Use of Random Regression Models for Repeated Measurements

Canadian Centre for Swine Improvement

1. Background

Current genetic evaluations of swine are based on the performance at 100 kg. In view of the requirement of heavier carcass weight for commercial pigs, a question has been raised whether the probe data should be recorded at and adjusted to 110 kg, 125 kg or even heavier weight. The question was raised and discussed in previous genetic committee meetings. One of the ways is to allow for more than one slaughter weight. It was suggested that random regression models (RRM) be used in genetic evaluation and allow breeders to test their pigs at various weights.

Repeated measurements have been taken on purebred pigs as part of the project on evaluating the effects of heavier carcass weight supported by ACAAF program. According to the suggestion of Genetic Committee, it is reviewed how random regression models can be applied to swine genetic evaluation for the purpose mentioned above based on the data collected at the Deschambault test station.

2. Random regression models

The random regression model is a method for analyzing the data of repeated measurements obtained during different life times of animals (longitudinal data).

Current genetic evaluations for probe and carcass traits are based on the data from single probing at the end of performance tests. For each animal, BLUP EBV for each selection trait is calculated. The EBV of an animal can take different values and can not be pre-determined, it is therefore assumed as a random variable in genetic evaluation.

If repeated measurements over time are available for an animal, the EBV of the animal changes with time, and therefore, is better described with a curve using a regression model. Since the curve for each animal can not be pre-determined and can be seen as a random event just as a BLUP EBV. Statistically, this random curve should be described using the random regression model. In random regression model, we estimate a set of genetic determined regression coefficients, instead of estimating a constant EBV. From the regression coefficients, we are able to estimate breeding values for any given time or days of age.

However, for the case of single measurement, RRM is not a method directly useful for genetic evaluation.

3. Examples of the uses of RRM

RRM is a powerful statistical tool, useful not only for genetic evaluation based on repeated measurements, but also for genetic parameter estimations, studies of animal growth rates and life processes. In this section, some examples are given.

3.1 Data

Purebred pigs from across Canada were tested at the Deschambault test station in Quebec. The pigs were received and tested in two batches. The first batch corresponds to test #17 and the second batch to #18. Repeated measurements were taken for body weight, backfat thickness, lean depth and feed intake. Ultrasound measurements are carried out on pigs during their growth. From 75 kg, the pigs were scanned every 2 weeks until they are slaughtered. In total, body weight and probing measurements are available on 472 animals (264 in #17 and 208 in #18). Feed intake data were collected every day for all the pigs (Table 1). 450 animals (245 in #17 and 205 in #18) have measurements over whole experimental period.

Table 1. The number of records classified in breeds and performance tests and the average number of records on each animal (in parenthesis)

Breed	Types of Records	Deschambault test number	
		#17	#18
Duroc	# Animal	69	37
	# weight	730 (10.6)	400 (10.8)
	# Backfat / Lean depth	390 (5.7)	193 (5.2)
	# Daily feed intake	7,302 (106)	3,346 (90)
Landrace	# Animal	65	63
	# weight	718 (11.0)	730 (11.2)
	# Backfat / Lean depth	379 (5.8)	365 (5.6)
	# Daily feed intake	7,026 (108)	6,107 (94.0)
Yorkshire	# Animal	130	102
	# Weight	1,343 (10.3)	1,191 (11.7)
	# Backfat / Lean depth	697 (5.4)	591 (5.8)
	# Daily feed intake	13,337 (103)	9,967 (97.7)
Lacombe	# Animal	0	6
	# Weight	0	72 (12.0)
	# Backfat / Lean depth	0	36 (6.0)
	# daily feed intake	0	607 (101)

The data from Yorkshire breed are presented as examples to show the potential uses of RRM.

3.2. Method

Random regression models used for the data analysis have the general form as follows:

$$y_{ijkl} = F_i + \sum_{n=0}^{p} \beta_n x_{jk}^n + \sum_{n=0}^{q} \beta_{pn} x_{jk}^n + \sum_{n=0}^{s} \beta_{an} x_{jk}^n + e_{ijkl}$$

y_{ijk} = observation

F_i = fixed effects

$\beta_n, \beta_{pn}, \beta_{an}$ = fixed, random permanent environmental and random genetic regression coefficients

x_{ij} = time covariate

e_{ijkl} = model residual effect

After model selection process, Legendre orthogonal polynomials were used for time covariates and the order of the regression function was set as 2 for all three regressions. Days of age were used as time covariates.

Software VCE 5 was used for the data analysis.

3.4 Results

Results from Yorkshire data are taken as an example. RRM can provide the results listed as follows.

3.4.1 Estimates of breeding values at different days of age

RRM can provide the random regression coefficients for each animal to be evaluated. From those regression parameters, it is able to estimate EBVs for every animals at different days of age. A list of EBVs of five animals estimated at different days of age is given in Table 1.

Table 1. EBVs for cumulative feed-intake for five animals

	EBVs for cumulative feed intake										
Age (days)	65	79	93	107	121	135	149	163	177	191	205
Av. Weight (kg)	28	39	51	62	73	84	96	107	118	129	140
1	1.33	1.57	2.11	2.95	4.11	5.57	7.34	9.42	11.81	14.50	17.50
2	1.32	-0.94	-2.73	-4.06	-4.93	-5.32	-5.26	-4.73	-3.74	-2.28	-0.36
3	0.50	0.01	-0.23	-0.20	0.07	0.61	1.39	2.42	3.72	5.26	7.07
4	-0.66	1.78	4.4	7.26	10.33	13.60	17.0	20.78	24.68	28.80	33.13
5	0.45	-1.39	-2.94	-4.18	-5.12	-5.77	-6.12	-6.17	-5.93	-5.38	-4.54

3.4.2 Heritability estimates

RRM can estimate the heritability for any given days of age within the time period of performance tests.
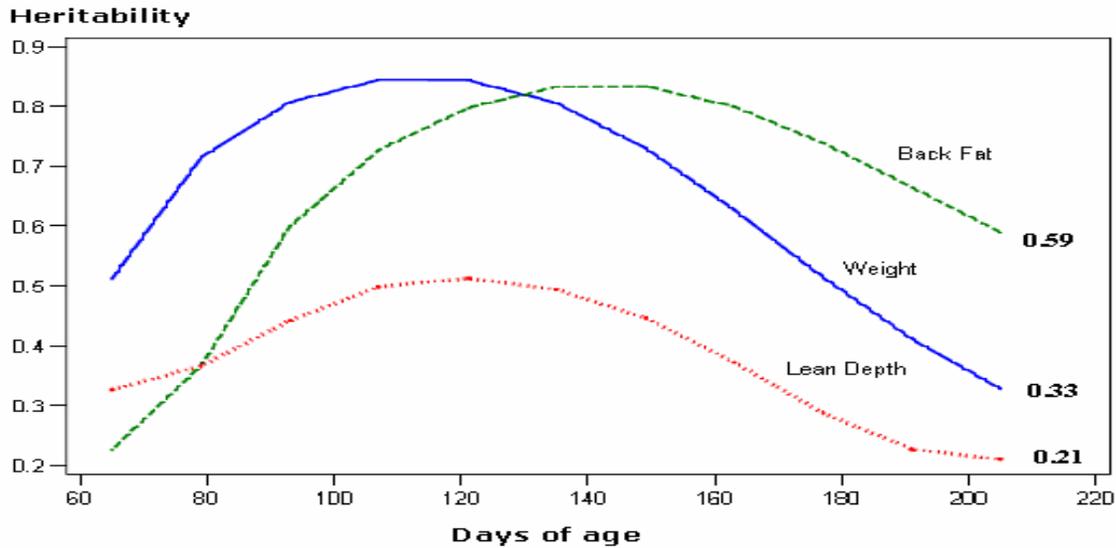

Figure 1. Heritabilities of backfat, weight and lean depth at different days of age

3.4.3 Genetic correlations for a trait measured at different days of age

Genetic correlations of the trait measurements among different days of age can be estimated from RRM analysis. Table 2 to Table 4 listed the genetic correlations for body weight, backfat and lean depth. The genetic correlations among different times are generally very high. The adjacent measurements of two weeks apart could be well seen as the same trait in some circumstance. Comparing the different traits, the correlations for backfat and body weight are higher than lean depth.

Table 2. Genetic correlations between body weights

	Genetic correlations for body weight										
Age (days)	65	79	93	107	121	135	149	163	177	191	205
Av. weight (kg)	28	39	51	62	73	84	96	107	118	129	140
65 (28)	1.00	0.83	0.70	0.63	0.578	0.54	0.51	0.48	0.44	0.40	0.35
79 (39)	0.83	1.00	0.98	0.95	0.927	0.90	0.87	0.84	0.79	0.73	0.65
93 (51)	0.70	0.98	1.00	0.99	0.983	0.97	0.95	0.91	0.87	0.81	0.73
107 (62)	0.63	0.95	0.99	1.00	0.997	0.99	0.97	0.95	0.91	0.85	0.78
121 (73)	0.58	0.93	0.98	0.997	1.00	0.997	0.98	0.97	0.94	0.89	0.82
135 (84)	0.54	0.90	0.97	0.99	0.997	1.00	0.996	0.98	0.96	0.92	0.86
149 (96)	0.51	0.87	0.95	0.97	0.99	0.996	1.00	0.99	0.98	0.95	0.90
163 (107)	0.48	0.84	0.91	0.95	0.97	0.98	0.995	1.00	0.99	0.97	0.93
177 (118)	0.44	0.79	0.87	0.91	0.94	0.96	0.98	0.99	1.00	0.99	0.97
191 (129)	0.40	0.73	0.81	0.85	0.89	0.92	0.95	0.97	0.99	1.00	0.99
205 (141)	0.35	0.65	0.73	0.78	0.82	0.86	0.90	0.93	0.97	0.99	1.00

(Left label spanning rows: Days of Age)

Table 3. Genetic correlations between backfat measurements

	Genetic correlations for backfat										
Age (days)	65	79	93	107	121	135	149	163	177	191	205
Av. weight (kg)	28	39	51	62	73	84	96	107	118	129	140
65 (28)	1.00	0.67	0.31	0.16	0.10	0.08	0.09	0.10	0.12	0.15	0.18
79 (39)	0.67	1.00	0.91	0.83	0.79	0.77	0.76	0.75	0.74	0.74	0.73
93 (51)	0.31	0.91	1.00	0.99	0.97	0.96	0.94	0.93	0.91	0.89	0.86
107 (62)	0.16	0.83	0.99	1.00	0.996	0.988	0.98	0.96	0.95	0.92	0.89
121 (73)	0.10	0.79	0.97	0.996	1.00	0.998	0.99	0.98	0.96	0.94	0.92
135 (84)	0.08	0.77	0.96	0.99	0.998	1.000	0.998	0.99	0.98	0.96	0.94
149 (96)	0.09	0.76	0.94	0.98	0.99	0.998	1.00	0.998	0.99	0.98	0.96
163 (107)	0.10	0.75	0.93	0.96	0.98	0.99	0.998	1.00	0.998	0.99	0.98
177 (118)	0.12	0.74	0.91	0.95	0.96	0.98	0.99	0.998	1.00	0.997	0.99
191 (129)	0.15	0.74	0.89	0.92	0.94	0.96	0.98	0.99	0.997	1.00	0.998
205 (141)	0.18	0.73	0.86	0.90	0.92	0.94	0.96	0.98	0.99	0.998	1.00

(Left label spanning rows: Days of Age)

Table 4. Genetic correlations between lean depth measurements

		Genetic correlations for lean depth										
Age (days)		65	79	93	107	121	135	149	163	177	191	205
Av. weight (kg)		28	39	51	62	73	84	96	107	118	129	140
	65 (28)	1.00	0.92	0.71	0.51	0.38	0.31	0.31	0.35	0.45	0.56	0.62
	79 (39)	0.92	1.00	0.93	0.81	0.71	0.65	0.64	0.65	0.67	0.66	0.57
	93 (51)	0.71	0.93	1.00	0.97	0.92	0.88	0.86	0.84	0.80	0.67	0.47
Days	107 (62)	0.51	0.81	0.97	1.00	0.988	0.97	0.95	0.92	0.84	0.65	0.37
of	121 (73)	0.38	0.71	0.92	0.988	1.00	0.995	0.98	0.95	0.85	0.64	0.34
Age	135 (84)	0.31	0.65	0.88	0.97	0.995	1.00	0.995	0.97	0.87	0.66	0.36
	149 (96)	0.31	0.64	0.86	0.95	0.98	0.995	1.00	0.99	0.91	0.72	0.43
	163 (107)	0.35	0.65	0.84	0.92	0.95	0.97	0.99	1.00	0.97	0.82	0.57
	177 (118)	0.45	0.67	0.80	0.84	0.85	0.87	0.91	0.97	1.00	0.94	0.76
	191 (129)	0.56	0.66	0.67	0.65	0.64	0.66	0.72	0.82	0.94	1.00	0.94
	205 (141)	0.62	0.57	0.47	0.37	0.34	0.36	0.43	0.57	0.76	0.94	1.00

The results show that the probe traits, such as backfat, body weight and lean depth are genetically highly correlated between 163, 177 and 191 days of age, which approximately correspond to 106 kg, 118 kg and 130 kg of body weight on average (Tables 2, 3 and 4). This indicates that the current selection for the performance at 100 kg may serve the breeding goal of heavier carcass weight well.

The genetic correlations among backfat, lean depth and body weight estimated from random regression on days of age are given in Table 5.

Table 5. Genetic correlations between weight, backfat and lean depth

	Genetic correlation		
Days of Age	Weight ~ Backfat	Weight ~ Lean depth	Backfat ~ Lean depth
65 (28)	0.88	0.39	0.31
79 (39)	0.81	0.48	0.38
93 (51)	0.73	0.58	0.34
107 (62)	0.69	0.62	0.28
121 (73)	0.67	0.61	0.22
135 (84)	0.66	0.58	0.16
149 (96)	0.66	0.54	0.10
163 (107)	0.66	0.47	0.04
177 (118)	0.67	0.38	-0.01
191 (129)	0.68	0.26	-0.05
205 (141)	0.69	0.16	-0.05

It is important to note that the above estimates are computed using age as the independent variable (time variable) in the random regression analysis. Therefore, the correlations between two traits are for a given age, e.g. there is a positive correlation between weight and backfat at a

given age. It is also possible to use weight as an independent variable (regressor) in the RRM. Then the correlations between the traits are for a given weight. Estimates of such correlations between backfat and lean depth are given in Table 6.

Table 6. Genetic correlations between backfat and lean depth at different weight, estimated from RRM with weight as regressor.

Weight (kg)	15	30	45	60	70	80	90	100	110	125	140
Genetic correlation	0.02	-0.00	-0.00	-0.49	-0.55	-0.55	-0.52	-0.35	-0.10	0.06	0.06

4. Discussion and conclusions

- RRM is a method for genetic evaluation based on repeated measurements over time. It is not a method for the data of single measurement. It can not make EBVs at 100 kg more comparable with the EBVs at other body weight than simple adjustment factors if animals are just probed once. It is necessary to test the pigs close to a standard body weight and adjust the variations to the standard body weight.

- If the repeated measurements are available for some reason, RRM will be very useful for genetic evaluation and can increase the accuracy of genetic evaluation. Now, the evaluation procedure for RRM genetic evaluation of multiple traits has been developed and pilot genetic evaluations can be provided if repeated measurements are available.

- Random regression models have been applied to test day data of dairy cattle. The collection of test day data in dairy cattle can simplify the milk recording. The main intention to estimate the total lactation yield while the milk test is done once a month, in stead of every day of the lactation. The situation in swine is quite different from that for dairy cattle. In swine, body weight, backfat and lean depth are cumulative traits that can be measured at the end of the performance test. Repeated probes for swine need extra effort and costs compared with single probing at the end of performance test.

- There are several studies of RRM for swine. These studies mainly focused on the evaluation of growth rate and feed intake. There are also a few investigations on the use of RRM for sow productivities across parities. Its practical application in a swine breeding program has not been found.

- The RRM could be routinely used in the future. It depends on how to make the repeated measurement available. There have been attempts to simplify multiple probes for swine. Doeschl-Wilson from PIC (2004) used visual image analysis to describe pig growth in term of size and shape, and applied RRM to evaluate pigs based on the visual images. RRM for evaluation of litter size is straightforward to use. The electronic feeding system can make recording feed intake simpler.