

Dealing with Errors in Data from Electronic Swine Feeders

D.S. Casey, research assistant,
and J.C.M. Dekkers, associate professor,
Department of Animal Science

ASL-R1777

Summary and Implications

Data from electronic swine feeders contain errors that must be identified, edited, and corrected. The objectives of this study were to develop comprehensive criteria to identify errors in feed intake data from FIRE electronic feeders and to compare the ability of five editing methods to accurately estimate daily feed intake (DFI) and average daily feed intake (ADFI). Data from FIRE feeders on 591 pigs from the National Pork Board's Maternal Line Genetic Evaluation Program were used. Errors in each visit were identified using 16 criteria. To create an error-free data set as a basis for comparison, data from 124 pigs with few errors were selected and visits with errors were replaced by error-free visits from the same pig. Resulting DFI and ADFI were assumed to be the true trait values. Error visits were then introduced, representative of field data. Data were edited using five methods (EM1-5). For EM1, a DFI record was deleted if $DFI < 1000$ g or >4500 g. For EM2-5, the 16 criteria were used to identify errors in each visit. For EM2 and 3, all DFI records with ≥ 1 and ≥ 2 error visits were deleted. For EM4-5, DFI was obtained by summing feed intake over error-free visits. For EM5, DFI records were then adjusted for the effects of presence of error visits on unadjusted DFI, which were estimated from a linear model analysis. For EM1-4, missing DFI records were replaced by linear regression estimates of DFI on test day for each pig. DFI and ADFI from the edited data sets were correlated to true values. Correlations were high (.90 to .99) for both traits for all editing methods except EM1. EM5 had the highest correlation for DFI (.99). EM2 and EM5 had the highest correlations for ADFI (.98 to .99). EM1 had the lowest correlations for both traits (.82 to .93). Results indicate that editing methods affect the accuracy of data from electronic feeders. EM5 is recommended for maximum accuracy for DFI and EM2 is recommended for ADFI for maximum accuracy and ease of implementation.

Introduction

Measuring feed intake in swine herds is beneficial for benchmarking a herd, evaluating a change in a management practice, and making genetic progress in feed efficiency. The first two can be done by measuring feed intake on groups but the last requires measuring feed intake on individual pigs. One way to measure individual feed intake is to house pigs separately. De Haer and Merks (2) and De Haer and De Vries (1), however, showed that pigs housed individually ate more, grew faster, and were fatter. Because performance was different and pigs in commercial herds are housed in groups, feed intake should be measured on pigs housed in groups. Measuring individual feed intake in this setting is difficult.

Electronic feeders make it possible to automatically measure individual feed intake on pigs housed in groups. These feeders weigh the feed trough frequently and only allow one pig to eat at a time. Each pig is identified by the feeder through an electronic ear tag. When a pig enters and exits the feeder, the time and weight of the feed trough are recorded. These data are used to calculate daily feed intake and feed intake over an entire test period. There are three brands of electronic feeders that are available commercially (FIRE-Osborne Industries Inc., IVOG-HokoFarm (3), and ACEMO 48)

A concern about using electronic feeders is that data from these feeders contain errors (5). These errors are caused by equipment malfunctions and by interactions of the animal with the feeder. Because each visit is recorded, data sets from electronic feeders are usually large, which means finding errors can be time consuming. To obtain an accurate measure of feed intake, editing methods are required that efficiently identify, edit, and correct errors. Several alternative editing methods have been developed. Eissen et al. (5) developed nine criteria to identify errors in each visit for the IVOG electronic feeder. Eissen et al. (4) used these criteria to identify errors and then developed a method to edit and correct them. In this method, any daily feed intake record from a pig was discarded if at least one visit during that day contained an error. The remaining daily feed intake records were regressed on test day and the resulting line was used to estimate daily feed intake for the

discarded records. Software used by the FIRE electronic feeder uses similar but simpler criteria to identify errors. Visits that contain errors are discarded and the remaining visits are summed per pig per day to calculate daily feed intake. This method assumes that no feed is consumed during error visits. The ability of different editing methods to identify and correct errors and to accurately estimate daily feed intake is unclear. In addition, the impact of error visits on true and estimated feed intake is unclear, in particular for the FIRE feeder, which is used extensively in the United States. The objectives of this study, therefore, were as follows:

1. Develop more comprehensive criteria to identify errors in feed intake data from FIRE electronic feeders.
2. Evaluate and compare the ability of five editing methods to accurately estimate daily feed intake and average daily feed intake.

Materials and Methods

Data used in this project were provided by the National Pork Board and came from the Maternal Line Evaluation

Table 1. Criteria used for identifying errors in feed intake data from electronic swine feeders.

Error Type	Feed Intake Variable	Visits Involved	Criteria Used for Classifying a Visit as an Error
1	Feed intake per visit (FIV)	All	FIV < -20 g
2		All	FIV > 2000 g
3		Occupation time per visit = 0 s	Abs (FIV) > 20 g
4	Occupation time per visit (OTV)	All	OTV < 0 s
5		All	OTV > 3600 s
6	Feeding rate per visit (FRV) ^a	0 < FIV < 50 g	FRV > 500 g/min
7		FIV ≥ 50 g, preceding or following a visit with FIV < -20 g	FRV > 110 g/min
8		FIV ≥ 50 g, not preceding or following a visit with FIV < -20 g	FRV > 170 g/min
9		FRV = 0 g/min	OTV > 500 s
10		All except FRV = 0 g/min	Abs (FRV) ≤ 2 g/min
11		Leading weight difference (LWD) ^b	All except last visit of each feeder in test period
12	All except last visit of each feeder in test period		LWD > 1800 g
13	Following weight difference (FWD) ^c	All except first visit of each feeder in test period	FWD < -20 g
14		All except first visit of each feeder in test period	FWD > 1800 g
15	Leading time difference (LTD) ^d	All except last visit of each feeder in test period	LTD < 0 s
16	Following time difference (FTD) ^e	All except first visit of each feeder in test period	FTD < 0 s

^aFeed rate per visit was not calculated for occupation time per visit ≤ 0 s.

^bLeading weight difference = entrance weight of trough of following visit - exit weight of trough of present visit.

^cFollowing weight difference = entrance weight of trough of present visit - exit weight of trough of preceding visit.

^dLeading time difference = entrance time of following visit - exit time of present visit.

^eFollowing time difference = entrance time of present visit - exit time of preceding visit.

Program (6). These data were from 591 crossbred pigs that represented six breeds (one sire line by six maternal lines), two sexes (barrows and gilts), and two replicates. Growth and feed intake were measured on each pig starting at an average weight of 112 lb and ending at an average weight of 252 lb. Pigs were housed in pens that contained an average of 13.7 pigs per pen. Feed intake data were collected using the FIRE electronic swine feeder.

Sixteen criteria were developed to identify errors in the feed intake data (Table 1). Development started by identifying five variables that could be affected by errors: feed intake per visit (FIV), occupation time per visit (OTV), feeding rate per visit (FRV), feed trough weight differences between adjacent visits in time (LWD and FWD), and time differences between adjacent visits in time (LTD and FTD). Within each category several criteria (usually high and low) were established to determine whether a visit contained an error (Table 1). Criteria were determined based on histograms, knowledge of the feeder, or based on Eissen et al. (5). All sixteen criteria were then used to identify visits that contained errors.

To evaluate the five different editing methods, a data set free of errors was created (SET1). SET1 was created in several steps in a manner to ensure that it was representative of field data. First, the following criteria were used to identify pigs with few errors for inclusion in SET1:

1. At least 85% of all daily feed intake records over the test period had no visits with errors
2. At least five error-free daily feed intake records during the first and last 7 days of the test period
3. No more than three daily feed intake records in a row that contain an error
4. Length of the test period ≥ 70 days

All four of these criteria were met by data from 124 of the 591 pigs. Next, any visits that contained an error were replaced with an error-free visit from the same pig and time period. Also, days for which all records were missing were replaced with a daily feed intake record from the same pig and time period. Daily feed intake (DFI) was calculated by summing FIV for each day within a pig. Average daily feed intake (ADFI) was calculated by averaging DFI for each

pig. These values were assumed to be the true trait values. There were 64,234 visits in SET1 of which 97.8% were original data. This indicates that SET1 represents field data.

The next step was to introduce errors into SET1, which then allowed the different editing methods to be evaluated with regard to their ability to correct for these errors. For each pig in SET1 a pig was chosen at random from the list of 591 pigs. Visits containing errors from that pig were used to replace visits in the same test day for the pig in SET1. This method of introducing errors created a data set (SET2) that was a true representative of the number of errors per day, the types of errors, the values of the errors, and the distribution of errors over time that occurs in field data.

Five editing methods (EM1-5) were then used to identify, edit, and correct errors in SET2. The different methods are described in Table 2. The 16 criteria, previously described, were used to identify errors in visits for EM2-5. Criteria described in Table 2 were used to identify errors in DFI records for EM1. For EM1 and 2, DFI records were discarded if the record contained an error. One error visit was allowed for EM3. For EM4 and 5, visits that contained errors were discarded and FIV from the remaining visits was summed. This was used as DFI for EM4, assuming no feed was consumed during error visits. For EM5 the sum was adjusted for the percentage of errors within that day on the basis of estimates from a linear model. The model included breed, sex, week, estimated live weight, and average daily gain as fixed effects, pig as a random effect, and covariates related to the 16 error types. The solutions for the covariates related to the error types were then used to adjust daily feed intake (feed intake summed over error-free visits) for errors. For all methods except EM5, days without DFI or with discarded DFI were replaced by estimates obtained from regressing DFI records on test day for each pig.

Data on DFI and ADFI obtained from each editing method were correlated to their true values obtained from SET1 to evaluate the ability of editing methods to accurately estimate DFI and ADFI.

Table 2. Description of editing methods used to identify, edit, and correct errors.

Editing Method	Identification of Errors	Editing of Errors	Correction of Errors
1	DFI < 1000, >4500, g	discard DFI record	linear regression
2	16 criteria ^a	discard DFI record with error	linear regression
3	16 criteria	discard DFI record with > 1 error	linear regression
4	16 criteria	discard visits with errors	linear regression
5	16 criteria	discard visits with errors	adjust for errors

^aCriteria are described in Table 1.

Results and Discussion

The original data set contained 290,073 visits from 591 pigs (491 visits per pig). A total of 16,288 (5.6%) visits contained errors, which is similar to the 5.7% found by Eissen et al. (5). Table 3 contains the number of visits that contained a specific error type. Errors ranged from 5 visits (0.00%) to 2,919 visits (1.01%). Error types 1, 10, 12, and 14 had the highest frequency. Of the 44,981 DFI records, 17.2% contained at least one visit with an error, which is less than the 29% found by Eissen et al. (5). This indicates that the number of errors was very similar but was more concentrated in fewer DFI records than what was found by Eissen et al. (5).

Table 3. Frequency of each error type.^a

Error Type	Frequency	Percentage ^b
1	2,319	0.80
2	1,523	0.53
3	23	0.01
4	5	0.00
5	151	0.05
6	76	0.03
7	640	0.22
8	2,046	0.71
9	444	0.15
10	2,918	1.01
11	2,244	0.77
12	2,780	0.96
13	2,236	0.77
14	2,919	1.01
15	1,325	0.46
16	1,136	0.39

^aError types are listed in Table 1.

^bThere were 290,073 visits.

Correlations of true DFI and ADFI (SET1) with estimated DFI, ADFI (SET2 after editing methods used) are shown in Figure 1. Correlations ranged from .82 to .96 for DFI and from .93 to .99 for ADFI. Editing method 5 had the highest correlation for both traits. Editing methods 2 and 3 had the next highest correlations and were very similar to correlations from EM5, especially for ADFI. Editing method 1 had the lowest correlations.

Results indicate that editing methods affect the accuracy of data that comes from electronic feeders. For the trait DFI, EM5 is recommended because of the highest correlation but for the trait ADFI, EM2 is recommended because of the high correlation and its ease of implementation.

Acknowledgments

We are grateful to the National Pork Board for providing the data used in this project.

References

- De Haer, L. C. M. and A. G. De Vries. 1993. Feed intake patterns of and feed digestibility in growing pigs housed individually or in groups. *Livestock Production Science* 33: 277–292.
- De Haer, L. C. M. and J. W. M. Merks. 1992a. Patterns of daily food intake in growing pigs. *Anim. Prod.* 54: 95–104.
- De Haer, L. C. M., J. W. M. Merks, H. G. Kooper, G. A. J. Buiting, and J. A. Van Hattum. 1992b. A note on the IVOG®-station: a feeding station to record the individual food intake of group-housed growing pigs. *Anim. Prod.* 54: 160–162.

4. Eissen, J. J., A. G. De Haan, and E. Kanis. 1999. Effect of missing data on the estimate of average daily feed intake of growing pigs. *J. Anim. Sci.* 77: 1372–1378.
5. Eissen, J. J., E. Kanis, J. W. M. Merks. 1998. Algorithms for identifying errors in individual feed intake data of growing pigs in group-housing. *Appl. Eng. Agric.* 14: 667–673.
6. Goodwin, R. and D. Boyd, Eds. (2000). Maternal Line National Genetic Evaluation Program Results. Des Moines, Iowa.

Figure 1. Accuracy of five editing methods at estimating daily feed intake (DFI) and average daily feed intake (ADFI).

