Uncertainty of national agricultural greenhouse gas emission inventories

MAF Technical Paper No: 2011/87

Report prepared for Ministry of Agriculture and Forestry
By Landcare Research CC MAF POL_0708-65
June 2008

Authors: FM Kelliher, GC Arnold, H Clark, RS Hegarty

ISSN 2230-2794 (online)
ISBN 978-0-478-38739-1 (online)

October 2011

Ministry of Agriculture and Forestry
Te Manatū Ahuwhenua, Ngāherehere

BIOSECURITY NEW ZEALAND

NEW ZEALAND FOOD SAFETY AUTHORITY

Disclaimer

The information in this publication is not government policy. While every effort has been made to ensure the information is accurate, the Ministry of Agriculture and Forestry does not accept any responsibility or liability for error of fact, omission, interpretation or opinion that may be present, nor for the consequences of any decisions based on this information. Any view or opinion expressed does not necessarily represent the view of the Ministry of Agriculture and Forestry.

Publisher

Ministry of Agriculture and Forestry
PO Box 2526
Pastoral House, 25 The Terrace
Wellington 6140
www.maf.govt.nz

Telephone: 0800 008 333

Facsimile: +64 4 894 0300

Uncertainty of National Agricultural Greenhouse Gas Emission Inventories

F.M. Kelliher
Landcare Research, PO Box 40, Lincoln 7640,
Lincoln University, Agriculture & Life Science Division, PO Box 84, Lincoln 7647

G.C. Arnold
Landcare Research, Private Bag 11052, Palmerston North 4442

H. Clark and R.S. Hegarty
AgResearch, Private Bag 11008, Palmerston North 4442

Landcare Research Contract Report: LC0708/

PREPARED FOR:
Ministry of Agriculture and Forestry
PO Box 2526
Wellington

Project Code: CC MAF POL_0708-65 (INVENT-08+09A)

DATE: 30 June 2008

ISO 14001

Reviewed by:	*Approved for release by:*
Dr Keith Lassey Scientist NIWA	Dr David Whitehead Science Team Leader Global Change Processes

Contents

Executive Summary

The project's two objectives were i) analyse the model used to determine grazing ruminant's energy requirement for maintenance and production and ii) assess uncertainties in the enteric methane (CH_4) and agricultural soils nitrous oxide (N_2O) emissions inventories.

We began by quantifying the sensitivity of an animal's energy requirement for maintenance to a change of live weight. For this report, maintenance means the basal metabolism excluding the metabolised energy, denoted ME, for grazing and ME used directly for production. Live weight (w) is activity data that affects feed intake and nitrogen (N) excretion rates, important variables in the CH_4 and N_2O emissions inventories. Contrasting case study analyses were done for mature milking cows and lambs. For the cow, w is relatively constant during a year, averaging 447 and 455 kg in 1990 and 2007, respectively. The sensitivity of the cow's annual (maintenance) energy requirement to a unit change of w (kg) was 29 MJ ME. From 1990 to 2007, this requirement increased by only 1% in correspondence with the increased (8 kg of average) live weight. For the inventories, w for a lamb at birth on 15 September is about 4 kg and it increases by nearly 8 times over the following 6 months until slaughter. Because the lamb's annual maintenance requirement could be interpreted as ambiguous, we confine the report to the lamb's w at slaughter that averaged 37.3 kg in 2007, 6 kg more than the corresponding 1990 value.

The cow requires energy for milk production, in addition to that required for its maintenance. Relations have been determined between the gross energy content of milk (measured in a calorimeter) and its concentration of fat and protein. Combining industry figures for milk fat and protein with the efficiency of ME use for lactation, including pasture ME content, calculations yielded 5.4 MJ ME/kg milk. Alternatively, we combined the total, annual ME requirement, the ME required for maintenance and milk production to compute 7.6 MJ ME/kg milk. In 1990, 56% of the total ME requirement was thus attributed to milk production including the ME required for grazing. In 2007, the percentage had increased to 62%.

For lamb meat production, we combined estimated ME requirements of the ewe for its lamb(s) during the 5 months from conception to birth and for the 3 months from birth until weaning as well as the lamb's 3 months of grazing between weaning and slaughter. In 1990, from birth until slaughter, 70% of the lamb's total ME requirement was attributed to meat production (live weight gain) including the ME required for grazing. In 2007, we estimated 77% of the lamb's total ME requirement was invested in meat production.

Uncertainty assessment followed the principles of good practice guidance developed by the Intergovernmental Panel for Climate Change (IPCC). Assessment began with the uncertainty of an emissions inventory during a given year, combining uncertainties in the emissions factor and activity data. Trend analyses were also done for activity and emissions inventory time series with independent and dependent data. Data dependence significantly affected the uncertainties of emissions and time trends. For example, incorrectly assuming independence, standard regression exaggerated the amount of information available in time series data and therefore understated the trend uncertainty. A lack of independence in time series data also meant standard linear regression “overstated” the accuracy of its predictions (forecasts). By generalising the linear regression to include a correlation between successive residuals, an autocorrelation, trend uncertainty can be appropriately measured and predictions considered more reliable. Contrasting New Zealand examples illustrated the statistical principles governing the uncertainties.

1. Introduction

This project's two goals were i) analyse the model used to determine grazing ruminant's energy requirement for maintenance and production and ii) assess uncertainties in the enteric methane (CH_4) and agricultural soils nitrous oxide (N_2O) emissions inventories.

Analyses to meet the project's first goal are described in Chapter 2. We begin by quantifying the sensitivity of an animal's energy requirement for maintenance to a change of live weight. Live weight (w) is activity data that affects feed intake and nitrogen (N) excretion rates, important variables in the CH_4 and N_2O emissions inventories. Contrasting case study analyses were done for cows and lambs.

The cow requires energy to produce milk, in addition to that required for its maintenance. Production data are also key activity data in the inventories and analyses quantified how a change in activity affects the cow's energy requirement. For the lamb, growth rate determines meat production so it was challenging to separate energy requirement for maintenance and production. This included the ewe's energy requirement for its lamb(s) from conception to birth as well as the period from birth until weaning. Analyses continued for the grazing lamb until slaughter.

Analyses to meet the project's second goal are described in Chapter 3. Uncertainty makes people uncomfortable and susceptibility to consoling beliefs includes the illusion of certainty. For policy analysts, a more sensible approach to uncertainty uses probabilistic information. We begin with uncertainty determination for an emissions inventory during a given year. Generically, the emissions are determined as the product of an emissions factor and activity data. Emissions uncertainty requires combination of uncertainties in the emissions factor and activity data. The emissions factor generally does not change from one year to another and it may be considered independent of the activity data. Hence, the trend of an emissions inventory time series may be determined solely by a change in the activity data from one year to another. Activity in one year may be independent of activity in another year. Alternatively, activity may depend on the level of activity in another year. We will analyse these alternatives using contrasting New Zealand examples for illustration.

2. Determining animal energy requirement in the CH_4 and N_2O emissions inventories

2.1 Energy requirement for maintenance: Sensitivity to live weight

Calculations are needed to quantify the sensitivity of an animal's ME requirement for maintenance (ME_m) to a change of live weight (w). This cannot be done without estimation and assumptions. To illustrate, contrasting case study analyses will be done for mature milking cows and lambs. In the inventories, for the cow, the average value of w is relatively constant. In contrast, for the lamb, w increases by nearly 8 times from birth until slaughter, 6 months later.

2.1.1 Mature milking cow

For a mature milking cow, on average including definition of the required terms for a daily calculation, the curvilinear relation between ME_m (MJ ME per day) and w (kg) is implemented in the inventory by the following expression:

$$ME_m = [K*0.28*(w^{0.75})*(2.718^{-0.03*a})]/k_m \quad (2.1)$$

Equation (2.1) yields the basal metabolism, excluding ME for grazing and ME used directly for production that are described in CSIRO (2007)(see their equation 1.20). For a mature milking cow, term K in equation (2.1) has a value of 1.4 and the age in years (term a) is 4. The efficiency of ME use for maintenance (k_m) is a function of the pasture (feed) energy density (denoted M/D following CSIRO (2007)) and the linear relation is given by

$$k_m = 0.019 \text{ M/D} + 0.503 \quad (2.2)$$

On average, for dairy cattle, M/D is 11.4 MJ ME/kg DM (DM is dry matter), so k_m is 0.72. Combining terms, and multiplying by 365 for conversion of ME_m to an annual basis (R_m, MJ ME per year), we may re-write equation (2.1) as

$$R_m = 175*(w^{0.75}) \quad (2.3)$$

In 1990, on average, a mature milking cow weighed 446 kg, so R_m was 16,984 MJ ME. If the cow had weighed 447 kg, R_m would have been 17,013 MJ ME. This comparison indicated the sensitivity of R_m to a unit change of w was 29 MJ of ME per year. In 2007, on average, a mature milking cow weighed 455 kg (8 kg more than the 1990 value), so R_m was 17,240 MJ ME (256 MJ ME or 1% more than the 1990 value).

The curvilinear relation between ME_m and w may be portrayed graphically as a straight line if both axes are transformed to logarithmic (base 10) scales. The power coefficient in Equation (1.3) indicates the line's slope is equal to 0.75, first verified for values of w across 4 orders of magnitude of data based on rats to steers by Kleiber (1932). Verification of the 0.75 value has since been done for w across 27 orders of magnitude (10^{-18} to 10^{10} kg; West and Brown, 2005). If the number of cells in an animal was proportional to w, a logarithmic plot of ME_m and w could be expected to have a slope of 1.0, so ME_m per unit of w was constant. However, because the line's slope is 0.75, it is evident that a heavier animal has a lower ME_m per unit of w than a lighter animal. This reflects an increasing energy supply challenge throughout the body for animals of increasing weight and complexity (West and Brown, 2005). To illustrate, consider an extreme comparison between a tiny, spherical microbe and the cow.

2.1.2 Lamb at the time of slaughter

As stated earlier, for the lamb, w at birth is about 4 kg (9% of the ewe's w) and it increases by nearly 8 times over the following 6 months until slaughter.

Accounting for the lamb's annual maintenance requirement could thus be interpreted as ambiguous. However, comparative calculations were done for illustration. In equation (2.1), terms K and a can be defined for the lab and the values were 1.0 and 0.5 (years), respectively. On average, for the lamb grazing pasture between weaning (15 December in the inventory) and slaughter (15 March), M/D is 9.8 MJ ME/kg DM, so k_m is 0.69. The average reflects a linear approximation over time for the differential equation that is a function of w. Calculations indicated this procedure was sufficient numerically (data not shown; we acknowledge the peer reviewer's criticism of our evaluation that prompted these calculations). For the lamb, by approximation, equation (2.1) was re-written onto an annual basis (MJ ME per year) as

$$R_m = 147*(w^{0.75}) \tag{2.4}$$

In 1990, on average, the lamb carcass at slaughter weighed 14.1 kg. As done in the inventory, this weight will be divided by 0.45 for a calculation of R_m. Inserting this value of w (31.3 kg) into equation (4), R_m was 1,947 MJ ME expressed on an annual basis. Again, w for the lamb is not constant and the use of annual units is only for comparative purposes. Increasing the lamb's live weight at slaughter to 32.3 kg (1 kg heavier), the corresponding R_m was 1,992 MJ ME. This comparison suggested the sensitivity of R_m to a unit change of w was 45 MJ of ME per year.

In 2007, on average, the lamb carcass at slaughter weighed 16.8 kg (2.7 kg more than the 1990 value). Inserting the 2007 value of w at slaughter (37.3

kg) into equation (4), R_m was 2,220 MJ ME per year. Consequently, from 1990 to 2007, R_m (based on w at the time of slaughter) increased by 273 MJ ME.

2.1.3 The lamb from birth to slaughter

In the inventory, the mature breeding ewe becomes pregnant in May, the lamb is born in September and it is weaned in December. In 1990, a population ratio of lambs and mature breeding ewes was 1.01, meaning on average there was 1.01 lambs for each breeding ewe, so the ewe was considered to bear a single lamb. In 1990, the mature breeding ewe's live weight was 46.4 kg. For equation (2.1), terms K and a will have values of 1.0 and 2 (years), respectively, and k_m is 0.69. Using equation (2.4) with this live weight over the period 15 May – 15 September 1990, the ewe's R_m was 1165 MJ. Over the same period, the ewe's total ME requirement was 1436 MJ. The difference between this total ME requirement and the corresponding value of R_m was thus equal to 271 MJ (1436 – 1165). This quantity was attributed to the lamb, so 271 MJ ME was an estimate of the ME required for its maintenance from conception until birth.

Estimating the energy required for a lamb's maintenance from conception until birth did not partition the energy required by the ewe for wool production. A synopsis of the inventory accounting procedure for the energy requirement of wool production begins with annual activity data. In 1990 and 2007, annual wool yield averaged 5.1 and 5.7 kg (greasy fleece weight) per sheep, respectively. Thus, given 365 days per year, the corresponding values of average daily (greasy) fleece production was 14.0 and 15.6 grams. Following CSIRIO (2007), the associated daily ME requirements were 1.0 and 1.2 MJ (0.13*[14.0 – 6] versus 0.13*[15.6 – 6]), respectively.

In 1990, at birth on 15 September, the lamb's live weight was 3.8 kg (9 % of the ewe's live weight at slaughter). Then, at slaughter on 15 March, the lamb's live weight was 31.3 kg. Consequently, over 181 days from birth until slaughter, the lamb's live weight increased by 27.5 kg. From birth until weaning on 15 December, the milk-fed lamb's live weight increased by 18.4 kg (stipulated in the inventory to be 67 % of 27.5 kg). During this 91-day-long period, the increase of live weight thus averaged 202 grams per day. From birth until weaning, the average value of w was 13.0 kg ([3.8 + 22.2]/2). Inserting this value of w into equation (2.4) yields ME_m equivalent to 2.8 MJ ME per day. Multiplying by 91 days, we estimate 255 MJ ME was required by the lamb for its maintenance from birth until weaning.

Pasture is grazed by the lamb for the 90 days between weaning and slaughter. In 1990, from weaning until slaughter, the average value of w was 26.8 kg ([22.2 + 31.3]/2). Inserting this value of w into equation (1.4) yields ME_m equivalent to 4.7 MJ ME per day. Multiplying by 90 days, we estimate 428 MJ ME was required by the lamb for its maintenance from weaning until slaughter.

Combining our three 1990 estimates, we deduce 954 MJ ME was required by the lamb for its maintenance from conception until slaughter.

In 2007, the population ratio of lambs and mature breeding ewes was 1.28, so each ewe was considered to bear 1.28 lambs. During May – September, the ewe's total ME requirement minus the corresponding estimate of R_m was 410 MJ. This quantity was attributed to 1.28 lambs, so 320 MJ ME (410/1.28) was an estimate of the metabolised energy required for a lamb's maintenance from conception until birth (49 MJ more than the 1990 value).

In 2007, the lamb's birth and slaughter live weights were 4.8 and 37.3 kg, respectively (increases of 1.0 and 6.0 kg above the 1990 values), a difference of 32.5 kg. From birth until weaning, the milk-fed lamb's live weight increased by 21.8 kg (67 % of 32.5 kg) at an average of 215 grams per day (33 grams per day more than in 1990). From birth until weaning, the average value of w was 15.7 kg ([4.8 + 26.6]/2)(2.1 kg more than in 1990). Inserting this value of w into equation (2.4) yields ME_m equivalent to 3.2 MJ ME per day. Again multiplying by 91 days, in 2007, we estimate 291 MJ ME was required by the lamb for its maintenance from birth until weaning (36 MJ more than in 1990).

In 2007, from weaning until slaughter, the average value of w was 32.0 kg ([26.6 + 37.3]/2). Inserting this value of w into equation (2.4) yields ME_m equivalent to 5.4 MJ ME per day. Multiplying by 90 days, we estimate 486 MJ ME was required by the lamb for its maintenance from weaning until slaughter (58 MJ more than in 1990).

Combining our three 2007 estimates, we deduce 1,097 MJ ME was required by a lamb for its maintenance from conception until slaughter (143 MJ less than in 1990).

2.2 Energy requirement for products: Sensitivities to milk and meat production

2.2.1 Mature milking cow

Relations have been determined between the gross energy content of milk (denoted evl, energy value of lactation, MJ/kg milk) and its concentration of fat (F, %) and protein (P, %). In New Zealand, the relations had different parameter values for Holstein-Friesian and Jersey breeds according to Grainger et al. (1983). For the Jersey breed's milk, on average, F was significantly larger (57.6% versus 49.4%), P slightly smaller (33.9 versus 34.6%) and evl significantly larger (3.88 versus 3.34 MJ/kg). For the Holstein – Friesian and Jersey breeds, evl = 0.381F + 0.284P + 0.482 (a typographical error in Grainger et al. (1983) for the multiplier of P has been corrected here) and evl = 0.291F + 0.337P + 1.059, respectively. In 1990, on average, F and P were 4.80 and 3.52% (4.92 and 3.69% in 2007), respectively. The corresponding values of evl were 3.31 and 3.64, averaging 3.48 MJ/kg milk. The Jersey breed's milk had a 10% higher value of evl. In the inventory, there is a single relation for dairy cattle that is evl = 0.376F + 0.209P + 0.948. In 1990, according to the inventory relation, evl was 3.49 MJ/kg milk and it was 2% larger in 2007. The efficiency of ME use for lactation (k_l) is a function of (M/D) given by k_l = 0.019 M/D + 0.42. On average for dairy cattle, M/D is 11.4 MJ ME/kg DM, so k_l is 0.64. In 1990, dividing evl (3.49) by k_l yielded 5.45 MJ ME/kg milk.

The cow requires ME to produce milk, in addition to that required for its maintenance. In 1990, for an average mature milking cow, the total, annual ME requirement was 38,595 MJ. The corresponding value for maintenance was 16,984 MJ. It may thus be estimated that milk production required 21,611 MJ ME (38,595 - 16,984)(56 % of the total, annual ME requirement). In 1990, on average, the cow produced 2829 kg of milk, so requiring 7.6 MJ ME to produce 1 kg of milk. The former estimate (5.45 MJ ME/kg milk) was less, reflecting our use of basal metabolism to define maintenance in the latter estimate.

In 2007, for an average mature milking cow, the total, annual ME requirement was 45,567 MJ (6,972 MJ more than in 1990). The corresponding value for maintenance was 17,240 MJ (256 MJ more than in 1990). It may thus be estimated that milk production required 28,327 MJ ME (45,567 - 17,240)(6,716 MJ more than in 1990). In 1990, on average, the cow produced 3757 kg of milk, so requiring 7.5 MJ ME to produce 1 kg of milk.

2.2.2 Lamb

In 1990, from weaning until slaughter, the lamb's total pasture ME requirement was 961 MJ. Earlier, for 1990, we estimated 428 MJ ME was required by the lamb for its maintenance from weaning until slaughter. Thus, by difference, we deduce 533 MJ ME was required for meat production from weaning until slaughter. This involved an increase of live weight equal to 9.1 kg (31.3 – 22.2), so 59 MJ ME per kg live weight.

From birth until weaning, live weight increase proceeded at twice the rate from weaning until slaughter (in 1990, from 3.8 kg up to 22.2 kg, so 18.4 kg increase of live weight). From above, 59 MJ ME was required per kg increase of live weight, so 1,086 MJ ME was required for the 18.4 kg increase of live weight from birth until weaning. Earlier, from birth until weaning in 1990, we estimated 255 MJ ME was required by the lamb for its maintenance. Hence, from birth until weaning in 1990, we estimate the lamb's total ME requirement was 1,341 MJ (1,086 + 255).

To summarise for 1990, from birth until slaughter, we estimated the lamb's total ME requirement was 2,302 MJ (1,341 + 961). The estimate of the ME required for the lamb's maintenance from conception until birth was 271 MJ. Thus, the lamb's total ME requirement from conception to slaughter was estimated to be 2,573 MJ. In 1990, on average, the lamb's carcass at slaughter weighed 14.1 kg. Using the carcass weight at slaughter as a measure, from conception until slaughter, meat production included a total ME requirement 182 MJ per kg.

From birth to slaughter in 1990, live weight gain was 27.5 kg (3.8 to 31.3 kg). Excluding maintenance, for this weight gain, the ME requirement was estimated to be 1,619 MJ (533 + 1,086)(70% of the total ME requirement estimated from birth to slaughter). By these measures, as stated, meat production included 59 MJ ME per kg of live weight gain.

In 2007, from weaning until slaughter, the lamb's total pasture ME requirement was 1,368 MJ (407 MJ more than in 1990). Earlier, for 2007, we estimated 486 MJ ME was required by the lamb for its maintenance from weaning until slaughter (58 MJ more than in 1990). Thus, by difference, we deduce 882 MJ ME was required for meat production from weaning until slaughter (349 MJ more than in 1990). This involved an increase of live weight equal to 10.7 kg (37.3 – 26.6), so 82 MJ ME per kg live weight. The 2007 value is 23 MJ per kg live weight more than the 1990 value (82 versus 59). In 2007, ewes bore significantly more and larger lambs, and these lambs grew faster than lambs did in 1990.

From birth until weaning, in 2007, live weight increased from 4.8 up to 26.6 kg, so by 21.8 kg. From above, 82 MJ ME was required per kg increase of live weight, so the ME requirement was 1,788 MJ. Earlier, from birth until weaning in 2007, we estimated 291 MJ ME was required by the lamb for its maintenance. Hence, from birth until weaning in 2007, we estimate the lamb's total ME requirement was 2,079 MJ (1,788 + 291).

To summarise for 2007, from birth until slaughter, we estimated the lamb's total ME requirement was 3,447 MJ (2,079 + 1,368). The estimate of the ME required for the lamb's maintenance from conception until birth was 320 MJ. Thus, the lamb's total ME requirement from conception to slaughter was estimated to be 3,767 MJ. In 2007, on average, the lamb's carcass at slaughter weighed 16.8 kg. Using the carcass weight at slaughter as a measure, meat production had a total ME requirement 224 MJ per kg (42 MJ more than the 1990 value).

3. Assessing uncertainties in the CH_4 and N_2O emissions inventories

3.1 Setting the scene: Uncertainty terminology and IPCC good practice guidance

Uncertainty is a range within which a quantity might lie. It has two components: the width of the range and the degree of certainty that the true value might lie within it. No analysis can provide absolute certainty; a range wide enough to provide high certainty may be too wide to limit the possible values of emission usefully. We work with the standard deviation because it has well defined statistical properties with at least an approximate relationship between the range and the degree of certainty. For example a range of one standard deviation will contain the true value about 2 times in 3 (68%), and a range of two standard deviations will contain it about 19 times in 20 (95%). Different degrees of certainty might be appropriate for different situations; using the standard error allows for a choice. In more formal statistical investigations there is a clearly defined population and estimates of its parameters come from randomly selected data. The standard deviation is a population parameter and estimates of it are termed the standard error. However in judging uncertainty in estimates of mean emissions there is rarely an opportunity to measure uncertainty in the variability estimates in any formal sense, so there is little point n making the distinction between standard error and standard deviation.

For many variables the uncertainty is proportional the variable's mean, that is the larger the variable the larger its uncertainty. The standard deviation divided by the mean is then constant. This ratio is called the coefficient of variation and denoted CV. Since it is constant for one particular class of measurement, in the absence of data it can often be estimated by expert judgement more easily than the standard deviation. The CV is also the natural measure of uncertainty to use when calculating the uncertainty of the product of two uncertain quantities, as illustrated in the following example.

The Intergovernmental Panel on Climate Change (IPCC) provides guidance in estimating and reporting uncertainties associated with national greenhouse gas emissions inventories. The 2006 guidelines advocate that analyses should assess national, annual emissions as well as trends over time. To guide inventory compilers, a structured approach illustrates methods to i) determine uncertainties in variables including activity data and emissions factors, ii) aggregating the component uncertainties to that of the emissions, iii) determining uncertainty in trends (time series) and iv) identifying significant sources of uncertainty to determine priorities for data collection and efforts to improve inventories. In this chapter, we focus on subjects i, ii and iii. Our methods will be consistent with IPCC good practice guidance, but adaptation was required for special cases presented by some of the New Zealand inventory issues analysed. For example, we carefully and explicitly distinguished between independent and dependent variables in emissions inventories and time series analyses. This recognised persistent concern about this issue in New Zealand and, apparently, guidance that has been available to policy analysts.

3.2 Uncertainty of emissions in a year and changes from one year to another

For our first analysis, we consider the emissions in a given year (E), the product of activity data (A) and an emissions factor (F) written

$$E = AF \qquad (3.1)$$

Suppose the best estimate of A is 2.00 and F 1.00 so that E = 2.00. Then suppose that the uncertainty of A (denoted CV[A]) and F (CV[F]) are 0.05 and 0.10, respectively. As a simplification the variables A and F will be considered independent. This means the value of F was determined quite independently from the value of A and vice versa. The fractional uncertainty of E [CV(E)] is then approximately equal to a root-mean-square combination of the variables' CVs. This may be written

$$CV[E] = (CV[A]^2 + CV[F]^2)^{½} \qquad (3.2)$$

Inserting values into equation (3.2), we obtain $CV[E] = ([0.05^2] + [0.10^2])^{½} = 0.11$. Multiplying CV[E] by the value of E, we obtain the uncertainty (standard deviation) of E = 0.22. Note that this approximation becomes more accurate as the CVs decrease, and again, A and F must be independent.

Our next two analyses will quantify the uncertainty of a change in the emissions from a base year to another thereafter. For the first example we assume that the activity changes but the emission factor remains the same. Activity and emissions will be designated by subscripts b and a, denoting the base year and year afterwards, respectively. Using equation (3.1), we write

$$E_a - E_b = [A_a - A_b]\ F \qquad (3.3)$$

For calculation, we set A_a and A_b to 2.20 and 2.00, respectively, and F to 1.00, so $(E_a – E_b) = 0.20$. For the first analysis of uncertainty in $(E_a - E_b)$, A_a and A_b will be assumed independent. For example, there is no common cause affecting the measurement error in A_a and A_b. Consequently, E_a will be independent of E_b. Combining equations (3.2) and (3.3), we write

$$CV[A_a - A_b)\ F] = (CV[(A_a - A_b)]^2 + CV[F]^2)^{½} \qquad (3.4)$$

The difference $A_a - A_b$ could potentially be quite small making the CV large. The relationship giving the CV of a product is an approximation that becomes worse as the CV increases. Equation (3.4) will only be a reasonable approximation if the increase in A is large relative to the SD of A.

If CV[F] remains 0.10, $CV[F]^2 = 0.01$. For the difference $(A_a - A_b)$, recall the standard result that the standard deviation (SD) of a difference between two independent variables is given by

$$SD[(A_a - A_b)]^2 = SD[A_a]^2 + SD[A_b]^2$$

Using this gives

$$CV[(A_a - A_b)]^2 = (SD[A_a]^2 + SD[A_b]^2)/([A_a – A_b]^2) \quad (3.5)$$

For A_a and A_b, the same CV as before will be applied (= 0.05). Given A_a is 2.20, SD(A_a) = 0.11 (0.05*2.20), and A_b is 2.00, so SD(A_b) = 0.10. Inserting the values into equation (2.5), $CV[(A_a - A_b)]^2 = 0.55$. Inserting the values for $CV[(A_a - A_b)]^2$ and $CV[F]^2$ into equation (3.4), $CV[(A_a - A_b) F] = 0.75$. Multiplying $CV[(A_a - A_b) F]$ by the value of $(E_a – E_b)$(= 0.20), we obtain the uncertainty of $(E_a – E_b) = 0.15$. This uncertainty estimate is considered a rough approximation because of the large value for $CV[(A_a - A_b)]^2$.

For the second example of uncertainty in $(E_a - E_b)$, A_a and A_b will assumed be dependent, so A_a will be determined according to the value of A_b. The value of A_a will depend on A_b according to a proportionality coefficient denoted Ω, so

$$A_a = \Omega A_b \quad (3.6)$$

Following equation (3.1), we write

$$E_b = A_b F \quad (3.7)$$

By combining equations (3.6) and (3.7), we can write

$$E_a = \Omega A_b F \tag{3.8}$$

Combining equations (3.7) and (3.8) leads to

$$E_a - E_b = (\Omega - 1) A_b F \tag{3.9}$$

The fractional uncertainty of ($E_a - E_b$) (denoted CV[$E_a - E_b$]) then comes from a root-mean-square combination of the variable's uncertainties. This may be written

$$CV[Ea - Eb] = (CV[\Omega - 1]^2 + CV[Ab]^2 + CV[F]^2)^{½} \tag{3.10}$$

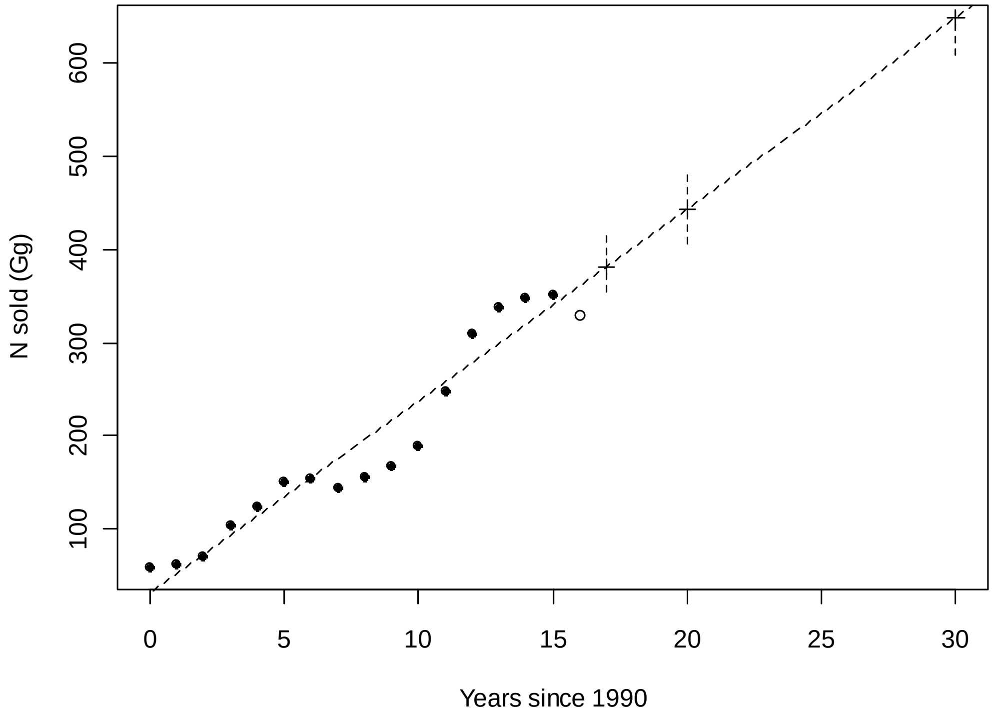
To illustrate by calculation, we set CV[Ω -1] to 0.05, while CV[A_b] and CV[F] remain 0.05 and 0.10, respectively. Inserting values into equation (3.2), we obtain CV[$E_a - E_b$] = $([0.05^2] + [0.05^2] + [0.10^2])^{½}$ = 0.12. Multiplying CV[$E_a - E_b$] by the value of ($E_a - E_b$)(= 0.20), we obtain the uncertainty of ($E_a - E_b$) = 0.024. Thus, the dependence of A_a and A_b corresponded with a large reduction in the uncertainty of ($E_a - E_b$) relative to the first example where A_a and A_b were independent.

3.3 Analysing time series of activity data and emissions including uncertainty

A common situation is an emissions trend determined by changes in activity, while the emissions factor stays constant. This presents a number of analysis problems including trend determination, trend uncertainty and trend projection to predict future values. In this section, we analyse representative, contrasting examples of activity data time series; namely, nitrogen (N) fertiliser sales and the number of breeding ewes at 30 June. The former is an important variable in the agricultural soils N_2O emissions inventory, while the latter represents a type of animal contributing to this inventory and the enteric CH_4 emissions inventory. Finally, including all types of animal, we analyse an enteric CH_4 emission inventory time series.

Our first example is activity data, a 1990 – 2006 series of nitrogen (N) fertiliser sales. These national-level data were compiled from fertiliser industry financial year (1 June – 31 May) records by FertResearch. Each data point had an uncertainty limit of ± 3% according to expert judgement (Dr. Hilton Furness, personal communication). For 13 of 17 years, N fertiliser sales increased from one year to the next but sales declined during years 6 + 8 and 15 + 16.

Figure 3.1: Nitrogen (N) fertiliser sales from 1990 (year 0) to 2006 (year 16). Circles are data, solid symbols used to fit a regression (dashed) line assuming independent data points. The open symbol (year 16 data) was not included, but used to verify predictions portrayed as horizontal solid lines with vertical dashed lines showing ± 1 SD prediction intervals.



The first analysis uses linear regression of fertiliser sales (Gg N) on time expressed as years since 1990 (years ranged from 0 – 15 with the data for year =16 excluded for forecast verification). Standard regression analysis gives the following information:

Equation of line: N = 31.5 (± 15.1) + 20.6 (± 1.7) Year SD = 31.7

The SD is the estimated standard deviation of points around the regression line, or the residual SD. It is therefore a measure of the uncertainty of predictions made for a single year, ignoring the uncertainty in estimating the trend line itself. The regression's slope, 20.6, ± 1.7 Gg N/y (± standard deviation), quantifies the average annual increase in N sales (time trend of the activity data). The regression's constant (31.5 ± 15.1 Gg N) is the regression estimate of the expected sales in year zero (1990) given the activity data time trend. It's SD (15.1 Gg N) is the uncertainty in the regression line when time = 0 (Year = 0). Additional to this is the uncertainty in the sales that year, the SD about the regression line, 31.7 Gg N. These combine according to root mean square combination to give a SD of 35.0 Gg N. The regression constant was 27.8 Gg N less than the corresponding sales data (59.3 Gg N for zero years after 1990, so 59.3 – 31.5 = 27.8). However, this difference is well within the "combined" SD (35 Gg N) and so is not evidence of poor fit of the regression to the data.

An important measure of the usefulness of a regression equation is how much it reduces the uncertainty of predictions beyond just quoting the mean N sales. For 1990 to 2005, the mean was 186 (Gg N/y) with SD 102. The SD about the regression line was 31.7 (Gg N/y), considerably smaller, representing a big drop in uncertainty. Conventionally, the difference between these two values is measured by the % variance explained, or the adjusted R^2. For the regression (equation given above), the adjusted R^2 was equal to 92%, very high.

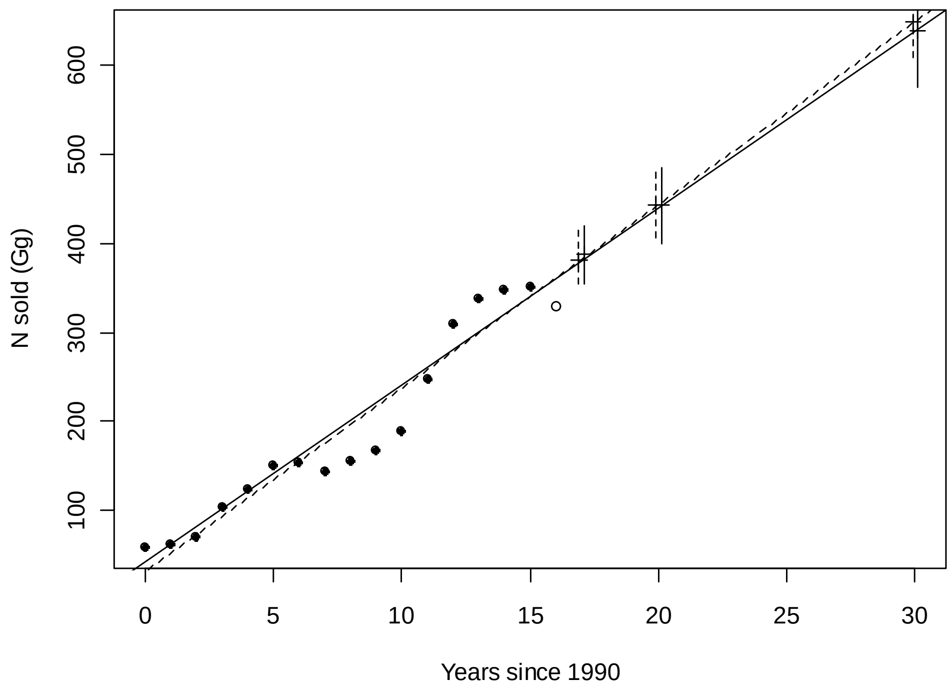
$$R^2 = 100\,(1 - ((\text{SD about line}) / (\text{SD about mean}))^2) \qquad (11)$$

Throughout the time series the SD about the line was much greater than the expert judgment of the uncertainty (± 3%) in each year's N sales. In 2005, expert judgment of the uncertainty obtained a maximum value of ± 11 Gg N. However, the SD about the line includes the departure of sales from a linear trend. Figure 3.1 shows that although the trend is approximately linear over the 16 years there are short term fluctuations much larger than any error in the measurement of sales. The uncertainty arising from these fluctuations are included in the SD about the line.

The 2006 data, not used in the regression analysis, were X = 16 yr and Y = 330 Gg N. For a value of 16 yr, the linear regression model yielded a forecast for 2006 of 361 ± 36 Gg N. (SD = 36 comes from a root-mean-square combination of the residual SD (31.7) and the corresponding value associated with the uncertainty in the regression line). Uncertainty in the regression line grows larger moving further from the mean year, 1997.5 for these data. The regression constant was a prediction for 1990, 7.5 years less than the mean year. Hence, for 2006, uncertainty associated with the regression line was ± 17 Gg N, exceeding ± 15.1 Gg N obtained for the regression constant. The 2006 forecast exceeded the corresponding sales data by 31 Gg N or 7 %, but not beyond the forecast's 1 SD uncertainty limit. For values of 20 and 30 yr, further extrapolation by the regression model yielded forecasts for 2010 and 2020 of 443 ± 39 and 648 ± 51 Gg N, respectively. The uncertainty in the regression line increases its effect as the value of X stretches beyond the data. These limits make no allowance for future events that change the nature of the linear trend.

After the assumption that the trend in linear, the most important assumption underlying the linear regression is that successive residuals are independent. This is rarely true in a time series. If external events cause one year's sales to be high it is very likely that the following year's sales will be high too. This effect is evident in some of the data shown in Figure 1; for example, compare sales from years 7 to 10 (averaging 164 ± 19 Gg N) and those from years 12 to 14 (averaging 336 ± 19 Gg N). This effect can be captured by generalising the linear regression to include a correlation between successive residuals, an autocorrelation. A lack of independence in the data means standard linear regression "overstates" the accuracy of its predictions. A times series model will now be fitted to the data allowing the residual at one year to be correlated with the residual of the next.

Figure 3.2: Fertiliser sales from 1990 (year 0) to 2006 (year 16). Circles are data, solid symbols used to fit two regression lines. The solid line allows autocorrelation, generalised regression, while the dashed line assumes independence (standard regression, portrayed in Figure 3.1). The open symbol (year 16 data) was not included, but used to verify predictions portrayed as horizontal solid lines with vertical dashed lines showing ± 1 SD prediction intervals.



The solid line and predictions are plotted in Figure 3.2 and the numerical results are:

Equation of line: N = 41.7 (± 24.1) + 19.9 (± 2.4) Year SD = 26.9

Correlation = 0.74

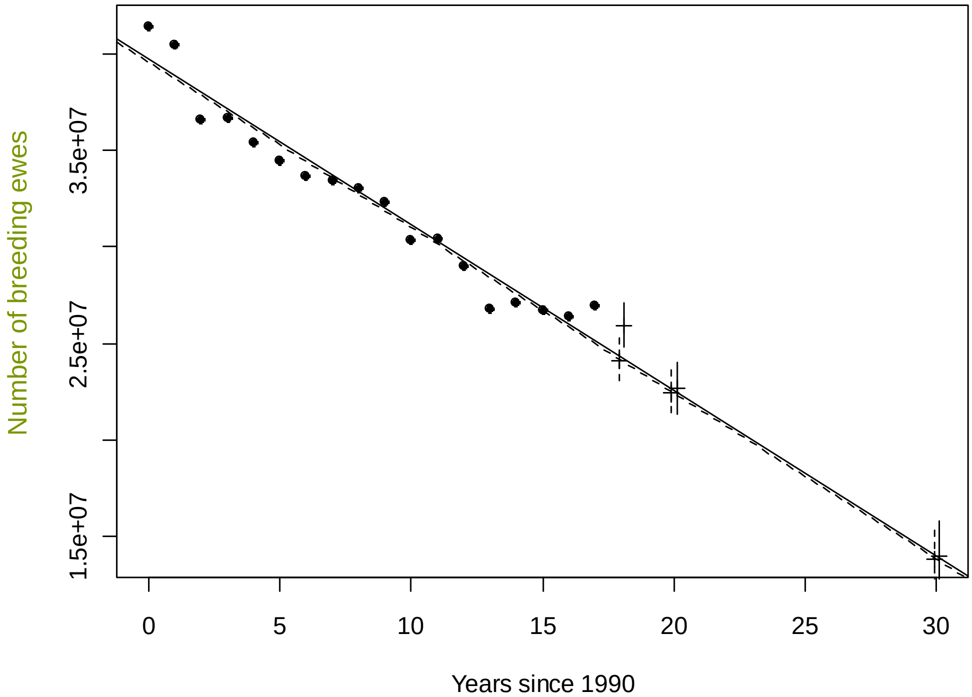
The generalised regression model yielded a slope estimate 8% less than standard regression (20.6 ± 1.6), but the two lines agreed well across the span of the data. Moreover, incorporating correlation does not greatly affect the estimation of means. Correctly accounting for data dependence by autocorrelation did, however, mean the regression slope's SD increased by 71%. By assuming that residuals are independent the standard regression exaggerated the amount of information available in the data and therefore understated the uncertainty.

The generalised regression's constant increased by 32% compared to standard regression (31.5 Gg N) and its SD increased by 77% (formerly ± 14.2 Gg N). Thus, uncertainty in the generalised regression line at time zero was 77% greater than for the standard regression line but, for these correlated data, the former is the correct value. For a value of 16 yrs the generalised regression model yielded a (2006) forecast of 368 ± 31 Gg N. Note that this is a larger value than would be given by simple substitution in the estimated equation of the line given above. The prediction includes the correlation between successive observations, so because the year 15 value was higher than average the value for year 16 is expected to be as well. This is in contrast to standard regression where, because observations are assumed independent, there is no information from the earlier observation and predictions lie on the fitted line. In Figure 3.2, with correlation, the forecast of 368 Gg N is above the fitted line. It happens that for 2006 this makes the estimate worse, 12% greater than the corresponding sales data. However, for most years successive points are on the same side of the line, so predictions are improved using the generalised regression model rather than standard regression.

For values of 20 and 30 y, the generalised regression model forecasts for 2010 and 2020 were 441 ± 44 and 638 ± 63 Gg N, respectively. These forecasts were close (5 – 6 % lower) to those according to standard regression, but the generalised model forecast's uncertainty limits were much greater. Again, incorrectly assuming independence exaggerated the information in the data and understated uncertainty in all the estimates. Thus, using a larger SD of the slope from the generalised regression model does increase uncertainty in the extrapolated predictions, but the smaller uncertainty by standard regression was an artefact of an inappropriate assumption about the data analysed.

A second example of activity data time series is the number of breeding ewes on 30 June from 1990 to 2007. Figure 3.3 shows the data with the prediction lines and intervals. The models were fitted to all the data and predictions made for 2008 for which data is not yet available.

Figure 3.3: Number of breeding ewes at 30 June from 1990 (year 0) to 2007 (year 17). The solid symbols are data used to fit two regression lines. The solid line allows autocorrelation, generalised regression, while the dashed line assumes independence (standard regression). Predictions are portrayed as horizontal solid lines with vertical dashed lines showing ± 1 SD prediction intervals.



Standard regression analysis gives the following information:

Equation of line:

Number = 39597×10^3 (± 504×10^3) – 860×10^3 (± 51×10^3) Year

SD = 1115×10^3

Regression accounted for 95% of the variance.

Allowing for autocorrelation, generalised regression analysis yielded:

Equation of line:

Number = 39741×10^3 (± 696×10^3) – 858 $\times10^3$ (± 67×10^3) Year

SD = 1029×10^3

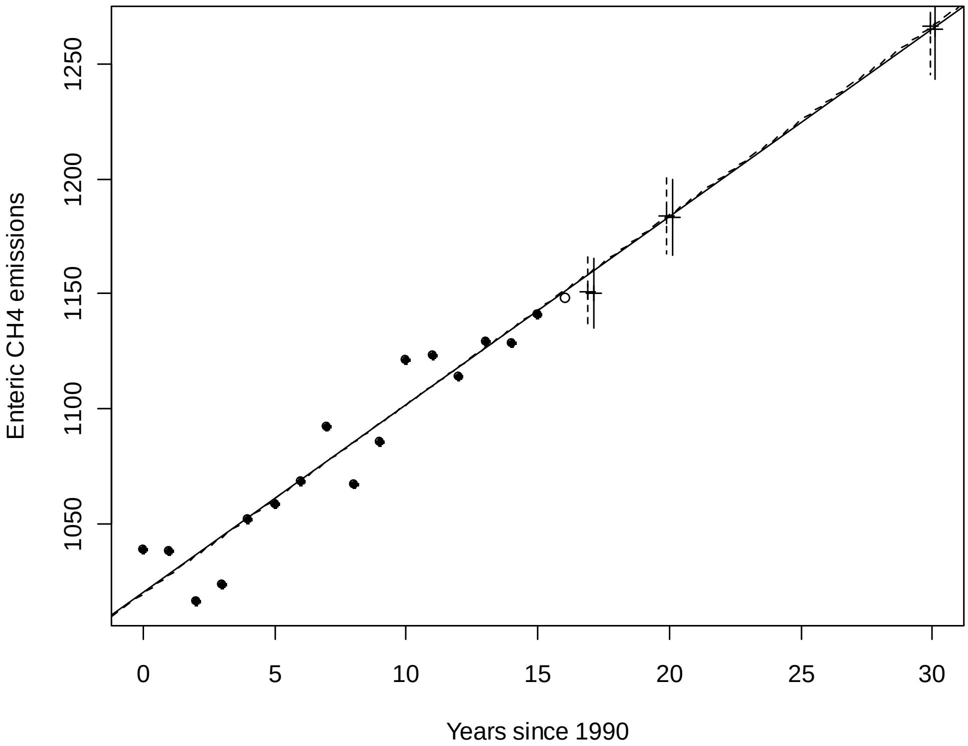
Correlation = 0.45

The two offsets were indistinguishable, but the slope from standard regression was 2% greater (more negative) than that by generalised regression. The standard regression predictions for 20 and 30 years were 22.4 ± 1.3 and 13.8 ± 1.6 million, respectively, very similar to 22.7 ± 1.3 and 14.0 ± 1.8 million predicted by generalised regression. The moderate correlation of 0.45 meant that the generalised regression gave a greater slope SD than generalised regression, by 30%, resulting in larger SDs with greater extrapolation. An interesting point however is that the generalised regressions' prediction for 2008 is 4% greater, 25082×10^3 ($\pm 1190\times10^3$) instead of 24121×10^3 ($\pm 1242\times10^3$). The 2007 (Year 17) data point is larger than the Year 16 data point as well as an average for Years 13 – 16. Consequently, because of the autocorrelation, the 2008 prediction is increased and correlation with the 2007 observation lowers uncertainty in the 2008 prediction.

This example shows that allowing for auto correlation can make small improvements to estimates a year or two ahead. However its main importance lies in correctly assessing the uncertainty, which is understated if independence is ignored. Independence was also seen to be important when assessing changes in emission in section 3.2.

Finally, including all types of animal, we analyse enteric methane (CH_4) emissions inventory data in the form of a 1990 – 2006 time series. These national-level data came from Ministry for the Environment (MfE) on 18 May 2008. As described by the MfE, each data point represents an inventory for a calendar year (1 January – 31 December) with an uncertainty limit of ± 50% according to numerical analysis. For 14 of 17 y, CH_4 emissions increased from one year to the next but they declined in years 2, 8 and 12.

Figure 3.4: Enteric CH_4 emissions inventory from 1990 (year 0) to 2006 (year 16). Circles are data, solid symbols used to fit two regression lines. The solid line allows autocorrelation, generalised regression, while the dashed line assumes independence (standard regression). The open symbol (year 16 data) was not included, but used to verify predictions portrayed as horizontal solid lines with vertical dashed lines showing ± 1 SD prediction intervals.



Standard regression analysis gives the following information:

Equation of line: CH_4 = 1019 (± 8) + 8.27 (± 0.78) Year SD = 14.3

Regression accounted for 89% of the variance.

Allowing for autocorrelation, generalised regression analysis yielded:

Equation of line: CH_4 = 1020 (±7) + 8.15 (± 0.78) Year SD = 12.6

Correlation = 0.14

The two analyses give very similar results because the autocorrelation, at 0.14, is small. The two values of residual SD are much lower than the estimated uncertainty limit of ± 50% for each data point, ranging from ± 519 to ± 671 Gg CH_4. The residual SD represents the uncertainty in year by year determinations in the CH_4 emissions inventory (the emissions trend over time). The ± 50% uncertainty refers to each year's determination. Inaccuracy in the determination process affects each year's CH_4 emissions inventory in the same way, moving the regression line but not affecting its fit to the data. The ± 50% therefore refers to position of the regression line as a whole.

For 2006, the data are X = 16 years and Y = 1148 Gg CH_4. Standard regression yielded a 2006 forecast of 1151 ± 16 Gg CH_4. The standard regression's forecasts for 2010 and 2020 were 1183 ± 17 and 1266 ± 23 Gg CH_4, respectively. After autocorrelation, the generalised regression's 2006, 2010 and 2020 forecasts were almost indistinguishable from those by standard regression alone at 1150 ± 15, 1183 ± 16 and 1265 ± 22 Gg CH_4, respectively. There were slight changes in the uncertainty limits, but no affect of the extent of the extrapolation. Again, this reflected the small value of the autocorrelation and small number of data points analysed.

Acknowledgements

We are grateful to Drs. Keith Lassey and Roger Hegarty for constructive critiques of the draft report.

References

CSIRO 2007. Nutrient requirements of domesticated ruminants. CSIRO Publishing. Collingwood, Victoria, Australia. 279 pages.

Grainger, C., Holmes, C.W., Bryant, A.M., Moore, Y. 1983. A note on the concentration of energy in the milk of Friesian and jersey cows. Animal production. 36: 307 – 308.

Kleiber, M. 1932. Body size and metabolism. Hilgardia 6: 315 – 353.

West, G.B. and J.H. Brown. 2005. The origin of allometric scaling laws in biology from genomes to ecosystems: Towards a quantitative unifying theory of biological structure and organization. The Journal of Experimental Biology 208: 1575 – 1592.