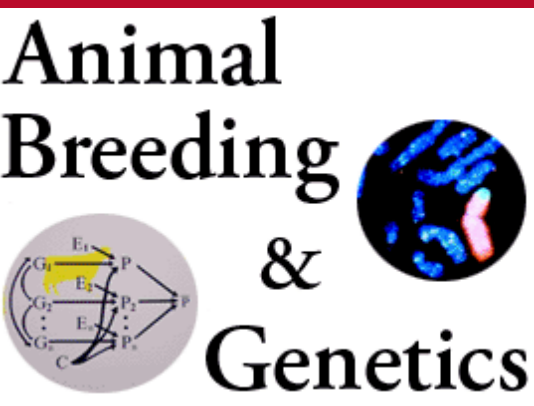


Inclusion of automated sensor data as a predictor of feed intake increases the variance explained by a random forest model.

C.J. Siberski¹, M.S. Mayes¹, P.J. Gorden², A. Copeland², M. Healey¹, B.M. Goetz¹, H. Beiki¹, L.M. Kramer¹, L.H. Baumgard¹, P. Dixon³, and J.E. Koltes¹

¹Department of Animal Science, Iowa State University, Ames 50011 – 1178, ²Veterinary Diagnostic & Production Animal Medicine, Iowa State University, Ames 50011 – 1178, ³Department of Statistics, Iowa State University, Ames 50011 - 1178



IOWA STATE UNIVERSITY

INTRODUCTION

- Commercial feed intake data is rare.
- Genetic selection for improved efficiency is limited due to this lack of commercial records.
- The current reliability of genomic prediction for feed efficiency is 13% across 1.6 million genotyped animals (Li et al., 2020, JDS 103: 2477-2486).

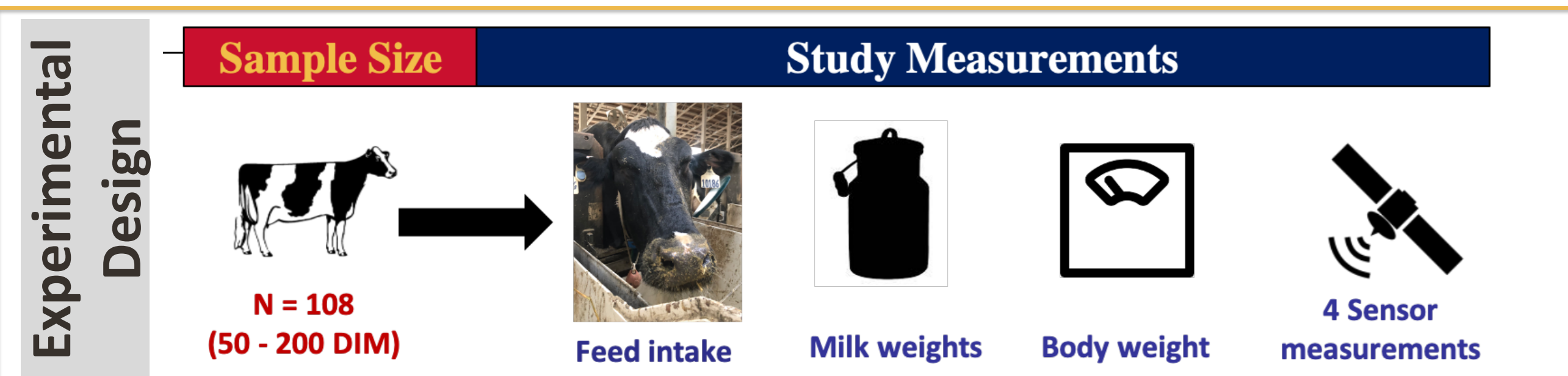


- Industry use of precision livestock technologies (PLTs) has increased.
- Phenotypes recorded by PLTs may explain some of the variation in feed intake observed between cows.
- Such measures may improve predictive ability of feed intake and health events.

OBJECTIVE

Examine the ability of automated sensor data to improve dry matter intake prediction.

MATERIALS & METHODS



Sensor Measures

- Ear tag 1**
 - Measures activity and inner-ear temperature
 - N = 92
- Ear tag 2**
 - Measures activity and rumination
 - N = 41
- Rumen Bolus**
 - Measures activity, rumen pH and temperature
 - N = 56
- Environmental**
 - Measures temperature, relative humidity and wind speed
 - N = 92

Random Forest Models

- Fit utilizing the Caret package in R, with 10-fold cross validation.
- All observations except the final study day were utilized for training, with the final day being utilized for validation.

Base model:
 $DMI_{ijkl} = Par_i + THI_j + HE_k + MBW_l + MY_l + MF_l + MP_l + ML_l + \epsilon_{ijkl}$

Sensor model:
 $DMI_{ijkl} = Par_i + THI_j + HE_k + MBW_l + MY_l + MF_l + MP_l + ML_l + ASM_l + \epsilon_{ijkl}$

Where: **DMI**: dry matter intake (kg); **Par**: parity; **THI**: temperature humidity index; **HE**: health event; **MBW**: metabolic body weight ($BW^{0.75}$); **MY**: milk yield (kg); **MF**: milk fat (kg); **MP**: milk protein (kg); **ML**: milk lactose (kg); **ASM**: adjusted sensor measures*

*Sensor measures adjusted for contemporary group and parity

RESULTS

Animals Included	Base Model MSE	Sensor Model MSE
Ear tag 1	9.86	9.70
Ear tag 2	12.50	12.32
Rumen bolus	13.25	12.53

Table 1: Mean square error (MSE) of the random forest models trained utilizing the caret package and 10-fold cross validation. At minimum, a small decrease in MSE is observed with the addition of sensor measures.

Inclusion of sensor measures increases the percent of DMI variance explained by random forest models

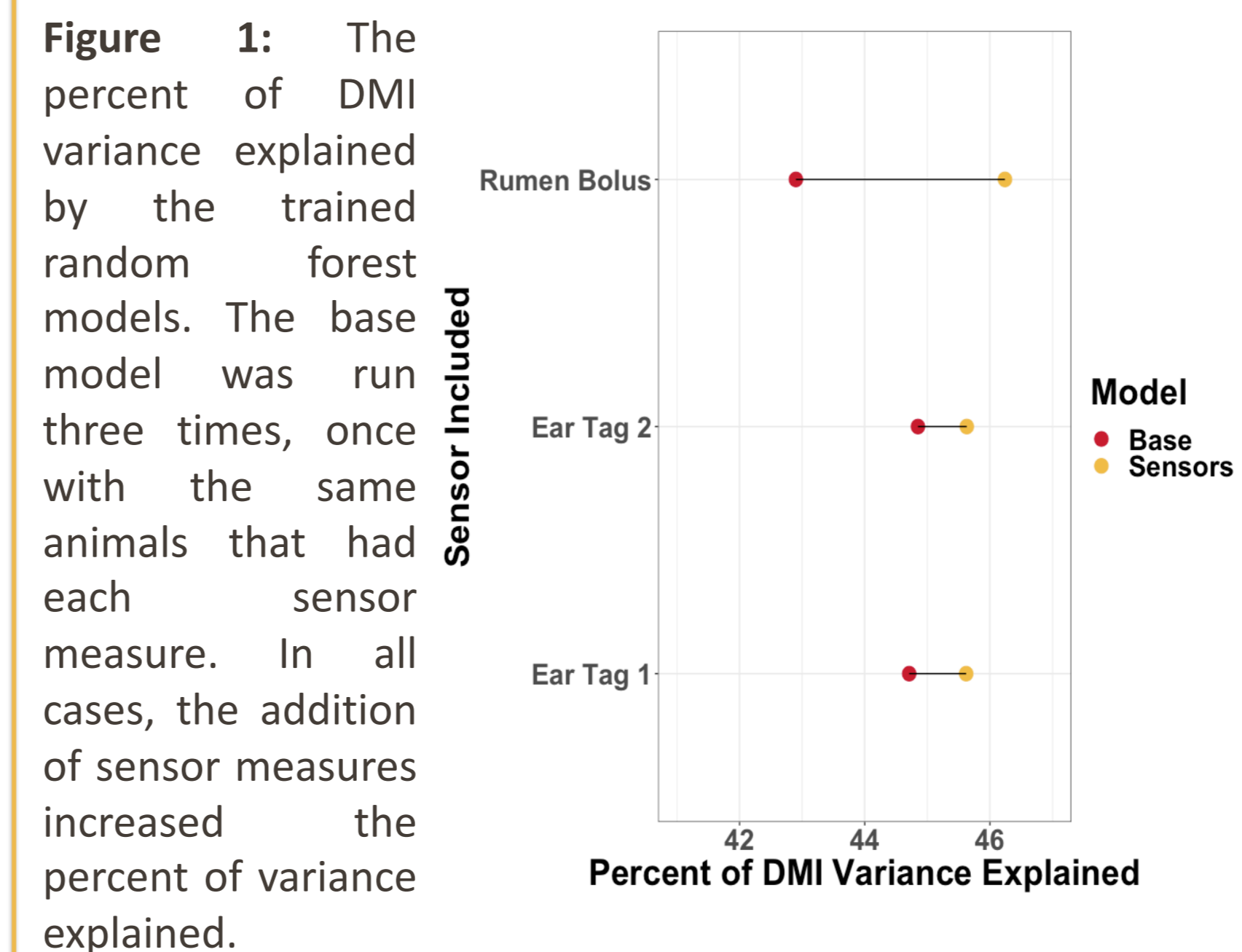
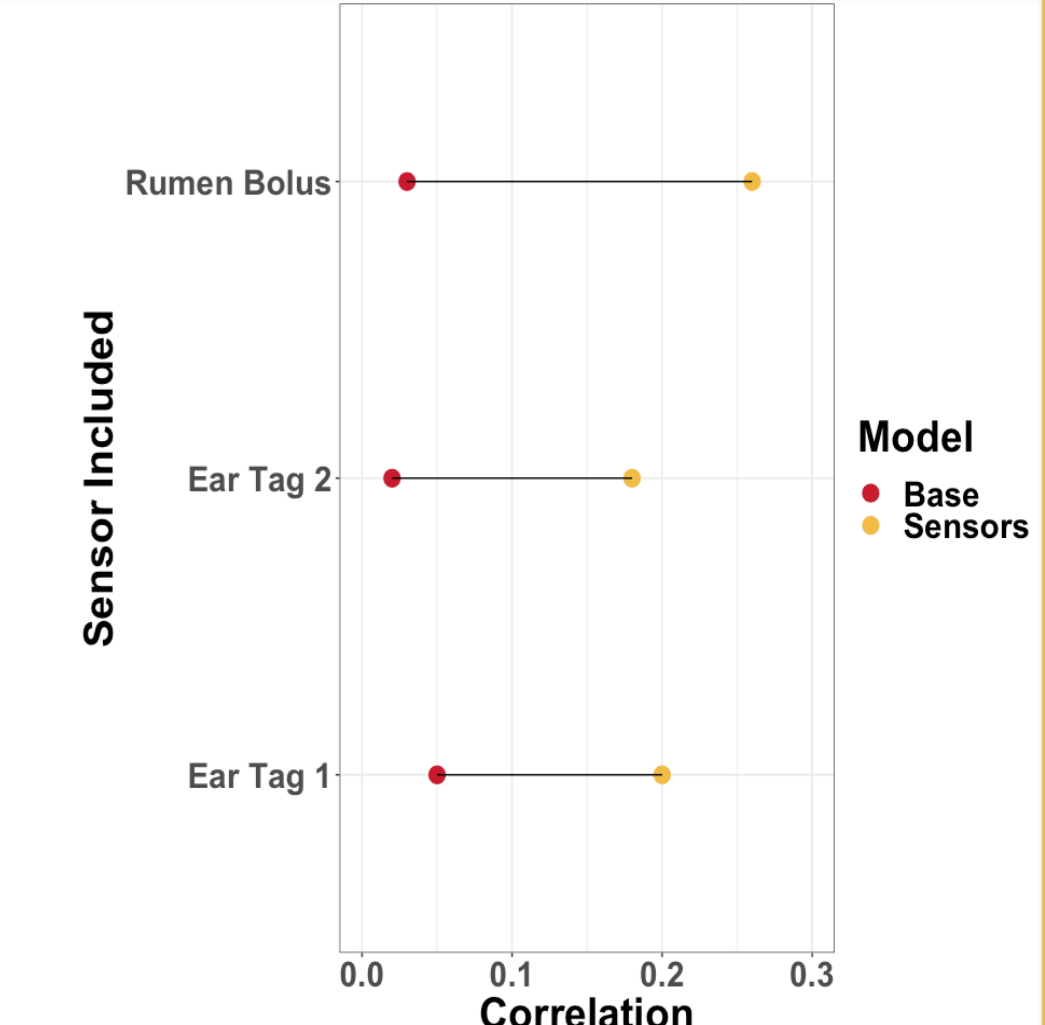


Figure 1: The percent of DMI variance explained by the trained random forest models. The base model was run three times, once with the same animals that had each sensor measure. In all cases, the addition of sensor measures increased the percent of variance explained.

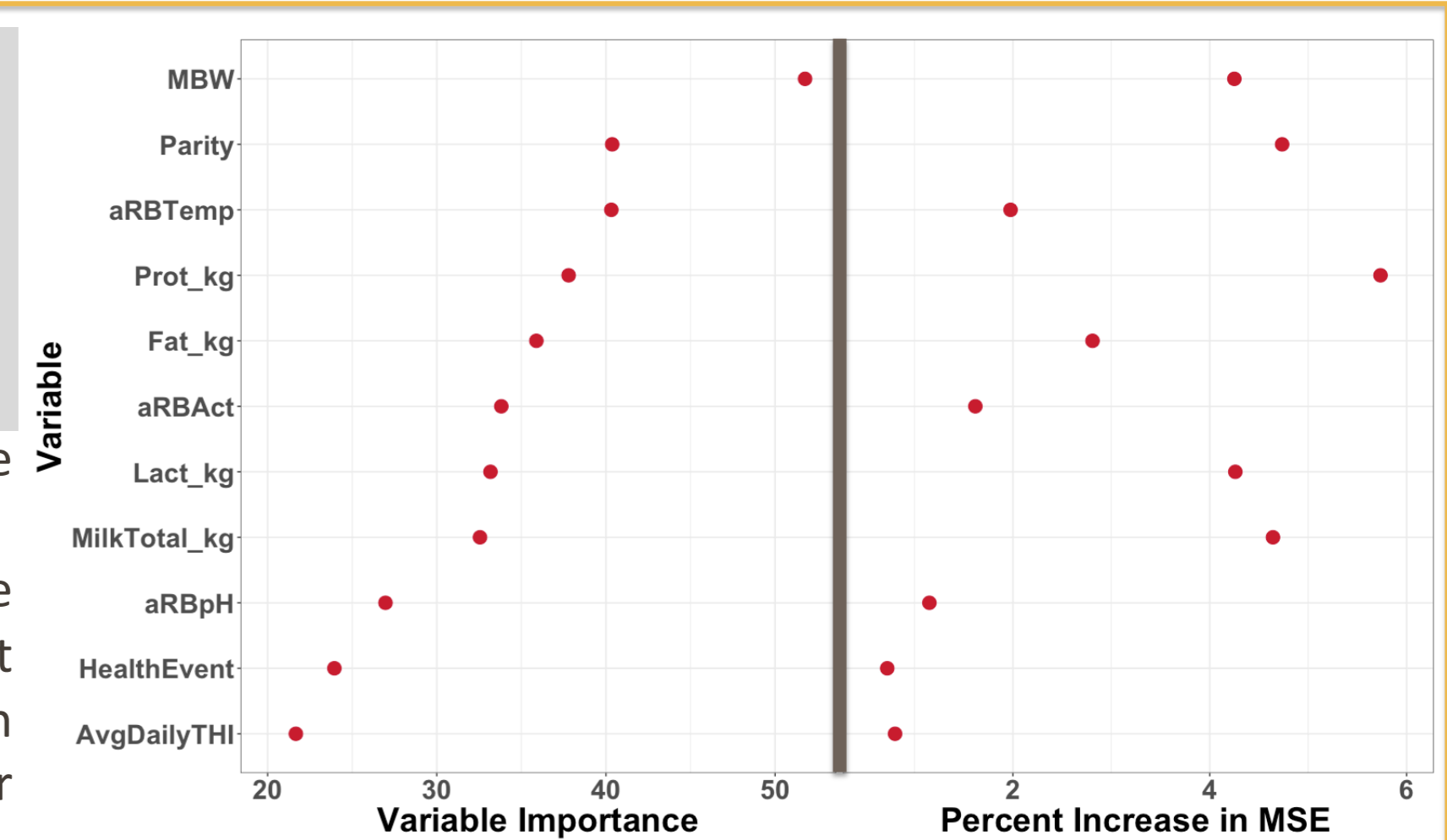
Correlation between actual and predicted DMI increases with the inclusion of sensor measures

Figure 2: The correlation between the predicted and actual DMI on the final day of the study. In all situations, a considerable increase in the correlation is observed when comparing models with and without sensor measures. This validates that sensor measures improve predictive ability in this population.



Variable importance and change in MSE for the rumen bolus model

Figure 3: The variable importance value and percent increase in mean square error (MSE) for the random forest model including rumen bolus measures. Notably, the exclusion of THI and/or health from the models caused slight increases in the percent increase in MSE in most variables, while few showed minimal changes (results not shown).



CONCLUSIONS

- The addition of sensor measures appears to explain additional variation in DMI that is not captured in the typical energy sinks utilized in the prediction of intake.
- The correlation between the predicted DMI using a random forest model and the actual DMI was higher when sensor measures were included as predictor variables.
- Results suggest that utilizing sensor measures to aid in the prediction of feed intake is beneficial.
- Validation of these findings in larger populations, as well as exploration into other predictive methods is needed.