UC Davis

UC Davis Previously Published Works

Title

Describing the distribution type of dry matter intake for dairy cow pens based on pen characteristics

Permalink

https://escholarship.org/uc/item/45c0n4vp

Authors

Lucey, PM
Rossow, HA

Publication Date

2023-06-01

DOI

10.1016/j.animal.2023.100888

Peer reviewed

animal: The international journal of animal biosciences

Describing the distribution type of dry matter intake for dairy cow pens based on pen characteristics
--Manuscript Draft--

Manuscript Number:	ANIMAL-23-90057R3
Full Title:	Describing the distribution type of dry matter intake for dairy cow pens based on pen characteristics
Article Type:	MODNUT22
Section/Category:	Supplements & Special/Themed Issues
Keywords:	precision feeding; diet formulation; dairy production; mathematical modeling
Corresponding Author:	Heidi Rossow, PhD University of California Davis Tulare, CA UNITED STATES
First Author:	Padraig Lucey, DVM
Order of Authors:	Padraig Lucey, DVM
	Heidi A Rossow, PhD
Manuscript Region of Origin:	UNITED STATES
Abstract:	In practice cows are fed by pen, but a diet is formulated to the nutrient requirements of a single cow. If the dry matter intake (DMI) of a pen were equal for all cows this approach would have no error, but cows are grouped into pens on pregnancy and other management factors creating a distribution of DMI. The goal of precision feeding is to meet the requirements of individual animals to increase efficiency and reduce environmental impact but is not achieved when a group is fed as if the individuals have uniform requirements and the DMI distribution is not normal. The hypothesis of this work is that the DMI of cow pens are not normally distributed and the total DMI from the best fit distribution shape for a cow pen will have lower percentage error to the observed DMI than a prediction of a single DMI that is fed at a uniform level and assumes a normal distribution. Our objective was to describe the distribution shape of DMI by week of lactation, and for different pen types. Pens were generated by randomly assorting cows by week of lactation from a database into different categories of pen for size and lactation period. These pens were fitted to the best distribution type, and its parameters were used to randomly generate distribution plots that predict the total DMI for each pen. A second predictive model estimated the DMI of each pen using an empirical equation of DMI that was multiplied by the number of cows in the pen to represent feeding of a uniform DMI quantity. The percentage error for the distribution shape model was significantly lower than the empirical model with pen errors being less than 1 %. The beta distribution type was the most common distribution to best represent the data of pen DMI. Describing the distribution and using it to predict a total pen DMI provides accurate estimates of feed quantity for a group. Reducing error by using the distribution of DMI for feed formulation, instead of the nutrient requirements of an individual animal can provide a precision nutrition approach to group feeding.

1   **Describing the distribution type of dry matter intake for dairy cow pens based**

2   **on pen characteristics**

3   *P. M. Lucey* [a]*, H. A. Rossow* [a]

4

5   [a] *Veterinary Medicine Teaching and Research Center, School of Veterinary Medicine,*

6   *University of California Davis, Tulare 93274, USA*

7

8   Corresponding author: Heidi Rossow. Email: harossow@ucdavis.edu

9   **Abstract**

10   In practice cows are fed by pen, but a diet is formulated to the nutrient requirements

11   of a single cow. If the dry matter intake (**DMI**) of a pen were ==equal for all cows== this

12   approach would have no error, but cows are grouped into pens on pregnancy and

13   other management factors creating a distribution of DMI. The goal of precision

14   feeding is to meet the requirements of individual animals to increase efficiency and

15   reduce environmental impact but is not achieved when a group is fed as if the

16   individuals have uniform ==requirements and the DMI distribution is not normal==. The

17   hypothesis of this work ==is that the DMI of cow pens are not normally distributed== and

18   the total DMI ==from the best fit distribution shape== for a cow pen will have lower

19   percentage error to the observed DMI than a prediction of a single DMI that is fed at

20   a uniform level ==and assumes a normal distribution.== Our objective was to describe the

21   distribution shape of DMI by week of lactation, and for different pen types. Pens were

22   generated by randomly assorting cows by week of lactation from a database into

23   different categories of pen for size and lactation period. These pens were fitted to the

24   best distribution type, and its parameters were used to randomly generate distribution

25   plots that predict the total DMI for each pen. A second predictive model estimated the

26   DMI of each pen using an empirical equation of DMI that was multiplied by the

27   number of cows in the pen to represent feeding of a uniform DMI quantity. The

28   percentage error for the distribution shape model was significantly lower than the

29   empirical model with pen errors being less than 1 %. The beta distribution type was

30   the most common distribution to best represent the data of pen DMI. Describing the

31   distribution and using it to predict a total pen DMI provides accurate estimates of feed

32   quantity for a group. Reducing error by using the distribution of DMI for feed

33 formulation, instead of the nutrient requirements of an individual animal can provide a

34 precision nutrition approach to group feeding.

35

36 **Keywords**

37 Precision feeding, diet formulation, dairy production, mathematical modeling.

38 **Implications**

39 Precision feeding in dairy production can improve nutrient efficiency and reduce feed

40 cost and environmental impact by meeting the requirements of each animal, but is

41 not achieved in group feeding systems. This work describes the distribution of feed

42 intake by pens of cows, and demonstrates that this calculates a pen feed quantity

43 with very low error. This estimates the feed quantity that meets the needs of a pen

44 and delivers population level precision feeding for commercial management that is

45 not achievable when applying individual animal data to a population.

46 **Introduction**

47 In practice, cows are fed by pen, and a diet is formulated to a single estimate of dry

48 matter intake (**DMI**) that represents the pen's mean cow. This is scaled to the number

49 of cows in the pen and a uniform quantity of feed per cow is provided. If the DMI of a

50 pen were constant, there would be 0 feed quantity error in this approach. However, if

51 DMI is not constant, the mean is not an appropriate parameter. It is accepted in the

52 dairy industry that milk production and dry matter intake are not constant within cow

53 pens. Feeding cows uniformly, consistently underestimates feed quantity. Then,

54 mathematical corrections are employed to ensure sufficient feed to high production

55 cows, such as feeding to the 83rd percentile nutrient requirements of the pen

56   (Stallings and McGilliard, 1984), or over-formulation where the estimated DMI

57   quantity is increased by a fixed proportion (Weiss, 2019). Precision feeding of dairy

58   cows requires accurate estimates of the pen feed quantity, as the goal of precision

59   feeding is to minimize nutrient waste through minimizing leftover feed. But, pen level

60   management prohibits individual cow feeding (Schulze, Spilke, and Lehner, 2007), as

61   diets formulated for an individual do not meet the requirements of a group.

62

63   Precision feeding addresses between-animal variation to deliver the correct amount

64   of feed to each animal and can reduce nutrient excretion, feed cost, environmental

65   impact and increase nutrient efficiency (Bewley, 2010; Pomar et al., 2011; Capper

66   and Cady, 2020). To employ precision feeding in a pen, the distribution of DMI by

67   pen cannot be represented by a constant as cows are not uniform across a pen. For

68   example, cows are sorted into pens based on pregnancy status, milk production level

69   (low and high) and nutrient efficiency (St-Pierre and Thraen 1999). Variation in DMI is

70   affected by milk yield (45 %), feed management (22 %), BW (17 %), climate (10 %)

71   and body condition score (6 %) (Roseler et al., 1997).x Grouping cows by these

72   same factors will create a distribution of DMI. The current assumption of uniformity

73   when formulating pen diets does not address between-cow variation in DMI. There is

74   concern in the applicability of individual cow studies to commercial operations where

75   cows are grouped in pens (St-Pierre, 2007). Empirically derived predictions of DMI

76   and nutrient requirements for an individual cow, such as Nutrient Requirements of

77   Dairy Cattle (NASEM, 2021) or the Cornell Net Carbohydrate and Protein System

78   (Van Amburgh et al., 2015), assume all cows in a pen eat the same amount, and do

79   not consider how distribution of DMI differs for groups of cows at different production

80   levels.

81

82 The hypothesis of this model is that feeding a uniform DMI does not represent the

83 distribution shape of DMI for cow pens, and the total DMI quantity for a pen

84 calculated as the density of the best fitting distribution shape will be closer to the true

85 DMI when compared to the empirically calculated DMI for the mean cow of the pen.

86 Using a database of individual cow DMI by week we will first describe the skew and

87 kurtosis of DMI of cows when grouped by week of lactation, then generate sample

88 pens of cows within a range of weeks of lactation, describe the distribution type and

89 parameters, and predict a total DMI. This will be compared to the observed total DMI,

90 the total DMI that is predicted by the mean DMI of a pen by the NASEM 2021

91 equation and the observed DMI.

92 **Material and methods**

93 No approval from the University of California, Davis, Animal Care and Use

94 Committee was needed for this study as it was conducted using only previously

95 collected data.

96 *Database description*

97 A database was constructed from datasets of nine research trials for this

98 mathematical modelling study. All trials were nutrition or management-based

99 interventions on lactating cows and seven of the nine have been published in peer

100 reviewed journals. Each trial includes data for different ranges of the lactation period,

101 with the recording period starting at 1 wk (first 7 d immediately after parturition) and

102 extended up to 44 wk (Table 1). All cows had individual milk yield and components

103 recorded at each milking daily, and DMI recorded daily using Calan gates (American

104 Calan, NH, USA). Data were averaged by week per cow for 426 cows in total.

Inclusion criteria

106 Every dataset needed to list individual records by unique cow identification number,

107 identify treatment assignment, parity, diet composition, days in milk, weekly milk

108 yield, weekly milk fat and weekly DMI consecutively for at least 4 wk. Some of the

109 datasets also included milk protein, BW, and body condition score.

Data cleaning

111 Cows with missing values for weekly observations were not removed. Datasets were

112 imported to a SQL database using an open database connector pipeline. Within the

113 database, a variable for fat corrected milk was created using the formula $3.5\% \, FCM =$

114 $(0.4324 * kg) + (16.216 * milkfat)$ (Tyrrell and Reid, 1965). We used fat corrected

115 milk instead of energy corrected milk because 1 dataset did not record milk protein

116 values. A variable for feed efficiency (**FE**) was created using the formula $FE =$

117 FCM/DMI (Korver, 1988). We used FE to compare the metabolic efficiency of cows

118 across trials despite differences in diet and environment to determine if data were

119 reasonable and if these cows could exist on a single dairy and be grouped together

120 into pens.

Database variance assessment

122 All datasets were compared for similarity of variance in DMI, milk yield and FE. The

123 variance of FE by dataset and treatment group within dataset was used as the main

124 measure of comparability as this represents metabolic efficiency and allows the

125 comparison of cows with different levels of feed intake and milk yield. Data were

126 retained in the overall database if 50% of the FE data overlapped with all other

127 datasets on boxplot analysis, and less than 20% of the weeks of lactation had a FE

128 mean significantly different from other datasets on analysis of variance (**ANOVA**)

129 analysis. The database was imported to R 4.2.1 (R Core Team, Vienna, Austria,

130 2018) and milk yield, DMI and FE were examined visually using scatter plots to

131 identify and remove outliers, and assess the range of the three variables (not shown).

132 The mean and median of milk yield, DMI and FE was plotted by week for each

133 dataset (Figure 1). Boxplots were examined at every week by dataset and treatment

134 within dataset for visual comparison of the mean and variance of milk yield, DMI and

135 FE (not shown). We used ANOVA models with R 4.2.1 base package to compare the

136 means of the variables of milk yield, DMI and feed efficiency with cow as the unit of

137 interest. Data were first evaluated using treatment group within dataset as the

138 subgroup for the outcomes of milk yield, DMI and FE. Next the database was

139 evaluated by ANOVA using dataset within database as the subgroup for the

140 outcomes of milk yield, DMI and FE. Using the equation $Y_{i,j} = \mu + \beta_i + \epsilon_{i,j}$, where $Y_{i,j}$

141 is the j-th observation of the i-th group (i = 1, 2, … ,9 datasets or i = 1, 2, … ,35

142 treatment within dataset), β_i is the i-th subgroup effect and $\epsilon_{i,j}$ is the random error

143 present in the j-th observation on the i-th treatment of these fixed effect models.

144 ***Model description***

145 *Skew and Kurtosis by week of lactation*

146 The skew, kurtosis and Shapiro-Wilk test for normality (Shapiro and Wilk, 1965) was

147 calculated based on DMI data for each week of lactation separately starting at 1 wk

148 to 44 wk using the Moments package of R 4.2.1 (Komsta, L., Novomestky, F., 2022).

149 This was performed on all cows in the database, then with primiparous cows only,

150 and then multiparous cows (Table 2). The null hypothesis of the Shapiro-Wilk test is

151 that the data come from a normal distribution. The null hypothesis was rejected at a

152 significance level of $P < 0.05$.

153 *Creation of pens to simulate DMI distributions*

154 The database was used to generate multiple cow pens for a set number of cows

155 within a range of weeks of lactation. Three pen types were created: fresh, high and

156 low. Fresh pens were lactation range 1 – 3 wk, high range 4 – 18 wk and low range

157 19 – 44 wk. We also generated a large and small dairy pen size for each type. We

158 chose pen size and lactation range as representative of management in the

159 California dairy industry (M. Wukadinovich, personal communication, November

160 2022). Fresh pens were set at 20 (**FRESH20**) and 60 cows (**FRESH60**), high at 50

161 (**HIGH50**) and 200 (**HIGH200**), and low at 50 (**LOW50**) and 150 (**LOW150**). One

162 additional high pen was created with 50 primiparous cows only (**HIGHP50**), resulting

163 in seven pen types (Table 3). To create replicates of each pen type, data was

164 prepared in longform with one row per cow per wk of data observation. The data

165 were subset into the week of lactation ranges for each pen type, and 100 unique

166 pens for each type were generated by selecting one weekly observation of DMI per

167 unique cow within that week range. This pen generating code was nested in a FOR

168 loop with a changing random seed number in R 4.2.1 for 100 replications.

169 Observations were replaced between each pen generation. This resulted seven sets

170 of 100 unique pen replicates via bootstrapping. Each bootstrapped pen contained

171 unique weekly DMI observations by cow within the lactation range for that pen, no

172 cow occurred at multiple time points as each pen represents a temporal snapshot of

173 potential grouping on a dairy. Each replicate selected a random weekly DMI

174 observation per cow from the data pool available for the pen constraints, and

175 uniqueness comes from the diversity of arrangements of different cow week

182 *Best distribution type fitting of DMI*

183 All 100 replicates of each seven pen types were fit to the best probability distribution

184 type using the ExtDist package in R 4.2.1 (Wu, Godfrey, and Pirikahu, 2020). The

185 distributions Normal, 4-parameter Beta, Weibull and Logistic were selected by

186 preliminary visual assessment of the distribution shape of DMI of replicated pens,

187 then tested with maximum log likelihood estimation. The value closest to zero

188 represented the best fit to the data of the four tested distributions. For each of the

189 100 replicates of the seven pen types, the best distribution type to the DMI data of

190 that pen replicate was calculated by this method. The number of occurrences of best

191 fit for each tested distribution type were summed for each pen type to calculate the

192 proportion of each's 100 replicates that were best fit by each distribution candidate.

193 The skew and kurtosis values of every replicate was calculated and the mean of both

194 values for each pen type is reported (Table 3). This method described the distribution

195 of DMI for multiple randomly generated pens to report the most common distribution

196 shape that best fit these pen types. The observed total DMI for each of the 700

197 replicated pens is known and calculated as the sum of observed DMI values per pen.

198 **Model results**

199 *Predicting total DMI for a distribution shape*

200   A total DMI for every replicate of the seven pen types was predicted using its best

201   fitting distribution. This distribution type had its appropriate parameters calculated

202   using the function for that distribution in the R 4.2.1 package ExtDist by maximum

203   likelihood estimation (Wu et al., 2020). For each distribution and its parameters, we

204   randomly generated DMI values that fit the shape of the curve (rBeta, rNormal,

205   rLogistic, rWeibull functions, base R 4.2.1). The sum of these values was considered

206   the prediction of DMI for each pen for that fitted distribution shape. This random

207   generation and summing of DMI values was performed 1000 times for each pen

208   replicate, then the mean value of total DMI for the 1000 iterations was taken as the

209   final total DMI estimate. This was performed because of variance between the

210   randomly predicted values for a given distribution's parameters, and to standardize

211   the prediction of total DMI across all replicates. (Table 4).

212   *Predicting total DMI by empirical equation*

213   The descriptive statistics of each of the 700 generated pen replicates of the seven

214   pen types were used to calculate the predicted DMI by the NASEM 2021 equation

215    $(DMI(kg/d) = [(3.7 + Parity \times 5.7) + 0.305 \times MilkE(Mcal/d) + 0.022 \times BW(kg) +$

216    $(-0.689 + Parity \times -1.87) \times BCS] \times [1 - (0.212 + Parity \times 0.136) \times e(-0.053 \times$

217    $DIM)])$ (de Souza et al., 2019; NASEM, 2021) in R 4.2.1. The mean values of parity,

218   milk energy (NRC, 2001), BW, body condition score and days in milk (DIM) for each

219   pen replicate were used as the input values. The predicted DMI was multiplied by the

220   n of cows in that pen replicate to produce a total DMI of the pen at an assumption of

221   uniform DMI.

222   *Comparing predicted DMI to observed and model diagnostics*

223 Every pen had predicted DMI compared to the observed DMI for percentage error

224 with the formula $PercentageError = (Value_{observed} - Value_{predicted})/Value_{observed} \times$

225 100. Observed total DMI was considered the sum of the observed DMI values of

226 every cow in each pen replicate. The predicted total DMI for a distribution shape, and

227 the predicted total DMI by empirical equation of every pen were both compared to the

228 observed total DMI for their percentage error in total DMI prediction. The mean error

229 for each of the seven pen types by both methods of prediction was compared for a

230 difference in means by student's t-test at a significance level of $P < 0.05$ (Table 4).

231 Both prediction models, distribution fitting vs. NASEM, were compared to the

232 observed data for their model fit to the data using mean squared prediction error

233 **(MSPE)** $(MSPE = ME^2 + VAR$, ME = mean error, VAR = variance) using Microsoft

234 Excel (Microsoft, WA, USA) and partitioned into error due to mean, slope and

235 random bias as described by Benchaar et al., 1998 and Theil, 1966).

236 **Results & Discussion**

237 The goals of this modeling study were to create a pool of cows that could be

238 assembled into virtual pens that represent possible grouping of cows as may occur

239 on a commercial dairy, determine the best fit DMI distribution, and then estimate DMI

240 of the pen using the distribution. With bootstrap sampling, we generated virtual pens

241 by selecting a different weekly observation per unique cow and assembling them into

242 pens of cows for a range of weeks of lactation. To observe the distribution in DMI of a

243 pen and represent how this truly may occur on a dairy, it was important to examine

244 many possible cow groupings for the consistency of distribution shape and estimate

245 the group level variables that defined the distribution. Because dairies are composed

246 of many different cows, variance between cows in the dataset was desired. But to

11

247 avoid generating pens of cows that would not realistically be grouped together, FE

248 was used to compare metabolism between cows. Cows are most commonly grouped

249 by reproductive status, lactation stage and milking performance (Contreras-Govea et

250 al., 2015), so we assumed FE should be comparable across the database as a

251 measure of the cow's ability to produce milk from feed, and to avoid skew by

252 introducing an artifact from combining metabolically dissimilar cows.

253 ***Database construction***

254 *Inclusion Criteria*

255 Three of the nine datasets were different ($P < 0.05$) in milk yield or DMI due to the

256 intervention in their publication. All datasets had overlap of at least 50 % of the data

257 on boxplot analysis by treatment group. Results of the ANOVA analysis showed

258 differences between treatment groups for multiple datasets at occasional weeks of

259 lactation. But there were no differences observed for a period of three or more

260 consecutive weeks, and differences did not occur for more than 20% of all weeks of

261 lactation of the trial. Therefore, all treatment levels of each dataset were retained.

262 Datasets were compared to each other visually using boxplots and statistically by

263 ANOVA with dataset within database as the level for the means of DMI, milk yield

264 and FE. Dataset seven had consistently lower DMI and milk yield than other

265 datasets, and the range of these did not overlap by at least 50 % of the data of other

266 datasets at all weeks of lactation. This dataset also had repeated weeks of milk yield,

267 DMI and FE that were different from other datasets ($P < 0.05$). All observations from

268 dataset seven were removed from the database, resulting in a total of 375 cows with

269 8,982 weekly DMI observations retained and used in this study.

270 *Database variance assessment*

271   Bootstrapping of pen groups allows the generation of many virtual cow pens, but as

272   the pen size approaches the total number of unique cows available, or number of

273   replicates increases, the variation between replicates will diminish. To generate

274   reasonable pens from this database, a high tolerance for variation in milk yield and

275   DMI within and between datasets was accepted, including reported significant

276   differences by the published studies. Dairies intentionally group cows of similar milk

277   production levels, and low milk yield accounts for 21 % of all cow culling removal

278   (USDA, 2014). To satisfy that the cows in the database could be grouped together on

279   a dairy, metabolic efficiency as FE, and the overlap in the variance of milk yield and

280   DMI was used to validate that generated virtual pens are reasonable, as dairies will

281   group cows to reduce variance in these variables.

282   ***Distribution of DMI by week of lactation***

283   Week of lactation was the grouping factor for this study and is an important

284   determinant of DMI as energy demand of milk production changes with DIM. To

285   understand how this contributes to the distribution of DMI for pens, we first described

286   DMI distribution at each week of lactation. Skewness is an asymmetry in a

287   distribution, and a normal distribution is considered to have no skew. Positive

288   skewness indicates the mass of the distribution is shifted to the lower value of the

289   variable with a right tail (right skew), and the median is lower than the mean.

290   Negative skewness indicates the mass of the distribution is shifted to the higher

291   value of the variable with a left tail (left skew), and the median is higher than the

292   mean. Kurtosis is a measure of the data in a distribution's tails. The normal

293   distribution has a kurtosis of three, with higher values indicating more data in the

294   tails.

295

296 Cows were first examined together, then the primiparous and multiparous separately

297 (Table 2) for skew, kurtosis, and Shapiro-Wilk test of normality for DMI. Best

298 distribution fit was not presented for these data as cows would not be grouped in

299 pens or fed by individual week on a dairy. For all cows together 2 – 10 wk of lactation

300 were non-normal except for 6 wk. Normality was observed from 11 – 35 wk for these

301 cow groups at almost every week, then non-normality was again observed at

302 occasionally between 37 – 44 wk. Almost all non-normal weeks were left skewed. For

303 primiparous, non-normality was observed at 20 and 28 wk with left skew at a n of 56

304 cows. For multiparous, non-normality was observed for the fresh and early high

305 period (2-6 wk), and 16 to 23, 36, 40 and 43 wk. For all three animal groups at most

306 weeks of lactation the DMI distribution shape was negatively skewed (left skew).

307

308 Under the central limit theorem, we expect that the random variable of DMI at each

309 week of lactation would assume a near-normal distribution for a large sample size.

310 This is only observed when looking at primiparous cows. For all cows in the early

311 lactation period the effects of parturition on energy balance, metabolism and

312 immunity (Pascottini, Leroy, and Opsomer, 2020) may be creating a non-normal

313 distribution of DMI. This distribution in the early lactation period is not considered in

314 empirical prediction equations of DMI. The NASEM 2021 and Cornell Net

315 Carbohydrate and Protein System are both calculated using observed data of cows

316 after their peak lactation milk yield. Normality was observed for early lactation

317 primiparous cows. Primiparous cows may be more uniform in their milking potential

318 and DMI as these animals still have growth energy demands and lower incidence of

319 production diseases or intramammary infections than older animals. Further non-

320 normality was observed in the multiparous cows after the milk yield peak (6 – 9 wk,

321 Figure 1), and this distribution may be introduced by factors such as the age

322 diversity, pregnancy status, disease history or management differences.

323

324 The mean and median of all cows were presented by week for the variables milk

325 yield, DMI and FE (Figure 1). There was a rapid increase in DMI and milk yield

326 across the early lactation period. This difference in DMI level early in lactation will

327 impact the DMI distribution of fresh pens. Later in lactation we see a gradual decline

328 in DMI. Pens at this period of the lactation usually include cows across a wider range

329 of weeks, and the number of cows at each week will influence the DMI distribution for

330 a pen. The median and mean lines of Figure 1 deviate when skew occurs, as the

331 mean is drawn away from the median by outliers. These factors, the construction of

332 pens across a range of the lactation, and skew in DMI within cows at the same week,

333 both highlight the concern in assuming uniformity of DMI for a pen.

334 ***Distribution fitting by pen type***

335 One hundred replicates of the seven pen types were randomly generated and each

336 one was fit to the best distribution. The 4-parameter Beta was the best distribution for

337 all the pen types. For pens FRESH20 and FRESH60 the Beta was the best fit for 85

338 and 87 % of replicates respectively. For HIGH50, HIGHP50 and HIGH200 the Beta

339 was the best fit for 77, 80 and 60 % of replicates respectively, and for LOW50 and

340 LOW150 it was the best fit for 62 and 36 % respectively (Table 3). The shape and

341 scale parameters of the beta distribution allow it to accommodate right and left

342 skewness in its shape and made it the best distribution to represent these data. For

343 FRESH20 and FRESH60, the mean skewness of all the replicates were close to 0,

15

344 indicating an even number of positive and negative skew values in the replicates, and

345 the mean kurtosis was less than three, describing a platykurtic distribution with flat

346 tails and small outliers. The HIGH50 and HIGH200 pens both had negative mean

347 skewness, indicating a left sided tail, and mean kurtosis close to three, describing

348 tails similar in weight to a normal distribution. The HIGHP50 pen had a negative

349 mean skew value and normal mean kurtosis. Both LOW50 and LOW150 had a mean

350 negative skewness and mean kurtosis values higher than three. This positive excess

351 kurtosis indicates a leptokurtic distribution with heavier weighted tails, describing

352 large outliers. This may be due to the largest week range of the LOW50 and

353 LOW150 pens at 19 – 44 wk. The skew values of all pen types led to the beta

354 distribution as the best overall fit. The pens FRESH20 and FRESH60 had a mean

355 skewness close to zero, but when examining each replicate separately pens were

356 equally dispersed as either heavily negative or heavily positive in skewness (not

357 shown), resulting in a mean skewness close to zero.

358

359 This demonstrates that in the early lactation fresh cow pens may have non-normal

360 distributions, with outliers to either the right or left side, depending on the cows in the

361 group. Pens later in the lactation, HIGH50, HIGH200, LOW50 and LOW150 were

362 consistently left skewed and best described by the beta distribution as it

363 accommodates this spread of data. The proportion of replicates best described by

364 normal was higher for HIGH200, LOW50 and LOW150 than other pens. The

365 HIGH200 pen may move closer to the normal as the n of this pen was the largest, but

366 beta was still a better distribution fit for most of these replicates. Pens LOW50 and

367 LOW150 had the normal as the best fit for 10 and 29% respectively, and the logistic

368 distribution as best fit for 24 and 34 % of replicates respectively. These pens were

369 spread across the longest lactation range, and at the lower milk production period,

370 when most cows that remain in the herd would be pregnant and at a lower intake

371 level so the decline in DMI is less steep for these weeks. The HIGHP50 pen was still

372 described best by the beta distribution but had a skewness and kurtosis close to the

373 normal. This pen only consisted of primiparous cows, and as only three of the eight

374 final datasets in the database included primiparous, this pen had less variation by

375 DMI. But the beta distribution was still most appropriate, indicating that skewness

376 was an important component of these pen groups at that lactation range.

377

378 The description of these pens as a beta distribution indicates that uniform feeding of

379 cows within a pen to a fixed DMI value is not appropriate. When a DMI value is

380 estimated for a hypothetical cow of a pen and that number is multiplied by the

381 number of cows, not all cows are eating at the same level. There is an assumption

382 that the symmetric tails of a normal distribution will correct for this feeding rate. By

383 feeding as if every cow will eat the same, it is expected that cows who eat less will

384 have their excess feed consumed by cows who require more DMI. But if distributions

385 are not normal, the pen will be over or underfed depending on how the distribution is

386 skewed. Correction factors are routinely applied to increase the pen DMI as this

387 method under-supplies the pen in practice. This is proven by the description of the

388 pen distributions in this study. All high and low yielding pens, except for HIGHP50,

389 had consistent left skew and non-normal distributions. With left sided outliers the

390 main mass of the data is in the upper range of values for these pens. The mean is

391 only an appropriate parameter of the normal distribution, and the mean is pulled

392 towards extreme outliers. For these pens, the mean value will be below the center of

17

393   mass of this distribution leading to an underestimate of total DMI when the pen is fed

394   at a uniform rate.

395   ***Total DMI prediction***

396   *Predicting total DMI for a distribution shape*

397   All 100 replicates of the seven pen types had a total DMI by pen predicted using the

398   best fit distribution shape for those data (Table 4). The observed DMI for every pen

399   replicate of the seven pen types was calculated by summing the observed DMI of the

400   cows in each replicate, and the pen DMI predicted by its best distribution shape was

401   compared to the observed for percentage error. A percentage error was calculated

402   for every replicate and the mean percentage error by pen type is presented. All

403   seven pens had a percentage error less than 1 % for the best distribution prediction

404   of DMI.

405   *Predicting total DMI by empirical equation*

406   Pen DMI was also predicted for all replicates by calculating a DMI value with the

407   NASEM 2021 equation with the inputs as the mean values of the pen for each, and

408   multiplying that value by the n of cows in the pen. These predicted values were

409   compared to the observed DMI by pen type for mean percentage error and ranged

410   from 11 to 22 %.

411

412   The mean squared prediction error of each pen type was calculated for both

413   prediction models to assess the fit of the predicted values to the observed (Table 4,

414   Figure 2). Every pen type had a lower mean squared prediction error for the

415   distribution shape prediction model as compared to the empirical equation model

416   showing a better fit to data of the distribution shape model. This analysis

18

417 demonstrates the model fit of a DMI predicted by an estimated distribution shape is

418 more appropriate than assuming uniform feeding. The model error was partitioned

419 into error due to bias, random variance and slope not equaling one. For the

420 distribution prediction model 77 – 98 % of error was due to random variation in the

421 data indicating that the model is well fitted to the data. For HIGHP50 the mean bias

422 error was 21 % suggesting systematic error in prediction of the DMI mean by the

423 model for this group. For the NASEM prediction model random error ranged from 6 –

424 68 % with the HIGHP50 pen type at 6 % error due to random variation and 93 % due

425 to mean bias. This HIGHP50 group was made up of primiparous only animal and

426 from a smaller subset of possible animals and both models demonstrated systematic

427 error in their prediction of mean producing this proportion of error in the model fit.

428

429 The NASEM 2021 empirical equation was derived using post milking peak cow data

430 so estimates for early lactation pens may be inaccurate. High and low lactation range

431 pens were all underestimated by this equation. High pen underestimation by the

432 empirical model of this study agrees with the observed practice of necessary over-

433 formulating diet on commercial dairies, as underfeeding cows can limit milk

434 production. Describing a distribution type for a pen DMI predicts a DMI closer to the

435 observed DMI of a pen, allowing diets to be calculated to a more accurate amount at

436 formulation rather than utilizing an imprecise correction factor. The practice of

437 formulating diets by increasing DMI by a fixed level is not precision feeding, and

438 ingredient composition may change if DMI is a binding constraint in the formulation.

439 ***Model assessment***

440 This modeling study was conducted by generating virtual pens with data of known

441 DMI values. But the application of this work relies on the future ability of predicting

442 the distribution shape of a cow pen with unknown DMI. For group housed dairy cows,

443 DMI estimation is crudely approximated by measuring the feed remaining after a

444 given time-period from the supply of a known feed quantity. Individual DMI

445 measurements, such as Calan gates, are not feasible for large groups as one gate

446 would be needed per cow, and would probably change inter-cow feeding behavior

447 (Seymour et al., 2019). Sensor technology attempts to measure individual cow DMI

448 with accelerometers and cameras at feed bunks, and if the data are validated could

449 contribute to pen calculation of total DMI.

450

451 Pens with known individual cow DMI values can describe the distribution shape and

452 parameters. This model generated virtual pens to build a labelled database of known

453 distribution shapes. For pens on commercial dairies where individual DMI is not

454 known, supervised learning could predict the distribution characteristics of a pen by a

455 model trained with a labeled database. The model presented in this study

456 demonstrates that predicting pen DMI with a distribution can provide accurate

457 estimates. Current group feeding approaches supply a uniform level of DMI and are

458 not precision feeding. Utilizing the distribution of DMI for each pen of cows is a

459 necessary factor in achieving precision feeding aims for group feeding methods.

460 **Conclusion**

461 Precision feeding for cow pens is not achieved by formulating a DMI value for an

462 individual cow and feeding it uniformly. Describing the distribution shape of DMI for

463 cow pens predicts a pen DMI with less error and can provide an approach to

464 precision feeding in this industry.

465 **Ethics approval**

466 Not applicable.

467 **Data and model availability statement**

468 None of the data were deposited in an official repository. The model was not

469 deposited in an official repository. The data used in this study is confidential. The

470 models that support the study findings are available to reviewers and authors upon

471 request.

472 **Author ORCIDs**

473 Pádraig Lucey 0000-0003-3885-2439

474 Heidi Rossow 0000-0002-3753-4263

475 **Author contributions**

476 **Pádraig Lucey:** Conceptualization, Methodology, Formal analysis, Data Curation,

477 Writing – Original Draft. **Heidi Rossow**: Conceptualization, Resources, Writing –

478 Review & Editing, Supervision, Funding acquisition.

479 **Declaration of interest**

480 None.

481 **Acknowledgements**

482 Thank you to the authors of the dataset publications used in this study.

487

488 **Copyright permission**

490

491 **References**

492 Bewley J., 2010. Precision Dairy Farming: Advanced Analysis Solutions for Future

493   Profitability. Proceedings of The First North American Conference on Precision Dairy

494   Management, 2 – 5 March 2010, Toronto, Canada, p. 16.

495 Benchaar C, Rivest J, Pomar C, Chiquette J. 1998. Prediction of methane production from

496   dairy cows using existing mechanistic models and regression equations1. Journal of

497   Animal Science. 76:617–27

498 Capper, JL., Cady, RA., 2020. The effects of improved performance in the U.S. dairy cattle

499   industry on environmental impacts between 2007 and 2017. Journal of Animal

500   Science. 98:skz291

501 Chalupa, W., Vecchiarelli, B., Galligan, DT., Ferguson, JD., Baird, LS., et al. 1996.

502   Responses of Dairy Cows Supplemented with Somatotropin During Weeks 5 Through

503   43 of Lactation. Journal of Dairy Science. 79:800–812

504 Contreras-Govea, F. E., V. E. Cabrera, L. E. Armentano, R. D. Shaver, P. M. Crump, D. K.

505   Beede, and M. J. VandeHaar. "Constraints for Nutritional Grouping in Wisconsin and

506   Michigan Dairy Farms." Journal of Dairy Science. 98:1336–44

507 Dann, HM., Varga, GA., Putnam, DE., 1999. Improving Energy Supply to Late Gestation and

508   Early Postpartum Dairy Cows. Journal of Dairy Science. 82:1765–78

509 DePeters, EJ., Smith, NE., Acedo-Rico, J., 1985. Three or Two Times Daily Milking of Older

510   Cows, and First Lactation Cows for Entire Lactations. Journal of Dairy Science.

511   68:123–32

512 de Souza, RA., Tempelman, RJ., Allen, MS., VandeHaar, MJ., 2019. Updating predictions of

513   dry matter intake of lactating dairy cows. Journal of Dairy Science. 102:7948–60

514 Dhiman, TR., Satter, LD., 1997. Yield Response of Dairy Cows Fed Different Proportions of

515   Alfalfa Silage and Corn Silage. Journal of Dairy Science. 80:2069–82

516 Greenfield, RB., Cecava, MJ., Johnson, TR., Donkin, SS., 2000. Impact of Dietary Protein

517   Amount and Rumen Undegradability on Intake, Peripartum Liver Triglyceride, Plasma

518   Metabolites, and Milk Production in Transition Dairy Cattle. Journal of Dairy Science.

519   83:703–10

520 Hamner, B., Frasco, M., 2018. Metrics: Evaluation Metrics for Machine

521   Learning. R package version 0.1.4, Retrieved on 10 January from https://CRAN.R-

522   project.org/package=Metrics.

523 Huyler, MT., Kincaid, RL., Dostal, DF., 1999. Metabolic and Yield Responses of Multiparous

524   Holstein Cows to Prepartum Rumen-Undegradable Protein. Journal of Dairy Science.

525   82:527–36

526 Komsta, L., Novomestky, F., 2022. moments: Moments, Cumulants, Skewness,

527   Kurtosis and Related Tests. R package version 0.14.1, Retrieved on 10 January 2023

528   from https://CRAN.R-project.org/package=moments.

529 Korver, S., 1988. Genetic aspects of feed intake and feed efficiency in dairy cattle: A review.

530   Livestock Production Science. 20:1–13

531 Livesey, CT., Harrington, T., Johnston, AM., May, SA., Metcalf, JA., 1998. The effect of diet

532   and housing on the development of sole haemorrhwordages, white line

533   haemorrhages and heel erosions in Holstein heifers. Animal Science. 67:9–16

534 National Academies of Sciences, Engineering, and Medicine (NASEM)., 2021. Nutrient

535   Requirements of Dairy Cattle: Eighth Revised Edition. Washington, DC: The National

536   Academies Press. https://doi.org/10.17226/25806.

537 National Research Council (NRC)., 2001. Nutrient Requirements of Dairy Cattle: Seventh

538   Revised Edition, 2001. Washington, DC: The National Academies Press.

539   https://doi.org/10.17226/9825.

540 Pascottini, OB., Leroy, JLMR., Opsomer, G., 2020. Metabolic Stress in the Transition Period

541   of Dairy Cows: Focusing on the Prepartum Period. Animals. 10:1419

542 Pomar, C., Hauschild, L., Zhang, GH., Pomar, J., Lovatto, PA., 2011. Precision feeding can

543   significantly reduce feeding cost and nutrient excretion in growing animals. In

544   Modelling nutrient digestion and utilisation in farm animals, (ed. Sauvant, D., Van

545   Milgen, J., Faverdin, P., Friggens, N.) Wageningen Academic Publishers,

546   Wageningen, Netherlands, pp. 327–34.

547   Roseler, DK., Fox, DG., Chase, LE., Pell, AN., Stone WC. 1997. Development and

548   Evaluation of Equations for Prediction of Feed Intake for Lactating Holstein Dairy

549   Cows. Journal of Dairy Science. 80:878–93

550   Schulze, C., Spilke, J., Lehner, W., 2007. Data modeling for Precision Dairy Farming within

551   the competitive field of operational and analytical tasks. Computers and Electronics in

552   Agriculture. 59:39–55

553   Sciences National Academy of, Engineering, Medicine., 2021. Nutrient requirements of dairy

554   cattle: Eighth revised edition. Washington, DC: The National Academies Press

555   Seymour, DJ., Cánovas, A., Baes, CF., Chud, TCS., Osborne, VR., et al. 2019. Invited

556   review: Determination of large-scale individual dry matter intake phenotypes in dairy

557   cattle. Journal of Dairy Science. 102:7655–63

558   Shapiro, S. S., and M. B. Wilk. "An Analysis of Variance Test for Normality (Complete

559   Samples)." Biometrika 52. 591–611.

560   Stallings, CC., McGilliard, ML., 1984. Lead Factors for Total Mixed Ration Formulation.

561   Journal of Dairy Science. 67:902–7

562   St-Pierre, NR., Thraen, CS., 1999. Animal grouping strategies, sources of variation, and

563   economic factors affecting nutrient balance on dairy farms. Journal of Animal

564   Science. 77:72–83

565   St-Pierre NR., 2007. Design and Analysis of Pen Studies in the Animal Sciences. Journal of

566   Dairy Science. 90:87–99

567   Theil H., 1966. Applied economic forecasting. Amsterdam: North-Holland Publishing

568   Company

569   Tyrrell, HF., Reid, JT., 1965. Prediction of the Energy Value of Cow's Milk. Journal of Dairy

570   Science. 48:1215–23

571   USDA., Dairy, 2014, "Trends in Dairy Cattle Health and Management Practices in the United

572   States, 1991-2014" USDA–APHIS–VS–CEAH-NAHMS. Fort Collins, CO. #711.0821

573   Van Amburgh, ME., Collao-Saenz, EA., Higgs, RJ., Ross, DA., Recktenwald, EB., et al.,

574   2015. The Cornell Net Carbohydrate and Protein System: Updates to the model and

575   evaluation of version 6.5. Journal of Dairy Science. 98:6361–80

576   Weiss, B., Setting Nutrient Specifications for Formulating Diets for Groups of Lactating Dairy

577   Cows. Retrieved on 10 January 2023 from https://dairy-cattle.extension.org/setting-

578   nutrient-specifications-for-formulating-diets-for-groups-of-lactating-dairy-cows

579   Wu, H., Godfrey, AJR., Govindaraju, K., Pirikahu, S., 2020. ExtDist: Extending the range of

580   functions for probability distributions. R package version 0.6-4, Retrieved on 10

581   January 2023 from https://CRAN.R-project.org/package=ExtDist

582   Wu, H., Godfrey, AJR., Pirikahu, S., 2020. Parameter estimation and distribution selection by

583   ExtDist. Retrieved on 10 January 2023 from

584   https://cran.microsoft.com/snapshot/2022-02-

585   19/web/packages/ExtDist/vignettes/ParaEst-and-DistSel-by-ExtDist.pdf

Table 1. Description of the trial datasets used in the database of this ==dairy cow== study

Trial	Index[2]	n[1] Total	n[1] Primi[3]	n[1] Multi[4]	Breed	Lactation range, wk[5]	Intervention[6]	Difference ($P < 0.05$)[7]
DePeters et al. 1985[8]	1	55	15	40	Holstein	1 – 44	Management	No
Dhiman and Satter 1997[9]	2	74	29	45	Holstein	1 – 36	Diet	Yes
Greenfield et al. 2000[10]	3	37	0	37	Holstein & Jersey	1 – 8	Therapy	Yes
Unpublished[10]	4	45	0	45	Not stated	3 – 18	Not stated	Not stated
Unpublished[10]	5	40	0	40	Not stated	1 – 21	Diet	Not stated
Huyler et al. 1999[10]	6	31	0	31	Holstein	1 – 10	Diet	No
Livesey et al. 1998[11]	7	51	0	51	Holstein	1 – 14	Diet & management	No
Chalupa et al. 1996[8]	8	36	12	24	Holstein	3 – 43	Diet	Yes
Dann et al. 1999[10]	9	57	57	0	Holstein	1 – 9	Diet	No
Total		426	113	313				

Primi = Primiparous

Multi = Multiparous

[1] Cow population per trial as the total, and primiparous and multiparous separately.

[2] Identifying index of each dataset.

[3] Primiparous dairy cows in their first lactation.

[4] Multiparous cows.

[5] The weeks of lactation for which each dataset recorded weekly individual cow milk yield and dry matter intake.

[6] The type of treatment intervention studied by each trial.

[7] Publication reported significant difference in the milk yield or dry matter intake of a single treatment group in the dataset.

[8] Dataset recorded weekly average ==dry matter intake==, milk, milkfat, milk protein and ==BW==.

[9] Dataset recorded weekly average ==dry matter intake==, milk, milkfat, and milk protein.

601   [10] Dataset recorded weekly average dry matter intake, milk, milkfat, milk protein, BW, and body

602   condition score.

603   [11] Dataset recorded weekly average dry matter intake and milk, and monthly milkfat, BW and body

604   condition score.

605 **Table 2 Skew, kurtosis and Shapiro-Wilk test of goodness of fit to normal distribution for dry matter intake values of all**

606 **dairy cows by parity**

	All cows				Primiparous				Multiparous		
Week[1]	n[2]	Skew[3]	Kurtosis[4]	Shapiro-Wilk P-value[5]	n	Skew	Kurtosis	Shapiro-Wilk P-value	n	Skew	Kurtosis	Shapiro-Wilk P-value
1	294	-0.10	2.70	0.5	101	-0.49	3.24	0.08	193	-0.04	2.31	0.1
2	294	-0.32	2.93	0.01	101	-0.012	2.61	0.2	193	-0.47	3.43	0.03
3	342	-0.34	3.02	0.05	102	0.10	2.28	0.2	240	-0.56	3.87	0.004
4	360	-0.45	3.22	0.001	109	0.10	2.38	0.5	251	-0.77	4.40	< 0.001
5	364	-0.30	3.24	0.06	110	-0.031	2.52	0.2	254	-0.47	4.04	0.001
6	374	-0.24	3.21	0.5	113	0.24	2.78	0.4	261	-0.42	4.20	0.008
7	375	-0.37	3.15	0.006	113	0.072	3.00	0.4	262	-0.29	3.19	0.1
8	373	-0.30	2.69	0.003	113	0.014	2.33	0.4	260	-0.23	2.79	0.005
9	337	-0.36	2.79	0.003	113	-0.022	2.37	0.3	224	-0.30	3.07	0.1
10	280	-0.11	2.33	0.02	56	0.063	2.69	0.6	224	-0.18	2.68	0.2
11	250	-0.21	2.61	0.09	56	-0.61	3.46	0.09	194	-0.16	2.79	0.3
12	250	-0.21	2.47	0.05	56	-0.13	2.46	0.5	194	-0.37	3.29	0.1
13	250	-0.15	2.91	0.8	56	-0.47	3.70	0.4	194	-0.11	3.37	0.5
14	250	0.052	2.65	0.9	56	-0.012	3.11	0.8	194	-0.071	3.21	0.7
15	250	-0.091	2.72	0.8	56	-0.26	2.58	0.7	194	-0.33	3.52	0.1
16	250	-0.27	3.05	0.3	56	0.35	3.06	0.4	194	-0.70	4.67	< 0.001
17	250	-0.22	2.78	0.4	56	-0.32	3.18	0.4	194	-0.55	3.68	0.01
18	250	-0.20	3.40	0.1	56	0.25	2.83	0.8	194	-0.65	4.71	< 0.001
19	205	-0.23	3.69	0.2	56	-0.12	3.06	0.7	149	-0.67	4.14	0.003
20	205	-0.42	3.96	0.002	56	-1.8	9.34	< 0.001	149	-0.66	3.93	0.003
21	205	-0.18	3.67	0.1	56	-0.40	3.16	0.3	149	-0.54	4.05	0.01
22	164	-0.16	3.65	0.04	56	0.21	3.03	0.9	108	-0.68	4.33	0.008
23	162	-0.16	4.01	0.2	56	-0.062	2.73	0.6	106	-0.59	4.59	0.02
24	162	0.34	2.93	0.2	56	-0.50	2.78	0.04	106	0.051	2.59	0.3
25	162	0.32	2.49	0.03	56	-0.052	2.45	0.5	106	-0.023	2.21	0.1
26	162	0.21	3.03	0.5	56	-0.53	3.32	0.4	106	-0.082	2.73	0.9
27	162	0.52	3.01	0.01	56	0.012	2.79	0.6	106	0.20	2.48	0.2

28	162	-0.082	3.46	0.6	56	-1.2	5.94	0.001	106	-0.081	2.63	0.7
29	162	0.16	2.58	0.2	56	0.19	2.29	0.3	106	-0.22	2.49	0.2
30	162	0.11	2.88	0.8	56	-0.17	3.63	0.4	106	-0.062	2.59	0.9
31	162	0.033	2.60	0.9	56	0.48	3.33	0.4	106	-0.23	2.43	0.3
32	162	0.11	2.74	0.2	56	0.68	3.55	0.05	106	-0.19	2.44	0.2
33	161	-0.18	2.66	0.4	56	-0.062	2.91	0.9	105	-0.39	2.45	0.03
34	161	-0.11	2.74	0.09	56	-0.13	3.55	0.5	105	-0.31	2.37	0.02
35	161	-0.21	2.84	0.1	56	0.17	3.47	0.2	105	-0.42	2.57	0.03
36	160	-1.2	8.10	< 0.001	56	-0.082	2.70	0.8	104	-1.4	7.31	< 0.001
37	86	-0.29	3.44	0.2	27	0.053	2.18	0.7	59	-0.29	3.07	0.3
38	85	0.094	4.16	0.08	26	-0.19	2.14	0.6	59	0.034	3.64	0.3
39	83	-0.27	4.15	0.07	26	-0.23	2.64	0.5	57	-0.22	4.10	0.1
40	82	-0.36	4.27	0.2	26	-0.20	2.07	0.4	56	-0.45	4.45	0.09
41	65	-0.46	3.72	0.3	21	-0.64	2.62	0.1	44	-0.39	4.14	0.5
42	62	-0.80	3.62	0.02	21	-0.86	2.48	0.01	41	-0.72	4.28	0.1
43	58	-0.87	3.92	0.01	21	-0.64	2.63	0.3	37	-0.97	5.19	0.03
44	26	-0.26	2.39	0.5	8	-0.091	1.68	0.5	18	-0.15	2.69	0.9

607 [1] The week of lactation of data observation

608 [2] The population of cows observed at the given week of lactation

609 [3] Measure of data skew. Negative values indicate left skew, positive values indicate right skew.

610 [4] Measure of data in tails of distribution. Kurtosis of 3 is considered normal.

611 [5] A $P < 0.05$ rejects the null hypothesis that the data are normally distributed.

Table 3 Proportion of best distribution fit for 100 replicates of each dairy cow pen class

Pen type index[3]	Cows, n	Lactation range[4], wk	Pen replicates[5], n	Mean DMI, cow[6] (kg)	DMI Distribution description[1]		Proportions of best distribution of DMI for all replicates[2] (%)			
					Skew	kurtosis	Normal	Beta	Weibull	Logistic
FRESH20	20	1 – 3	100	17.06	0.07	2.76	8	85	1	6
FRESH60	60	1 – 3	100	16.97	-0.03	2.49	8	87	0	5
HIGH50	50	4 – 18	100	22.71	-0.2	2.99	8	77	0	15
HIGHP50	50	4 – 18	100	20.51	0.05	2.87	7	80	8	5
HIGH200	200	4 – 18	100	22.97	-0.2	3.12	28	60	0	12
LOW50	50	19 - 44	100	20.90	-0.2	3.22	10	62	4	24
LOW150	150	19 – 44	100	20.86	-0.2	3.43	29	36	1	34

FRESH20 = Pens of 20 random cows within the lactation range 1 – 3 week.

FRESH60 = Pens of 60 random cows within the lactation range 1 – 3 week.

HIGH50 = Pens of 50 random cows within the lactation range 4 – 18 week.

HIGHP50 = Pens of 50 random primiparous cows within the lactation range 4 – 18 week.

HIGH200 = Pens of 200 random cows within the lactation range 4 – 18 week.

LOW50 = Pens of 50 random cows within the lactation range 19 – 44 week.

LOW150 = Pens of 150 random cows within the lactation range 19 – 44 week.

[1] The mean distribution statistics of skew and kurtosis for all replicates of each pen type.

[2] The proportion of each distribution type occurring as the best fit for DMI of a pen across all replicates of pen types.

[3] Index of generated virtual pen type for given lactation range and population size.

[4] The range of weeks of lactation the pen type was constructed across.

[5] The number of replicates generated for each pen type with a random and unique selection of cows in each replicate.

625   [6] The mean cow DMI of all replicates of each pen type.

Table 4 DMI predictions by best distribution type and by NASEM equation of all pen replicates of each dairy cow pen class and model fit statistics

Pen type index[3]	Observed pen DMI (kg)[4]	Predicted pen DMI by distribution[1]						Predicted pen DMI by NASEM[2]					
		DMI (kg)	% Error of distribution[5]	MSPE[6]	% Bias[7]	% Slope not 1[8]	% Random variation[9]	DMI (kg)	% Error of NASEM	MSPE	% Bias	% Slope not 1	% Random variation
FRESH20	341.09	342.43	+ 0.4[a]	55.64	0	5	95	291.66	- 14.37[b]	211.4	0	32	68
FRESH60	1 018.17	1 019.8	+ 0.09[a]	60.42	0	2	98	876.96	- 13.82[b]	641.6	0	42	58
HIGH50	1 135.61	1 136.6	+ 0.09[a]	43.63	0	2	98	937.45	- 17.42[b]	545.9	0	35	65
HIGHP50	1 025.69	1 026.2	+ 0.05[a]	55.28	21	2	77	921.81	- 10.11[b]	7 612	93	1	6
HIGH200	4 593.08	4 595.1	+ 0.04[a]	182.8	2	2	96	3761.0	- 18.11[b]	1 573	0	33	67
LOW50	1 044.79	1 045.6	+ 0.08[a]	47.44	0	2	98	810.67	- 22.38[b]	587.1	0	64	36
LOW150	3 127.82	3 131.1	+ 0.1[a]	81.99	0	3	97	2 427.9	- 22.37[b]	803.2	0	82	18

DMI = Dry matter intake

MSPE = Mean squared prediction error

FRESH20 = Pens of 20 random cows within the lactation range 1 – 3 week.

FRESH60 = Pens of 60 random cows within the lactation range 1 – 3 week.

HIGH50 = Pens of 50 random cows within the lactation range 4 – 18 week.

HIGHP50 = Pens of 50 random primiparous cows within the lactation range 4 – 18 week.

HIGH200 = Pens of 200 random cows within the lactation range 4 – 18 week.

LOW50 = Pens of 50 random cows within the lactation range 19 – 44 week.

LOW150 = Pens of 150 random cows within the lactation range 19 – 44 week.

[1] Distribution type prediction model. The mean total pen DMI of 100 replicates of the given pen type from the predictive model of the best fitted distribution shape for each replicate.

[2] Empirical predication model (NASEM, 2021). The mean total pen DMI of 100 replicates of the given pen type from the predictive model of a single empirical DMI estimate that is fed uniformly to the population of cows.

642　[3] Index of generated virtual pen type for given lactation range and population size.

643　[4] The observed mean total pen DMI for 100 replicates of each pen type.

644　[5] Mean percentage error of each predicted pen DMI estimate from the observed value for each replicate. Means of percentage error for each pen replicate

645　were compared across both prediction models using Students t-test.

646　[6] Model fit evaluated as the predicted total pen DMI of both prediction models to the observed values of each replicate. Mean square prediction error to

647　evaluate the error in model fit.

648　[7] Proportion of model error as MSPE due to bias (Benchaar et al., 1998).

649　[8] Proportion of model error as MSPE due to the slope of data not equal to 1 (Benchaar et al., 1998).

650　[9] Proportion of model error as MSPE due to random variation in the data (Benchaar et al., 1998).

651　[a-b] Values within a row with different superscripts differ significantly at $P < 0.05$ for t-test comparison of mean percentage error of both prediction models.

652 **Figure captions**

653 **Fig. 1.** Line plots of the mean and median of milk yield, DMI and feed efficiency of

654 dairy cows by week of lactation and dataset. Abbreviations: DMI = Dry matter intake.

655   **A.** Mean and median milk yield (kg) of dairy cows by week of lactation for each

656   dataset in this study separately.

657   **B.** Mean and median DMI (kg) of dairy cows by week of lactation for each

658   dataset in this study separately.

659   **C.** Mean and median feed intake (milk kg/DMI kg) of dairy cows by week of

660   lactation for each dataset in this study separately.

661 **Fig. 2.** Scatter plots of observed versus predicted model diagnostics for each of the 7

662 dairy cow pen types for both the Distribution and NASEM model. Abbreviations: DMI

663 = Dry matter intake. FRESH20 = Pens of 20 random cows within the lactation range

664 1 – 3 week. FRESH60 = Pens of 60 random cows within the lactation range 1 – 3

665 week. HIGH50 = Pens of 50 random cows within the lactation range 4 – 18 week.

666 HIGH50Lact1 = Pens of 50 random primiparous cows within the lactation range 4 –

667 18 week. HIGH200 = Pens of 200 random cows within the lactation range 4 – 18
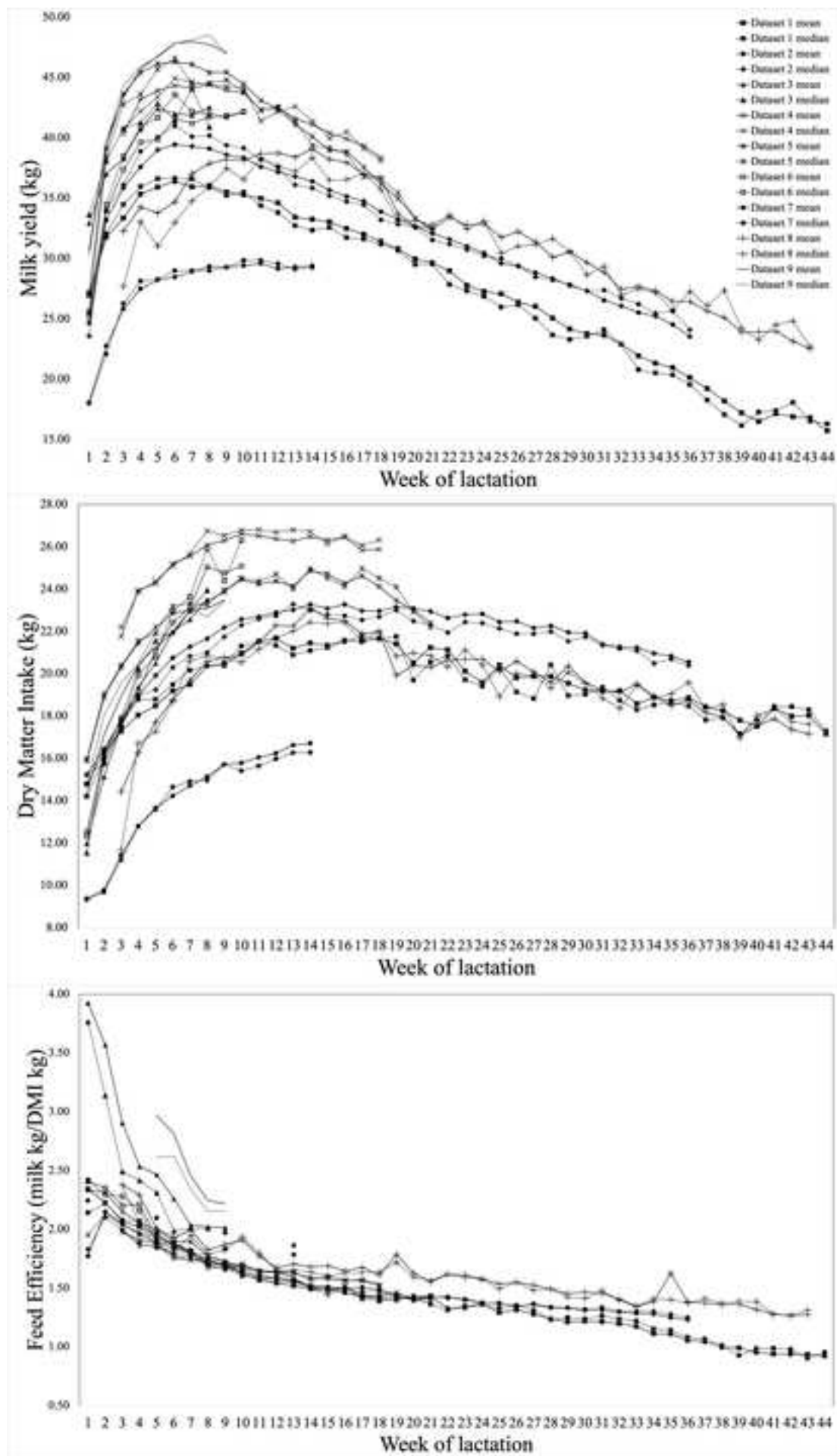
668 week. LOW50 = Pens of 50 random cows within the lactation range 19 – 44 week.
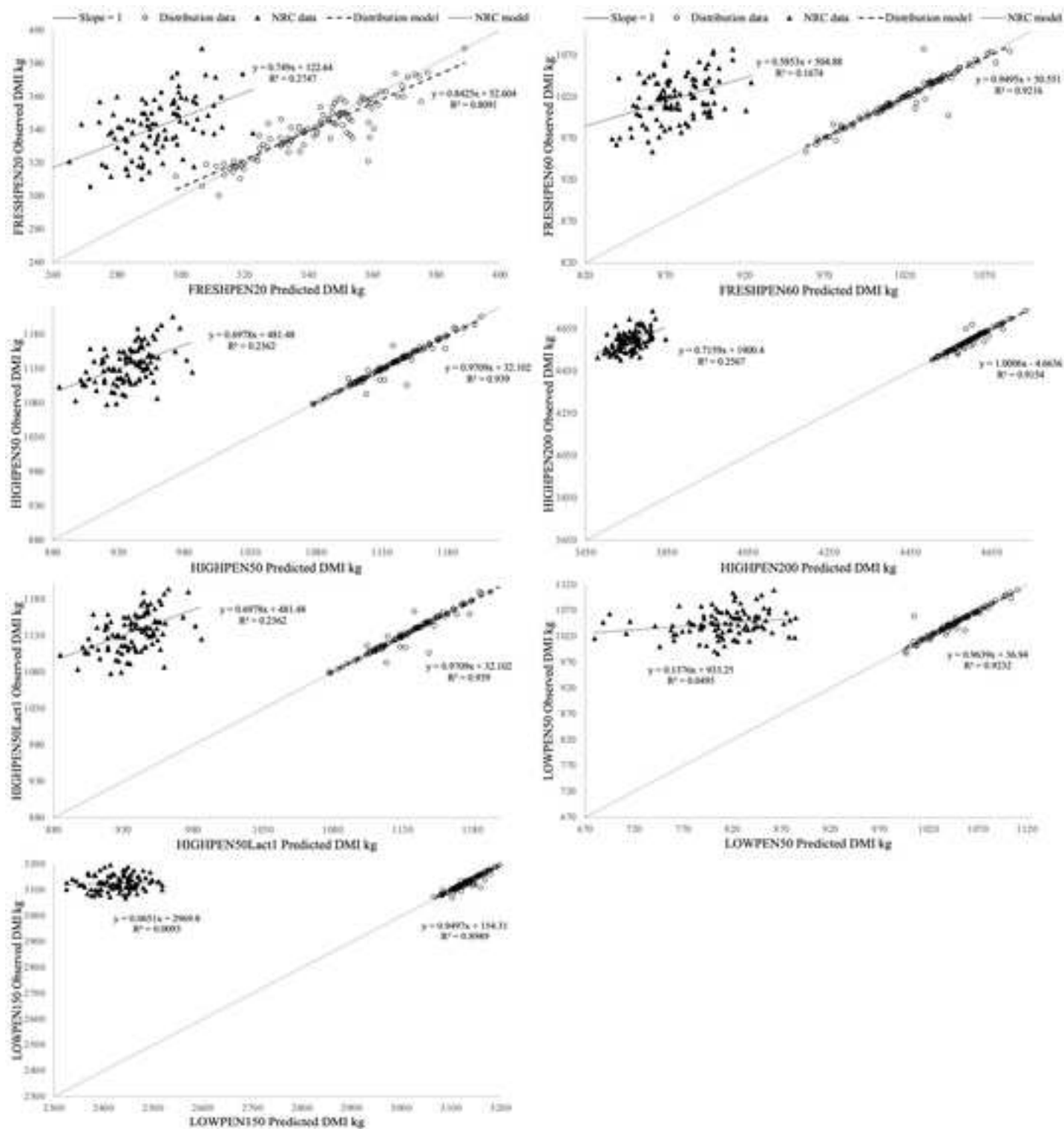
669 LOW150 = Pens of 150 random cows within the lactation range 19 – 44 week.

670

671

1. Is the feed quantity for a pen calculated accurately when using the mean cow?
2. Dry matter intake is not normally distributed for dairy cows by week of lactation.
3. The beta distribution shape best described dry matter intake for pens of cows.
4. The beta had lower error than calculating the pen's mean cow dry matter intake.
5. Increasing precision and accuracy will improve the formulation of pen diets.

Figure1

Click here to access/download;Figure;Figure 1.tiff

Figure2

Click here to access/download;Figure;Figure 2.tiff

animal minor technical revision checklist
Last updated December 2022

Manuscript number: 23-90057R2
Title in Editorial Manager: Describing the distribution type of dry matter intake for dairy cow pens based on pen characteristics
Corresponding author: Heidi Rossow

Please modify your manuscript so that it meets the following requirements, and tick when done.
Pay particular attention to items in bold font. Please highlight all changes made to the manuscript.

Discoverability issues

Tick when done	Requirements
☒	Highlights o Highlights are now requested for research articles, review articles, short communications and opinion papers. o Highlights are a collection of 5 bullet points **(maximum 85 characters, including spaces, per bullet point)** that convey the core messages of your article. o Use the Implications section as the basis to write the Highlights of your paper o We recommend that the bullet points include 1) a presentation of the research question, 2) your main results, and 3) the potential applications of your own results and their field of application for the livestock industry. o No author-defined abbreviations o Please submit Highlights as a <u>separate file</u>, and in Editorial Manager select 'highlights' as the file type.

Formatting issues
Consult Instructions for Authors at <u>https://animal-journal.eu/instructions-and-policies</u> for further guidance.

Tick when done	Requirements
☒	References o **Please give the volume number for all journal citations.** o Please refer to the reference checklist appended below for specific formatting instructions.
☒	Figures o **Please provide Figures 1 and 2, as they are currently missing.**
☒	Illustrations o Double-check the table/figure captions and notes for spelling, typos and grammatical errors, and revise them where relevant. We also advise you to proofread your work carefully to ensure your text is error-free.

☒	Copyright permission
	 o If your manuscript contains copyrighted material (such as figures, illustrations, charts, tables, photographs, and text excerpts), you must
	  1. Obtain permission from the owner to use that material
	  2. Supply the written permission as a 'Copyrights and Permissions' item type when submitting
	  3. Include the corresponding credit in your article
	See guidance and details at https://www.elsevier.com/about/policies/copyright/permissions

☒	Information declared in the submission system (Editorial Manager) should not differ from the contents of the submission: list of authors, title, abstract, keywords, total word count.
☒	Authorship change For any changes in authorship, please contact the Editorial Office at <u>editorialoffice@animal-journal.eu</u> to get the formal request form.

Appendix: Reference checklist

Reference type	Example
Authors' information and publication year *All authors must be listed.*	Author, A., Author, B., Author, C.D., Author, E., Year.
Journal article *Journal names are given in full, not in abbreviated form.* *Issue numbers are not required.*	Author(s), Year. Article title. Full Journal name Volume number, first-last page numbers. • Martin, C., Morgavi, D.P., Doreau, M., 2010. Methane mitigation in ruminants: from microbe to the farm scale. Animal 4, 351-365. • Berry, D.P., Wall, E., Pryce, J.E., 2014. Genetics and genomics of reproductive performance in dairy and beef cattle. Animal 8 (suppl. 1), 115–121. • Knowles, T.G., Kestin, S.C., Haslam, S.M., Brown, S.N., Green, L.E., Butterworth, A., Pope, S.J., Pfeiffer, D., Nicol, C.J., 2008. Leg disorders in broiler chickens: prevalence, risk factors and prevention. PLoS ONE 3, e1545. • Pérez-Enciso, M., Rincón, J.C., Legarra, A., 2015. Sequence- vs. chip-assisted genomic selection: accurate biological information is advised. Genetics Selection Evolution 47, 43. doi:10.1186/s12711-015-0117-5. • When the article is online but not yet printed, the right format is: Zamaratskaia, G., Squires, E.J., 2008. Biochemical, nutritional and genetic effects on boar taint in entire male pigs. Animal, doi:10.1017/S1751731108003674, Published online by Cambridge University Press 17 December 2008.
Book (or official report) *If a publisher is based in more than one place, use only the first one.* *If multiple publishers are list, it is acceptable to use only the first one.*	Author(s)/Editor(s)/Institution, Year. Book title, volume number if more than 1, edition if applicable. Publisher's name, City, State (2-letter abbreviation) for US places, Country. • Association of Official Analytical Chemists (AOAC), 2004. Official methods of analysis, volume 2, 18th edition. AOAC, Arlington, VA, USA. • Littell, R.C., Milliken, G.A., Stroup, W.W., Wolfinger, R.D., 1996. SAS system for mixed models. Statistical Analysis Systems Institute Inc., Cary, NC, USA. • Martin, P., Bateson, P., 2007. Measuring behaviour. Cambridge University Press, Cambridge, UK. • National Research Council (NRC), 2012. Nutrient requirements of swine, 11th edition. National Academy Press, Washington, DC, USA. • Statistical Analysis Systems Institute, 2002. SAS user's guide, version 9.00. SAS Institute Inc., Cary, NC, USA.
Book chapter (or official report part) *If a publisher is based in more than one place, use only the first one.* *If multiple publishers are list, it is acceptable to use only the first one.*	Author(s), Year. Chapter title. In Title of book (ed. Editor, A., Editor, B.). Publisher's name, City, State (2-letter abbreviation) for US places, Country, pp. first-last page numbers. • Nozière, P., Hoch, T. 2006. Modelling fluxes of volatile fatty acids from rumen to portal blood. In Nutrient digestion and utilization in farm animals (ed. Kebreab, E., Dijkstra, J., Bannink, A., Gerrits, W.J.J., France, J.). CABI Publishing, Wallingford, UK, pp. 40–47.
Proceedings (or Conference paper) *The conference location and date must be given.* *If proceedings are published in a journal, the article should be formatted as for a journal article.* *If they have been published as chapters in a book, the article should be formatted as for a chapter in a book.*	Author(s), Year. Paper title. Proceedings of the (or Paper presented at the) XXth Conference title, date of the conference, location of the conference, pp. first-last page numbers or poster/article number. • Bispo, E., Franco, D., Monserrat, L., González, L., Pérez, N., Moreno, T., 2007. Economic considerations of cull dairy cows fattened for a special market. Proceedings of the 53rd International Congress of Meat Science and Technology, 5-10 August 2007, Beijing, China, pp. 581–582. • Martuzzi, F., Summer, A., Malacarne, M., Mariani, P., 2001. Main protein fractions and fatty acids composition of mare milk: some nutritional remarks with reference to woman and cow milk. Paper presented at the 52nd Annual Meeting of the European Association for Animal Production, 26-29 August 2001, Budapest, Hungary.
Website *The accessed date must be given in DD Month YYYY format*	Author(s)/Institution, Year. Document/Page title. Retrieved on DD Month YYYY (i.e. accessed date) from http://www.web-page address (URL). • Bryant, P., 1999. Biodiversity and Conservation. Retrieved on 4 October 1999 from http://darwin.bio.uci.edu/~sustain/bio65/Titlpage.htm
Thesis	Author, A.B., Year. Thesis title. Type of thesis, University with English name, City, State (2-letter abbreviation) for US places, Country (i.e. location of the University). • Vlaeminck, B., 2006. Milk odd- and branched-chain fatty acids: indicators of rumen digestion for optimisation of dairy cattle feeding. PhD thesis, Ghent University, Ghent, Belgium.
Publisher/Conference/University location	Publisher/Organisation's name, City, State (2-letter abbreviation) for US places, Country • AOCS Press, Champaign, IL, USA • International Organization for Standardization, Geneva, Switzerland

Appendix: Example of table format

Table 1

Effect of x and y on z in w; tables should be sufficiently detailed for being understood without any reference to the text, but do not give details of the Material and methods.

| Items | Heading | | | | | | |
| | Heading[1] | | Heading | | | | |
	ColH1[2]	ColH2	ColH3	ColH4	ColH5	RMSE	P-value
Row heading (units)							
Row subheading							
Row sub-subheading	Value	Value	Value	Value	Value	Value	Value
Row sub-subheading	Value	Value	Value	Value	Value	Value	Value
RSSH	Value	Value	Value	Value	Value	Value	Value
Row subheading							
Row sub-subheading	Value	Value	Value	Value	Value	Value	Value
Row sub-subheading	Value[a]	Value[ab]	Value[bc]	Value[c]	Value	Value	Value
Row heading[3]	Value	Value	Value	Value	Value	Value	Value
Row heading	Value	Value	Value	Value	Value	Value	Value
Row heading	Value	Value	Value	Value	Value	Value	Value

Abbreviations: ColH1 = Column heading 1; ColH2 = Column heading 2; ColH3 = Column heading 3; ColH4 = Column heading 4; ColH5 = Column heading 5; RSSH = Row sub-subheading.

[1] Footnote explaining heading.

[2] Footnote explaining column heading 1.

[3] Footnote explaining row heading.

[a,b] Values within a row with different superscripts differ significantly at $P<0.05$.