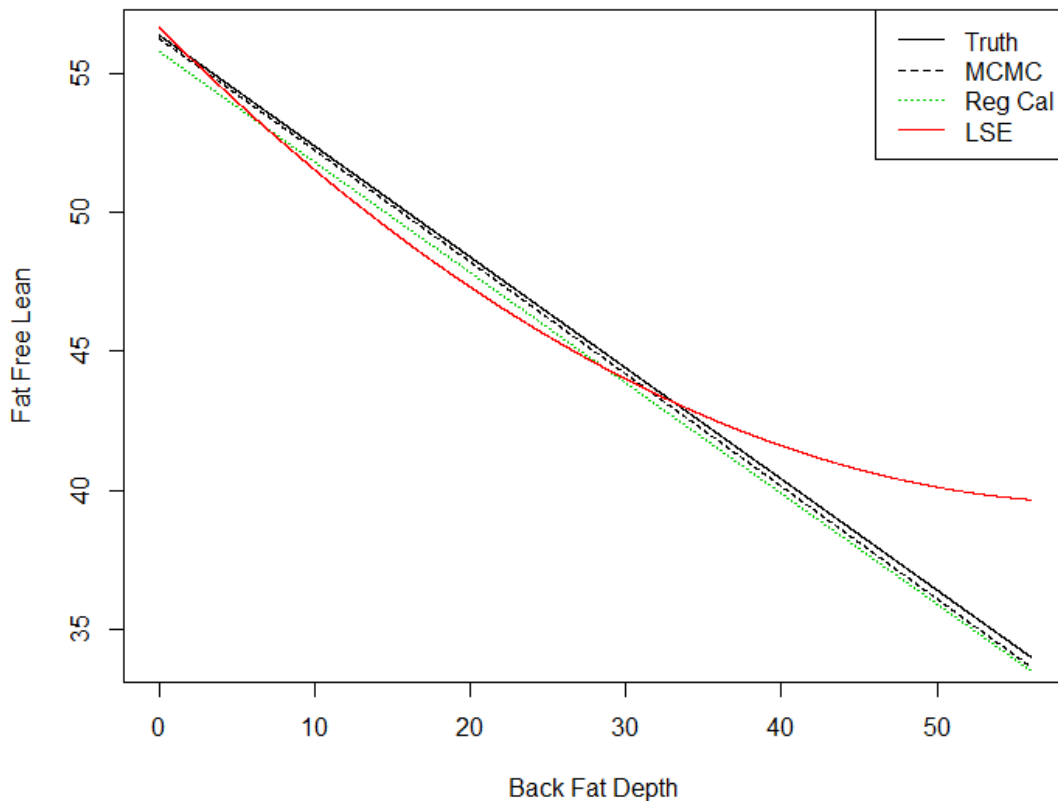
Figure 3 C – MCMC Posterior Fit

Figure 4 – Comparison of Fit Relationship for Linear Data, NPigs=250, K=0.03

References

[1] Boland, B.A., E.P. Berg, J.T. Akridge, and J.C. Forrest. "The Impact of Operator Error Using Optical Probes to Estimate Pork Carcass Value." Review of Agricultural Economics, Vol. 17, No. 2 (May, 1995), pp 193-204.

[2] Johnson, R.K., E.P. Berg, R. Goodwin, J.W. Mabry, R.K. Miller, O.W. Robison, H. Sellers, and M.D. Tokach. "Evaluation of procedures to predict fat-free lean in swine carcasses." Journal of Animal Science, 2004, Vol. 82, No. 8, pp 2428-2441.

[3] Johnson, R.K., E.P. Berg, R. Goodwin, J.W. Mabry, R.K. Miller, O.W. Robison, H. Sellers, and M.D. Tokach. "Authors' response to the Letter to the Editor by Schinckel." Journal of Animal Science, 2005, Vol. 83, pp 2721-2722.

[4] Jones, S.D.M., W.M. Robertson, and T. Coupland. "The Effect of Hog Carcass Rib Number, Probe Distance from the Chine Bone and Probe Orientation on Predicted Lean Yield and Carcass Index." Unpublished manuscript, Agriculture Canada, 1993.

[5] Olsen, E.V., M. Candek-Potokar, M. Oksama, S. Kien, D. Lisiak, and H. Busk. "On-line measurements in pig carcass classification: Repeatability and variation caused by the operator and the copy of instrument." Meat Science, 2007, Vol. 75, No 1, pp 29-38.

[6] Schinckel, A.P., J.R. Wagner, J.C. Forest, and M.E. Einstein. "Evaluation of the prediction of alternative measures of pork carcass composition by three optical probes." Journal of Animal Science, 2001, Vol. 79, No. 5, pp 1093-1119.

[7] Schinckel, A.P. "Critique of 'Evaluation of procedures to predict fat-free lean in swine carcass'". Journal of Animal Science, 2005, Vol. 83, pp 2719-2720.

[8] Schinckel, A.P., M.E. Einstein, K. Foster, and B.A. Craig. "Evaluation of the Impact of Errors in the Measurement of Backfat Depth on the Prediction of Fat-Free Lean Mass." Journal of Animal Science, 2007, Vol. 85, No. 8, pp 2031-2042.

[9] Schinckel, A.P. and C.P. Rusk. "The Need for Accurate Prediction Equations for the Carcass Lean Content of Pigs." Journal of Extension, 2012, Vol. 50, No. 3, Article 3RIB5.

[10] Carroll, R.J., D. Ruppert, L.A. Stefanski, C.M. Crainiceanu. *Measurement Error in Nonlinear Models: A Modern Perspective*. Boca Raton: Chapman & Hall/CRC, 2006. Print.

[11] Christensen, R., W. Johnson, A. Branscum, and T.E. Hanson. *Bayesian Ideas and Data Analysis: An Introduction for Scientists and Statisticians*. Boca Raton: CRC Press, 2011. Print.

[12] Kass, R.E., and L. Wasserman. *A Short Course on Applied Bayesian Statistics*. Pittsburgh: Self Published, 1995. Print Course Materials.