Journal of
Animal Breeding and Genetics

J. Anim. Breed. Genet. ISSN 0931-2668

ORIGINAL ARTICLE

Accounting for unknown foster dams in the genetic evaluation of embryo transfer progeny

M.J. Suárez[1], S. Munilla[2] & R.J.C. Cantet[2,3]

1 Advanta Semillas SAIC-Nutrisun Business Unit, Balcarce Biotechnology Center, Balcarce, Argentina
2 Departamento de Producción Animal, Facultad de Agronomía, Universidad de Buenos Aires, Buenos Aires, Argentina
3 Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina

Keywords
Measurement error model; missing recipient dam; records from ET calves.

Correspondence
R.J.C. Cantet, Departamento de Producción Animal, Facultad de Agronomía, Universidad de Buenos Aires, 1417, Ciudad Autónoma de Buenos Aires, Argentina.
Tel: 54 11 4524 8000 extension 8184;
Fax: 54 11 4524 8735;
E-mail: rcantet@agro.uba.ar

Received: 30 May 2014;
accepted: 8 September 2014

Summary

Animals born by embryo transfer (ET) are usually not included in the genetic evaluation of beef cattle for preweaning growth if the recipient dam is unknown. This is primarily to avoid potential bias in the estimation of the unknown age of dam. We present a method that allows including records of calves with unknown age of dam. Assumptions are as follows: (i) foster cows belong to the same breed being evaluated, (ii) there is no correlation between the breeding value (BV) of the calf and the maternal BV of the recipient cow, and (iii) cows of all ages are used as recipients. We examine the issue of bias for the fixed level of unknown age of dam (AOD) and propose an estimator of the effect based on classical measurement error theory (MEM) and a Bayesian approach. Using stochastic simulation under random mating or selection, the MEM estimating equations were compared with BLUP in two situations as follows: (i) full information (FI); (ii) missing AOD information on some dams. Predictions of breeding value (PBV) from the FI situation had the smallest empirical average bias followed by PBV obtained without taking measurement error into account. In turn, MEM displayed the highest bias, although the differences were small. On the other hand, MEM showed the smallest MSEP, for either random mating or selection, followed by FI, whereas ignoring measurement error produced the largest MSEP. As a consequence from the smallest MSEP with a relatively small bias, empirical accuracies of PBV were larger for MEM than those for full information, which in turn showed larger accuracies than the situation ignoring measurement error. It is concluded that MEM equations are a useful alternative for analysing weaning weight data when recipient cows are unknown, as it mitigates the effects of bias in AOD by decreasing MSEP.

Introduction

Calves that are born by embryo transfer (ET) procedures are usually the progeny from selected bulls and cows. Genetic evaluation for preweaning growth of these animals requires specific models to disentangle additive genetic and environmental sources of direct and maternal effects. As a standard, Van Vleck (1990) proposed a maternal animal model that requires handling equations for the 'three parents': the sire, the donor dam and the foster dam. In addition, proper model specification of the data from ET calves requires the knowledge of the breed and the age of the foster dam (Schaeffer & Kennedy 1989). The former requirement was related to the use of dairy or crossbred cows as surrogate dams during the stage of dissemination of ET in beef cattle. Nowadays, however, the use of grade or commercial cows of the same

  doi:10.1111/jbg.12121

breed is the rule. The latter requirement instead may become an issue, especially in purebred operations with a sizable fraction of calves born using ET techniques where identification of foster cows is usually not recorded. For example, 7.25 and 3.87% of all calves evaluated in Brangus and Braford, respectively, in Argentina are calves born from ET, and none of them has their foster dam recorded. Leaving those animals out of the evaluation is out of the question, as they are usually the most profitable.

If maternal effects of foster dams are assumed to be uncorrelated with the breeding values of their ET calves, the analysis of weaning records from ET calves with unknown foster cows is afflicted by the missing value of age of dam (AOD), usually a fixed effect. The goal of this research was to present a way of attenuating this problem. In the mixed linear model, the expected value of the data vector $\boldsymbol{y}$ is calculated as

$$E(\boldsymbol{y}|\boldsymbol{X},\ \boldsymbol{Z}),$$

assuming that the incidence matrices of fixed ($\boldsymbol{X}$) and random effects ($\boldsymbol{Z}$) are known. However, when some individuals have unknown AOD, a column of $\boldsymbol{X}$ will relate records from those animals to the parameter representing the level 'unknown AOD'. This situation is quite rare in the statistical literature, as the universal situation is when the entire values of the covariate are unknown, or the classification variable is completely misclassified. Therefore, we approach the problem as a classical measurement error model (MEM, Fuller 1987; Buzas *et al.* 2005; Carroll 2005), and for easiness of presentation, the consequence of measurement error on a covariate is presented first from a simple linear regression model. Later on, we deal with the case of the mixed model with measurement error in a set of fixed effects, and we obtain estimators of fixed effects and predictors of breeding values by maximizing the posterior density of a MEM discussed by Buonaccorsi *et al.* (2000). Finally, we evaluate the effects of the predictors obtained using simulated data from a stochastic simulation experiment, and we illustrate the procedure using preweaning data from Brangus calves.

Material and methods

Theory

The MEM and the covariate AOD

To simplify the mathematics of MEM, consider the following linear regression model for weaning weight records on age of dam (AOD) with some of the records having missing AOD. The latter variable is considered continuous and is measured in years. Then, the model equation is equal to

$$y_i = z_i\,\beta_z + \ x_i\beta_X + e_i, \tag{1}$$

where y_i ($i = 1,\ldots, n$) represents the weaning weight record of animal i, z_i and x_i are observed and unobserved values of AOD, respectively, and β_Z and β_X are the corresponding regression coefficients. Errors are such that $e_i \sim N\left(0, \sigma_e^2\right)$, and assumed independent of both, z_i and x_i.

As x_i is unobserved, a surrogate variable w_i will be used instead. In our particular setting, w_i will be an indicator variable pointing out to those records of calves with unknown age of dam. In the classical MEM (Fuller 1987; Buzas *et al.* 2005), this surrogate variable w_i is modelled as the true (unobserved) covariate x_i plus a measurement error u_i, such that

$$w_i = x_i + u_i, \tag{2}$$

where $u_i \sim N(0, \sigma_U^2)$ and assumed independent of $x_i \sim N(\mu_X, \sigma_X^2)$. Under this formulation, w_i is an unbiased estimator of a realized value of x_i. This is seen as follows:

$$\begin{aligned} \mathrm{E}(w_i|x_i) &= \mathrm{E}\,(w_i) + \frac{\mathrm{Cov}(w_i,\ x_i)}{\mathrm{Var}(x_i)}\left[x_i - \mathrm{E}(x_i)\right] \\ &= \mathrm{E}(x_i + u_i) + \frac{\mathrm{Cov}(x_i + u_i, x_i)}{\sigma_X^2}(x_i - \mu_X) \\ &= \mu_X + \frac{\sigma_X^2}{\sigma_X^2}(x_i - \mu_X) = \mu_X + \ x_i - \mu_X = x_i. \end{aligned} \tag{3}$$

Replacing now the unobserved covariate x_i in equation (1) by the surrogate variable w_i as defined in equation (2), the following model is obtained:

$$y_i = z_i\,\beta_Z^* + \ w_i\beta_X^* \ + \varepsilon_i, \tag{4}$$

with $\mathrm{E}(\varepsilon_i) = 0$. The asterisks emphasize the fact that the regression coefficients have been redefined and that the original ones are to be estimated from an approximate model [i.e. model (4)]. As we will show next, in this situation the least-squares estimator of β_X will be biased.

An important remark is in place before going on. For the procedure to be valid, the measurement error for w_i must be 'non-differential' (Buzas *et al.* 2005). Formally, this occurs when

$$f\,(y_i|z_i, x_i, w_i) = f(y_i|z_i, x_i). \tag{5}$$

Stated in words, the error is non-differential whenever the proxy w_i does not introduce any additional knowledge in case the (unobserved) covariate x_i is known and included in the model. In our setting, the proxy w_i has non-differential measurement error when: 1. it can occur at any age of dam; 2. individuals

that belong to the missing AOD group do not deviate genetically from the rest. For the first condition, this is generally the rule with ET records, as cows of any age are used at the time the embryo has to be implanted. However, the second condition could become a problem whenever ET calves belong to a particularly selected group. In our experience with field data from breeds with an open policy of registration, this problem is somehow alleviated as usually there is an important proportion of females served by AI or natural mating whose age is unknown. This latter fact entails a randomization of the level 'missing AOD' that renders measurement error non-differential, thus attenuating any association of the missing AOD group with the potentially higher genetic mean of ET calves.

As defined so far and given the conditions just described, the observed z_i and the unobserved x_i are uncorrelated as long as the assumption of non-differential error holds; thus, $\mathrm{Cov}(z_i, x_i) = 0$. Moreover,

$$\begin{aligned}\mathrm{Cov}(e_i, w_i) &= \mathrm{Cov}(e_i, x_i + u_i)\\ &= \mathrm{Cov}(e_i, x_i) + \mathrm{Cov}(e_i, u_i) = 0.\end{aligned} \tag{6}$$

Using these results, the covariance between a record and the surrogate variable w_i is equal to

$$\begin{aligned}\mathrm{Cov}(y_i, w_i) &= \mathrm{cov}[z_i\beta_Z + x_i\beta_X + e_i, w_i]\\ &= \mathrm{cov}[z_i\beta_Z + x_i\beta_X + e_i, x_i + u_i]\\ &= \mathrm{cov}[z_i\,\beta_Z, x_i] + \mathrm{cov}[z_i\,\beta_Z, u_i]\\ &\quad + \mathrm{cov}[x_i\,\beta_X, x_i] + \mathrm{cov}[x_i\,\beta_X, u_i] + \mathrm{cov}[e_i, w_i]\\ &= \mathrm{cov}[z_i\,, x_i]\,\beta_Z + \mathrm{cov}[z_i\,, u_i]\beta_Z + \mathrm{cov}[x_i, x_i]\,\beta_X\\ &\quad + \mathrm{cov}[x_i, u_i]\beta_X + \mathrm{cov}[e_i, w_i]\\ &= 0 + 0 + \beta_X\sigma_X^2 + 0 + 0 = \beta_X\sigma_X^2\end{aligned} \tag{7}$$

Also, the variance of y_i under the usual assumption that known levels of AOD are fixed equals to

$$\begin{aligned}\mathrm{Var}(y_i) &= \mathrm{Var}(y_i|z_i) = \mathrm{Var}[z_i\,\beta_Z + x_i\,\beta_X + e_i|z_i]\\ &= \mathrm{Var}[x_i\,\beta_X + e_i] = \beta_x^2\sigma_X^2 + \sigma_e^2.\end{aligned} \tag{8}$$

With all previous results, we are now in position to obtain the expected values of the least-squares estimators of the regression coefficients β_X^* and β_Z^* in model (4). For the former,

$$E\left(\hat{\beta}_X^*\right) = E\left(\hat{\beta}_X^*|w_i\right) = \frac{\mathrm{cov}[y_i, w_i]}{\mathrm{Var}(w_i)}.$$

And, on using expression (7), we obtain

$$\frac{\mathrm{cov}[y_i, w_i]}{\mathrm{Var}(w_i)} = \frac{\beta_X\sigma_X^2}{\mathrm{Var}(x_i + u_i)} = \frac{\beta_X\sigma_X^2}{\sigma_X^2 + \sigma_U^2} = \beta_X\lambda, \tag{9}$$

where

$$\lambda = \frac{\mathrm{Var}(x_i)}{\mathrm{Var}(w_i)} = \frac{\sigma_X^2}{\sigma_X^2 + \sigma_U^2} \tag{10}$$

is termed the 'attenuation factor' or 'reliability ratio' (Fuller 1987; Buzas *et al.* 2005; Carroll 2005). The inverse of λ, denoted as θ, is the linear correction for attenuation and is formally defined as

$$\frac{1}{\lambda} = \frac{\sigma_X^2 + \sigma_U^2}{\sigma_X^2} = \theta. \tag{11}$$

Equation (9) shows that measurement error tends 'to shrink' β_X towards zero as $0 < \lambda \leq 1$, with λ being the attenuation (or 'attenuation to the null'). Moreover, β_X^* is a biased estimator of β_X. However, if λ (or equivalently θ) is known or there exists a reasonable estimate of it, a simple correction for attenuation of the bias can be applied as follows:

$$\hat{\beta}_X = \hat{\beta}_X^*\,\hat{\theta}. \tag{12}$$

(Fuller 1987; expression (1.1.7) in page 5). In turn, given that z_i and x_i are uncorrelated, $\hat{\beta}_Z^*$ is an unbiased estimator of β_Z as shown by Carroll (2005, page 15):

$$\begin{aligned}E\left(\hat{\beta}_Z^*\right) &= \beta_Z + \beta_Z\beta_X(1-\lambda)\frac{\mathrm{cov}\,(x_i, z_i)}{\mathrm{Var}(z_i)}\\ &= \beta_Z + \beta_Z\beta_X(1-\lambda)\,\frac{0}{\mathrm{Var}(z_i)} = \beta_Z.\end{aligned} \tag{13}$$

Finally, we will obtain the variance for a record subject to measurement error. By conditioning on the observed variable w_i, this variance is equal to

$$\begin{aligned}\mathrm{Var}(y_i\,|w_i) &= \mathrm{Var}(y_i)\left(1 - \rho_{YW}^2\right)\\ &= \sigma_Y^2\left(1 - \frac{\mathrm{cov}(y_i\,, w_i)^2}{\mathrm{Var}(y_i)\mathrm{Var}(w_i)}\right)\\ &= \sigma_Y^2\left(1 - \frac{\beta_X^2\left(\sigma_X^2\right)^2}{\sigma_Y^2\mathrm{Var}(w_i)}\right)\\ &= \sigma_Y^2 - \sigma_Y^2\frac{\beta_X^2\left(\sigma_X^2\right)^2}{\sigma_Y^2\mathrm{Var}(w_i)} = \sigma_Y^2 - \frac{\beta_X^2\left(\sigma_X^2\right)^2}{\mathrm{Var}(w_i)}\\ &= \sigma_Y^2 - \frac{\beta_X^2\left(\sigma_X^2\right)^2}{\sigma_X^2 + \sigma_U^2}\\ &= \sigma_Y^2 - \beta_X^2\sigma_X^2\left(\frac{\sigma_X^2}{\sigma_X^2 + \sigma_U^2}\right) = \sigma_Y^2 - \beta_X^2\sigma_X^2\lambda\end{aligned} \tag{14}$$

where σ_Y^2 represents the variance of y_i obtained in equation (8). On replacing with the latter in (14) yields

$$\begin{aligned}\mathrm{Var}(y_i|w_i) &= \left(\beta_X^2\sigma_X^2+\sigma_e^2\right) - \beta_X^2\sigma_X^2\lambda = \beta_X^2\sigma_X^2(1-\lambda)+\sigma_e^2\\ &= \beta_X^2\sigma_X^2\left(1-\frac{\sigma_X^2}{\sigma_X^2+\sigma_U^2}\right)+\sigma_e^2\\ &= \beta_X^2\sigma_X^2\left(\frac{\sigma_X^2+\sigma_U^2-\sigma_X^2}{\sigma_X^2+\sigma_U^2}\right)+\sigma_e^2\\ &= \beta_X^2\sigma_X^2\left(\frac{\sigma_U^2}{\sigma_X^2+\sigma_U^2}\right)+\sigma_e^2\\ &= \beta_X^2\sigma_U^2\left(\frac{\sigma_X^2}{\sigma_X^2+\sigma_U^2}\right)+\sigma_e^2 = \beta_X^2\sigma_U^2\lambda+\sigma_e^2\end{aligned} \tag{15}$$

Expression (15) shows that the variance of any observation with measurement error is increased by an amount $\beta_X^2\sigma_U^2\lambda$, compared with the variance of any observation for which x_i is directly observed. Thus, observing w_i is always less informative than observing x_i.

Estimating equations for mixed models with measurement error in the fixed effects

The case of the mixed model with measurement error in some of the fixed effects is now considered. Estimators of fixed effects and predictors of breeding values are obtained by maximizing the log of the posterior density of fixed and random effects. Using a frequentist approach, Buonaccorsi (2010), section 11.3.3, page 382) called the procedure 'pseudo maximum likelihood'. Estimates of the (co)variance components and λ are assumed to be available, and estimators and predictors are calculated using the estimated dispersion parameters in place of the true ones.

Without loss of generality, the vector of observations is ordered such that records with known AOD ($\boldsymbol{y}_o$) precede those with unknown AOD ($\boldsymbol{y}_u$). Accordingly, let

$$\begin{bmatrix}\boldsymbol{X}_o & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{X}_u\end{bmatrix}$$

be the matrix relating records with the fixed parameters of the factor AOD. As matrix $\boldsymbol{X}_u$ is unobserved, we replace it with the estimated conditional mean of $\boldsymbol{X}_u$ given $\boldsymbol{W}_u$. In our setting, $\boldsymbol{W}_u$ is a vector with all elements equal to one in the rows of animals with records and unknown AOD. The conditional mean of $\boldsymbol{X}_u$ given $\boldsymbol{W}_u$ is equal to

$$\begin{aligned}&\mathrm{E}(\boldsymbol{X}_u|\boldsymbol{W}_u)=\mathrm{E}(\boldsymbol{X}_u)\\ &+\mathrm{Cov}(\boldsymbol{X}_u,\boldsymbol{W}_u)(\mathrm{Var}(\boldsymbol{W}_u))^{-1}(\boldsymbol{W}_u-\mathrm{E}(\boldsymbol{W}_u))\\ &=\mathrm{E}(\boldsymbol{X}_u)+\left(\boldsymbol{I}\sigma_X^2-\left(\boldsymbol{I}\sigma_W^2\right)^{-1}\right)(\boldsymbol{W}_u-\mathrm{E}(\boldsymbol{W}_u))\\ &=\mathrm{E}(\boldsymbol{X}_u)+\boldsymbol{I}\left(\frac{\sigma_X^2}{\sigma_X^2+\sigma_u^2}\right)(\boldsymbol{W}_u-\mathrm{E}(\boldsymbol{W}_u))\\ &=\mathrm{E}(\boldsymbol{X}_u)+\lambda(\boldsymbol{W}_u-\mathrm{E}(\boldsymbol{W}_u))=(1-\lambda)\mathrm{E}(\boldsymbol{X}_u)+\lambda\boldsymbol{W}_u\end{aligned} \tag{16}$$

In turn, the conditional variance is equal to

$$\begin{aligned}&\mathrm{Var}(\boldsymbol{X}_u|\boldsymbol{W}_u)=\mathrm{Var}(\boldsymbol{X}_u)\\ &-\mathrm{Cov}(\boldsymbol{X}_u,\boldsymbol{W}_u)(\mathrm{Var}(\boldsymbol{W}_u))^{-1}\mathrm{Cov}(\boldsymbol{W}_u,\boldsymbol{X}_u)\\ &=\boldsymbol{I}\sigma_X^2-\boldsymbol{I}\sigma_X^2\left(\boldsymbol{I}\sigma_W^2\right)^{-1}\boldsymbol{I}\sigma_X^2\\ &=\boldsymbol{I}\sigma_X^2\left(1-\frac{\sigma_X^2}{\sigma_X^2+\sigma_u^2}\right)=\boldsymbol{I}\sigma_X^2(1-\lambda)\end{aligned} \tag{17}$$

Let $\boldsymbol{f}=\hat{\mathrm{E}}(\boldsymbol{X}_u|\boldsymbol{W}_u)=\left(1-\hat{\lambda}\right)\widehat{\mathrm{E}}(\boldsymbol{X}_u)+\hat{\lambda}\boldsymbol{W}_u$ with the hat denoting an estimator of the parameter. Then, an animal model with missing AOD in a fraction of records is written as

$$\begin{bmatrix}\boldsymbol{y}_o\\ \boldsymbol{y}_u\end{bmatrix}=\boldsymbol{X}_1\beta_1+\begin{bmatrix}\boldsymbol{X}_o & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{f}\end{bmatrix}\begin{bmatrix}\beta_o\\ \beta_u\end{bmatrix}+\boldsymbol{Z}\boldsymbol{a}+\begin{bmatrix}\boldsymbol{e}_o\\ \boldsymbol{e}_u\end{bmatrix} \tag{18}$$

In (18), matrices $\boldsymbol{X}_1$, $\boldsymbol{X}_o$ and $\boldsymbol{Z}$ relate records to fixed effects measured with certainty, to known classes of AOD and to breeding values, respectively, with corresponding vectors $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_o$ and $\boldsymbol{a}$, whereas vector $\boldsymbol{f}$ relates records with unknown AOD to the parameter β_u for the average value of the class missing AOD. Random effects in model (18) are assumed to follow a multivariate normal distribution with

$$\mathrm{Var}\begin{bmatrix}\boldsymbol{a}\\ \boldsymbol{e}_o\\ \boldsymbol{e}_u\end{bmatrix}\sim N\begin{bmatrix}\boldsymbol{G} & \boldsymbol{0} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{I}\sigma_e^2 & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{I}\left(\sigma_e^2+\sigma_X^2(1-\lambda)\right)\end{bmatrix} \tag{19}$$

The increased variability present in $\mathrm{Var}(\boldsymbol{e}_u)$ of (19) is due to using $\hat{E}(\boldsymbol{X}_u|\boldsymbol{W}_u)$ instead of $\boldsymbol{X}_u$, and is taken from expression (17). Estimating equations for model (18) and equations (19) can be obtained using a Bayesian approach such as in Dempfle (1977). On writing

$$\boldsymbol{R}=\begin{bmatrix}\boldsymbol{I}\sigma_e^2 & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{I}\left(\sigma_e^2+\sigma_X^2(1-\lambda)\right)\end{bmatrix}$$

and on assuming normality, the likelihood is equal to

$$\begin{aligned}&f(\boldsymbol{y}|\beta,\beta_o,\beta_u,\boldsymbol{u},\boldsymbol{R})\\ &\qquad=N(\boldsymbol{X}_1\beta_1+\boldsymbol{X}_o\beta_o+\boldsymbol{f}\beta_u+\boldsymbol{Z}\boldsymbol{u},\boldsymbol{Z}\boldsymbol{G}\boldsymbol{Z}'+\boldsymbol{R}),\end{aligned}$$

whereas the prior distribution of all $\boldsymbol{\beta}$'s is proportional to a constant (K, say), that is $f(\beta,\beta_o,\beta_u)\propto K$. Finally, the prior of $\boldsymbol{u}$ is equal to $f(u)=N(0,G)$. Thus, the joint distribution is proportional to

$$f(\beta,\ \beta_o,\beta_u\ ,\ \boldsymbol{u}\,|\,\boldsymbol{y},\ \boldsymbol{R})\propto f(\boldsymbol{y}|\beta,\beta_o,\beta_u,\boldsymbol{u},\boldsymbol{R})f(\boldsymbol{u}|\boldsymbol{G})\boldsymbol{K}.$$

Estimating equations for all linear parameters are now obtained by maximizing the log of this posterior density (and denoted by F):

$$\log[f(\beta, \beta_o, \beta_u, \boldsymbol{u}|\boldsymbol{y}, \boldsymbol{R})] \propto$$
$$-\frac{1}{2}\left[\left(\boldsymbol{y} - \boldsymbol{X}_1\beta_1 - \boldsymbol{X}_o\beta_o - \boldsymbol{f}\beta_u - \boldsymbol{Zu}\right)'\boldsymbol{R}^{-1}\right.$$
$$\left.\left(\boldsymbol{y} - \boldsymbol{X}_1\beta_1 - \boldsymbol{X}_o\beta_o - \boldsymbol{f}\beta_u - \boldsymbol{Zu}\right) + \boldsymbol{u}'\boldsymbol{G}^{-1}\boldsymbol{u}\right] = F$$

Now, let

$$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{X}_1 & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{X}_o & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{f} \end{bmatrix} \text{ and } \beta = \begin{bmatrix} \beta_1 \\ \beta_o \\ \beta_u \end{bmatrix}$$

Taking derivatives of F with respect to $\boldsymbol{\beta}$ and $\boldsymbol{u}$ produces

$$\frac{\partial F}{\partial \beta} = \boldsymbol{X}'\boldsymbol{R}^{-1}(\boldsymbol{y} - \boldsymbol{X}\beta - \boldsymbol{Zu})$$

$$\frac{\partial F}{\partial u} = \boldsymbol{Z}'\boldsymbol{R}^{-1}(\boldsymbol{y} - \boldsymbol{X}\beta - \boldsymbol{Zu}) - \boldsymbol{G}^{-1}\boldsymbol{u}$$

And, after equating them to zero, the following system of equations is obtained

$$\begin{bmatrix} \boldsymbol{X}_1'\boldsymbol{R}^{-1}\boldsymbol{X}_1 & \boldsymbol{X}_1'\boldsymbol{R}^{-1}\boldsymbol{X}_o & \boldsymbol{X}_1'\boldsymbol{R}^{-1}\boldsymbol{f} & \boldsymbol{X}_1'\boldsymbol{R}^{-1}\boldsymbol{Z} \\ \boldsymbol{X}_o'\boldsymbol{R}^{-1}\boldsymbol{X}_1 & \boldsymbol{X}_o'\boldsymbol{R}^{-1}\boldsymbol{X}_o & \boldsymbol{X}_o'\boldsymbol{R}^{-1}\boldsymbol{f} & \boldsymbol{X}_o'\boldsymbol{R}^{-1}\boldsymbol{Z} \\ \boldsymbol{f}'\boldsymbol{R}^{-1}\boldsymbol{X}_1 & \boldsymbol{f}'\boldsymbol{R}^{-1}\boldsymbol{X}_o & \boldsymbol{f}'\boldsymbol{R}^{-1}\boldsymbol{f} & \boldsymbol{f}'\boldsymbol{R}^{-1}\boldsymbol{Z} \\ \boldsymbol{Z}'\boldsymbol{R}^{-1}\boldsymbol{X}_1 & \boldsymbol{Z}'\boldsymbol{R}^{-1}\boldsymbol{X}_o & \boldsymbol{Z}'\boldsymbol{R}^{-1}\boldsymbol{f} & \boldsymbol{Z}'\boldsymbol{R}^{-1}\boldsymbol{Z} + \boldsymbol{G}^{-1} \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_o \\ \hat{\beta}_u \\ \hat{u} \end{bmatrix} = \begin{bmatrix} \boldsymbol{X}_1'\boldsymbol{R}^{-1}\boldsymbol{y} \\ \boldsymbol{X}_o'\boldsymbol{R}^{-1}\boldsymbol{y} \\ \boldsymbol{f}'\boldsymbol{R}^{-1}\boldsymbol{y} \\ \boldsymbol{Z}'\boldsymbol{R}^{-1}\boldsymbol{y} \end{bmatrix} \quad (20)$$

System (20) looks like a set of mixed model equations in which $\boldsymbol{f}$ replaces the unobserved matrix $\boldsymbol{X}_u$, whereas $\boldsymbol{R}$ takes into account the extra variability in records from animals with missing AOD, but estimators and predictors are no longer unbiased. Henderson (1984, chapter 9) observed that 'biased predictors and estimators exist that have smaller mean-squared errors than BLUE and BLUP'. Later on, the mean-squared error of prediction (MSEP) of BV from equations (20) is examined by means of stochastic simulation.

Estimation of σ^2_X and σ^2_U

The variance components related with the MEM part of the model, that is σ^2_X and σ^2_U, can be estimated from the data at hand. To avoid any interference of genetic effects, records are assumed to be precorrected by predictions of breeding values from the last genetic evaluation ($\hat{u}$ say), in a vector $\boldsymbol{y}_*$ such that $y_* = y - Z\hat{u}$. The model for analysis then is equal to

$$\begin{bmatrix} y_{*o} \\ y_{*u} \end{bmatrix} = X_1\beta_1 + \begin{bmatrix} X_o \\ 0 \end{bmatrix}\beta_o + \begin{bmatrix} \varepsilon_o \\ \varepsilon_u \end{bmatrix}$$

Subindices are equal as in model (18), but now effects of AOD are assumed to be random such that $\beta_o \sim N(0, I\sigma^2_X)$. Also the error variance is equal to $\begin{bmatrix} I\sigma^2_e & 0 \\ 0 & I\sigma^2_{eu} \end{bmatrix}$ and $\sigma^2_{eu} = \sigma^2_e + \sigma^2_X + \sigma^2_U$. In this formulation, σ^2_X is estimated from the variability of known AOD classes, whereas σ^2_U is estimated by contrasting error variances from records of unknown and known AOD. Then, a REML or a Bayesian algorithm such as Gibbs sampling can be employed to estimate σ^2_X, σ^2_e and σ^2_{eu}. Finally, measurement error variance is estimated from $\hat{\sigma}^2_U = \hat{\sigma}^2_{eu} - \hat{\sigma}^2_e - \hat{\sigma}^2_X$.

Simulation of data

The base programme to simulate the data was originally written by Cantet *et al.* (2000) for an animal population with overlapping generations under selection, a trait measured in both sexes, a yearly mating season and females having single progeny in any given year. For the current application, the programme was modified to introduce maternal effects and the presence of ET. The base population consisted of 10 sires and 200 dams. The number of breeding animals was kept constant during 5 years (or selection events), to obtain a data set of at most 2000 records per replicate. The trait simulated was weaning weight. Fixed effects in the model were classification variables of sex and AOD, whereas age at weaning was a covariate. The simulated value of the regression coefficient of the trait on age (b) was equal to 0.750 kg/day. Sex was assigned at random with equal probability, and males weighted 20 kg more than females. Five classes of AOD were simulated for ages 2, 3, 4, 5 to 8, and 9 to 12 years old. Corresponding values simulated for the five classes were 4, 12, 14, 18 and 12 kg. A 20% loss in AOD identification was randomly simulated for every replicate. Breeding values of animals in the base generation were simulated by premultiplying a 2 × 1 vector of $N(0,1)$ random variables by the Cholesky decomposition $\left(G_0^{-1/2}\right)$ of the 2 × 2 covariance matrix

$$\boldsymbol{G}_0 = \begin{bmatrix} 150.0 & -37.5 \\ -37.5 & 150.0 \end{bmatrix}$$

The error variance was set equal to 487.5 kg^2 in order for direct and maternal heritabilities to be equal to 0.30. The genetic correlation between direct and maternal effects was equal to −0.25. All of these values can be characterized as typical of the estimates found for weaning weight of beef cattle (Koots *et al.* 1994). For the animals born in years 2–5, direct and

maternal breeding values were obtained by adding the Mendelian residuals to half the sum of paternal and maternal breeding values. Mendelian residuals for direct and maternal effects were simulated as $\sqrt{\frac{1}{2}\left[1-\frac{F_S+F_D}{2}\right]}G_0^{-1/2}\begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$ (Cantet *et al.* 1992; pages 213–214), where z_i is a $N(0,1)$ random variable and F_S and F_D are the inbreeding coefficients of the sire and the dam of the individual, respectively. All animals had records, and, except for the base individuals, they have both biological parents known. Inbreeding coefficients were calculated by the algorithm of Quaas (1976).

Matings of breeding animals were at random while avoiding sire-daughter and dam-son matings. Embryo transfers were randomly chosen to 10% of the matings, and 8 progenies were produced in each of them. The three oldest males and the 20 oldest females were culled each generation. The replacements were from progeny born during the same year and selected on predicted direct breeding values calculated with either regular mixed model equations, or equations (20). To obtain a solution, the estimate of the AOD class for the 2-year-old females was set equal to zero. All records were used to build the estimating equations. Either random mating or truncation selection using the greatest predicted direct BV was practiced. A total of 1000 replicates were run for each combination of the following factors: (i) random mating versus selection; (ii) a) 'Full': complete information on AOD versus b)'Missing': 20% of the records with the value of AOD missing at random plus in all data from ET calves, and the data analysis ignores the presence of measurement errors versus c) 'MEM': the pattern of missing records is the same as in b) but the analysis *accounts* for the presence of measurement errors. To evaluate the effects of MEM on AOD, the following parameters were computed: (i) empirical average bias for the effects (a) age of calf, (b) difference between sexes and (c) AOD: estimable functions between the difference among pairs of effects; (ii) empirical average bias for direct BV; (iii) mean-square error of prediction (MSEP) of direct BV; (iv) empirical accuracy of prediction, calculated as the correlation between true and predicted BV from (a) mixed model equations for the complete data, (b) mixed model equations for data with missing AOD and *ignoring* measurement errors and (c) predictions from expression (20) for data with missing AOD and *accounting* for measurement errors.

For each of the four situations simulated, σ^2_X and σ^2_U were estimated from a data set created by running 50 replicates. All data were precorrected with the predicted direct or maternal BVs, except for the records from ET calves that were only corrected for direct BVs of the individuals. The model of estimation had fixed effects of age of calf and sex, and random effects of AOD from dams with known age. As discussed before, heterogeneity of residual variance was assumed for data with either known or unknown AOD. The method of estimation was REML (Patterson & Thompson 1971) using the EM algorithm as presented by Henderson (1984). Thus, the variance σ^2_X was estimated from the variability among AOD subclasses from data of animals with known AOD. Once the algorithm converged, the variance σ^2_U was estimated by subtracting from the error variance of the data with missing AOD both, the error variance from data with known AOD as well as the estimate of σ^2_X. To write equations (20), the vector $\boldsymbol{f}$ was of order equal to the number of records with missing AOD, such that $f' = [1 \quad .. \quad 1 \quad .. \quad 1]f$. The value of the scalar $f = \hat{E}(x|w)$ was estimated from (16) as

$$f = \left(1-\hat{\lambda}\right)\hat{\mathrm{E}}(\mathrm{x}) + \hat{\lambda}\mathrm{w} \qquad (21)$$

Analysis of the Brangus data set

Preweaning growth records (birth and weaning weights) of 137 304 beef cattle born from 1974 to 2010 were used to illustrate the procedure. The records were from herds enrolled in the genetic evaluation programme (ERBra) of the Argentinean Brangus Association. There were 96 470 animals with known AOD, and 40 834 individuals with unknown AOD, either from ET, or simply from missing AOD. Two analyses were run under a multiple trait animal model, and the vector of breeding values was enlarged to accommodate maternal effects for weaning weight: (i) conventional analysis; (ii) MEM by solving equations (20) (for both birth and weaning weight). Dispersion parameters of the MEM process were estimated in the same manner as indicated for the stochastic simulation.

Results

Estimates of dispersion parameters for the MEM process obtained from the reduced data set (50 replicates) were $\hat{\sigma}^2_X = 5.7$, $\hat{\sigma}^2_e$ (from known AOD) = 518.3, and $\hat{\sigma}^2_{eU}$ (from unknown AOD) = 618.3. Therefore, $\hat{\sigma}^2_U = 618.3 - 518.3 - 5.7 = 94.3$, so that $\hat{\lambda} = \hat{\sigma}^2_X/(\hat{\sigma}^2_X + \hat{\sigma}^2_U) = 0.053926$. From (16), f was estimated to be equal to

$$f = (1 - 0.053926)(0.44) + 0.053926(1) = 0.4702$$

Table 1 Bias, mean-square error and accuracy of prediction from simulated data

	Random mating			Selection		
	Full	Missing	MEM	Full	Missing	MEM
Bias in estimable functions						
Effect of calf age	0.040	0.033	0.001	0.037	0.034	0.001
Difference of sexes	0.546	0.447	0.123	0.218	0.082	−0.070
AOD missing – AOD 6	–	0.714	0.683	–	0.336	0.811
AOD 2 – AOD 6	0.586	1.375	1.929	0.641	1.221	1.855
AOD 3 – AOD 6	−0.248	0.648	1.341	−0.430	0.479	1.289
AOD 4 – AOD 6	−0.138	0.527	1.175	−0.181	0.357	1.121
AOD 5 – AOD 6	−0.451	0.410	0.787	−0.484	−0.138	0.721
Bias of DBV	−0.07	0.217	0.471	0.081	0.332	0.641
MSEP of DBV	147.55	231.64	122.56	148.91	248.67	125.15
Average accuracy	0.42	0.29	0.49	0.47	0.30	0.53

AOD = age of dam; DBV = direct breeding value; Full, Missing, MEM = see text for definition. MSEP = mean-squared error of prediction. The value of the parameter for calf age was 0.75 kg/day, and for the difference of sexes 20 kg.

The value of $\hat{\mathrm{E}}(x) = 0.44$ is the estimated expected value of AOD, which was obtained from the data with known AOD. The value 0.44 is the standardized average age of dam (4.4 years) in the scale of $w = 1$, that is 4.4/10. The value of the denominator (10) is the maximum possible value of AOD in the simulation.

The results of the simulation are seen in Table 1.

Whereas the patterns of bias for fixed effects with regard to size and direction were not clearly defined, biases and MSEP for BVs were smaller for random mating than for selection. As expected, however, empirical bias was smaller for the estimates of fixed effects (calf age, difference between sexes, AOD) from the control in which AOD was known completely, than from the situation ignoring measurement error. In turn, ignoring measurement error produced smaller biases in the estimates of fixed effects than the MEM. The same order was found for predictions of BV: predictions from the full information situation had the smallest empirical average bias followed by predictions of BV obtained without taking measurement error into account, and finally MEM. However, the order was the opposite for MSEP: MEM displayed the smallest MSEP, for either random mating or selection, followed by full information, whereas ignoring measurement error produced the largest MSEP. Interestingly enough, as a consequence from the smallest MSEP with a relatively small bias, empirical accuracies observed from the MEM were larger than those for the full information, which in turn displayed much larger accuracies than ignoring measurement error. A word of caution is in order: BLUP minimizes variance in the class of unbiased predictors, whereas predictors calculated from equations (20) belong to a different class, that is the class of biased predictors seemingly minimizing MSEP = square bias + variance.

Dispersion parameter estimates for measurement error in the Brangus data estimated with REML-EM were equal to $\hat{\sigma}^2_{\mathrm{X}} = 0.75\,\hat{\sigma}^2_e$ and $\hat{\sigma}^2_{\mathrm{U}} = 0.16\,\hat{\sigma}^2_e$, so that the inverse of the ratio $\hat{\sigma}^2_{\mathrm{X}}/(\hat{\sigma}^2_{\mathrm{X}} + \hat{\sigma}^2_{\mathrm{U}})$ was equal to 1.214. Estimable functions for AOD effects under the conventional animal model and the animal model with measurement error in AOD (MEM) are presented in Table 2.

Estimable functions in both models were quite similar except for the difference between missing AOD and the highest level of AOD (cows > 8 year old). In the latter case, the estimate from the MEM was larger in magnitude than for the conventional animal model, for both birth weight (0.279 versus 0.264) and weaning weight (|−3.548| versus |−1.227|). A look at the estimates in Table 2 suggests that the average age of cows with missing AOD was close to 4.5 years old, a value similar to the one found in the simulation (4.4).

Table 2 Estimable functions for AOD (in kg) under the conventional animal model, and for the animal model with measurement error in AOD (MEM) in Brangus

	Birth weight		Weaning weight	
	Regular animal model	MEM	Regular animal model	MEM
AOD missing – AOD 6	0.264	0.279	−1.227	−3.548
AOD 2 – AOD 6	−0.723	−0.724	−9.519	−9.520
AOD 3 – AOD 6	−0.393	−0.394	−7.521	−7.522
AOD 4 – AOD 6	0.156	0.156	−3.694	−3.695
AOD 5 – AOD 6	0.478	0.478	1.324	1.325

Discussion

Guidelines for evaluating ET calves have been originated by research of Schaeffer & Kennedy (1989) and Van Vleck (1990), and have never been upgraded. Breeders have put many emphases in avoiding bias, a reassuring point of view for the 'frequentist' statistical school. Bias can only come from effects whose expectation is different from zero, such as fixed effects or breeding values with different expectation due to selection and loss of additive relationships. However, the distribution needed to calculate those expected values is conditional on the incidence matrix of the effects involved being correct, that is in the classical notation $\boldsymbol{X}$ and $\boldsymbol{ZQ}$, respectively. From time to time, field data are such that some rows of $\boldsymbol{X}$ and $\boldsymbol{ZQ}$ become uncertain (random) as the elements are imprecisely measured (with error) in: (i) a classification variable, or (ii) a covariate or (iii) the incorrect assignment of a genetic group. The first one is the case that we have considered in this research: missing AOD. Editing of data permits sorting some errors, but the edit cannot be as restrictive as to leave out an entire category of animals from the genetic evaluation. This is the case of records from ET calves with missing identification of the recipient dam. National shows preclude calves without EPDs to enter into the contest, and a great amount of calves at the shows are born from ET techniques. On the other hand, those animals born from ET have the potential to excel in selection differential. It is not the purpose of this research to suggest that foster dams of ET calves should not be identified. On the contrary, this is a sound practice for genetic evaluation purposes and, whenever possible, the model of Schaeffer & Kennedy (1989) to deal with foster dams and known age of dam should be used. However, if for any compelling reason identification and age of foster dams cannot be recorded, the procedure presented here allows evaluating the weaning weights from ET calves. The solution we have proposed uses classical measurement error theory (Fuller 1987; Buzas *et al.* 2005; Carroll 2005) and consists of replacing the unobserved value in the column of $\boldsymbol{X}_{\mathrm{u}}$ by the estimated conditional mean given that the animal has missing AOD, that is $\boldsymbol{W}_{\mathrm{u}}$. This approach to dealing with measurement error is referred to in the literature as 'regression calibration' (RC, Buzas *et al.* 2005; Buonaccorsi 2010; section 6.10). Freedman *et al.* (2008) used stochastic simulation to compare RC with similar techniques that replace $\boldsymbol{X}_{\mathrm{u}}$ by linear functions of $\boldsymbol{W}_{\mathrm{u}}$. These authors employed a multivariate normal model, like the one used in the simulation conducted in the current research, and observed that under non-differential measurement error, RC displayed minimum MSEP. This advantage of RC was observed even in situations of more drastic losses of information about $\boldsymbol{X}$ than the conditions simulated here, such as larger values of σ^2_{U} in relation to σ^2_e, or all values of AOD missing (i.e. $\boldsymbol{X} = \boldsymbol{X}_{\mathrm{u}}$), scenarios very unlikely to occur in animal breeding data. It should be stressed that RC methods as used here to deal with bias in AOD rely on the assumption of non-differential measurement errors, an assumption that can be essentially formulated as $\mathrm{cov}(y_i, w_i) = 0$ under multivariate normality. In section Theory, we have identified two conditions that must be satisfied to ensure non-differential measurement error. These conditions can now be restated in more practical terms as follows: 1. all recipient females should be from the same breed and from any age category, and the pattern of missing AOD should not be exclusively associated to them; 2. The BV for maternal effects of recipient cows must be uncorrelated to the direct BV of the calves producing the records. Violation of these conditions would introduce a non-zero value for the $\mathrm{cov}(y_i, w_i)$. Condition 1 may require discarding older data on ET calves when breed of foster cow is unknown, and there should also exist records from cows mated naturally or by artificial insemination with missing AOD. In practice, contemporary groups in which all records have missing AOD should be edited, as there will be no contrast among AOD classes. Condition 2 is most likely to hold in beef cattle as it is most difficult to synchronize donor cow pregnancies and flushing schedules to recipient cow heat with embryos of potentially higher BV. To solve equation (20), a programme that solves the mixed model equations needs to be modified to account for a missing category by (i) replacing the 1 in $\boldsymbol{X}$ that contributes to terms such as $\boldsymbol{X}'\boldsymbol{R}^{-1}\boldsymbol{X}$ or $\boldsymbol{Z}'\boldsymbol{R}^{-1}\boldsymbol{X}$ by

$$f = \hat{E}(x|w) = \left(1 - \hat{\lambda}\right)\hat{E}(x) + \hat{\lambda}w$$

as done in the simulation; (ii) keeping a diagonal covariance matrix of error terms, but using $\hat{\sigma}^2_e + \hat{\beta}^2_{\mathrm{X}}\hat{\sigma}^2_{\mathrm{U}}\hat{\lambda}$ rather than $\hat{\sigma}^2_e$ for the error variance of records with missing age of dam. In addition, the parameter β_{X} can be estimated by the value of the missing AOD in equations (20).

The stochastic simulation included two situations in comparison with the MEM estimator and predictors from equation (20). The full information situation differs from MEM in data, model and estimator of fixed effects and predictors of BV, and it is not an option to analyse ET data when information on foster dams is missing. If we had had completed information, we

would have used the full information model of Schaeffer & Kennedy (1989). Conversely, ignoring AOD effects from unknown foster dams differs in model and estimators, and results in a sizeable amount of bias and a concomitant reduction in empirical accuracy. Results of the simulation suggest that MEM had higher bias but lower MSEP of direct breeding values than full information or missing AOD. This is not surprising, as Efron (1975) indicated 'that certain deliberately induced biases can dramatically improve estimation properties when there are several parameters to be estimated'. Moreover, Henderson (1984) observed that 'biased predictors and estimators exist that have smaller mean-squared errors than BLUE and BLUP'. Although at first the higher empirical accuracy obtained for MEM when compared to full information seems to be striking, the smaller MSEP of MEM than the full information explains the result. Accuracy was calculated as the correlation between true and predicted BV. Therefore, if in the denominator the square root of the variance of the predictor (i.e. MSEP minus squared bias) decreases, accuracy increases. However, this higher accuracy in the MEM situation is for a model with less information and less random variables to predict (the BVs of foster dams), and not strictly comparable to the full information case. We conclude by insisting on the use of regular BLUP under the model of Schaeffer & Kennedy (1989) if information on foster dams is complete. However, MEM through equations (20) is a useful alternative when recipient cows are unknown, as it mitigates the effects of bias in AOD in those data by decreasing MSEP as compared with treating the class for missing AOD as a regular level of a fixed effect.

Acknowledgements

Funding for this research was provided by grants of Secretaría de Ciencia y Técnica, UBA, (UBACyT G861/2011) and CONICET (PIP 0833/2010). The authors would like to thank Asociación Argentina de Brangus for providing the data for the study.

References

Buonaccorsi J., Demidenko E., Tosteson T.D. (2000) Estimation in longitudinal random effects models with measurement error. *Stat. Sin.*, **10**, 855–903.

Buonaccorsi J.P. (2010) Measurement Error: Models. Methods and Applications. Chapman & Hall, New York.

Buzas J.S., Tosteson T.D., Stefansky L.A. (2005) Measurement Error. Handbook of Epidemiology. Springer, Berlín, Germany, pp. 729–765.

Cantet R.J.C., Fernando R.L., Gianola D., Misztal I. (1992) Genetic grouping for direct and maternal effects with differential assignment of groups. *Genet. Sel. Evol.*, **24**, 211–223.

Cantet R.J.C., Birchmeier A.N., Santos-Cristal de Sivak M.G., de Avila V.E.S. (2000) Comparison of restricted maximum likelihood and method *R* for estimating heritability and predicting breeding value under selection. *J. Anim. Sci.*, **78**, 2554–2560.

Carroll R.J. (2005) Measurement error in epidemiologic studies. *Encyclopedia Biostat.*, **5**, doi: 10.1002/0470011815.b2a03082/abstract

Efron B. (1975) Biased versus unbiased estimation. *Adv. Math. (N Y)*, **16**, 259–277.

Freedman L.S., Midthune D., Carroll R.J., Kipnis V. (2008) A comparison of regression calibration, moment reconstruction and imputation for adjusting for covariate measurement error in regression. *Stat. Med.*, **27**, 5195–5216.

Fuller W.A. (1987) Measurement Error Models. Wiley, New York, USA.

Henderson C.R. (1984) Applications of linear models in animal breeding. University of Guelph, Guelph, ON, Canada.

Koots K.R., Gibson J.P., Smith C., Wilton J.W. (1994) Analyses of published genetic parameter estimates for beef production traits. 1. *Heritability. Anim. Breed. Abstr.*, **62**, 309–338.

Koots K., Gibson J.P., Wilton J.W. (1994) Analyses of published parameter estimates for beef production traits 2. Phenotypic and genetic correlations. *Anim. Breed. Abstr.*, **62**, 825–853.

Patterson H.D., Thompson R. (1971) Recovery of interblock information when block sizes are unequal. *Biometrika*, **58**, 454–554.

Quaas R.L. (1976) Computing the diagonal elements and inverse of a large numerator relationship matrix. *Biometrics*, **32**, 949–953.

Schaeffer L.R., Kennedy B.W. (1989) Effects of embryo transfer in beef cattle on genetic evaluation methodology. *J. Anim. Sci.*, **67**, 2536–2543.

Van Vleck L.D. (1990) Alternative animal model with maternal effects and foster dams. *J. Anim. Sci.*, **68**, 4026–4038.