

MODELLING ANIMAL SYSTEMS RESEARCH PAPER

Bayesian simultaneous equation models for the analysis of energy intake and partitioning in growing pigs

A. B. STRATHE¹*, H. JØRGENSEN², E. KEBREAB¹ AND A. DANFÆR²

¹Department of Animal Science, University of California, Davis 95616, CA, USA

²Department of Animal Science, Faculty of Science and Technology, Aarhus University, Blichers Allé 20, 8830 Tjele, Denmark

(Received 26 September 2011; revised 23 January 2012; accepted 28 February 2012;
first published online 4 April 2012)

SUMMARY

The objective of the current study was to develop Bayesian simultaneous equation models for modelling energy intake and partitioning in growing pigs. A key feature of the Bayesian approach is that parameters are assigned prior distributions, which may reflect the current state of nature. In the models, rates of metabolizable energy (ME) intake, protein deposition (PD) and lipid deposition (LD) were treated as dependent variables accounting for residuals being correlated. Two complementary equation systems were used to model ME intake (MEI), PD and LD. Informative priors were developed, reflecting current knowledge about metabolic scaling and partial efficiencies of PD and LD rates, whereas flat non-informative priors were used for the remainder of the parameters. The experimental data analysed originate from a balance and respiration trial with 17 cross-bred pigs of three genders (barrows, boars and gilts) selected on the basis of similar birth weight. The pigs were fed four diets based on barley, wheat and soybean meal supplemented with crystalline amino acids to meet or exceed Danish nutrient requirement standards. Nutrient balances and gas exchanges were measured at c. 25, 75, 120 and 150 kg body weight (BW) using metabolic cages and open circuit respiration chambers. A total of 56 measurements were performed. The sensitivity analysis showed that only the maintenance component was sensitive to the prior specification, and hence the maintenance estimate of 0.91 MJ ME/kg^{0.60} per day (0.95 credible interval (CrI): 0.78–1.09) should be interpreted with caution. It was shown that boars' ability to deposit protein was superior to that of barrows and gilts, as these had an estimated maximum PD (PD_{max}) of 250 g/day (0.95 CrI: 237–263), whereas the barrows and gilts had a PD_{max} of 210 g/day (0.95 CrI: 198–220). Furthermore, boars reached PD_{max} at 109 kg BW (0.95 CrI: 93.6–130), whereas barrows and gilts maximized PD at 81.7 kg BW (0.95 CrI: 75.6–89.5). At 25 kg BW, the boars partitioned on average 5–6% more of the ME above maintenance into PD than barrows and gilts, and this was progressively increased to 10–11% more than barrows and gilts at 150 kg BW. The Bayesian modelling framework can be used to further refine the analysis of data from metabolic studies in growing pigs.

INTRODUCTION

Simultaneous equation statistical models are an attractive method for analysing interactions among components of animal systems directly. An application of these models is in bio-energetics of animal growth, because it is preferable to analyse protein deposition (PD) and lipid deposition (LD) data using simultaneous equations statistical models (Koong 1977; van Milgen

& Noblet 1999; Strathe *et al.* 2010a). However, they require utilization of additional parameters, which is computationally costly and hence these models have been applied mainly to larger datasets. For instance, Strathe *et al.* (2010a) analysed 384 energy balances, which can be considered a large dataset for metabolic studies. These studies require expensive equipment and labour, so they often produce sparse datasets. As a consequence, parameter identifiability becomes an issue, which leads to poor convergence of the non-linear estimation routine, i.e. parameter

* To whom all correspondence should be addressed. Email: abstrathe@ucdavis.edu

estimates with large standard errors (s.e.) and point estimates of key parameters that may be scientifically unreasonable.

In Bayesian paradigm, prior information can be synthesized through formal statements of probability because model parameters are viewed as random variables with associated probability distributions, so-called 'priors' to the parameters (Gelman *et al.* 2004). The inclusion of previous knowledge is a fundamental and integral part of the modelling process – for example, parameters describing partial efficiencies of utilizing metabolizable energy (ME) above maintenance (ME_M) for protein deposition (PD; k_p) and LD (k_l) and the exponent scaling the metabolic body size (b), are known *a priori* because the bioenergetics of growth have been studied for decades. The prior distributions are subsequently updated with regard to the data at hand. The resulting, so-called 'posterior probability distributions', or 'posteriors' for short, are consistent with both the experimental data and the priors, as the posteriors are derived from the product of the likelihood of the data and the prior probability of the parameters. A Bayesian approach has recently been described for the analysis of energy balance data in dairy cows as a robust framework for integration and extraction of information (Strathe *et al.* 2011).

The objective of the current study was to develop Bayesian simultaneous equation models for modelling energy intake and partitioning in growing pigs.

MATERIALS AND METHODS

All experimental procedures complied with the Danish Ministry of Justice, Law No. 382 (10 June 1987) and Act No. 726 (9 September 1993), regarding animal experimentation and care.

Experimental procedure

The pigs used in the study were crosses of Yorkshire × Landrace sows and Duroc boars. Eighteen pigs of three genders (barrow (male pig castrated before puberty), boar and gilt), originating from six litters were used and hence there were six pigs per gender. The pigs were obtained from the swine herd at Foulum Research Centre (Denmark). The first 12 pigs were subjected to three balance periods, whereas the last six pigs were subjected to four periods, covering the growth phase from 20 to 150 kg body

weight (BW). During the growth phase, the pigs were fed four diets (described later in this section) in the corresponding intervals: 25–45, 45–65, 65–100 and 100–150 kg BW. The balance periods were planned to be done in the middle or last part of each interval. All balance periods lasted 7 days in metabolism cages, with days 4 and 5 in the respiration chambers. One barrow was omitted from the dataset due to illness in the first balance period. A total of 56 energy balances including 16 for barrows, 20 for boars and 20 for gilts were measured. During the collection and balance periods, the pigs were kept in stainless steel metabolism cages in the respiration chambers. Feed allowance approximated *ad libitum* feed intake. Faeces and urine were collected quantitatively during the 7-day collection period. Gas exchange and heat production (HP) were measured for 48 h on days 3 and 4. BW over each balance period was calculated as the mean of initial and final BW during the period. The calculation of ME included energy losses in faeces, urine, methane and hydrogen. The average daily HP was calculated according to Brouwer (1965). Total energy gain over the balance period was obtained as the difference between ME intake (MEI) and HP; its partition between protein and lipid gain was calculated by assuming that protein gain ($6.25 \times$ nitrogen (N) retention) contained 23.8 kJ/g (ARC 1981). Nitrogen retention was calculated as the difference between N intake and N losses in faeces and urine. The experimental procedure of conducting respiration and energy balance trials at Foulum Research Centre is described in full by Jørgensen *et al.* (1996).

The diets were based on barley, wheat and soybean meal, and were formulated to meet nutrient requirements according to Danish nutritional standards. Their ingredient and chemical compositions are presented in Table 1. The pigs were individually housed under thermoneutral conditions and given *ad libitum* access to feed and water when they were not subjected to determination of nutrient and energy balance.

Statistical model

It was assumed that multiple observations from the pig were independent because diets were changed during the consecutive balance periods. Also, a preliminary analysis showed that variance components associated with structural model parameters were poorly estimated, because it was difficult to estimate pig specific MEI, PD and LD curves with only three observations

Table 1. *Nutrient and energy composition of experimental diets*

| Item | Diet 1 25–45 kg | Diet 2 45–65 kg | Diet 3 65–100 kg | Diet 4 100–150 kg |
|------------------------------------|-----------------|-----------------|------------------|-------------------|
| Ingredient composition (g/kg) | | | | |
| Barley | 335.11 | 360.07 | 500.00 | 500.00 |
| Wheat | 335.11 | 360.07 | 305.10 | 378.85 |
| Wheat bran | 0.25 | 0.18 | 0.20 | 0.08 |
| Soybean meal | 225.76 | 221.41 | 144.20 | 92.60 |
| Fishmeal | 13.88 | – | – | – |
| Animal fat | 60.00 | 30.00 | 20.00 | – |
| Lysine | 2.38 | 1.82 | 2.44 | 2.00 |
| Methionine | 0.66 | 0.48 | 0.52 | 0.22 |
| Threonine | 0.99 | 0.31 | 0.80 | 0.35 |
| Limestone | 10.99 | 10.46 | 9.75 | 10.33 |
| Dicalcium phosphate | 8.86 | 9.21 | 11.00 | 9.55 |
| Salt | 4.00 | 4.00 | 4.00 | 4.00 |
| Vitamin mixture | 2.00 | 2.00 | 2.00 | 2.00 |
| Chemical composition | | | | |
| Dry matter (DM: g/kg of diet) | 889 | 883 | 886 | 888 |
| Net energy (MJ/kg of DM)* | 8.96 | 8.41 | 8.26 | 8.11 |
| Components and nutrients (g/kg DM) | | | | |
| Crude protein | 213 | 204 | 179 | 162 |
| Crude fat | 86.3 | 54.5 | 44.4 | 26.8 |
| Dietary fibret | 176 | 171 | 178 | 157 |
| Lysine | 12.4 | 10.8 | 9.51 | 7.88 |
| Methionine | 3.80 | 3.34 | 3.02 | 2.68 |
| Cystine | 3.58 | 3.46 | 3.26 | 3.13 |
| Threonine | 8.48 | 7.31 | 6.74 | 5.75 |
| Minerals (g/kg DM) | | | | |
| Ash | 56.2 | 53.0 | 50.3 | 45.5 |
| Calcium | 8.7 | 8.2 | 7.9 | 7.7 |
| Phosphorus | 6.0 | 5.9 | 6.1 | 5.7 |

* Estimated according to Just (1982), Just *et al.* (1983) and Boisen & Fernandez (1997).

† Non-starch polysaccharides + lignin.

per pig and only 17 pigs. The distribution of the data conditional on model parameters is

$$y_{ijk} | \beta_i, k_p, k_f, b, \Sigma \sim \text{MVN}(f(\beta_i, k_p, k_f, b, \text{BW}_{ij})_k, \Sigma) \quad (1)$$

where y_{ijk} denotes the k th observed response (MEI, PD and LD, all expressed in MJ/day) related to the i th gender (barrow, boar and gilt) at the j th BW (1, 2, ..., n_i); $f(\beta_i, k_p, k_f, b, \text{BW}_{ij})_k$ denotes the expected MEI, PD and LD values given by the simultaneous equations, which are described in detail below. The vector of structural parameters is specific to the i th gender and denoted by β_i . The elements in β_i are also specific to the equation format, equating the expected values, and its attributes and dimensions are presented below. Parameters k_p , k_f and b are the partial efficiencies and metabolic exponent parameter, respectively, which are dimensionless quantities. The

variance–covariance matrix Σ describes the residual variability in the three responses. The introduction of a multivariate normal (MVN) distribution assumes that the responses are correlated (Strathe *et al.* 2010a). Moreover, measurement errors that might have occurred in the determination of ME, PD or LD are assumed to be correlated. This is an important model feature in the calculation of energy balances and their interrelationships, and it represents an extension of the previous framework proposed by Strathe *et al.* (2010a).

Simultaneous equations describing energy partitioning

MEI (MJ/d) is described by an exponential function, which is used frequently to model intake in growing pigs (e.g. ARC 1981). The function has two parameters

that can be interpreted as the asymptotic (maximum) MEI M (MJ/day) and the fractional change in MEI due to a change in BW, k (1/kg), respectively. The energy partitioning is modelled as follows: ME available for growth (ME_G , MJ/day) is calculated as the difference between MEI and ME for maintenance (ME_M , MJ/day). The fraction (F , dimensionless) of ME_G available for PD is assumed not to be constant, but a linear function of BW. Multiplication of ME_G by F and by $1 - F$ gives ME for PD (ME_{PD} , MJ/day) and ME for LD (ME_{LD} , MJ/day), which are used with efficiencies k_p and k_f , respectively. Combination of these equations proposed by van Milgen & Noblet (1999) with the equation for predicting MEI results in the following simultaneous equations:

$$\begin{aligned}
 f(\beta_i, BW_{ij})_1 &= ME_{ij} = M_i(1 - e^{-k_i \times BW_{ij}}) \\
 f(\beta_i, BW_{ij})_2 &= PD_{ij} = k_p(c_i - d_i(BW_{ij} - 25)) \\
 &\quad \times (ME_{ij} - a_i BW_{ij}^b) \\
 f(\beta_i, BW_{ij})_3 &= LD_{ij} = k_f(1 - (c_i - d_i(BW_{ij} - 25))) \\
 &\quad \times (ME_{ij} - a_i BW_{ij}^b)
 \end{aligned}
 \tag{2}$$

Maintenance is specified as $ME_M = a_i BW^b$ where a_i is the gender-specific maintenance requirement (MJ ME/(kg BW^b × day)) and b is the metabolic exponent (dimensionless). It is assumed that F decreases linearly with BW, i.e. $F_{ij} = c_i - d_i(BW_{ij} - 25)$. Here c_i is the fraction of ME_G used for PD at 25 kg BW and d_i (1/kg) is the change in F per kg BW change. The gender-specific parameter vector is $\beta_i = [M_i, k_i, a_i, c_i, d_i]^T$.

Another set of MEI, PD and LD equations, which provides additional insight into the partitioning of ME during growth, is presented below. This approach requires a functional specification of the PD curve. The concept of an upper limit for PD in growing pigs is widely accepted, as pigs are limited by either their genetic or feed intake capacity (at earlier stages of growth). The Gompertz function is often the preferred function for representing the pattern of maximum PD (PD_{max}) in pigs (Whittemore & Green 2002) and it can be parameterized to express the BW at maximum rate of PD (BW_{PDmax} , kg) and the associated PD_{max} (MJ/day) (Strathe *et al.* 2010a), when PD is expressed as a function of BW. Strathe *et al.* (2010a) used the specialized Gompertz function in combination with the Michaelis–Menten function to model effects of ME supply on growth. However, the pigs used in the present study were all fed close to *ad libitum* and thus the effect of energy supply on PD cannot be separated

from the stage of growth. The equation set is given as

$$\begin{aligned}
 f(\beta_i, BW_{ij})_1 &= ME_{ij} = M_i(1 - e^{-k \times BW_{ij}}) \\
 f(\beta_i, BW_{ij})_2 &= PD_{ij} = \frac{PD_{max_i}}{BW_{PDmax_i}} BW_{ij} \\
 &\quad \times \log\left(\frac{BW_{PDmax_i} \times e}{BW_{ij}}\right) \\
 f(\beta_i, BW_{ij})_3 &= LD_{ij} = k_f(ME_{ij} - a_i \times BW_{ij}^b - PD_{ij}/k_p)
 \end{aligned}
 \tag{3}$$

The gender-specific parameter vector is $\beta_i = [M_i, k_i, a_i, PD_{max_i}, BW_{PDmax_i}]^T$ and e is the exponential to 1. Equation sets (2) and (3) require the same number of parameters to be estimated, but equation set (3) provides complementary information about energy utilization in growing pigs, i.e. PD_{max} estimation. The fraction of ME_G partitioned for PD can be calculated from equation set (3) as

$$F(BW) = \frac{\frac{PD_{max_i}}{k_p \times BW_{PDmax_i}} BW \times \log\left(\frac{BW_{PDmax_i} \times e}{BW}\right)}{MEI(BW) - a_i \times BW^b}
 \tag{4}$$

Clearly, F changes non-linearly during the course of growth. The quantity F is suggested here as a measure of the priority for PD in energy terms. To facilitate interpretation (when needed), the results are reported in grams per day by assuming 23.8 kJ/g of protein and 39.6 kJ/g of lipid (ARC 1981).

Specification of priors

The prior specification is concerned with assigning probability distributions to β_i , Σ , k_p , k_f and b . Priors for β_i and Σ are

$$\beta_i \sim MVN(\mu, H); \Sigma \sim IW(R, \rho)
 \tag{5}$$

where $IW(\cdot, \cdot)$ denotes the inverse-Wishart distribution. At this point, numerical values for μ , H , R and ρ must be stated. Here, $\mu = [0, \dots, 0]^T$ and $H = 100^2 \times I$ are used, where I is the identity matrix. The prior for β_i is minimally informative, leading to no restriction on the partitioning of ME between growth and maintenance. The inverse-Wishart distribution was selected to represent the prior for the residual variance–covariance because it is the only closed form distribution that naturally imposes the appropriate constraint, i.e. positive definiteness. Hence, $R = I$ and $\rho = 3$, which specifies the vaguest possible proper prior for Σ .

The above equation systems cannot be solved properly without utilization of numerical information

Table 2. *Compilation of literature values for the partial efficiencies k_p and k_f*

| k_f | k_p | Reference |
|-----------|-----------|----------------------------------|
| 0.72–0.88 | 0.52–0.63 | Strathe <i>et al.</i> (2010a)* |
| 0.75 | 0.56 | van Milgen <i>et al.</i> (2000)* |
| 0.77–0.82 | 0.58–0.60 | van Milgen & Noblet (1999)* |
| 0.84 | 0.62 | Noblet <i>et al.</i> (1999)† |
| 0.76 | 0.54 | NRC (1998)† |
| 0.60 | 0.52 | Tess <i>et al.</i> (1984)† |
| 0.74 | 0.56 | ARC (1981)† |

* Partial efficiencies derived from multivariate modelling approaches.

† Partial efficiencies derived from the factorial approach (multiple linear regression).

about some of the parameters. Table 2 represents a compilation of values from the literature for the partial efficiencies k_p and k_f , which suggest that the efficiencies of PD and LD are c. 0.6 and 0.8, respectively. Hence, conservative prior distributions can be deduced from Table 2. Normal and uniform distributions are specified by assuming that $k_p \sim N(0.60, 0.10^2)$ or $k_p \sim U(0.40, 0.80)$ and $k_f \sim N(0.80, 0.10^2)$ or $k_f \sim U(0.60, 1.00)$. The 0.95 confidence intervals (CIs) for parameters k_p and k_f specified by the normal distributions are (0.40–0.80) and (0.60–1.00), respectively. However, the biological interpretation of the assigned prior distributions is different because the normal distribution states that it is more likely that k_p is 0.6 than 0.4, whereas the uniform distribution states that these values are equally likely. Likewise, the prior for b is $b \sim N(0.60, 0.10^2)$ or $b \sim U(0.40, 0.80)$. Finally, the two different prior distributions also constitute a sensitivity analysis because two sets of posterior distributions are produced, and hence the conclusions may or may not differ between the choices of priors (Gelman *et al.* 2004).

Convergence and model selection

Two chains were run with different initial over-dispersed values. To assess convergence, four formal convergence tests at the core of the convergence diagnostic and output analysis (CODA) package were used (Best *et al.* 1995), which are the Geweke, Heidelberger–Welch, Raftery–Lewis and Gelman–Rubin diagnostic tools.

Statistics, such as the mean, median, percentiles and 0.95 credible interval (CrI) can be calculated from the samples making up the posterior distribution. The CrI is

the Bayesian version of the traditional CI and it represents a non-parametric interval of the posterior distribution. In Bayesian methodology, a parameter is considered a random variable and thus the CrI may be interpreted as the 0.95 probability that the parameter lies within the interval (lower–upper) given the observed data and prior distribution (Gelman *et al.* 2004). In Bayesian analysis, there is no standard model selection/reduction criterion such as the likelihood ratio test. In the current analysis, three criteria for model selection were used. Firstly, a reduction in the residual variance for three response variables fitted to the two sets of equations. Secondly, the precision of the estimated parameter/effect, which is judged by its 0.95 CrI and computation of the proportion of posterior samples that are different from zeros. This proportion may be interpreted similar to a P -value of the underlying hypothesis in traditional statistics. Thirdly, the Deviance Information Criteria (DIC), which is a general tool to assess the trade-off between model fit (deviance: $-2 \log$ likelihood) and complexity (number of effective parameters) was used. The notion that ‘smaller is better’ is preserved in the DIC (Spiegelhalter *et al.* 2002). Differences of 5 and 10 DIC units are considered a tendency and a substantive improvement of fit to data, respectively (Spiegelhalter *et al.* 2002).

Prediction procedure

The prediction procedure in the current paper followed Bayesian modelling convention (Gelman *et al.* 2004) and used the observed data (\mathbf{Y}) and posterior parameters ($\boldsymbol{\theta} = [\boldsymbol{\beta}_i, k_p, k_f, b]^T$) obtained from the model estimation to make inferences about a predicted quantity \mathbf{Y}_p , e.g. PD. Hence, the distribution of \mathbf{Y}_p was obtained conditional on the observed data and the model parameter (i.e. the posterior predictive distribution of \mathbf{Y}_p) as

$$p(\mathbf{Y}_p|\mathbf{Y}) = \int p(\mathbf{Y}_p|\boldsymbol{\theta})p(\boldsymbol{\theta}|\mathbf{Y})d\boldsymbol{\theta} \quad (6)$$

The first of the two factors in the integral is the MVN distribution for the future observations given the value of $\boldsymbol{\theta}$ (Eqn (1)). The second factor is the posterior distribution of $\boldsymbol{\theta}$ given the observed data. The posterior predictive distribution of \mathbf{Y}_p can thus be thought of as an average of the conditional predictions over the posterior distribution of $\boldsymbol{\theta}$. The integral in Eqn (6) is not analytically tractable, but it can be approximated by

Table 3. Parameter values estimated by equation set (2)* presented as posterior means and 0.95 CrI in parenthesis

| Parameter† | Barrow | | Boar | | Gilt | |
|----------------------|-----------------------|-----------------------|-----------------------|-----------------------|----------------------|----------------------|
| | Normal‡ | Uniform‡ | Normal‡ | Uniform‡ | Normal‡ | Uniform‡ |
| <i>c</i> | 0.45 (0.37–0.53) | 0.45 (0.37–0.54) | 0.50 (0.42–0.59) | 0.50 (0.43–0.59) | 0.46 (0.38–0.54) | 0.46 (0.38–0.54) |
| <i>d</i> (kg) | 0.0015 (0.0008–0.002) | 0.0017 (0.0008–0.003) | 0.0012 (0.0004–0.002) | 0.0012 (0.0005–0.002) | 0.0017 (0.001–0.003) | 0.0018 (0.001–0.003) |
| <i>M</i> (MJ/d) | 43.9 (40.8–47.4) | 43.9 (40.8–47.4) | 48.4 (44.7–52.8) | 48.4 (44.7–52.8) | 45.3 (42.0–49.1) | 45.3 (42.0–49.1) |
| <i>K</i> (BW) | 0.022 (0.018–0.028) | 0.022 (0.018–0.028) | 0.017 (0.014–0.021) | 0.017 (0.014–0.021) | 0.020 (0.016–0.025) | 0.020 (0.016–0.025) |
| <i>a</i> (MJ/d/kg) | 0.91 (0.77–1.09) | 0.83 (0.70–0.97) | 0.96 (0.81–1.15) | 0.88 (0.75–1.03) | 0.91 (0.78–1.10) | 0.84 (0.71–0.98) |
| <i>b</i> | Normal‡ | | Normal‡ | | Uniform‡ | |
| <i>k_p</i> | 0.60 (0.54–0.65) | | 0.60 (0.54–0.65) | | 0.62 (0.57–0.68) | |
| <i>k_f</i> | 0.59 (0.53–0.65) | | 0.59 (0.53–0.65) | | 0.54 (0.48–0.60) | |
| DIC§ | 641.6 | | 641.6 | | 642.3 | |

* See Materials and methods section for details.

† Here *a* is ME requirement for maintenance, *c* is the fraction of ME for production (*F*) used for PD at 25 kg BW and *d* is the change in *F* per kg BW change. Parameters *M* and *k* can be interpreted as the asymptotic MEI and the fractional change in MEI due to a change in BW. Parameters *k_p*, *k_f* and *b* are the partial efficiencies and metabolic exponent parameter, respectively.

‡ Normal and uniform refer to the priors assigned to parameters *b*, *k_p* and *k_f*.

§ Deviance information criteria.

using Monte Carlo sampling from the posterior and summarize the samples in median and 0.95 CrI.

All models were implemented in the general purpose software for Bayesian modelling, WinBUGS (Lunn *et al.* 2000).

RESULTS

Parameter estimates and their uncertainties are presented in Tables 3 and 4 for the two equation sets under two different prior settings. Both equation sets and both priors predicted the same MEI at various stages of growth as the parameter estimates were identical. The asymptotic MEI was estimated in the range of 44–48 MJ/day, which was not significantly different between genders (*P* > 0.05). Parameters (*c*, *d*, PD_{max} and BW_{PDmax}) describing energy partitioning between PD and LD were unaffected by the different priors as the estimates were numerically identical and thus the representation of the prior setting was not important. Figure 1 presents the posterior predictive distribution of PD data with 0.95 CrI. Both models covered the observed variability in the PD data. However, equation set (3) provided a more realistic trend for the PD data, as PD_{max} occurred at later stages of growth, which suggests that the assumption of a linear decline in *F* may not hold when growth is studied from 25 to 150 kg BW. This was supported by a better fit to the data with equation set (3), i.e. a DIC of 637 v. 642.

The estimates of the maintenance requirement (*a*, MJ ME/(kg BW^{*b*} × day)) were dependent on the specification of the prior distributions, and the uniform distribution produced lower point estimates of *a* (Tables 3 and 4). ME maintenance requirement point estimates ranged from 0.83 to 0.96 MJ ME/(kg BW^{*b*} × day) and they were highly correlated to estimates of *b*, *k_p* and *k_f*. These parameters did not change much from their prior location of 0.60, 0.60 and 0.80, respectively, but their precision increased because the standard errors (s.e.) were reduced by two-thirds, i.e. s.e. ≈ 0.1 (prior) v. s.e. ≈ 0.03 (posterior). Hence, the estimates of *a*, *b*, *k_p* and *k_f* derived here were not unique as they were dependent on the setting of the priors and thus they should be interpreted with caution. Moreover, gender-specific maintenance requirements should be reduced to a single parameter. Parameter estimates and CrI given in Tables 3 and 4 showed that barrows and gilts could not be separated, and hence these groups were combined in the further analysis resulting in two groups, i.e. barrows and

Table 4. Parameter values estimated by equation set (3)* presented as posterior means and 0.95 CrI in parenthesis

| Parameter† | Barrow | | | Gilt | | |
|---------------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
| | Normal‡ | Uniform‡ | Normal‡ | Uniform‡ | Normal‡ | Uniform‡ |
| PD _{max} (g/day) | 213 (198–230) | 213 (198–229) | 250 (237–263) | 250 (237–263) | 209 (192–221) | 206 (192–220) |
| BW _{PDmax} (kg) | 81.7 (73.1–94.1) | 82.3 (73.3–95.0) | 109 (93.4–133) | 109 (93.7–130) | 82.8 (74.1–95.2) | 83.1 (74.2–94.7) |
| M (MJ/d) | 44.1 (41.1–45.2) | 44.2 (41.1–48.1) | 48.6 (44.9–53.5) | 48.6 (44.8–53.5) | 45.9 (42.5–50.4) | 46.0 (42.6–50.4) |
| k (kg) | 0.022 (0.017–0.027) | 0.022 (0.017–0.027) | 0.017 (0.014–0.021) | 0.017 (0.014–0.021) | 0.019 (0.015–0.024) | 0.019 (0.015–0.024) |
| a (MJ/d/kg) | 0.88 (0.65–1.15) | 0.85 (0.68–1.04) | 0.93 (0.69–1.23) | 0.90 (0.73–1.08) | 0.88 (0.65–1.16) | 0.85 (0.69–1.03) |
| <i>b</i> | Normal‡ | | | Uniform‡ | | |
| | 0.60 (0.56–0.66) | | | 0.61 (0.57–0.67) | | |
| <i>k_p</i> | 0.59 (0.53–0.65) | | | 0.58 (0.50–0.68) | | |
| <i>k_r</i> | 0.78 (0.72–0.85) | | | 0.76 (0.70–0.87) | | |
| DIC‡ | 637.4 | | | 637.9 | | |

* See Materials and methods section for details.

† Here *a* is ME requirement for maintenance, maximum rate of PD is denoted by PD_{max} and the associated BW at that point BW_{PDmax}. Parameters *M* and *k* can be interpreted as the asymptotic MEI and the fractional change in MEI due to a change in BW. Parameters *k_p*, *k_r* and *b* are the partial efficiencies and metabolic exponent parameter, respectively.

‡ Normal and uniform refer to the priors assigned to parameters *b*, *k_p* and *k_r*.

§ Deviance information criteria.

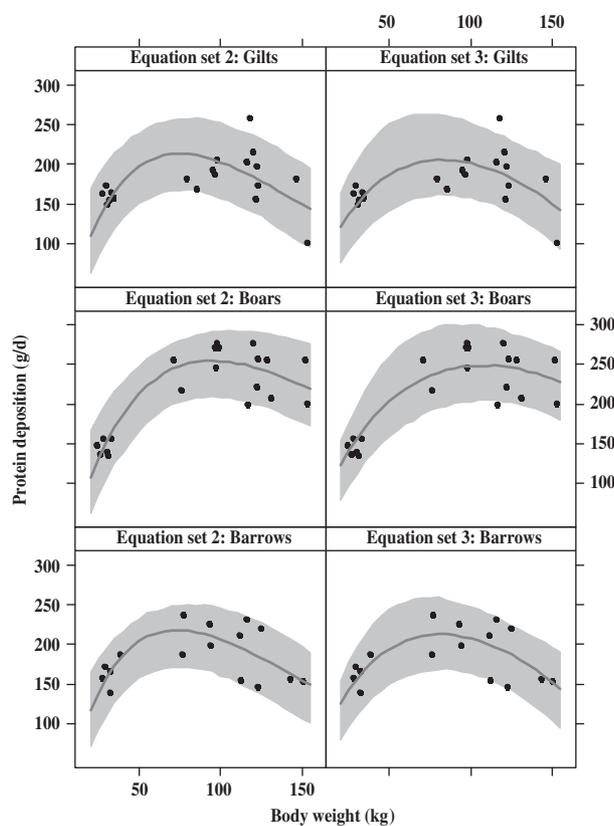


Fig. 1. Posterior predictive distribution for PD (g/day) given by two different equation sets and three genders. The line represents the posterior median and the grey shaded area is the 0.95 CrI for the predictions.

barrows+gilts. The DIC statistic for the two reduced models (Table 5) resulted in simpler models that were obtained by grouping barrows and gilts. The associated DICs were 633 and 626 for the reduced versions v. DICs of 642 and 637 for the original versions of equation sets (2) and (3). Still, equation set (3) was the preferred model, i.e. a DIC of 626.

Parameter values obtained by fitting the reduced models to data and use normal priors are presented in Table 5 for the two equation systems. Parameters describing energy partitioning above ME_M between PD and LD in equation set (2) were not statistically different between the two groups as posterior probabilities for the expressions $C_{\text{boar}} - C_{\text{barrow+gilt}} > 0$ and $d_{\text{boar}} - d_{\text{barrow+gilt}} > 0$ were 0.24 and 0.10, respectively. A large numerical difference was observed, as boars partitioned 5% more of ME_G towards PD at 25 kg BW. In general, energy was diverted from PD towards LD with increasing BW because estimates of *d* for both groups were significantly different from zero ($P < 0.05$). Evaluation of PD_{max} values estimated by equation set (3) suggests that the boars deposited more

Table 5. Parameter values and 0.95 CI in parenthesis estimated by two different equation sets*. Parameter contrasts are computed for boars against barrows + gilts

| Parameter† | Estimate | P-value‡ |
|--|------------------------|----------|
| Equation set (2)* | | |
| $C_{\text{barrow+gilt}}$ | 0.45 (0.40–0.51) | |
| C_{boar} | 0.50 (0.42–0.59) | 0.240 |
| $d_{\text{barrow+gilt}}$ (/kg) | 0.0016 (0.0013–0.0019) | |
| d_{boar} (/kg) | 0.0012 (0.0009–0.0015) | 0.100 |
| $M_{\text{barrow+gilt}}$ (MJ/d) | 44.5 (42.2–46.9) | |
| M_{boar} (MJ/d) | 48.5 (44.9–53.0) | 0.070 |
| $k_{\text{barrow+gilt}}$ (/kg) | 0.021 (0.018–0.025) | |
| k_{boar} (/kg) | 0.017 (0.014–0.021) | 0.120 |
| Equation set (3)* | | |
| $PD_{\text{max, barrow+gilt}}$ (g/d) | 210 (198–220) | |
| $PD_{\text{max, boar}}$ (g/d) | 250 (237–263) | <0.001 |
| $BW_{PD_{\text{max, barrow+gilt}}}$ (kg) | 81.7 (75.6–89.5) | |
| $BW_{PD_{\text{max, boar}}}$ (kg) | 109 (93.6–130) | 0.004 |
| $M_{\text{barrow+gilt}}$ (MJ/d) | 45.0 (42.7–47.7) | |
| M_{boar} (MJ/d) | 48.4 (44.6–53.2) | 0.160 |
| $k_{\text{barrow+gilt}}$ (/kg) | 0.020 (0.017–0.024) | |
| k_{boar} (/kg) | 0.017 (0.014–0.021) | 0.200 |

* See Materials and methods section for details.
 † Here c is the fraction of ME for production (F) used for PD at 25 kg BW and d is the change in F per kg BW change. Parameters M and k can be interpreted as the asymptotic MEI and the fractional change in MEI due to a change in BW. Maximum rate of PD is denoted by PD_{max} and the associated BW at that point $BW_{PD_{\text{max}}}$.
 ‡ The proportion of posterior samples that are different from zero computed for the contrast boars against barrows + gilts.

protein ($P < 0.001$), i.e. 250 g/day (0.95 CrI: 237–263) v. 210 g/day (0.95 CrI: 198–220). Furthermore, the PD_{max} was reached at later stages of growth in boars because $BW_{PD_{\text{max}}}$ was 109 kg (0.95 CrI: 93.6–130) for boars and 81.7 kg (0.95 CrI: 75.6–89.5) for barrows and gilts. The quantity F was suggested here as a measure of the priority for PD in energy terms. Hence, a contrast in F between boars and the combination of barrows and gilts was computed at different stages of growth (Fig. 2). Although F was related non-linearly to BW in equation set (3), both equation sets estimated the numerical difference between the two groups to be c. 5–6% at 25 kg BW increasing to c. 10–11% at 150 kg BW. Moreover, boars partitioned significantly more ME above ME_M towards PD c. 50–60 kg BW as the 0.95 CrI did not overlap zero, which was also shown in both equation systems.

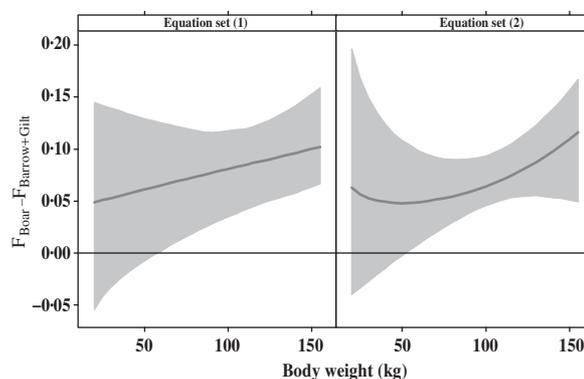


Fig. 2. Posterior predictive distribution for the comparison between boars and barrows + gilts in their priority for PD computed at different stages of growth. The quantity ($F_{\text{boars}} - F_{\text{barrows+gilts}}$) describes the difference between the two groups in the fraction of ME above maintenance allocated for PD as estimated by two different equation sets. The line represents the posterior median and the grey shaded area is the 0.95 CrI for the predictions.

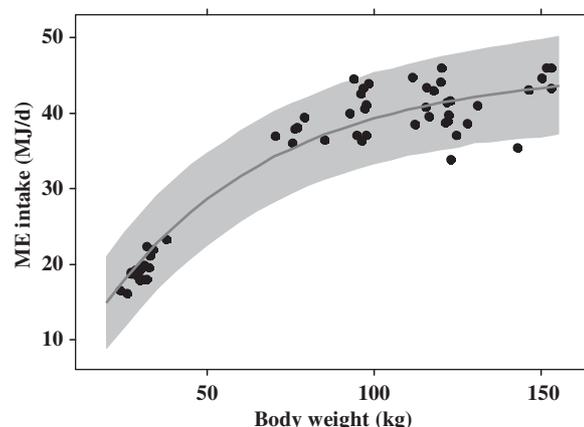


Fig. 3. Posterior predictive distribution for MEI (MJ/day) plotted as a function of BW. The line represents the posterior median and the grey-shaded area is the 0.95 CrI for the predictions.

DISCUSSION

MEI

Similar to the current study, the ARC (1981) and the NRC (1998) used asymptotic equations to relate energy intake to BW. Figure 3 presents the posterior predictive distribution for MEI (MJ/day) plotted as a function of BW. The line represents the posterior median and the grey shaded area is the 0.95 CrI for the predictions. The plot shows clearly that adopting an asymptotic function was justified, but MEI was c. 10–15% less in the current study compared with those reported by the ARC and NRC

(assuming 0.95 ME to DE ratio). The experiment was conducted in metabolism cages, which may reduce feed intake compared to that measured when pigs are housed individually in pens. Therefore, care should be taken when extrapolating the current results to other situations, although it should be noted that most modern genotypes will probably consume less than those reported by ARC (1981) and NRC (1998).

None of the previous multivariate approaches used to analyse PD and LD data included simultaneous modelling of MEI (van Milgen & Noblet 1999; Strathe *et al.* 2010a). Moreover, measured MEI in the preceding studies was used as an independent variable and assumed to be free of error. Some consequences of ignoring measurement error include masking important features of the data, losing the power to detect relationships among variables and introducing bias in function/parameter estimation (Carroll *et al.* 2006). The severity of the problem depends on the magnitude of error in the independent variables. Dhanoa *et al.* (2007) presented alternative regression approaches to the ordinary least square for modelling energy components; however, their study was concerned with estimating k_g (efficiency of utilizing ME for energy gain) and ME_M , which were derived from regressing energy gain on MEI. Their results showed that accounting for measurement error in MEI increased both k_g and ME_M , suggesting that these parameters were biased downwards when the ordinary least squares method was used (Dhanoa *et al.* 2007). The functional forms that were utilized in the present study were non-linear, and therefore 'classical' regression models were not applicable. Shortcomings of multiple linear regression models have been covered by Koong (1977) and van Milgen & Noblet (1999), including reversion of the relationship between responses (PD and LD) and driver (MEI) and statistical issues related to collinearity between independent variables. Use of Markov Chain Monte Carlo (MCMC) techniques for parameter estimation in these complex models was straightforward, and hence measurement error in the determination of MEI could be modelled. The BW variable was assumed to be free of error, which is a realistic assumption as measurement error in the determination BW is small as fluctuations in BW over time are animal driven, i.e. 'animal intrinsic variability' (Strathe *et al.* 2009, 2010b) and hence, utilizing the measured value can be justified. It would be straightforward to include BW as a fourth trait in the recursive equations, modelling the growth curve,

but then several more parameters would have to be estimated and therefore it was decided not to do this in the current study.

Partitioning of MEI into protein and lipid

Published estimates of the maximum PD for growing pigs vary considerably from c. 80 to more than 200 g/day. Tauson *et al.* (1998) investigated PD patterns in boars of three breeds (Landrace, Duroc and Hampshire, Danbred lines) and found that the breeds differed in the capacity for PD, with Duroc and Hampshire being superior to Danish Landrace boars. These Duroc and Hampshire boars of high genetic potential had a capacity for PD_{max} of c. 227 g/day and there was a significant quadratic relationship between PD and metabolic BW, showing that the shape of the PD curve was non-linear. These results agree well with the current results. In a later study, litter mates were serially slaughtered, and a preliminary analysis of these data has been presented by Danfaer & Strathe (2012). The barrows and gilts in that study reached their PD_{max} at c. 120 days of age, whereas the boars' PD rate continued to increase until c. 150 days of age. The predicted maximum rates of PD were 229, 197, 186 g/day at 113, 79 and 80 kg BW for boars, gilts and barrows, respectively. The small discrepancies between the results on PD_{max} from Tauson *et al.* (1998), from the present study and those obtained from the serial slaughter study can be explained by differences in experimental methodology. It is well documented that the N balance technique overestimates the amount of PD because the deposition is calculated as the remainder after correcting for N loss in urine and faeces. These two components are more likely to be under- than overestimated, and hence the N balance tends to be overestimated (e.g. Quiniou *et al.* 1995).

With increasing BW, an increasing part of MEI above maintenance will be designated towards LD for the three genders as the parameter $d > 0$, i.e. energy partitioned towards PD declined linearly with increasing BW. van Milgen & Noblet (1999) reported that lean genotypes partitioned 0.49 of ME above ME_M towards PD at 20 kg BW, which agrees well with the current estimates (Table 5). In addition, van Milgen & Noblet (1999) observed that extremely lean boars maintained a constant partitioning of energy within the observed BW range (20–100 kg). This could not be confirmed in the current study, perhaps due to the limited amount of information between 25 and 75 kg BW.

Utilization of prior information

In the current data analysis, informative prior distributions for parameters b , k_p and k_f were used and this gave numerical information crucial for estimation of the model. The priors were derived from earlier data analysis and hence represented the traditional way of using priors in the Bayesian framework. In addition, the prior distributions used in the current study did not supply any controversial information, but were strong enough to pull the data away from biologically inappropriate inferences, which might have been consistent with the likelihood (Strathe *et al.* 2011). However, care also needs to be taken in implementing these analyses to account for biologically supported parameter spaces, because consideration must be given to whether the pig population in the new dataset resembles the population(s) from which the prior information is borrowed – for example, priors for parameters describing PD potential should be not be taken from Meishan pigs when aiming at modelling PD data from Pietran pigs.

The prior information in the current study was structured into two different proper prior distributions, for purposes of sensitivity analysis, which is an important part of Bayesian analysis (Gelman *et al.* 2004). It was shown that only the ME requirement for maintenance was sensitive to the statement of prior belief. The final estimate of the maintenance component 0.91 MJ ME/(kg BW^b × day) (0.95 CrI: 0.78–1.09) should be interpreted with caution. Nonetheless, introduction of prior information in the data analysis eases comparison with literature values for maintenance requirements in growing pigs and the final estimates were similar to those reported by van Milgen & Noblet (1999) for lean meat-type pigs. An older Danish investigation reported an ME requirement for maintenance to be 0.93 MJ ME/(kg BW^{0.60} × day) (Just *et al.* 1983), which aligned very well with the current estimate(s).

The Bayesian framework enables the development of robust probabilistic analysis of error and uncertainty in model predictions by explicitly accommodating measurement error, parameter uncertainty and model structure imperfection. The current analysis presents a Bayesian formulation for simultaneous calibration of an energy-based pig growth model, with prior precisions of model parameters, data collected, measurement error or inter-animal variability. The model was developed initially within the statistical framework presented by Strathe *et al.* (2010a), but it was noted that

parameter solutions were sensitive to the starting values provided to the routine. This was discovered by means of a grid search and hence it was decided not to report these results for comparison. The intent was to illustrate how this novel approach can be used to transfer knowledge in time (i.e. past to present), and therefore used to calibrate an energy-based pig growth model to sparse data, collected in an intensive metabolic study. Moreover, it was illustrated that the analysis was strengthened through integration of prior knowledge in the modelling process. Finally, it must also be noted that the value of the Bayesian modelling approach increases with the complexity of the structural model.

CONCLUSION

Bayesian framework is a tool well-suited to modelling MEI, PD and LD curves, when these traits are considered as dependent variables. Utilization of prior knowledge could be used directly in the data analysis, which may be important when sparse datasets are analysed due to issues related to parameter identifiability. The concepts presented in the current paper may be extended to form the basis of a complete nutritional pig growth model in which all parameters are expressed in terms of distributions that can be updated/calibrated using MCMC methods.

Funding from the Pig Research Centre, Denmark and the Sesnon Endowed Chair Fund (UC Davis) is acknowledged.

REFERENCES

- ARC (1981). *The Nutrient Requirements of Pigs*. Slough, England: CABI.
- BEST, N. G., COWLES, M. K. & VINES, S. K. (1995). *CODA Manual Version 0.30*. Cambridge, UK: MRC Biostatistics Unit.
- BOISEN, S. & FERNANDEZ, J. A. (1997). Prediction of the total tract digestibility of energy in feedstuffs and pig diets by *in vitro* analyses. *Animal Feed Science and Technology* **68**, 277–286.
- BROUWER, E. (1965). Report of sub-committee on constants and factors. In *Energy Metabolism: Proceedings of the 3rd Symposium, held at Troon, Scotland, May 1964* (Ed. K. L. Blaxter), pp. 441–443. EAAP publication No. 11. London: Academic Press.
- CARROLL, R. J., RUPPERT, D., STEFANSKI, L. A. & CRAINICEANU, C. (2006). *Measurement Error in Nonlinear Models: A Modern Perspective*. 2nd edn. Monographs on Statistics and Applied Probability 105. New York: Chapman & Hall.

- DANFAER, A. & STRATHE, A. (2012). Quantitative and physiological aspects of pig growth. In *Nutrition and Physiology of the Pig* (Ed. K. E. Bach Knudsen), e-learning book. Denmark: Pig Research Centre.
- DHANOVA, M. S., SANDERSON, R., LOPEZ, S., DIJKSTRA, J., KEBREAB, E. & FRANCE, J. (2007). Alternative regression approaches when modelling energy components. In *Energy and Protein Metabolism and Nutrition* (Ed. I. Ortigues-Marty), pp. 593–594. EAAP publication No. 124. Wageningen, The Netherlands: Wageningen Academic Publishers.
- GELMAN, A., CARLIN, J. B., STERN, H. S. & RUBIN, D. B. (2004). *Bayesian Data Analysis. Texts in Statistical Science*. London: Chapman & Hall.
- JØRGENSEN, H., ZHAO, X. Q. & EGGUM, B. O. (1996). The influence of dietary fibre and environmental temperature on the development of the gastrointestinal tract, digestibility, degree of fermentation in the hind-gut and energy metabolism in pigs. *British Journal of Nutrition* **75**, 365–378.
- JUST, A. (1982). The net energy value of balanced diets for growing pigs. *Livestock Production Science* **8**, 541–555.
- JUST, A., JØRGENSEN, H. & FERNÁNDEZ, J. A. (1983). Maintenance requirement and the net energy value of different diets for growth in pigs. *Livestock Production Science* **10**, 487–506.
- KOONG, L. J. (1977). A new method for estimating energetic efficiencies. *Journal of Nutrition* **107**, 1724–1728.
- LUNN, D. J., THOMAS, A., BEST, N. & SPIEGELHALTER, D. (2000). WinBUGS – a Bayesian modelling framework: concepts, structure, and extensibility. *Statistics and Computing* **10**, 325–337.
- NOBLET, J., KAREGE, C., DUBOIS, S. & VAN MILGEN, J. (1999). Metabolic utilization of energy and maintenance requirements in growing pigs: effects of sex and genotype. *Journal of Animal Science* **77**, 1208–1216.
- NRC (1998). *Nutrient Requirements of Swine*. 10th edn. Washington, DC: National Academy Press.
- QUINIQU, N., DUBOIS, S. & NOBLET, J. (1995). Effect of dietary crude protein level on protein and energy balances in growing pigs: comparison of two measurement methods. *Livestock Production Science* **41**, 51–61.
- SPIEGELHALTER, D. J., BEST, N. G., CARLIN, B. P. & VAN DER LINDE, A. (2002). Bayesian measures of model complexity and fit. *Journal of Royal Statistical Society: Series B* **64**, 583–640.
- STRATHE, A. B., SØRENSEN, H. & DANFÆR, A. (2009). A new mathematical model for combining growth and energy intake in animals: the case of the growing pig. *Journal of Theoretical Biology* **261**, 165–175.
- STRATHE, A. B., DANFÆR, A., CHWALIBOG, A., SØRENSEN, H. & KEBREAB, E. (2010a). A multivariate nonlinear mixed effects method for analyzing energy partitioning in growing pigs. *Journal of Animal Science* **88**, 2361–2372.
- STRATHE, A. B., DANFÆR, A., NIELSEN, B., KLIM, S. & SØRENSEN, H. (2010b). Population based growth curve analysis: a comparison between models based on ordinary or stochastic differential equations implemented in a nonlinear mixed effect framework. In *Modelling Nutrient Digestion and Utilisation in Farm Animals* (Eds D. Sauvant, J. van Milgen, P. Faverdin & N. Friggens), pp. 22–30. Wageningen, The Netherlands: Wageningen Academic Publishers.
- STRATHE, A. B., DIJKSTRA, J., FRANCE, J., LOPEZ, S., YAN, T. & KEBREAB, E. (2011). A Bayesian approach to analyze energy balance data from lactating dairy cows. *Journal of Dairy Science* **94**, 2520–2531.
- TAUSON, A. H., CHWALIBOG, A., JAKOBSEN, K. & THORBEB, G. (1998). Pattern of protein retention in growing boars of different breeds, and estimation of maximum protein retention. *Archives of Animal Nutrition* **51**, 253–262.
- TESS, M. W., DICKERSON, G. E., NIENABER, J. A., YEN, J. T. & FERRELL, C. L. (1984). Energy costs of protein and fat deposition in pigs fed *ad libitum*. *Journal of Animal Science* **58**, 111–122.
- VAN MILGEN, J. & NOBLET, J. (1999). Energy partitioning in growing pigs: the use of a multivariate model as an alternative for the factorial analysis. *Journal of Animal Science* **77**, 2154–2162.
- VAN MILGEN, J., QUINIQU, N. & NOBLET, J. (2000). Modeling the relation between energy intake and protein and lipid deposition in growing pigs. *Animal Science* **71**, 119–130.
- WHITTEMORE, C. T. & GREEN, D. M. (2002). The description of the rate of protein and lipid growth in pigs in relation to live weight. *Animal Science* **138**, 415–423.