

Subpopulations and accuracy of prediction in pig carcass classification

B. Engel^{1†}, W. G. Buist¹, M. Font i Furnols² and E. Lambooij¹

¹Animal Sciences Group, Wageningen UR, PO Box 65, 8200 AB Lelystad, The Netherlands

²IRTA-Meat Technology Centre, Granja Camps i Armet. E-17121 Monells, Girona, Spain

† E-mail : Bas.Engel@wur.nl

Abstract

Classification of pig carcasses in the European Community (EC) is based on the lean meat percentage of the carcasses. The lean meat percentage is predicted from instrumental carcass measurements, such as fat and muscle depth measurements, obtained in the slaughterline. The prediction formula for an instrument is derived from the data of a dissection experiment. When the relationship between percentage lean and instrumental carcass measurements differs between subpopulations, such as sexes or breeds, accuracy of prediction may differ between these subpopulations. In particular for some subpopulations predicted lean meat percentages may be systematically too low and for other subpopulations systematically too high. Producers or buyers that largely specialize in subpopulations where the percentage lean is underestimated, are put at a financial disadvantage.

The aim of this paper is to gain insight, on the basis of real data, into the effects of differences between subpopulations on the accuracy of the predicted percentage lean meat of pig carcasses. A simulation study was performed based on data from dissection trials in The Netherlands, comprising gilts and castrated males, and trials in Spain, comprising different genetic types. The possible gain in accuracy, i.e. reduction of prediction bias and mean squared prediction error, by the use of separate prediction formulae for (some of) the subpopulations was determined.

We concluded that marked bias in the predicted percentage lean meat may occur between subpopulations when a single overall prediction formula is employed. Systematic differences in predicted percentage lean between subpopulations that are overestimated and underestimated may exceed 4% and for selected values of instrumental measurements may run up to 6%. Bias between subpopulations may be eliminated, and prediction accuracy may be markedly improved, when separate prediction formulae are used. With the use of separate formulae the root mean squared prediction error may be reduced by 13 to 26% of the expected value when a single prediction formula is used for all pig carcasses.

These are substantial reductions on a national scale. This suggests that there will be a commercial interest in the use of separate prediction formulae for different subpopulations. In the near future, when the use of implants becomes more reliable, subpopulations will be recognized automatically in the slaughterline and use of different prediction formulae will become practically feasible. Some possible consequences for the EC regulations and national safeguards for quality of prediction formulae are discussed.

Keywords: carcasses, classification, lean, pigs, populations, regression analysis, sampling.

Introduction

A prediction formula for the lean meat percentage of pig carcasses from a particular population, as derived by linear regression, is unbiased. This means

that averages of predictions and true lean meat percentages of large samples of pigs with comparable instrumental measurements are similar. In particular the average true lean meat percentage and the average predicted lean meat percentage over

all pigs in the population will be alike. Since we are talking about averages, lean meat percentages of some carcasses will be underestimated and others will be overestimated. This is not necessarily a problem, unless carcasses which are underestimated belong to a relevant subpopulation, such as breed or sex. When the relationship between the percentage lean and the instrumental carcass measurements differs between subpopulations, e.g. due to differences in conformation, unbiasedness will not hold for the separate subpopulations. For some of the subpopulations the average predicted percentage lean will be too low, while for other subpopulations it will be too high. This is a problem when producers or buyers largely specialize in some subpopulations. Some producers or buyers will financially be put at a disadvantage when a single overall prediction formula is used. For those producers or buyers the use of separate prediction formulae for subpopulations ensures more fair payment.

Significant differences between sexes were reported by Engel and Walstra (1991b and 1993), Daumas *et al.* (1994) and Gispert *et al.* (1996). In France separate equations are used for the sexes (Daumas *et al.*, 1998). In a recent enquiry (EUPIGCLASS, 2000) most European Community (EC) member states considered their pig population to be genetically heterogeneous, with up to six genetic subpopulations. Italy was the first country to get EC approval for a separate prediction formula for its so called 'heavy pigs' (EC, 1989b).

In this paper we studied by simulation the extent to which bias and accuracy of prediction may differ when a single overall prediction formula is used compared with the use of separate prediction formulae for subpopulations. The simulation was based on data from dissection trials in The Netherlands and Spain. In 1990 a dissection trial was carried out throughout the EC as a first step towards harmonization of methods for pig grading. The proposal for this trial (EC, 1989a) specifically mentioned a possible interest in differences between subpopulations. In The Netherlands at that time there was an interest in differences between gilts and castrated males (hereafter referred to as castrates). Separate samples were taken for the two sexes. The instrumental carcass measurements were a fat and muscle depth measured with the Hennessy Grading Probe (HGP) (Walstra, 1986). Separate prediction formulae were calculated for the sexes and found to be significantly, although not markedly, different (B. Engel, personal communication; Engel and Walstra, 1993). The Dutch data of the harmonization trial, supplemented with recent data from Dutch

slaughterhouses, provided the basis for the start of the simulation study. In addition, data were simulated for two and four subpopulations, based on recent data from dissection trials in Spain with eight genetic types (from now on referred to as breeds). Again, the instrumental carcass measurements were fat and muscle depth measurements, this time obtained with the Fat-O-Meater (FOM) (European Commission, 1988 and 1994). Our simulation does not exactly reflect practice and accuracy of carcass grading in The Netherlands and Spain. It does however mimic a situation that might easily occur within the EC. It is not unlikely that in some member states differences between subpopulations are more pronounced than in our simulation study.

Carcasses for dissection are usually not chosen at random, but with over-sampling of carcasses with more extreme instrumental carcass measurements to increase the accuracy of prediction. When subpopulations are of interest, separate samples per subpopulation may be taken. In each subsample, for a more accurate comparison between subpopulations, carcasses may be selected, say, with over-sampling of smaller and larger fat depth measurements. When significant differences between separate regressions per subpopulation are found, subsequent inference will have to take this into account. When the use of separate formulae is not feasible and a single overall prediction formula is required, regression of the percentage lean on the instrumental carcass measurements, simply ignoring the subpopulations, is not covered by standard regression theory. This is because the distribution of the subpopulations over the instrumental measurements in the sample is not representative for the population. Therefore, in the harmonization trial, for the Dutch dissection data, a method was developed by Engel and Walstra (1993) to take proper account of the subpopulations. In the present simulation study, prediction formulae were derived by linear regression, ignoring the subpopulations, and by the Engel and Walstra (1993) method, to account for the subpopulations. Carcasses for dissection were sampled according to four different sampling schemes, including the most commonly used sampling schemes in the EC. In all four schemes carcasses were selected with over-sampling of smaller and larger fat depth measurements. The performance of the prediction formulae, as obtained under the different sampling schemes and by the different statistical methods was determined. Performance of the formulae was measured in terms of bias and mean square prediction error for the separate subpopulations and for the pig population as a whole.

Material and methods

The statistical model

As a first step towards simulation, we have to specify a statistical model for the instrumental measurements and the percentage lean of pig carcasses. Here, the main components of this model are introduced and illustrated for the Dutch data, where the subpopulations of interest are gilts and castrates. From the fat and muscle depth measurements taken with the HGP, we used the fat depth measurements only, since this simplifies the presentation of the results. Later on for the Spanish data we will consider two prediction variables, i.e. fat and muscle depth measurements, and four subpopulations, as well.

For each of the two sexes, a linear regression model was assumed that relates the percentage lean y to the fat depth measurement x of a carcass:

$$y = \mu_1(x) + e = a_1 + b_1x + e, \text{Var}(e) = \sigma_{1,y,x}^2 \quad \text{for gilts} \quad (1a)$$

and

$$y = \mu_2(x) + e = a_2 + b_2x + e, \text{Var}(e) = \sigma_{2,y,x}^2 \quad \text{for castrates} \quad (1b).$$

Here, e is a normally distributed random error term representing variation in lean meat percentage between carcasses with the same fat depth measurement. The size of the error terms is quantified by the residual variances $\sigma_{1,y,x}^2$ and $\sigma_{2,y,x}^2$ for gilts and castrates respectively. A linear relationship was assumed because marked departures from linearity within sexes or breeds were not apparent for the Dutch and Spanish data.

The model was completed by introducing $\pi_1 = 0.5$ as the proportion of gilts and $\pi_2 = 0.5$ as the proportion of castrates in the population and by assuming a normal distribution for fat depth x with means μ_1 and μ_2 and variances σ_1^2 and σ_2^2 for gilts and castrates respectively.

In Figure 1 the separate regression lines for the sexes are shown. For an impression of the variation around the lines, the lean meat percentages from a simulated random sample of 120 carcasses (the minimal sample size for a dissection experiment as required by the EC regulations) were shown as well. The values for constants a_1 and a_2 , coefficients b_1 and b_2 , residual variances $\sigma_{1,y,x}^2$ and $\sigma_{2,y,x}^2$, mean fat depths μ_1 and μ_2 and variances σ_1^2 and σ_2^2 , that were used to construct Figure 1, are reproduced in Table 1a. These parameter values were derived from the EC harmonization trial (B. Engel, personal communication) and a recently obtained large data

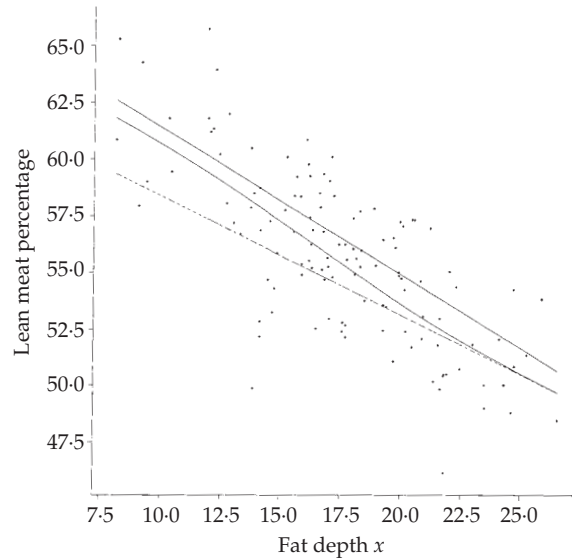


Figure 1 The separate regressions for gilts (solid line) and castrates (dotted line) and the overall mean curve. A random sample (comprising both sexes) of size 120 (solid dots) is presented as well. Input parameters are based on the Dutch data (Table 1a).

set from Dutch slaughterhouses comprising fat and muscle depth measurements obtained with the HGP. Although these parameter values are not an accurate reflection of the present situation in The Netherlands, since we lack recent information about the sexes, they mimic a situation that might occur within the EC. The simulated sample of 120 carcasses was easily generated by computer employing standard statistical software. In this paper all calculations were performed with GenStat (Genstat Committee, 2000).

A first look at prediction

For a quicker understanding of the impact of subpopulations on the accuracy of lean meat prediction, at first some simplifying assumptions will be made. These assumptions are for illustrative purposes only and will be dropped later on.

For the moment, imagine a perfect separation between the sexes on the basis of the fat depth measurements. So, there exists a boundary value c , such that fat depth x is below c for all gilts and above c for all castrates. This is formalized with the following indicators for gilts and castrates:

$$p_1 = 1, \text{ when } x < c \text{ (a gilt), and } 0 \text{ when } x > c \text{ (a castrate)}$$

and

$$p_2 = 1, \text{ when } x > c \text{ (a castrate), and } 0 \text{ when } x < c \text{ (a gilt).}$$

Table 1 Input parameters in the simulations

(a) Parameters for two sexes based on Dutch data and two genetic types (referred to as breeds) based on Spanish data

	Based on	
	Dutch data	Spanish data
μ_1 = mean fat depth subpopulation 1 (mm)	15.49	14.58
μ_2 = mean fat depth subpopulation 2 (mm)	19.11	12.77
σ_1 = s. d. fat depth x subpopulation 1 (mm)	3.15	2.75
σ_2 = s. d. fat depth x subpopulation 2 (mm)	3.95	2.76
a_1 = constant in regression of y on x subpop. 1	67.94	64.36
b_1 = coefficient x in regression of y on x subpop. 1	-0.6531	-0.619
$\sigma_{1,y,x}^2$ = residual variance regression y on x subpop. 1	4.09	6.754
a_2 = constant in regression of y on x subpop. 2	63.65	74.06
b_2 = coefficient x in regression y on x subpop. 2	-0.5276	-1.022
$\sigma_{2,y,x}^2$ = residual variance regression y on x subpop. 2	5.32	6.954
π_1 = proportion of subpop. 1 in the population	0.5	0.44
π_2 = proportion of subpop. 2 in the population	0.5	0.56

(b) Parameters for three maternal lines and a pool of five sire lines based on Spanish data

	Subpopulation ($i = 1 \dots 4$)			
	1	2	3	4
μ_{i1} = mean fat depth (mm)	14.58	14.08	14.44	14.10
μ_{i2} = mean muscle depth (mm)	53.21	54.82	54.38	59.86
σ_{i1} = standard deviation fat depth (mm)	2.75	3.28	2.91	3.65
σ_{i2} = standard deviation muscle depth (mm)	6.29	7.49	7.84	7.39
ρ_{i2} = correlation fat with muscle depth	0.059	0.030	0.232	-0.197
a_i = constant in regression of y on x_1, x_2	53.15	63.65	61.08	60.17
b_{i1} = coeff. x_1 in regression of y on x_1, x_2	-0.6494	-0.8839	-0.6731	-0.9122
b_{i2} = coeff. x_2 in regression of y on x_1, x_2	0.2189	0.1000	0.1087	0.1947
$\sigma_{y,x}^2$ = residual variance regression y on x_1, x_2	4.90	5.05	4.54	5.51
π_i = proportion in the population	0.17	0.25	0.20	0.38

To emphasize that p_1 and p_2 depend on x they will be denoted by functions $p_1(x)$ and $p_2(x)$.

As a further simplification, in addition to the perfect separation between the sexes, we also temporarily assume equal residual variances $\sigma_{1,y,x}^2 = \sigma_{2,y,x}^2 = \sigma_{12,y,x}^2$ and equal regression coefficients $b_1 = b_2 = b$ for the two sexes. Now, we define the predicted lean meat percentage for a carcass with fat depth measurement x as follows (predictions are indicated by a hat):

$$\hat{y} = p_1(x)a_1 + p_2(x)a_2 + bx = a_2 + (a_1 - a_2)p_1(x) + bx \quad (2).$$

The last expression follows from $p_1(x) + p_2(x) = 1$ (a carcass is either from a gilt or a castrate). It is readily checked that \hat{y} equals $a_1 + bx$ for a gilt and $a_2 + bx$ for a castrate. Effectively we are using separate prediction formulae for the sexes. From the expression on the right-hand side we see that basically $p_1(x)$ is just an extra explanatory variable,

derived from x , with values 0 or 1 indicating whether the carcass is from a castrate or from a gilt. Function $p_1(x)$ is a step function that starts with value 1 for lean pigs and at $x = c$ steps down to value 0, as shown by the bold solid line in Figure 2. The unknown parameters a_2 , $(a_1 - a_2)$ and b in expression (2) may be replaced by their least squares estimates. Unfortunately, there is no such thing as a perfect separation between the sexes on the basis of fat depth measurements. However, we do know that for small fat depth measurements a carcass is more likely to be from a gilt and for large fat depth measurements from a castrate. Therefore, we replace $p_1(x)$ and $p_2(x)$ by the probabilities for a gilt or castrate respectively, given the fat depth measurement x . Probability $p_1(x)$ for a gilt will decrease for increasing fat depth, as illustrated by the solid curve in Figure 2 that is based on the parameter values in Table 1. In the next section we will have a closer look at these probabilities and show how they can be calculated. With a small overlap between the

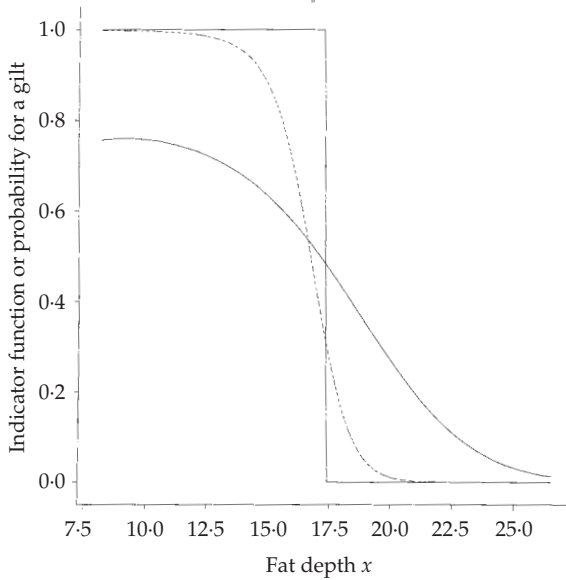


Figure 2 Function $p_1(x)$ as a step function for a gilt, under assumption of a perfect separation between the sexes with respect to fat depth x (straight line) and as a probability for a gilt (solid curve). Input parameters are based on the Dutch data (Table 1a). The dotted curve is added to show that it is possible to approximate a step function, when sexes are more separated.

sexes with respect to fat depth measurements, function $p_1(x)$ will almost be a step function as illustrated by the dotted line in Figure 2. The prediction formula presented in Engel and Walstra (1993) is essentially expression (2) but without the simplifying assumptions:

$$\hat{y} = p_1(x) \hat{y}_1 + p_2(x) \hat{y}_2, \quad (3)$$

where

$$\hat{y}_1 = A_1 + B_1 x \text{ and } \hat{y}_2 = A_2 + B_2 x.$$

Here, $p_2(x) = 1 - p_1(x)$ is the probability for a castrate and \hat{y}_1 and \hat{y}_2 are the separate predictions for the two sexes, for an observed fat depth measurement x . A_1 , B_1 , A_2 and B_2 are the least squares estimates for a_1 , b_1 , a_2 and b_2 obtained from regression of y on x for each of the sexes separately. The two predictions for the sexes are combined into a weighted average, where weights are equal to the probabilities for the sexes and depend on the observed fat depth. For small fat depth measurements x , the probability for a gilt will be large and \hat{y} will be closer to the prediction \hat{y}_1 for a gilt, for large x , the probability for a castrate will be large and \hat{y} will be closer to the prediction \hat{y}_2 for a castrate.

A closer look at the probabilities for subpopulations

The following expression for $p_1(x)$ can be derived:

$$\text{logit}(p_1(x)) = \gamma_0 + \gamma_1 x + \gamma_2 x^2.$$

Here, $\text{logit}(p) = \log(p/(1-p))$ is a log odds ratio, commonly referred to as the logit (transformation) of probability p . The logit 'stretches' probabilities from numbers between 0 and 1 to numbers between minus and plus infinity. For these 'stretched' probabilities we have a simple quadratic relationship with fat depth x . Expressions for the constant γ_0 and coefficients γ_1 and γ_2 can be derived from the model assumptions (**Appendix A1**). In practice γ_0 , γ_1 and γ_2 can be estimated from additional data by logistic regression (Cox and Snell, 1989), as shown in Engel and Walstra (1993). Logistic regression is specifically developed for binary response variables. For the additional data, sex coded as 1 for a gilt and 0 for a castrate is the binary response variable and the instrumental carcass measurements and their quadratic and product terms (when two or more instrumental measurements are involved) are the explanatory variables (Engel and Walstra, 1993, Table 3 and Figure 1). Standard software for logistic regression is widely available. Collection of the data requires knowledge of subpopulation membership but does not involve costly dissections. Therefore, quite a large set of data can be collected. The additional random sample employed for logistic regression in Engel and Walstra (1993) comprised 134158 carcasses. We will therefore assume that γ_0 , γ_1 and γ_2 are effectively known.

Prediction and bias between subpopulations

For the lean meat percentage y of a carcass with a given fat depth x and *unknown* sex:

$$y = \mu(x) + e \text{ and } \text{Var}(e) = \sigma_{y,x}^2.$$

Under the simplifying assumption of equal coefficients of fat depth and equal residual variances for gilts and castrates, the overall mean $\mu(x)$ and residual variance $\sigma_{y,x}^2$ are (**Appendix A2**):

$$\mu(x) = p_1(x) a_1 + p_2(x) a_2 + bx$$

$$\text{and } \sigma_{y,x}^2 = \sigma_{12,y,x}^2 + p_1(x) p_2(x) (a_1 - a_2)^2.$$

The difference between $\sigma_{y,x}^2$ and $\sigma_{12,y,x}^2$ where the latter is the residual variance for *known* sex and the former is the residual variance for *unknown* sex, will be small when $(a_1 - a_2)$ is small or when $p_1(x)$ is close to 0 or 1, i.e. when subpopulations hardly differ or when subpopulations can be well separated on the basis of measurements of x . Suppose that $a_1 > a_2$. When we compare the overall mean $\mu(x)$ with the separate means $\mu_i(x) = a_i + bx$, $i = 1, 2$, for the sexes:

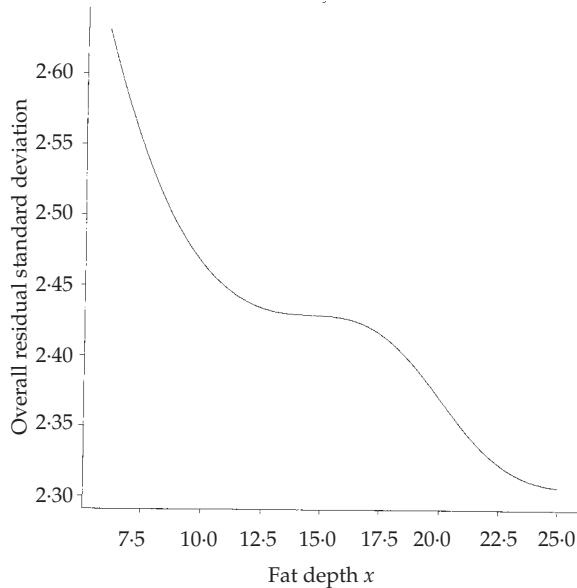


Figure 3 The overall residual standard deviation for the parameter values based on the Dutch data (Table 1a).

$$\begin{aligned} \mu(x) - \mu_1(x) &= -p_2(x)(a_1 - a_2) < 0 \\ \text{and } \mu(x) - \mu_2(x) &= p_1(x)(a_1 - a_2) > 0. \end{aligned}$$

A prediction obtained with a single overall prediction formula for both sexes is basically an estimate of $\mu(x)$. So, predictions for one subpopulation are systematically too low, while predictions for the other subpopulation are

systematically too high. Averaged over all animals in the population there will be no bias since:

$$p_1(x) \{\mu(x) - \mu_1(x)\} + p_2(x) \{\mu(x) - \mu_2(x)\} = 0.$$

When no further information about subpopulations is available in the slaughterline, the bias between subpopulations is unavoidable. Should such a bias be unacceptable, the only remedy is to find additional carcass measurements to discriminate between the subpopulations such that $p_1(x)$ will resemble a step function (as illustrated in Figure 2), or to use separate prediction formulae by, for example, reading subpopulation membership from an implanted chip. Similar comments hold when the simplifying assumptions are dropped. The more general expressions for the mean and residual variance, without the simplifying assumptions, are (Appendix A2):

$$\mu(x) = p_1(x) \mu_1(x) + p_2(x) \mu_2(x) \quad (4a)$$

and

$$\sigma_{y,x}^2 = p_1(x) \sigma_{1,y,x}^2 + p_2(x) \sigma_{2,y,x}^2 + p_1(x) p_2(x) \frac{\{\mu_1(x) - \mu_2(x)\}^2}{\mu_1(x) + \mu_2(x)} \quad (4b)$$

where $\mu_1(x)$ and $\mu_2(x)$ are the linear parts (without the error terms e) from the regressions in expressions (1a) and (1b). Note that prediction from expression (3) is an estimate of $\mu(x)$ from expression (4a).

Because $\mu(x)$ should be closer to the regression line for gilts for low fat depth x and closer to the

Table 2 Sampling schemes S_1 , S_2 , S_3 , S_4

Rows 1, 2 and 3 for x below $m-v$ (low), x between $m-v$ and $m+v$ (middle) and x above $m+v$ (high values). Here, m and v are the overall mean and standard deviation of fat depth x for schemes S_1 , S_3 and S_4 and the separate means μ_1 , μ_2 and standard deviations σ_1 , σ_2 of x for the two subpopulations respectively for scheme S_2 . Selection is within subpopulations for S_1 and S_2 and over subpopulations, i. e. ignoring sexes or breeds, for S_3 . S_4 is similar to S_3 except that a fixed number of each subpopulation is selected. In rows 4, 5, 6 and 7 it is shown which calculations are performed under schemes S_1 , S_2 , S_3 and S_4 in the simulation: simple linear regression on x , polynomial regression (up to degree 3), approach of Engel and Walstra (EW-method).

Sampling scheme	Selection within subpopulations				Selection over subpopulations	
	S_1		S_2		S_3	S_4
Subpopulation	1	2	1	2	-	-
Fat class						
Low	24	24	24	24	48	48
Middle	12	12	12	12	24	24
High	24	24	24	24	48	48
Involving:						
Simple regression	yes		yes		yes	yes
Polynomial regression	yes		yes		yes	yes
EW-method	yes		yes		no	no

regression line for castrates for large fat depth x , it is not a simple linear function of x . This is illustrated in Figure 1 for the parameter values based on the Dutch data (Table 1a). The curve for $\mu(x)$ is better approximated by a cubic polynomial in x than by a straight line. As illustrated in Figure 3, the overall residual standard deviation $\sigma_{y,x}$ is not constant but depends on x . The residual standard deviation increases when the difference between the separate regression lines becomes more pronounced, i.e. for small fat depth, or when there is a marked overlap between the sexes, i.e. for intermediate values of fat depth.

Selection of the carcasses

In the simulation four sampling schemes were studied, referred to as S_1 , S_2 , S_3 and S_4 . Each sampling scheme involved 120 carcasses. In the first two sampling schemes (S_1 and S_2) separate samples of size 60 were taken for each sex. Carcasses were selected on the basis of fat depth in three classes. The numbers in these classes correspond to 40, 20 and 40% of the sample size and are shown in Table 2. The 40-20-40% scheme is quite common in dissection experiments. The boundary values for the fat depth classes were either the same for the two sexes (S_1) or different (S_2). In the former case the overall mean and standard deviation for fat depth were used to construct boundary values $\mu \pm \sigma$, where (Appendix A3):

$$\mu = \pi_1 \mu_1 + \pi_2 \mu_2 \text{ and } \sigma^2 = \pi_1 \sigma_1^2 + \pi_2 \sigma_2^2 + \pi_1 \pi_2 (\mu_1 - \mu_2)^2 \quad (5).$$

In the latter case the separate means and standard deviations for the sexes were employed: $\mu_1 \pm \sigma_1$ and $\mu_2 \pm \sigma_2$. In the third sampling scheme (S_3) 120 carcasses were selected employing the same fat classes as for S_1 , but ignoring the sexes (Table 2). In the fourth sampling scheme (S_4) sexes were also largely ignored, but care was taken that an equal number of both sexes was present in the sample.

Note that S_1 and S_4 are quite different sampling schemes with respect to the distribution of the sexes over the fat depth measurements. Schemes S_3 and S_4 are frequently used in practice. Although there is a mild form of selection on the sexes under S_4 , it is generally assumed that adverse effects are negligible. The simulation study will show that this assumption is indeed true.

For the subpopulations based on the Spanish data, the same sampling schemes were used, with equal numbers in the sample for each subpopulation for schemes S_1 and S_2 and an obvious modification for unequal proportions π_i for scheme S_4 .

Simulation of a dissection experiment

For selection schemes S_1 and S_2 , fat depths x were sampled from the distributions for each sex. Those values that fitted into the selection scheme were accepted, until sample sizes of 60 for each sex were reached. The corresponding lean meat percentages y were generated according to regressions (1a) and (1b) with the parameter values from Table 1a. For scheme S_3 , first the sex of a carcass was determined with probabilities π_1 and π_2 , then the fat depth x was generated according to the appropriate normal distribution and when x fitted into the scheme, the lean meat percentage y was generated according to (1a) or (1b). For scheme S_4 , fat depth values x were alternately generated for gilts and castrates, until 120 carcasses were found that fitted into the selection scheme. Lean meat percentages were alternately generated by (1a) and (1b).

One thousand samples of size 120 were generated according to each of the four selection schemes. Each generated sample represents a dissection experiment, under one of the four selection schemes, such as might have occurred in practice. For all samples of 120 carcasses, for all four selection schemes, we calculated an overall prediction formula by regression of the lean meat percentage y on the fat depth x , ignoring the sexes and an overall formula according to polynomial regression also ignoring the sexes, for reasons explained in the next section. For schemes S_1 and S_2 we also calculated an overall prediction formula by the method of Engel and Walstra (1993). Which calculations that were performed under which scheme is summarized in Table 2.

Polynomial regression

We concluded (Figure 1) that the overall mean $\mu(x)$ is not a simple linear function of fat depth x but resembles a cubic polynomial in x . However, in view of the variation in the data (Figure 1), it is doubtful whether departures from linearity will be noticeable in a sample of 120 dissected carcasses. So far no marked departures from linearity in pig carcass grading have been reported. Therefore, in the simulation we had a particular interest in the t tests for the coefficients of the quadratic (x^2) and cubic (x^3) terms in the polynomial, as a test for lack of fit. In the simulation both simple (straight line) regression and polynomial regression were included. In both cases homogeneous residual variances were assumed, supposing that the heterogeneity of variances (Figure 3) goes unnoticed. When the coefficient of the cubic term x^3 was not significantly different from 0 it was dropped from the polynomial regression model. Subsequently, when the coefficient of the quadratic term x^2 was not significant it was discarded as well.

Evaluation of performance of prediction formulae

To assess the performance of a prediction formula we used a set of 1600 'standard' carcasses with pre-chosen values for fat depth x . These standard carcasses are representative for the pig population and mimic a large random sample. The number of standard carcasses per subpopulation was proportional to the number in the population as reflected by proportions π_i , i.e. 800 gilts and 800 castrates in the simulation based on the Dutch data. When two instrumental variables were involved in the simulation, e.g. a fat and a muscle depth

measurement, the standard carcasses were chosen with pre-chosen fat and muscle depth values. Details are presented in **Appendix A4**. For the four subpopulations based on the Spanish data, the number of standard carcasses was changed to 1521, since this number could more conveniently be combined with the proportions π_1, π_2, π_3 and π_4 (Table 1b).

For each sampling scheme, for each of the 1000 simulated dissection experiments, a true lean meat percentage y was generated for each standard carcass

Table 3 Bias and mean squared prediction error (MSPE) for two sexes

	MSPE _{av}	MSPE ₁	MSPE ₂	Bias _{av}	Bias ₁	Bias ₂	P ₂	P ₃
Average bias and MSPE (Bias _{av} and MSPE _{av}) and bias and MSPE per sex (Bias ₁ , Bias ₂ , MSPE ₁ and MSPE ₂ for gilts and castrates respectively) for sampling schemes S ₁ .. S ₄ . Results for linear regression (ignoring sexes) and the EW-method (S ₁ and S ₂ only). Proportion of 1000 simulations with a significant coefficient for cubic term (P ₂) and quadratic term (P ₃), after the cubic term has been removed, in the polynomial regression. Input parameters are from Table 1a and based on the Dutch data.								
Sampling scheme S ₁								
Simple regression								
Mean	5.94	5.54	6.38	-0.10	-1.16	0.96	0.04	0.04
s. d.	0.25	0.56	0.48	0.20	0.21	0.21		
2.5% percentile	5.48	4.63	5.52	-0.47	-1.55	0.55		
97.5% percentile	6.47	6.70	7.35	0.29	-0.74	1.38		
EW-method								
Mean	5.74	4.91	6.61	0.01	-0.84	0.87		
s. d.	0.22	0.41	0.47	0.20	0.21	0.24		
2.5% percentile	5.34	4.23	5.76	-0.38	-1.26	0.41		
97.5% percentile	6.20	5.84	7.58	0.42	-0.41	1.32		
Sampling scheme S ₂								
Simple regression								
Mean	5.80	5.12	6.52	-0.01	-0.98	0.96	0.03	0.13
s. d.	0.23	0.48	0.51	0.21	0.22	0.23		
2.5% percentile	5.40	4.31	5.63	-0.42	-1.41	0.54		
97.5% percentile	6.27	6.16	7.55	0.41	-0.56	1.41		
EW-method								
Mean	5.74	4.96	6.56	-0.01	-0.87	0.85		
s. d.	0.22	0.43	0.48	0.21	0.22	0.23		
2.5% percentile	5.34	4.22	5.70	-0.41	-1.30	0.42		
97.5% percentile	6.17	5.88	7.56	0.41	-0.45	1.30		
Sampling scheme S ₃								
Simple regression								
Mean	5.96	5.62	6.35	-0.13	-1.19	0.94	0.06	0.06
s. d.	0.26	0.65	0.52	0.23	0.25	0.24		
2.5% percentile	5.48	4.50	5.47	-0.58	-1.67	0.47		
97.5% percentile	6.48	7.01	7.47	0.30	-0.70	1.40		
Sampling Scheme S ₄								
Simple regression								
Mean	5.95	5.55	6.37	-0.10	-1.17	0.96	0.04	0.06
s. d.	0.25	0.57	0.49	0.20	0.22	0.22		
2.5% percentile	5.48	4.52	5.53	-0.52	-1.60	0.53		
97.5% percentile	6.46	6.79	7.36	0.29	-0.74	1.38		

and compared with the predicted percentage \hat{y} from the formula. The true lean meat percentages y of the standard carcasses were simulated with the same model and parameter values as the samples that were used for linear regression. The differences $(\hat{y} - y)$ and squared differences $(\hat{y} - y)^2$ were averaged over the 1000 simulated experiments and plotted against fat depth x to obtain an impression of the bias and mean squared prediction error (MSPE). In addition to these plots, differences and squared differences were averaged over the standard carcasses to obtain an average bias and MSPE per subpopulation and an overall bias and MSPE over the total population. Some details are in **Appendix A4**.

Results

Two sexes from the Dutch data

Results for the simulation based on the Dutch data are presented in Table 3. In the last column of the table the proportions of the 1000 simulated dissection experiments are shown where the coefficient of the quadratic and cubic term in the polynomial regression was significant according to a t test. Clearly, departures from a simple straight line could hardly be detected in any of the four sampling schemes, since these proportions barely exceeded the nominal significance level of 0.05. There was no apparent gain in accuracy in the use of polynomials (results not shown).

Table 4 Bias and mean squared prediction error (MSPE) for two breeds

	MSPE _{av}	MSPE ₁	MSPE ₂	Bias _{av}	Bias ₁	Bias ₂	P ₂	P ₃
Average bias and MSPE (Bias _{av} and MSPE _{av}) and bias and MSPE per breed (Bias ₁ , Bias ₂ , MSPE ₁ and MSPE ₂ for breeds 1 and 2 respectively) for sampling schemes S ₁ . . S ₄ . Results for linear regression (ignoring breeds) and EW-method (S ₁ and S ₂ only). Proportion of 1000 simulations with a significant coefficient for the cubic term (P ₃) and quadratic term (P ₂), after the cubic term has been removed, in the polynomial regression. Input parameters are from Table 1a and based on the Spanish data.								
Sampling scheme S ₁								
Simple regression								
Mean	11.80	10.77	12.60	-0.46	1.90	-2.29	0.03	0.05
s. d.	0.51	1.01	1.28	0.24	0.25	0.25		
2.5% percentile	10.85	8.90	10.36	-0.96	1.39	-2.79		
97.5% percentile	12.84	12.85	15.28	0.01	2.39	-1.80		
EW-method								
Mean	11.04	12.59	9.85	-0.01	2.10	-1.65		
s. d.	0.41	1.16	0.94	0.25	0.26	0.27		
2.5% percentile	10.28	10.44	8.23	-0.52	1.56	-2.19		
97.5% percentile	11.82	14.98	11.92	0.47	2.59	-1.13		
Sampling scheme S ₂								
Simple regression								
Mean	11.31	11.51	11.15	-0.27	1.98	-2.01	0.11	0.33
s. d.	0.44	1.10	1.10	0.26	0.27	0.26		
2.5% percentile	10.49	9.48	9.12	-0.76	1.44	-2.49		
97.5% percentile	12.22	13.85	13.32	0.25	2.49	-1.50		
EW-method								
Mean	11.04	12.58	9.85	-0.01	2.10	-1.65		
s. d.	0.41	1.18	0.89	0.26	0.27	0.26		
2.5% percentile	10.28	10.43	8.23	-0.49	1.57	-2.11		
97.5% percentile	11.88	15.15	11.68	0.49	2.61	-1.12		
Sampling scheme S ₃								
Simple regression								
Mean	11.57	11.99	11.25	-0.17	2.16	-1.97	0.06	0.06
s. d.	0.53	1.50	1.57	0.33	0.33	0.35		
2.5% percentile	10.64	9.33	8.68	-0.84	1.51	-2.66		
97.5% percentile	12.74	15.16	14.60	0.48	2.81	-1.33		
Sampling scheme S ₄								
Simple regression								
Mean	11.48	12.16	10.95	-0.12	2.21	-1.92	0.05	0.05
s. d.	0.46	1.19	1.11	0.25	0.26	0.26		
2.5% percentile	10.62	9.92	8.89	-0.63	1.69	-2.42		
97.5% percentile	12.42	14.65	13.29	0.38	2.72	-1.39		

For schemes S_1 and S_2 , the method by Engel and Walstra (1993), from now on referred to as the EW-method, performed better than linear regression ignoring sexes, with respect to overall bias ($Bias_{av}$)

and mean squared prediction error ($MSPE_{av}$). Also the variation in $MSPE_{av}$ over the 1000 simulated experiments, as reflected by the 2.5 and 97.5 percentile points, was lower for the EW-method.

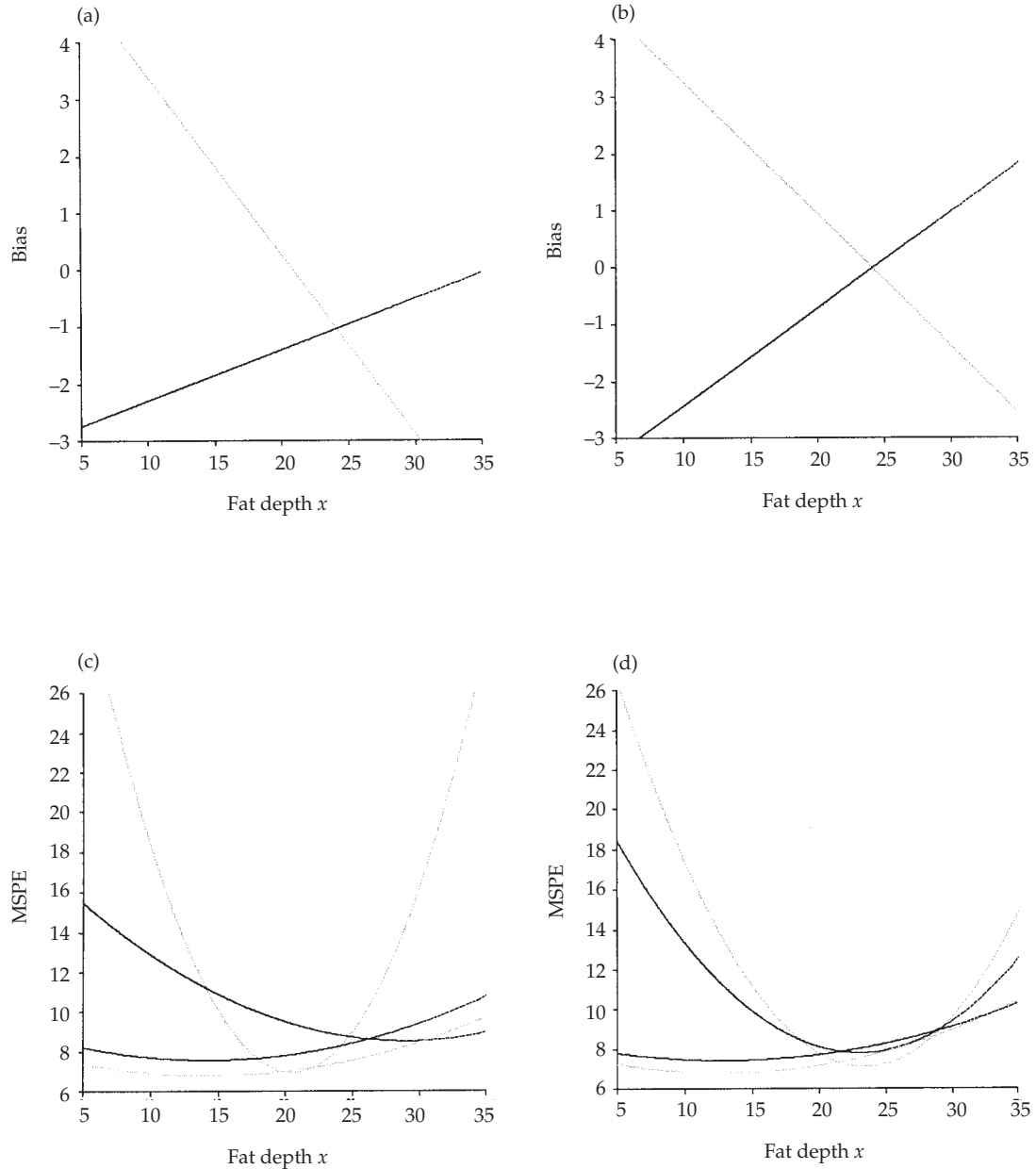


Figure 4 The bias for each of two breeds from the Spanish data (input parameters from Table 1a) for sample schemes S_2 and S_4 (Figure 4a and b). A black line for breed 1 and a grey line for breed 2. Bias for separate regression lines per breed is nearly 0. The mean squared prediction error (MSPE) for each of two breeds from the Spanish data (input parameters from Table 1a) for sample schemes S_2 and S_4 (Figure 4c and d) MSPEs for separate regression lines per breed are depicted as well (relatively flat curves near the horizontal axis).

Table 5 Bias and mean squared prediction error (MSPE) for four subpopulations

Average bias and MSPE (Bias_{av} and MSPE_{av}) and bias and MSPE per breed (Bias_v, MSPE_v for subpopulations 1...4) for sampling schemes S₁...S₄. Results for linear regression (ignoring subpopulations) and EW-method (S₁ and S₂ only). Proportion P of 1000 simulations with a significant result for the F-test for departures from linearity (test on quadratic terms and product term). Input parameters are from Table 1b and based on the Spanish data.

	MSPE _{av}	MSPE ₁	MSPE ₂	MSPE ₃	MSPE ₄	Bias _{av}	Bias ₁	Bias ₂	Bias ₃	Bias ₄	P
Sampling scheme S ₁											
Simple regression											
Mean	6.05	6.32	5.39	4.96	7.22	-0.21	1.09	0.36	-0.56	-1.14	0.06
s. d.	0.25	0.71	0.43	0.47	0.69	0.21	0.24	0.23	0.24	0.24	
2.5% percentile	5.58	5.05	4.60	4.07	5.97	-0.61	0.61	-0.06	-0.98	-1.60	
97.5% percentile	6.59	7.75	6.28	5.93	8.61	0.21	1.57	0.83	-0.10	-0.68	
EW-method											
Mean	5.91	6.39	5.49	4.95	6.28	-0.01	1.11	0.48	-0.50	-0.74	
s. d.	0.23	0.72	0.45	0.47	0.54	0.22	0.24	0.24	0.24	0.28	
2.5% percentile	5.46	5.07	4.66	4.05	5.30	-0.45	0.63	0.02	-0.96	-1.27	
97.5% percentile	6.39	7.83	6.45	5.90	7.48	0.43	1.57	0.95	-0.05	-0.18	
Sampling scheme S ₂											
Simple regression											
Mean	5.96	6.45	5.37	5.06	6.68	-0.13	1.07	0.42	-0.54	-0.99	0.15
s. d.	0.24	0.74	0.44	0.49	0.61	0.22	0.25	0.24	0.25	0.26	
2.5% percentile	5.48	5.12	4.59	4.17	5.60	-0.55	0.60	-0.01	-0.99	-1.47	
97.5% percentile	6.45	8.03	6.35	6.12	7.97	0.35	1.57	0.91	-0.04	-0.44	
EW-method											
Mean	5.91	6.40	5.51	4.96	6.28	-0.01	1.11	0.47	-0.50	-0.74	
s. d.	0.23	0.74	0.46	0.48	0.53	0.23	0.25	0.25	0.25	0.27	
2.5% percentile	5.45	5.07	4.68	4.06	5.31	-0.48	0.60	-0.01	-0.99	-1.30	
97.5% percentile	6.42	7.92	6.50	6.00	7.41	0.47	1.61	0.99	0.01	-0.18	
Sampling scheme S ₃											
Simple regression											
Mean	5.98	6.81	5.52	4.85	6.71	-0.02	1.26	0.55	-0.38	-0.95	0.07
s. d.	0.24	0.84	0.47	0.46	0.64	0.23	0.27	0.25	0.25	0.26	
2.5% percentile	5.54	5.41	4.67	4.02	5.55	-0.48	0.73	0.04	-0.91	-1.45	
97.5% percentile	6.46	8.72	6.51	5.83	8.04	0.43	1.78	1.03	0.11	-0.47	
Sampling scheme S ₄											
Simple regression											
Mean	5.98	6.87	5.53	4.86	6.68	-0.01	1.27	0.57	-0.36	-0.94	0.08
s. d.	0.23	0.86	0.48	0.45	0.62	0.24	0.27	0.25	0.26	0.26	
2.5% percentile	5.52	5.35	4.68	4.04	5.51	-0.45	0.77	0.07	-0.84	-1.43	
97.5% percentile	6.42	8.85	6.49	5.82	7.91	0.45	1.85	1.09	0.16	-0.43	

However, the gain in accuracy was small and fairly unimportant.

In all instances the overall bias ($Bias_{av}$) was practically negligible. The average bias within the sexes ($Bias_1$ and $Bias_2$) was about -1.1% for gilts and 0.96% for castrates. For small values of fat depth x (not shown) bias may run up to about -1.5 and 1.5% respectively. Hence, the use of a single overall prediction formula induced a systematic difference between the sexes of 2.1% in percentage lean, while for selected values of fat depth this difference may run up to 3% . When separate prediction formulae are used this bias will vanish.

As noted in the **Introduction**, linear (and polynomial) regression under schemes S_1 and S_2 is not covered by standard regression theory, because of the selection on fat depth within sexes. Nevertheless, for simple regression the $MSPE_{av}$ under S_2 was the smallest among the sampling schemes. The average intercept and slope for the fitted regression lines for S_1 , S_3 and S_4 were close to the averages of the intercepts a_1 and a_2 and slopes b_1 and b_2 (Table 1) for the two sexes, i.e. fitted lines for S_1 , S_3 and S_4 were on average in between the two regression lines for the sexes. Under S_2 there were relatively more gilts for low fat depth and relatively

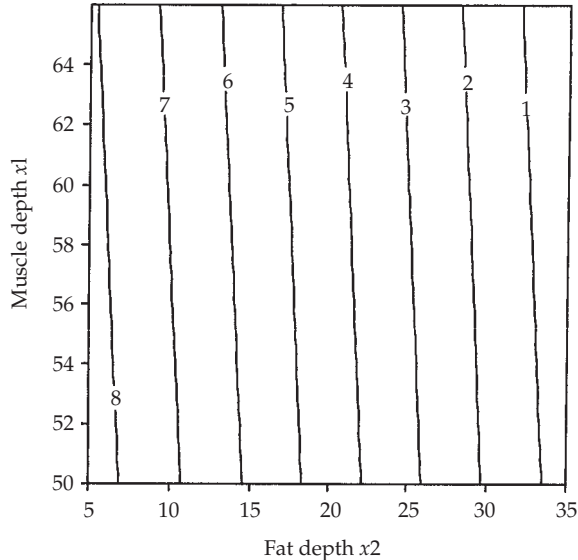


Figure 5 Contour plot for sampling scheme S_3 of the relative bias between subpopulations 1 and 4 for the simulation with four subpopulations. Relative bias as a function of fat and muscle depth measurements. Input parameters are from Table 1b and based on the Spanish data. Contours numbered 1 . . . 8 correspond to a relative bias of $-3, -2, \dots, 4\%$ respectively.

more castrates for high fat depth compared with S_1 , S_3 and S_4 . Consequently, for S_2 the fitted lines had a larger intercept, declined more steeply and tended to be closer to the non-linear curve for the overall mean (Figure 1) than for S_1 , S_3 and S_4 . This illustrates that for linear regression, ignoring the sexes, the differences between the sampling schemes were largely determined by the way lack of fit was incorporated in the predictions through the configurations of the instrumental measurements. Despite the fact that there was no apparent lack of fit in the simulated data.

For regression under sampling schemes S_3 and S_4 , the schemes commonly used in dissection experiments, the root of the MSPE was 2.4 and 2.5% for gilts and castrates respectively. For small values of fat depth x the root MSPE may run up to 2.5 and 3% (not shown) for gilts and castrates respectively. When separate prediction formulae are used, the root MSPE will be close to 2.0 and 2.3%, the residual standard deviations for gilts and castrates respectively, when at least moderately sized samples, say 60 carcasses, are collected for each sex. So, although differences between the sexes were not marked, there was some room for improvement with respect to bias and MSPE.

Two breeds from the Spanish data

For the two Spanish breeds the proportions π_1 and π_2 in the population differed. Consequently, the numbers of standard carcasses for the breeds also differed: 697 and 903 for breeds 1 and 2 respectively. The numbers of dissected carcasses in the simulation were still equal to 60 for each breed. The simulation results are presented in Table 4.

Again differences between linear regression and polynomial regression (not shown) were small. There was an increased probability of detecting departures from a straight line under scheme S_2 . Under S_3 and S_4 , the sampling schemes commonly used in practice, the probabilities were close to the nominal significance level of 0.05.

Differences between regression ignoring breeds and the EW-method were more pronounced for the Spanish breeds compared with the Dutch sexes. The average bias ($Bias_{av}$) for simple regression ignoring the breeds was -0.5% under scheme S_1 . For the EW-method the average bias was negligible, as it should be on theoretical grounds (Engel and Walstra, 1993). Under scheme S_1 , the reduction in $MSPE_{av}$ for the EW-method was about 6% of the MSPE for simple regression, with some loss of accuracy for breed 1 and a relatively larger gain in accuracy for breed 2.

Under S_3 and S_4 , the average bias within each breed was about 2.2 and -2.0% for linear regression for breeds 1 and 2 respectively. Bias under S_2 clearly differed from the other sampling schemes. This is illustrated in Figure 4 where bias is plotted against fat depth for S_2 and S_4 (plots for S_1 and S_3 are similar to S_4). Under S_3 and S_4 bias may run up to 3.5% and -2.5% for breeds 1 and 2 respectively for low fat depth. So, with a single overall prediction formula, there was an average difference between the breeds of some 4%, which ran up to 6% for selected values of fat depth.

Under S_3 and S_4 , for simple linear regression, the root MSPE was about 3.4% for breeds 1 and 2 (square root of $MSPE_{av}$ in Table 4). When separate prediction formulae are used, the root MSPE will be close to 2.6%, the residual standard deviation for the two breeds, when at least moderately sized samples, 60 carcasses say, are collected for each breed. In Figure 4 the MSPE per breed for simple regression for schemes S_2 and S_4 is plotted against fat depth for a single overall prediction formula and for separate formulae based on 60 dissected carcasses each. Again there was a marked difference between S_2 and the other schemes. Under S_3 and S_4 the root MSPE for linear regression may run up to some 4% for lean pigs.

For simple regression, the $MSPE_{av}$ under S_2 was the smallest among the sampling schemes. The explanation is the same as for the sexes in the Dutch data. Also, for scheme S_2 the probability of detecting departures from linearity was distinctly larger than the nominal 0.05 level. Results with the EW-method were still superior to linear regression under S_2 . The EW-method under schemes S_1 and S_2 performed better than simple regression under S_1 . S_4 with respect to $MSPE_{av}$. Differences between S_3 and S_4 were minor. Apparently, there is no harm, and perhaps some slight gain, in forcing the numbers of animals in the sample to be proportional to the numbers in the population.

So, accuracy was appreciably increased by the use of the EW-method, and markedly increased by the use of separate prediction formulae compared with simple linear regression under schemes S_1 .. S_4 .

Four subpopulations from the Spanish data

In Table 1b input parameters are shown for four subpopulations derived from the Spanish data. The first subpopulation is also the first subpopulation in the aforementioned study with two Spanish breeds. The first three subpopulations are maternal lines which are used in the majority of crosses in Spain. The fourth subpopulation consists of five sire lines,

pooled together. Both the FOM fat and muscle depth measurements were employed. Multivariate normality for the lean meat percentage and instrumental measurements was assumed within each breed. Selection of carcasses was performed on fat depth only in conformance with Table 2.

To test for lack of fit, quadratic terms and a product term were added to the regression model. Various significance tests were performed for these additional terms. The most powerful test was the F-test on the quadratic terms and product term. Although the P-values (Table 5) were larger than the nominal 0.05 level, the power to detect departures from linearity was low.

The EW-method performed better than linear regression under schemes S_1 and S_2 , but the gain was small.

The relative bias between subpopulations 1 and 4 was 2.2% for S_3 and S_4 . The relative bias may run up to 3 to 4% for selected values of (mostly) fat and muscle depth (Figure 5). The root $MSPE_{av}$ for S_3 and S_4 was about 2.4%, and root MSPE ranged from 2.2 to 2.6% for the separate subpopulations. By contrast the residual standard deviations for the separate regressions per subpopulation ranged from 2.1 to 2.3%. The largest improvements were for S_1 and S_4 . Note that compared with the aforementioned study with two Spanish breeds, the use of both a fat and muscle depth measurement offered a considerable reduction in the MSPE. Consequently the largest gain in root MSPE to be made by employing a separate formula is about 13% of its value for a single common formula. The use of separate prediction formulae eliminates the relative bias between the subpopulations and increases the accuracy of prediction as reflected by the MSPE.

Discussion

We decided on a simulation study because this was by far the simplest way to handle the different sampling strategies under the four sampling schemes that were considered. It was assumed that within subpopulations lean meat percentages and instrumental measurements are linearly related. This was supported by the Dutch and Spanish data. As a consequence, the induced relationship between instrumental carcass measurements and lean meat percentages over the subpopulations, i.e. for the pig population as a whole, was not linear (Figure 1). Serious departures from linearity in dissection studies have not been reported so far. The sampling schemes used in dissection experiments are similar to schemes S_3 and S_4 in our study. For these schemes we found that the probability of detecting departures

from linearity, by testing the significance of quadratic and cubic terms in a polynomial regression, hardly exceeded the nominal significance level of 0.05. The choice of model in the simulation therefore seems an entirely reasonable one. To simplify the calculations, multivariate normality was assumed for instrumental carcass measurements and lean meat percentages. Although in practice some variables, e.g. fat depth, may be skewed, this is not expected to interfere with our conclusions to any great extent. Moreover, serious departures from normality suggest marked departures from linearity in regression and, again, such departures have not been reported yet. The results can be expected to apply to double regression (Engel and Walstra, 1991a and b; Causeur and Dhorne, 1998), a cost saving alternative for ordinary linear regression, as well.

In Engel *et al.* (2003) the effect was studied of selection of carcasses on variables which are not all included in the final prediction formula. Here, in schemes S_1 and S_2 , the indicator variables for the subpopulations are involved in the selection of the carcasses for dissection, but not in the simple regression for the final formula. In Engel *et al.* (2003), it was shown that adverse effects of selection on continuous variables, that are not involved in regression later on, can be considerable. Notably, the residual variance can be seriously inflated. Here, only minor differences between the residual standard deviations under the four schemes were observed (not shown). Under schemes S_1 and S_2 , the method by Engel and Walstra (1993) offered an improvement over simple regression, when differences between subpopulations were marked.

When a single overall prediction formula was used, we found a bias in the predicted lean meat percentage of 2% between gilts and castrates in the Dutch data and of 4% between some of the Spanish breeds. For lean pigs, the systematic differences ran up to 3 and 6.5% for the sexes and breeds respectively. This bias can be eliminated and the mean squared prediction error can be markedly reduced, either by introduction of extra prediction variables that are able to discriminate between subpopulations or by the use of separate prediction formulae for the subpopulations. Our simulation mimics conditions that might easily occur within the EC. Although results do not completely reflect the accuracy of the HGP and FOM as used in The Netherlands and Spain, we concluded that for some subpopulations it can be worthwhile to find extra prediction variables or to use separate prediction formulae. When the Hampshire breed was introduced in Sweden, to reduce problems with PSE,

Hampshire crosses were found to have meatier hams than the conventional White pigs. However, the fat depth, as measured by the HGP, was the same for these subpopulations. Fortunately, introduction of an extra muscle depth measurement of the *m. longissimus dorsi* largely resolved the problem (Ingemar Hansson, personal communication). In general, extra instrumental carcass measurements that markedly increase discriminative power between subpopulations are hard to find, but the use of separate prediction formulae will become feasible in the near future with the advance of technology for implants. This would ensure a fairer payment for producers who happen to specialize in subpopulations with a sizeable negative bias and would substantially increase the accuracy of prediction.

Possibly, a requirement for a minimal sample size for subpopulations has to be included in the EC-regulations. Otherwise, any gain in accuracy by use of separate formulae may be lost due to the extra variation originating from small samples of carcasses per subpopulation. If so, what the minimum sample size per subpopulation should be will be a point of discussion. When several subpopulations are considered, the total sample size required may possibly exceed the present minimum sample size of 120 carcasses as required by the EC-regulations (EC, 1994). In the EUPIGCLASS research project (EUPIGCLASS, 2000), the scope for indirect methods based on imaging technology to determine the lean meat percentage of pig carcasses is investigated. Although such indirect techniques are too time consuming and costly to apply in the slaughterline, they may offer a reliable and much cheaper alternative to the present EC reference dissection method (Walstra and Merkus, 1996). In that case the cost of calculating separate prediction formulae for subpopulations would not be a serious impediment. A need for additional practical measures can be anticipated. For instance, suppose that use of a separate prediction formula is proposed for a subpopulation at the initiative of one or more producers. It will probably have to be decided whether the introduction of a separate formula is acceptable. The considerations, at a national level and in general, will include both the problem concerning a period of transition and the estimation of a separate formula for the remaining pigs in the population. On a national level it is advisable for dissection experiments to be performed under the supervision of one or more approved agencies, that are recognized for their technical skills, in cooperation with certified statisticians. Otherwise, when new formulae are proposed more frequently in

the future, quality of sampling, use of proper statistical methodology, proper handling of outliers and missing data and correct and critical model validation cannot be sufficiently guaranteed. The increasing use of instruments that collect large numbers of carcass measurements and the complexity of the statistical methodology that is required for proper handling of this data, make the need for expert certified supervision even more pressing.

Acknowledgements

This research was conducted under the EC project G6RD-CT-1999-00127 EUPIGCLASS of the Measurements and Testing activity (Competitive and Sustainable Growth programme).

References

- Causeur, D. J. and Dhorne, T. J. 1998. Finite sample properties of a multivariate extension of double regression. *Biometrics* **54**: 1591-1601.
- Cox, D. R. and Snel, E. J. 1989. *Analysis of binary data, second edition*. Chapman and Hall, London.
- Daumas, G., Causeur, D., Dhorne, T. and Schollhammer, E. 1998. The new pig carcass grading methods in France. *Proceedings of the 44th ICoMST, Barcelona, Spain*.
- Daumas, G., Dhorne, T. and Gispert, M. 1994. Accounting for the sex effect on prediction of pig carcass lean meat percentage in the Community. *Proceedings of the fourth ICoMST, The Hague, The Netherlands*.
- Engel, B., Buist, W. G., Walstra, P., Olsen, E. and Daumas, G. 2003. Accuracy of prediction of percentage lean meat and authorization of carcass measurement instruments: adverse effects of incorrect sampling of carcasses in pig classification. *Animal Science* **76**: 199-209.
- Engel, B. and Walstra, P. 1991a. Increasing precision or reducing expense in regression experiments by using information from a concomitant variable. *Biometrics* **47**: 13-20.
- Engel, B. and Walstra, P. 1991b. A simple method to increase precision or reduce expense in regression experiments to predict the proportion of lean meat of carcasses. *Animal Production* **53**: 353-359.
- Engel, B. and Walstra, P. 1993. Accounting for subpopulations in prediction of the proportion of lean meat of pig carcasses. *Animal Production* **57**: 147-152.
- European Commission. 1988. *Commission decision authorizing methods for grading pig carcasses in Spain (88/479/EEC)*.
- European Commission. 1994. *Commission decision amendment to the decision authorizing methods for grading pig carcasses in Spain (94/337/EEC)*.
- European Community. 1989a. EC document VI/3860/89, proposal for research concerning the harmonisation of methods for grading pigs in the Community.
- European Community. 1989b. EC decision no. 89/53/CEE related to the approval of classification methods of pig carcasses in Italy.
- European Community. 1994. EC regulation no. 3127/94, amending regulation (EC) no. 2967/85 laying down detailed rules for the application of the Community scale for grading pig carcasses.
- EUPIGCLASS. 2000. *Standardisation of pig carcass classification in the EU through improved statistical procedures and new technological developments. Program funded under the Growth Programme*. www.cordis.lu/en/home.html
- GenStat Committee. 2000. *The guide to GenStat* (ed. R. Payne). VSN International Ltd, Oxford.
- Gispert, M., Dhorne, T., Diestre, A. and Daumas, G. 1996. Efecto del sexo en la predicción del contenido en magro de la canal porcina en Espana, Francia y Países Bajos. *Investigación Agraria, Producción y Sanidad Animales* **11**: 19-28.
- McCullagh, P. and Nelder, J. A. 1989. *Generalized linear models, second edition*. Chapman and Hall, London.
- Walstra, P. 1986. *Assessment of the regression formula for estimation of the lean meat percentage by HGP-measurements in The Netherlands*. EC working paper, Brussels VI/4849/86.
- Walstra, P. and Merkus, G. S. M. 1996. *Procedure for assessment of the lean meat percentage as a consequence of the new EU reference dissection method in pig carcass classification*. Research report 96-014 Institute for Animal Science and Health, (ID-Lelystad), Lelystad, The Netherlands.

(Received 1 April 2003—Accepted 18 September 2003)

Appendix

Expressions are presented for T subpopulations and, to keep notation simple, for one prediction variable x only. Generalization to several prediction variables is straightforward.

A1. Conditional probabilities for subpopulations

$$p_i(x) = P(\text{subpopulation } t \mid x) = P(x \mid \text{subp. } t) \pi_t / \sum_i \pi_i P(x \mid \text{subp. } i) = \{\pi_t \exp(-1/2(x-\mu_t)^2/\sigma_t^2)/\sigma_t\} / \{\sum_i \pi_i \exp(-1/2(x-\mu_i)^2/\sigma_i^2)/\sigma_i\},$$

so

$$\log(p_i(x)/p_1(x)) = \{\log(\pi_i) - \log(\pi_1) + \log(\sigma_1) - \log(\sigma_i)\} + \{\mu_i/\sigma_i^2 - \mu_1/\sigma_1^2\} x + \{(\sigma_1^2 - \sigma_i^2)/(2\sigma_1^2\sigma_i^2)\} x^2.$$

For $t = 1$ and $T = 2$: $\log(p_i(x)/p_1(x)) = \text{logit}(p_i(x)) = \gamma_0 + \gamma_1 x + \gamma_2 x^2$, where $\gamma_0 = \log(\pi_1/\pi_2) + \log(\sigma_2/\sigma_1) + 1/2 \mu_2^2/\sigma_2^2 - 1/2 \mu_1^2/\sigma_1^2$, $\gamma_1 = \mu_1/\sigma_1^2 - \mu_2/\sigma_2^2$ and $\gamma_2 = (\sigma_1^2 - \sigma_2^2)/(\sigma_1^2\sigma_2^2)$. For more than two subpopulations, the logistic regression model can be replaced by a multinomial logistic model (Cox and Snell, 1989). The latter model can be fitted with standard software for log-linear models for count data (McCullagh and Nelder, 1989, §6.4.2).

A2. Means and variances for percentage lean over subpopulations

$$\begin{aligned} E(y \mid x) &= \mu(x) = \sum_i p_i(x) E(y \mid x, \text{subp. } t) = \sum_i p_i(x) \mu_i(x) = \sum_i p_i(x) (a_i + b_i x), \\ \text{Var}(y \mid x) &= E_i(\text{Var}(y \mid x, \text{subp. } t)) + \text{Var}_i(E(y \mid x, \text{subp. } t)) = \\ &= \sum_i p_i(x) \sigma_{i,y,x}^2 + \sum_i p_i(x) (\mu_i(x) - \mu(x))^2. \end{aligned}$$

This yields expressions (4a) and (4b) for $T = 2$.

A3. Means and variances for fat depth over subpopulations

$$\begin{aligned} E(x) &= \mu = E_i(E(x \mid \text{subpopulation } t)) = \sum_i \pi_i \mu_i, \\ \text{Var}(x) &= \sigma^2 = E_i(\text{Var}(x \mid \text{subp. } t)) + \text{Var}_i(E(x \mid \text{subp. } t)) = \sum_i \pi_i \sigma_i^2 + \sum_i \pi_i (\mu_i - \mu)^2, \end{aligned}$$

where E_i and Var_i denote expectation and variance with respect to subpopulation membership. This yields expression (5) for $T = 2$ subpopulations.

A4. Choice and use of standard carcasses

Let N be the total number of standard carcasses and $n_i = \pi_i N$ be the number per subpopulation, $t = 1 \dots T$. For the t th subpopulation, n_i measurements x_{it} for x were generated according to $x_{it} = \Phi^{-1}(i/(n_i + 1))\sigma_i + \mu_i$, $i = 1 \dots n_i$, where Φ^{-1} denotes the inverse of the cumulative probability function of the standard normal distribution (the probit function). This parallels the construction of a random sample from a normal distribution, replacing random numbers from a homogeneous distribution on the interval (0,1) by their expected order statistics as approximated by $i/(n_i + 1)$. When two prediction variables were involved standard carcasses were chosen to mimic a random sample from a bivariate normal distribution in parallel with the Box-Muller algorithm (Engel *et al.*, 2002; **Appendix A1**).

Differences $\hat{y} - y$ between predicted and true lean meat percentages and squared differences $(\hat{y} - y)^2$ were averaged to obtain the bias (Bias_i , $\text{Bias}_{2, \dots}$) and mean squared prediction error (MSPE) (MSPE_i , $\text{MSPE}_{2, \dots}$) per subpopulation. The overall bias (Bias_{av}) and MSPE (MSPE_{av}) were calculated by averaging over all standard carcasses. In the latter case, for each new assessment of a prediction formula, for each standard carcass, sex or breed were sampled according to the conditional probabilities (**Appendix A1**) corresponding to the standard instrumental measurements. This is because overall bias and MSPE are conditional upon (standard) carcass measurements only, while bias and MSPE per subpopulation are, evidently, conditional upon sex or breed as well.