


ORIGINAL ARTICLE

Effects of correcting missing daily feed intake values on the genetic parameters and estimated breeding values for feeding traits in pigs

Tetsuya ITO,^{1,2}  Kazuo FUKAWA,¹ Mai KAMIKAWA,¹ Satoshi NIKAIDOU,³ Masaaki TANIGUCHI,⁴ Aisaku ARAKAWA,⁴ Genki TANAKA,¹ Satoshi MIKAWA,⁴ Tsutomu FURUKAWA⁵ and Kensuke HIROSE⁶

¹Central Research Institute for Feed and Livestock, ZEN-NOH (National Federation of Agricultural Cooperative Associations), Kamishihoro, Hokkaido, ²Graduate School of Agriculture, ³Faculty of Agriculture, Tokyo University of Agriculture, Atsugi, Kanagawa, ⁴Technical Section, ZEN-NOH LIVESTOCK CO., LTD., Koto Ward, Tokyo, ⁵Institute of Livestock and Grassland Science, National Agriculture and Food Research Organization, Tsukuba, Ibaraki, and ⁶East Japan GGP Farm, ZEN-NOH LIVESTOCK CO., LTD., Shizukuishi, Iwate, Japan

ABSTRACT

Daily feed intake (DFI) is an important consideration for improving feed efficiency, but measurements using electronic feeder systems contain many missing and incorrect values. Therefore, we evaluated three methods for correcting missing DFI data (quadratic, orthogonal polynomial, and locally weighted (Loess) regression equations) and assessed the effects of these missing values on the genetic parameters and the estimated breeding values (EBV) for feeding traits. DFI records were obtained from 1622 Duroc pigs, comprising 902 individuals without missing DFI and 720 individuals with missing DFI. The Loess equation was the most suitable method for correcting the missing DFI values in 5–50% randomly deleted datasets among the three equations. Both variance components and heritability for the average DFI (ADFI) did not change because of the missing DFI proportion and Loess correction. In terms of rank correlation and information criteria, Loess correction improved the accuracy of EBV for ADFI compared to randomly deleted cases. These findings indicate that the Loess equation is useful for correcting missing DFI values for individual pigs and that the correction of missing DFI values could be effective for the estimation of breeding values and genetic improvement using EBV for feeding traits.

Key words: estimated breeding values, feeding traits, genetic parameter, pig, regression equation.

INTRODUCTION

Feed accounts for approximately 60% of pig production costs and has recently increased in price in Japan; thus, improving feed efficiency is of great importance for the pork industry. The feed conversion ratio (FCR) and/or residual feed intake (RFI) are used as indices of feed efficiency (Koch *et al.* 1963). However, to calculate feed efficiency traits, it is necessary to record the daily feed intake (DFI), which is difficult to measure accurately for an individual pig due to the high labor and time requirements. Electronic feeder systems have recently become available, allowing individual feeding traits, such as individual feed intake, feeding time, and the number of feeds, to be recorded. Thus, these systems are useful for collecting DFI data for individual pigs to

estimate genetic parameters and improve feed efficiency traits (Gilbert *et al.* 2007; Cai *et al.* 2008; Bunter *et al.* 2010; Young *et al.* 2011; Do *et al.* 2013; Jiao *et al.* 2014).

However, unfortunately, these systems often malfunction for functional, environmental and operative reasons (De Haer *et al.* 1992; Eissen *et al.* 1998; Casey *et al.* 2005), resulting in missing DFI records, which makes it difficult to calculate accurate feed efficiency values. Therefore, the correction of missing feed

Correspondence: Tetsuya Ito, Central Research Institute for Feed and Livestock, ZEN-NOH (National Federation of Agricultural Cooperative Associations), Kamishihoro, Hokkaido 080-1406, Japan.
(Email: itou-tetsuya@zennoh.or.jp)

Received 1 December 2016; accepted for publication 10 July 2017.

intake values is imperative for the effective use of these systems. Mathematical analysis methods are used across various fields to optimize the estimation or prediction of missing data, with both parametric and nonparametric polynomial regression models being particularly useful in biological studies. For example, fetal weight can be estimated using polynomial regression (Schild *et al.* 2000; Salomon *et al.* 2007), and growth curves can be analyzed using polynomial, Gompertz, logistic and von Bertalanffy equations (Köhn *et al.* 2007; Koivula *et al.* 2008; Jiao *et al.* 2014). Therefore, it is likely that missing DFI values for individual pigs could be estimated using a polynomial regression equation that incorporates their age, because individual feeding traits change according to age and individual characteristics.

In this study, we determined: (i) the correction methods on missing feed intake values using parametric and nonparametric regression equations; (ii) the effects of the missing and corrected DFI on genetic parameters and estimated breeding values (EBV) for average DFI (ADFI); and (iii) the heritability and genetic correlations for ADFI and production traits based on actual measured DFI values including missing or corrected DFIs in a Duroc pig population.

MATERIALS AND METHODS

Animals

A purebred Duroc pig population originating from the East Japan Great Grand Parent (GGP) Farm of ZEN-NOH Livestock Co., Ltd. (Iwate, Japan; Farm 1) and the Central Research Institute for Feed and Livestock ZEN-NOH (Hokkaido, Japan; Farm 2) was used in this study. This population was derived from 28 boars and 52 gilts and was taken to the sixth generation using a closed nucleus breeding system in Farm 2 from 2005, as described by Hirose *et al.* (2009, 2011).

Data collection

Feed Intake Recording Equipment (FIRE; Osborne Industries Inc., Osborne, KS, USA) electronic feeders were used to measure the individual feed intake and weight gain of Duroc pigs. Only one pig could take feed from each feeder at any one time due to its shape. Individual feed intake data were recorded under an *ad libitum* feeding regime. In total, 40 pens in Farm 1 were equipped with a FIRE electronic feeder. Each pen contained 4–6 pigs with free access to feed and water at all times. The pigs increased in weight from approximately 30 to 100 kg body weight (BW) during the test period, with the feed composition being changed in the latter period of growth (when they reached 60–70 kg BW). The growth and feed intake of 1622 pigs (1462 boars and 160 gilts) were measured throughout the test period from 2010 to 2014.

In addition, average daily gain (ADG) was calculated on the basis of gained weight from 30 to 100 kg BW. When the pigs reached approximately 100 kg BW, back fat thickness (BFT), loin eye muscle area (EMA), and intramuscular fat content (IMF) were measured at a half body length position using a real-time ultrasound machine (SSD-500; Aloka Co., Ltd., Tokyo, Japan) in Farm 1, and the ultrasound images were analyzed using SigmaScan Pro 5.0 (Biotronics Inc., Ames, IA, USA). ADG, BFT, EMA, and IMF data from Farm 1 were collected from 1811, 1824, 1824 and 579 animals, respectively (Table 1). All data were used for the genetic parameter estimates.

Data editing for feed intake

Four criteria were used by the FIRE software to filter out erroneous data: feed intake per visit (FIV) <−300 g and >1500 g; and occupation time per visit (OTV) <0 s and >2 h. In addition, any DFI data of <0 g and >10 kg were excluded as measurement errors due to these being physically impossible.

Correction of missing DFI data

To estimate the DFI for each individual pig, parametric regression models were fitted using the quadratic polynomial and fourth degree orthogonal polynomial equations, as follows:

$$Y = \beta_0 + \beta_1 X^1 + \beta_2 X^2$$

$$Y = \beta_0 + \beta_1 \Phi_1(X) + \beta_2 \Phi_2(X) + \beta_3 \Phi_3(X) + \beta_4 \Phi_4(X)$$

where Y is the estimated DFI at test age X , β_{0-4} are function parameters, and $\Phi_n(X)$ are orthogonal polynomials of age X .

In addition, a nonparametric regression model was fitted using a locally weighted regression equation (Loess). The Loess procedure was fitted to the DFI with two degrees of freedom and a linear model according to the following model (Cleveland 1979):

$$y_i = g(x_i) + \varepsilon_i$$

where y_i is the fitted DFI value of the first degree polynomial fitted to the data using weighted least squares with weights of $w(x_i)$ at each age x_i , g is a smoothing function, and ε_i are random variables with mean 0 and a constant scale. The weighted function $w(x_i)$ was calculated using the following model:

$$w(x_i) = \left(1 - \left(\frac{|x - x_i|}{\max|x - x_i|} \right)^3 \right)^3$$

where x defines the predictor position where the above model is being evaluated, x_i defines the

Table 1 Average and standard deviation for production traits in Duroc pigs

Trait†	Unit	Number	Males	Females	Total
BW	kg	1856	102.16 ± 6.45	101.75 ± 6.89	102.11 ± 6.51
ADG	g/day	1722	1021.40 ± 106.40	921.11 ± 98.01	1012.11 ± 109.57
BFT	cm	1811	1.93 ± 0.30	2.20 ± 0.34	1.96 ± 0.32
EMA	cm ²	1811	34.31 ± 3.34	34.77 ± 3.41	34.37 ± 3.35
IMF	%	559	2.98 ± 0.62	–	2.98 ± 0.62

†BW, body weight in individuals at the measured production traits; ADG, average daily gain in individuals from 30 to 100 kg body weight; BFT, back fat thickness; EMA, loin eye muscle area; IMF, intramuscular fat content.

measurement position, $\max|x-x_i|$ is the maximum distance between a predictor point and the measurement point within the fit region, and x_i is the nearest neighbor of x within the fit region.

All model fitting was conducted with the statistical software GenStat 17.1 (VSN International Ltd., Hemel Hempstead, UK).

Simulation of missing DFI records

To compare the different methods for correcting missing DFI records, randomly deleted DFI datasets were generated from true-measured DFI datasets based on 902 pigs with non-missing measured DFI records (total 60 182 records). A proportion of 5–50% of the true-measured DFI records was randomly deleted in 5% increments, with the random deletion repeated twice at each level to give 20 random deletion DFI datasets in total. Each function was then fitted to the corrected DFI records for each pig using each randomly deleted dataset.

Estimation of genetic parameter and breeding value

To verify the effect of correcting missing DFI records, the genetic parameters and breeding values for ADFI were estimated for each of the 902 pigs from the true-measured, randomly deleted and corrected DFI cases. Further, we estimated genetic parameters of the following five groups: the true-measured ADFI group (tru-ADFI), which included pigs with no missing DFI records ($n = 902$); missing ADFI group (mis-ADFI), which included pigs with missing DFI records ($n = 720$); and corrected ADFI group (cor-ADFI), which used the regression equation-corrected values from the mis-ADFI ($n = 720$). In addition, a raw-ADFI group (raw-ADFI) and an imputed-ADFI group (imp-ADFI) were included in this analysis. The raw-ADFI used both tru-ADFI and mis-ADFI values obtained from all sampled individuals, including missing values ($n = 1622$). The imp-ADFI used both tru-ADFI and cor-ADFI values obtained from all sampled individuals, including corrected values ($n = 1622$). The ADG, BFT, EMA and IMF phenotypes were mentioned earlier. The pedigree

information for 17 198 animals in both farms was used to estimate the genetic parameters. The genetic and residual variances for each individual trait were estimated using the average information restricted maximum likelihood (AI-REML) procedure in AIR-EMLF90 program software (Misztal *et al.* 2002), according to the following model:

$$Y = X_b + Z_u + e$$

where Y is the phenotypic value of each trait, b is a vector of fixed effect (sex from male, female, and castrated; and birth year and month, from May 2010 to November 2014), u is the random additive genetic effect, X and Z are incidence matrices relating b and u to y , and e is the random residual effect. This model was used for ADG, IMF and ADFI, while the models for BFT and EMA also included BW as a covariate.

The genetic variance components for ADFI and production traits were estimated using a bivariate animal model with the same fixed effects as described above.

Estimated breeding values for the true-measured, randomly deleted, and corrected ADFI cases of 902 pigs were predicted by best linear unbiased predictor (BLUP) of single trait animal model using BLUPF90 program (Misztal *et al.* 2002). Genetic parameter for EBV of each ADFI case was used as mentioned earlier.

Evaluation methods of missing and corrected DFI values

To compare the suitability of the three equations to each randomly deleted dataset, total feed intake (TFI) for each of the 902 pigs was calculated from the true-measured and corrected DFI values for each dataset. The correlation coefficient (r), Akaike's Information Criterion (AIC), and Bayesian Information Criterion (BIC) of each corrected TFI from three equations to the true-measured TFI were calculated by using GenStat 17.1 software (VSN International Ltd.). The true-measured TFI is defined as the objective variable, and each corrected TFI from the three equations is defined as the explanatory variable; AIC and BIC were calculated by the mixed regression

model. Similarly, to determine the effect of DFI that has been randomly deleted or corrected on EBV for ADFI, the information criteria are calculated between true-measured EBV and both randomly deleted and corrected cases. EBV from the true-measured ADFI is defined as the objective variable, and each EBV from the randomly deleted and corrected cases is defined as the explanatory variable; AIC and BIC were calculated using the mixed regression model. These AICs and BICs to the true-measured values were calculated using the following model:

$$\text{AIC} = 2\ln L + 2k$$

$$\text{BIC} = -2\ln L + k\ln(n)$$

where $\ln L$ is the maximum log likelihood for REML mixed model regression estimation and k is the number of parameters.

Also, to evaluate effects of the presence or absence of a missing DFI correction on change in EBV ranking for ADFI, Spearman's rank correlation coefficient (r_s) was calculated between EBV for true-measured ADFI and for randomly deleted ADFI or for corrected ADFI.

r_s was converted into Fisher's z transformation (Z_r) using the following model:

$$Z_r = \frac{1}{2} \ln \left(\frac{1+r}{1-r} \right)$$

And then, the test statistic values Z were calculated using the following formula:

$$Z = \frac{Z_{r1} - Z_{r2}}{\sqrt{\frac{1}{n_1-3} + \frac{1}{n_2-3}}}$$

Z was used to compare the r_s of randomly deleted and r_s of corrected cases in the same deleted DFI proportion. Significance tests were performed using a two-tailed test from standard normal distribution; differences were considered significant at $P < 0.05$.

RESULTS

Proportion of missing DFI records

In terms of the proportion of missing DFI records, 2233 missing DFI records were included in the missing DFI group (Table 2). The missing DFI records showed an average of 3.1% per pig in the missing DFI group. The percentage of individuals who had at least one error was 44.4% of the total of 1622 individuals. The non-missing group with no missing DFI records had DFI collection data for 34–100 days per pig, with a mean test period of 66.7 days, while the mean test period for the missing DFI group was 3.2 days longer.

A histogram of the number of missing DFI records is presented in Figure 1. The maximum number of

Table 2 Summary statistics for daily feed intake (DFI) values in Duroc pigs

Item	Non-missing†	Missing†	Total
No. of DFI records	60 182	50 294	110 476
No. of DFI collection pigs	902	720	1622
Minimum test period (days)	34	26	26
Maximum test period (days)	100	112	112
Mean test period (days)	66.7	69.9	68.1

†Non-missing, based on the measured DFI obtained from 902 individuals with no missing DFI records. Missing, based on the measured DFI obtained from individuals with missing values ($n = 720$).

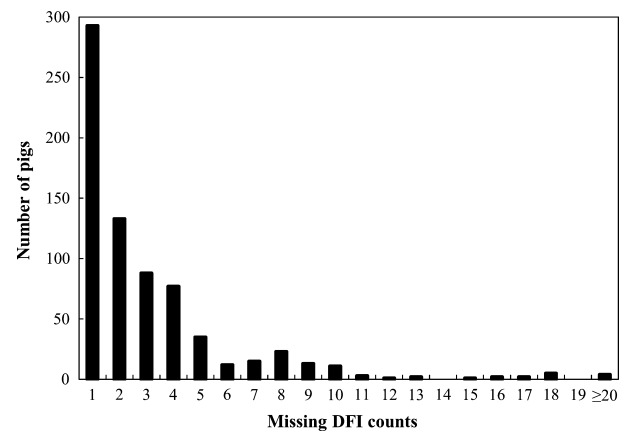


Figure 1 Histogram for the number of missing daily feed intake (DFI) records in Duroc pigs. The vertical axis indicates the number of pigs having missing DFI records and the horizontal axis indicates the missing DFI counts per pig.

missing DFI records per pig was 35, which was equivalent to approximately 50% of the mean number of DFI records.

Goodness of fit of corrected DFI values

According to the results of a simulation where data were randomly deleted from the true-measured 902 pigs with no missing DFI records, the correlation coefficients between the true-measured and the corrected TFIs during the test period ($n = 902$) ranged from 0.981 to 0.999 and decreased with an increased deletion rate (Table 3). The Loess equation showed the minimum values of AIC and BIC to the true-measured TFI among the equations in all randomly deleted datasets. Consequently, the Loess-corrected DFI values were used for the subsequent genetic parameters and breeding values estimation of feeding traits for the missing records. Figure 2 shows a comparison of the actual observed DFI values for

Table 3 Correlation and information criteria of daily feed intake (DFI) between true-measured values with no missing records and corrected values by three different equations from randomly deleted datasets using three different classification criteria over differing proportions of DFI deleted values

Equation†	Classification criteria§	Proportion of DFI values deleted‡									
		5%	10%	15%	20%	25%	30%	35%	40%	45%	50%
Loess	<i>r</i>	0.999	0.998	0.997	0.996	0.994	0.994	0.990	0.989	0.985	0.983
	AIC	873	1619	2034	2241	2523	2651	3012	3147	3383	3535
	BIC	878	1623	2039	2246	2528	2656	3016	3152	3387	3539
Orthogonal	<i>r</i>	0.999	0.998	0.997	0.996	0.994	0.993	0.990	0.988	0.984	0.981
	AIC	890	1663	2062	2285	2565	2727	3076	3228	3488	3601
	BIC	895	1667	2066	2290	2569	2732	3081	3233	3492	3606
Quadratic	<i>r</i>	0.999	0.998	0.997	0.996	0.994	0.993	0.990	0.988	0.985	0.982
	AIC	916	1654	2068	2272	2561	2673	3048	3175	3421	3584
	BIC	920	1659	2073	2277	2566	2678	3053	3179	3426	3589

†Loess, locally weighted regression; orthogonal, fourth-degree orthogonal polynomial equation; quadratic, quadratic polynomial equation.
‡Two randomly deleted DFI datasets were generated by randomly deleting a proportion of the records from the 902 true-measured pigs with no missing values. The table reflects the two randomly deleted datasets averaged for each proportion within the table. There were 18 040 iterations per equation. §*r*, correlation coefficient; AIC, Akaike's Information Criterion; BIC, Bayesian Information Criterion.

two pigs and the corrected values using the Loess equation.

Effects of missing and corrected DFI to the variance components and the EBV for ADFI

Genetic variances for ADFI ranged from 0.024 to 0.025 in both randomly deleted and Loess-corrected cases, similar to the 0% deleted true-measured value (0.024; Fig. 3). Residual variances for ADFI ranged from 0.032 to 0.035 in both cases, similar to the 0% true-measured value (0.033). The heritability ranged from 0.41 to 0.44 in both cases, similar to the 0% true-measured value (0.42). There were no differences in the standard errors of variance components and heritability among the cases.

With respect to the EBV for ADFI, r_s of both randomly deleted and Loess-corrected cases declined as the proportion of deleted DFI increased (Table 4). A comparison of r_s of EBV for ADFI between randomly deleted and Loess-corrected cases in the same deleted proportion showed that Loess-corrected cases had significantly higher values than randomly deleted cases for all deletion rates ($P < 0.05$). Compared with the information criteria of EBV for ADFI between the true-measured, no missing DFI and randomly deleted or Loess-corrected DFI cases, both AIC and BIC in Loess-corrected cases had smaller values than those of randomly deleted cases throughout the DFI deleted proportion.

Variance components, heritability and genetic correlations

The average and standard deviation for ADFI values from the five groups (tru-ADFI, mis-ADFI, cor-ADFI, raw-ADFI and imp-ADFI) had no significant

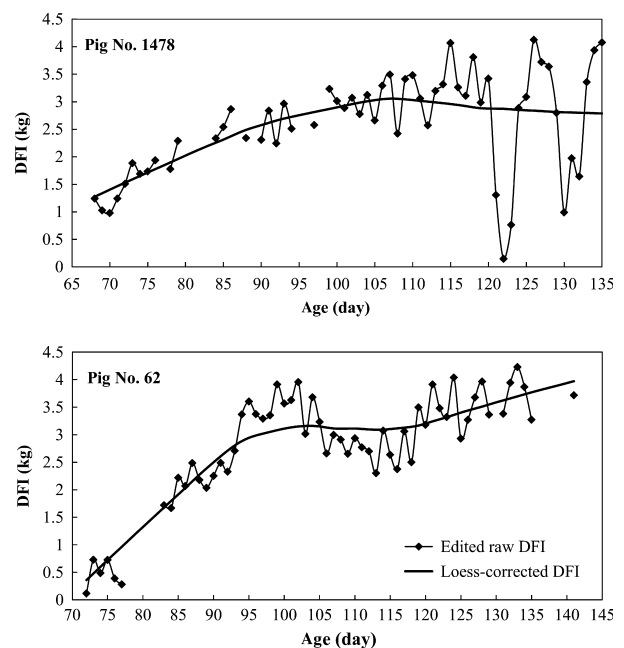


Figure 2 Edited serial daily feed intake (DFI; kg) datasets containing missing records and corrected serial DFI datasets using the Loess function plotted against age (days) for two Duroc pigs. The scatter diagrams represent the edited raw DFI against age for two pigs with missing DFI records, while the lines represent the corrected DFI by the Loess function with age. The vertical axis indicates DFI (kg) and the horizontal axis indicates age (day).

differences, regardless of whether they had missing DFI correction (Table 5). Based on the results of estimation of variance components and heritability for ADFI from each of the five groups, tru-ADFI exhibited both the lowest value for residual variance and the highest value for heritability among all groups (Table 6). There were no changes in variance components and heritability based on the correction of

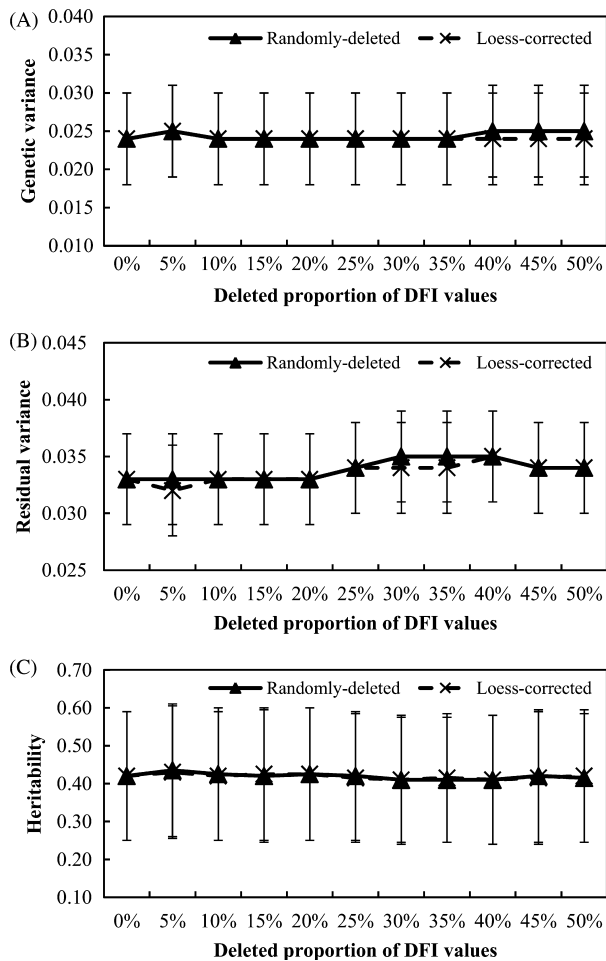


Figure 3 Genetic variance, residual variance, and heritability relative to proportion of deleted daily feed intake (DFI) values for both randomly deleted DFI and Loess-corrected DFI cases. In the randomly deleted DFI cases, variance components and heritability for ADFI were estimated from each randomly deleted DFI dataset. In locally weighted regression equation-corrected (Loess-corrected) cases, variance components and heritability for ADFI were estimated using correction values from the same dataset as the randomly deleted cases. Average values obtained from the two datasets in each deleted DFI proportion are shown. Triangles with lines represent variance components and heritability of the randomly deleted cases, whereas cross marks with dotted lines represent those of the Loess-corrected cases. (A) Genetic variance components were estimated. (B) Residual variance components were estimated. (C) Heritabilities were estimated.

missing DFI. Among the five groups, the standard errors for both variance components and heritability showed the lowest values in raw-ADFI and imp-ADFI. Both of the raw-ADFI and imp-ADFI had the largest number of individuals. Inversely, mis-ADFI and cor-ADFI showed the highest values for standard error in both variance components and heritability with the lowest number of individuals. The standard

errors showed intermediate values in the tru-ADFI, which had a moderate number of individuals. Heritability for ADFI from the five groups was moderate to high.

In the genetic correlations between feeding and production traits, highly positive genetic correlations were estimated between ADFI and ADG (Table 7). Between ADFI and BFT and between ADFI and IMF, moderately positive genetic correlations were estimated. Conversely, the value between ADFI and EMA was moderately negative. The genetic correlation between feeding traits and production traits showed no significant difference between raw-ADFI and imp-ADFI.

DISCUSSION

Regarding missing DFI values using electronic feeder systems, it has previously been shown that the FIRE system results in 17%–50% of DFI records containing at least one error (Casey *et al.* 2005) compared with 29%–35% for IVOG[®] feeding stations (Eissen *et al.* 1998). Similarly, our results demonstrated that over 40% of individuals had missing DFI values. Although these studies used different editing algorithms to identify errors, they indicate that at least 10% or more individuals had missing values.

In this study, the maximum number of missing DFI records per pig was approximately 50%, and many individuals had over 5% missing DFI records. Therefore, we evaluated using randomly deleted DFI datasets with 5–50% to represent all individuals with missing DFI records. Eissen *et al.* (1999) reported that the first-degree polynomial equation is more effective for estimating missing DFI values than the third degree polynomial equation and the nonlinear equation, based on correlations between true ADFI and estimated ADFI values. Our results indicate that the nonparametric Loess equation is more suitable than the parametric quadratic and orthogonal polynomial equations by both AIC and BIC. Loess is a regression analysis approach that uses local fitting (Cleveland & Devlin 1988), and the key strength of nonparametric smoothers is their high degree of flexibility (Jacoby 2000). Since we used different fitting methods for the missing DFI values from Eissen *et al.* (1999), it is impossible to compare our findings with theirs. However, our results on actual missing DFI proportions indicate that the Loess equation is one of the most useful methods for correcting missing DFI values in growing pigs and Loess correction can cover most individuals having missing DFI data.

Compared with randomly deleted DFI cases, the correction of missing DFI data did not show significant changes in estimated variance components, heritability and standard errors for ADFI, even though the deleted DFI proportion increased. On the other

Table 4 Rank correlation and information criteria of estimated breeding value (EBV) for average daily feed intake (ADFI) between true-measured daily feed intake (DFI) values with no missing records and both randomly deleted and Loess-corrected DFI values using three different classification criteria over differing proportions of DFI deleted values

Data†	Classification criteria§	Proportion of DFI values deleted‡									
		5%	10%	15%	20%	25%	30%	35%	40%	45%	50%
Randomly deleted	r_s	0.996	0.994	0.991	0.988	0.982	0.979	0.976	0.971	0.961	0.955
	AIC	716	1177	1485	1775	2068	2237	2416	2536	2791	2903
	BIC	721	1182	1490	1780	2073	2242	2421	2541	2796	2908
Loess-corrected	r_s	0.998	0.996	0.994	0.993	0.991	0.990	0.985	0.983	0.976	0.972
	AIC	296	800	1101	1245	1509	1617	1968	2034	2329	2468
	BIC	301	805	1105	1250	1514	1622	1972	2039	2334	2473

†Randomly deleted, the randomly deleted DFI datasets generated from the true-measured 902 pigs with no missing values were used; Loess-corrected, the locally weighted regression equation-corrected values from the same data set as the randomly deleted cases were used. ‡Two randomly deleted DFI datasets were generated by randomly deleting a proportion of the records from the true-measured 902 pigs with no missing values. The table reflects the two randomly deleted datasets averaged for each proportion within the table. There were 18 040 iterations per data. § r_s , Spearman's rank correlation coefficient; AIC, Akaike's Information Criterion; BIC, Bayesian Information Criterion. Each criterion is calculated to the true-measured 902 pigs with no missing values.

Table 5 Average and standard deviation for the average daily feed intake (ADFI) in pigs

Group†	Unit	Males	Females	Total
tru-ADFI	kg/day	2.57 ± 0.25	2.61 ± 0.31	2.57 ± 0.25
mis-ADFI	kg/day	2.59 ± 0.25	2.72 ± 0.30	2.61 ± 0.26
cor-ADFI	kg/day	2.59 ± 0.25	2.70 ± 0.29	2.60 ± 0.26
raw-ADFI	kg/day	2.58 ± 0.25	2.69 ± 0.31	2.59 ± 0.26
imp-ADFI	kg/day	2.58 ± 0.25	2.68 ± 0.31	2.59 ± 0.26

†tru-ADFI, based on the measured daily feed intake (DFI) obtained from 902 individuals with no missing DFI records; mis-ADFI, based on the measured DFI obtained from individuals with missing values ($n = 720$); cor-ADFI, based on the Loess-corrected DFI obtained from mis-ADFI ($n = 720$); raw-ADFI, based on tru-ADFI and mis-ADFI obtained from all sampled individuals including missing values ($n = 1622$); imp-ADFI, based on tru-ADFI and cor-ADFI obtained from all sampled individuals with corrected values ($n = 1622$).

Table 6 Variance components and heritability for the average daily feed intake (ADFI) in pigs (mean ± SE)

Group†	Genetic variance	Residual variance	Heritability
tru-ADFI	0.0238 ± 0.0056	0.0325 ± 0.0040	0.42 ± 0.17
mis-ADFI	0.0226 ± 0.0074	0.0420 ± 0.0056	0.35 ± 0.19
cor-ADFI	0.0230 ± 0.0074	0.0405 ± 0.0055	0.36 ± 0.19
raw-ADFI	0.0219 ± 0.0044	0.0389 ± 0.0031	0.36 ± 0.12
imp-ADFI	0.0216 ± 0.0044	0.0387 ± 0.0031	0.36 ± 0.12

†tru-ADFI, based on the measured daily feed intake (DFI) obtained from 902 individuals with no missing DFI records; mis-ADFI, based on the measured DFI obtained from individuals with missing values ($n = 720$); cor-ADFI, based on the Loess-corrected DFI obtained from mis-ADFI ($n = 720$); raw-ADFI, based on tru-ADFI and mis-ADFI obtained from all sampled individuals including missing values ($n = 1622$); imp-ADFI, based on tru-ADFI and cor-ADFI obtained from all sampled individuals with corrected values ($n = 1622$).

hand, rank correlations of EBV for ADFI obtained by the correction of missing DFI records to the true-measured EBV with no missing DFI records were significantly higher than those of the randomly deleted cases, even in the case of 5% deleted DFI records. Moreover, the Loess-corrected EBV for ADFI demonstrated lower values in both AIC and BIC than the uncorrected cases relative to true-measured EBV. Our results indicate that correcting missing DFI values is necessary for breeding value estimation and improves the selection accuracy using EBV for feeding traits. In the actual datasets using the raw-ADFI and imp-ADFI with 1622 individuals, we confirmed that EBV ranking exhibited a change by the

correction of missing DFI, even for an average of 3% missing DFI proportion (data not shown). Furthermore, standard errors for both estimated variance components and heritability for ADFI in 1622 individuals, including 720 pigs with missing DFI records, were smaller than those in the 902 true-measured individuals without missing DFI records. Satoh *et al.* (1999) reported that as the number of sampled individuals increased, the root-mean-square errors of estimated additive genetic variance decreased in a closed pig strain. Moreover, Satoh (2009) suggested that a gap may be generated in the estimated value of genetic variance if fewer than 1000 individuals are sampled. Similar to these reports, our research

Table 7 Heritability of production traits (diagonal in bold) and the genetic correlations between average daily feed intake (ADFI) and the other production traits (below diagonal) in pigs (mean \pm SE)

Trait†	ADG	BFT	EMA	IMF	imp-ADFI	raw-ADFI
ADG	0.45 \pm 0.13					
BFT	0.35 \pm 0.18	0.53 \pm 0.15				
EMA	-0.23 \pm 0.19	-0.01 \pm 0.20	0.35 \pm 0.10			
IMF	0.15 \pm 0.32	0.07 \pm 0.32	-0.20 \pm 0.31	0.19 \pm 0.14		
imp-ADFI	0.78 \pm 0.08	0.40 \pm 0.18	-0.20 \pm 0.21	0.29 \pm 0.32	0.36 \pm 0.12	
raw-ADFI	0.79 \pm 0.08	0.41 \pm 0.18	-0.20 \pm 0.21	0.29 \pm 0.32	—	0.36 \pm 0.12

†ADG, average daily gain in individuals with 30–100 kg body weight (BW); BFT, back fat thickness; EMA, loin eye muscle area; IMF, intramuscular fat content; imp-ADFI, average daily feed intake based on true-measured daily feed intake values with no missing and Loess-corrected daily feed intake values obtained from all sampled individuals with 30–100 kg BW ($n = 1622$); raw-ADFI, average daily feed intake based on measured daily feed intake values obtained from all sampled individuals with 30–100 kg BW including missing values ($n = 1622$).

showed the standard error for genetic parameter estimation for ADFI lowered as the number of individuals increased using pigs generated by a closed nucleus breeding system. Therefore, the correction of missing DFI records enables more accurate estimation of genetic parameters and breeding values by increasing the number of individuals with feeding trait data.

As described earlier, we used the values from the raw-ADFI, which included missing DFI records, and the imp-ADFI, which included corrected DFI values, for the estimation of genetic correlations between ADFI and production traits because both groups had the highest number of individuals and indicated the lowest standard error values among the groups. Our estimated ADFI heritability from 1622 individuals was within the range of values being similar to the range of 0.23–0.66 that has previously been reported for swine (Mrode & Kennedy 1993; Johnson *et al.* 1999; Suzuki *et al.* 2005; Hoque *et al.* 2009; Do *et al.* 2013; Jiao *et al.* 2014). With regard to the genetic correlations between ADFI and production traits, the estimated values did not appear to change significantly between raw-ADFI containing missing values and imp-ADFI with correction DFI similar to the heritability. There was a positive genetic correlation between ADFI and ADG, and BFT in this study, supporting previous findings (Mrode & Kennedy 1993; Von Felde *et al.* 1996; Hoque *et al.* 2009; Do *et al.* 2013; Jiao *et al.* 2014). The genetic correlation between ADFI and IMF in our study was not considered in previous studies (Cai *et al.* 2008; Jiao *et al.* 2014), and few studies have considered the genetic correlation between ADFI and IMF. Therefore, further investigation will be required to better understand this.

Here, we demonstrated that the nonparametric Loess equation is one of the most useful methods for correcting missing DFI values from individual pigs and that the correction of missing DFI values is important for accuracy during estimation of genetic parameters and breeding values. The improved

accuracy of EBV for ADFI through corrected DFI cases compared to uncorrected cases promotes genetic improvement in regard to feeding traits when missing DFI for individuals within a dataset. The correction of missing DFI is capable of effectively using the valuable DFI data including missing values because collecting DFI data for individual pigs requires a long time and higher expenses. Therefore, our study is also expected to improve the breeding for FCR and RFI based on DFI data.

ACKNOWLEDGMENTS

This work was supported by a grant from the Ministry of Agriculture Forestry and Fisheries of Japan (Research Program on Innovative Technologies for Animal Breeding, Reproduction, and Vaccine Development, AGB2002).

REFERENCES

- Bunter KL, Cai W, Johnston DJ, Dekkers JC. 2010. Selection to reduce residual feed intake in pigs produces a correlated response in juvenile insulin-like growth factor-I concentration. *Journal of Animal Science* **88**, 1973–1981.
- Cai W, Casey DS, Dekkers JC. 2008. Selection response and genetic parameters for residual feed intake in Yorkshire swine. *Journal of Animal Science* **86**, 287–298.
- Casey DS, Stern HS, Dekkers JC. 2005. Identification of errors and factors associated with errors in data from electronic swine feeders. *Journal of Animal Science* **83**, 969–982.
- Cleveland WS. 1979. Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association* **74**, 829–836.
- Cleveland WS, Devlin SJ. 1988. Locally weighted regression: an approach to regression analysis by local fitting. *Journal of the American Statistical Association* **83**, 596–610.
- De Haer LCM, Merks JWM, Kooper HG, Buiting GAJ, van Hattum JA. 1992. A note on the IVOG[®]-station: a feeding station to record the individual food intake of group-housed growing pigs. *Animal Production* **54**, 160–162.

- Do DN, Strathe AB, Jensen J, Mark T, Kadarmideen HN. 2013. Genetic parameters for different measures of feed efficiency and related traits in boars of three pig breeds. *Journal of Animal Science* **91**, 4069–4079.
- Eissen JJ, Kanis E, Merks JWM. 1998. Algorithms for identifying errors in individual feed intake data of growing pigs in group-housing. *Applied Engineering in Agriculture* **14**, 667–673.
- Eissen JJ, de Haan AG, Kanis E. 1999. Effect of missing data on the estimate of average daily feed intake of growing pigs. *Journal of Animal Science* **77**, 1372–1378.
- Gilbert H, Bidanel JP, Billon Y, Lagant H, Guillouet P, Sellier P, *et al.* 2007. Correlated responses in sow appetite, residual feed intake, body composition, and reproduction after divergent selection for residual feed intake in the growing pig. *Journal of Animal Science* **90**, 1097–1108.
- Hirose K, Nakamura M, Takizawa T, Fukawa K, Ito T, Ueda M, *et al.* 2009. An insertion/deletion variant of a thymine base in exon 2 of the porcine beta 3-adrenergic receptor gene associated with loin eye muscle area. *Animal Science Journal* **80**, 624–630.
- Hirose K, Takizawa T, Fukawa K, Ito T, Ueda M, Hayashi Y, *et al.* 2011. Association of an SNP marker in exon 24 of a class 3 phosphoinositide-3-kinase (PIK3C3) gene with production traits in Duroc pigs. *Animal Science Journal* **82**, 46–51.
- Hoque MA, Katoh K, Suzuki K. 2009. Genetic associations of residual feed intake with serum insulin-like growth factor-I and leptin concentrations, meat quality, and carcass cross sectional fat area ratios in Duroc pigs. *Journal of Animal Science* **87**, 3069–3075.
- Jacoby WG. 2000. Loess: a nonparametric, graphical tool for depicting relationships between variables. *Electoral Studies* **19**, 577–613.
- Jiao S, Maltecca C, Gray KA, Cassady JP. 2014. Feed intake, average daily gain, feed efficiency, and real-time ultrasound traits in Duroc pigs: I. Genetic parameter estimation and accuracy of genomic prediction. *Journal of Animal Science* **92**, 2377–2386.
- Johnson ZB, Chewning JJ, Nugent RA 3rd. 1999. Genetic parameters for production traits and measures of residual feed intake in large white swine. *Journal of Animal Science* **77**, 1679–1685.
- Koch RM, Swiger LA, Chambers D, Gregory KE. 1963. Efficiency of Feed Use in Beef Cattle. *Journal of Animal Science* **22**, 486–494.
- Köhn F, Sharifi AR, Simianer H. 2007. Modeling the growth of the Goettingen minipig. *Journal of Animal Science* **85**, 84–92.
- Koivula M, Sevón-Aimonen ML, Strandén I, Matilainen K, Serenius T, Stalder KJ, *et al.* 2008. Genetic (co)variances and breeding value estimation of Gompertz growth curve parameters in Finnish Yorkshire boars, gilts and barrows. *Journal of Animal Breeding and Genetics* **125**, 168–175.
- Misztal I, Tsuruta S, Strabel T, Auvray B, Druet T, Lee DH. 2002. *BLUPF90 AND RELATED PROGRAMS (BGF90)*. University of Georgia, Athens, GA.
- Mrode RA, Kennedy BW. 1993. Genetic variation in measures of food efficiency in pigs and their genetic relationships with growth rate and backfat. *Animal Production* **56**, 225–232.
- Salomon LJ, Bernard JP, Ville Y. 2007. Estimation of fetal weight: reference range at 20–36 weeks' gestation and comparison with actual birth-weight reference range. *Ultrasound in Obstetrics & Gynecology* **29**, 550–555.
- Satoh M. 2009. Effects of selection and population size on estimates of variance components and breeding values in pigs. *Japanese Journal of Swine Science* **46**, 17–24.
- Satoh M, Furukawa T, Ishii K. 1999. Effects of change in number of animals measured on genetic parameter and breeding value estimates obtained by single-trait REML on a closed strain in pigs. *Japanese Journal of Swine Science* **36**, 130–135.
- Schild RL, Fimmers R, Hansmann M. 2000. Fetal weight estimation by three-dimensional ultrasound. *Ultrasound in Obstetrics & Gynecology* **16**, 445–452.
- Suzuki K, Kadowaki H, Shibata T, Uchida H, Nishida A. 2005. Selection for daily gain, loin-eye area, backfat thickness and intramuscular fat based on desired gains over seven generations of Duroc pigs. *Livestock Production Science* **97**, 193–202.
- Von Felde A, Roeheha R, Looftb H, Kalm E. 1996. Genetic association between feed intake and feed intake behaviour at different stages of growth of group-housed boars. *Livestock Production Science* **47**, 11–22.
- Young JM, Cai W, Dekkers JCM. 2011. Effect of selection for residual feed intake on feeding behavior and daily feed intake patterns in Yorkshire swine. *Journal of Animal Science* **89**, 639–647.