Title

Development of mathematical models to predict volume and nutrient composition of fresh manure from lactating Holstein cows

Journal

Animal Production Science, 54(12)

ISSN

1836-0939

Authors

Appuhamy, JAD Ranga Niroshan
Moraes, LE
Wagner-Riddle, C
et al.

Publication Date

2014

DOI

10.1071/an14533

Peer reviewed

CSIRO PUBLISHING

Animal Production Science, 2014, **54**, 1927–1938
http://dx.doi.org/10.1071/AN14533

Development of mathematical models to predict volume and nutrient composition of fresh manure from lactating Holstein cows

J. A. D. Ranga Niroshan Appuhamy[A,E], *L. E. Moraes*[A], *C. Wagner-Riddle*[B], *D. P. Casper*[C], *J. France*[D] *and E. Kebreab*[A]

[A]Department of Animal Science, University of California, Davis, CA 95616, USA.
[B]School of Environmental Sciences, University of Guelph, Guelph, ON, N1G 2W1, Canada.
[C]Department of Dairy Science, South Dakota State University, Brookings, SD 57007, USA.
[D]Centre for Nutrition Modelling, Department of Animal and Poultry Science, University of Guelph, Guelph, ON, N1G 2W1, Canada.
[E]Corresponding author. Email: jaappuhamy@ucdavis.edu

Abstract. Organic compounds in dairy manure undergo a series of reactions producing pollutants such as ammonia and methane. Because various organic compounds have different reaction rates, the emissions could be accurately determined if amounts and concentrations of individual nutrients in manure are known. A set of empirical models were developed for predicting faecal and urinary water, carbon (C), nitrogen (N), acid detergent fibre and neutral detergent fibre output (kg/day) from lactating Holstein cows. Dietary nutrient contents, milk yield and composition, bodyweight, age and days in milk were used with or without dry matter intake (DMI) as potential predictor variables. Multi-collinearity, goodness of fit, model complexity, and random study and animal effects were taken into account during model development, which used 742 measured faecal or urinary nutrient output observations (kg/day). The models were evaluated with an independent dataset ($n = 364$). When DMI was used as a predictor variable, the models predicted faecal and urinary nutrient outputs successfully with root mean square prediction error as a percentage of average observed values (RMSPE%) ranging from 9.1% to 20.7%. All the predictions except urine output had RMSPE% ranging from 18.3% to 24.6% when DMI was not used. The nutrient output predictions were in reasonable agreement with observed values throughout the data range (systematic bias <14% of total bias). Fresh manure C : N ratio predictions were acceptable (RMSPE% = 14.3–15.2%) although the systematic bias were notable (17.1–20.7% of total bias). The models could be integrated successfully with process-based manure or soil models to assess nutrient transformation in dairy production systems.

Additional keywords: faeces, dairy cows, nutrient excretion, prediction models, urine.

Received 1 May 2014, accepted 20 June 2014,

Introduction

Manure from dairy farms has been recognised as a major source of air quality pollutants and greenhouse gases such as nitrous oxide, methane (CH_4), and ammonia (NH_3) (Külling *et al.* 2001; Hristov *et al.* 2011). The rise in milk production and the expansion of dairy herds have increased the need for strategies aimed at mitigating these negative environmental effects (Wilkerson *et al.* 1997; St-Pierre and Thraen 1999). Quantitative tools for determining manure nutrient output and rate of nutrient release to the environment are of much importance in designing, implementing and evaluating such strategies.

Soon after excretion, nitrogenous and other organic compounds in manure undergo a series of reactions including denitrification and decomposition (DNDC), nitrification, hydrolysis, ammonia volatilisation and fermentation under different environmental conditions. Manure and soil models (e.g. the Manure-DNDC model (Li *et al.* 2012)) attempt to represent quantitatively the effects of substrate availability and environmental factors on these reactions and thereby predict greenhouse gas emissions and NH_3 volatilisation from dairy farms at barn, manure storage and land application levels (Li *et al.* 2012). Carbon (C), nitrogen (N) and water content in fresh manure from dairy cows are major input variables to the Manure-DNDC model and are estimated using factors such as milk yield and diet composition. However, calculations in Manure-DNDC do not appear to address satisfactorily the significant variation associated with C, N and water outputs. For example, manure N output is estimated assuming that a constant fraction of the N fed to dairy cows is secreted in milk. Moreover, a detailed characterisation of manure organic matter, particularly fibre composition would improve manure nutrient transformation predictions. Külling *et al.* (2002) and Hindrichsen *et al.* (2005) showed that increasing neutral detergent fibre (NDF) and acid detergent fibre (ADF) concentrations in fresh dairy manure decreased CH_4 emissions from corresponding manure when stored as slurry. Hashimoto *et al*. (1981) observed a negative relationship between cellulose to hemicellulose ratios in cattle manure slurry and CH_4 emissions. Furthermore, Amon *et al*. (2006) showed that dairy manure nutrient composition; specifically crude protein (CP) and lignin concentrations, were significantly related to CH_4 yields from anaerobic digesters. Estimates of manure organic matter composition can also assist in effective bio-based uses of dairy manure, such as biogas and monosugars production, effectively reducing environmental liabilities and providing an economic incentive to dairy producers (Liao *et al.* 2004; Schievano *et al.* 2008). Acid hydrolysis is the typical process used to treat and help convert lignocellulosic materials to sugars, yields of which are significantly related to cellulose, hemicellulose and lignin contents of dairy manure (Liao *et al.* 2005).

Although several models have been constructed to predict manure and nutrient excretions from dairy cows (Wilkerson *et al.* 1997; Jonker *et al.* 1998; Bannink *et al.* 1999; Kauffman and St-Pierre 2001; ASAE 2005; Nennich *et al.* 2005, 2006), they do not allow for detailed manure characterisation and are limited to dry matter (DM), N or mineral excretion prediction. Moreover, several of these extant models include potentially correlated predictor variables, e.g. both DM intake (DMI) and milk yield or both DMI and N intake. Correlated variables in a model are often associated with computational issues such as multi-collinearity. Overall model fit may not be affected by multi-collinearity but parameter estimates of incorrect sign and implausible magnitude may be produced, which can lead to inaccurate and non-generalisable interpretations (Mason and Perreault 1991). Furthermore, variable and model selection procedures for extant models are based primarily on model goodness of fit and little attention has been given to model complexity, which could also significantly affect the generalisability or extrapolation power of the prediction model (Myung 2000). The objective of the study was to construct a set of empirical models allowing for more detailed characterisation of fresh manure output from lactating dairy cows. The equations were developed with and without DMI as a predictor variable while taking multi-collinearity, goodness of model fit and model complexity into consideration.

Materials and methods

Data sources

A total of 1106 measured faecal or urinary nutrient outputs (kg/day) by individual cows, and related DMI and dietary nutrient composition, milk yield and milk composition, days in milk (DIM), bodyweight (BW) and age, were obtained from energy balance trials conducted at the former Energy Metabolism Unit (EMU), USDA-Beltsville. Each energy balance trial consisted of a metabolism trial (7–10 days), where faeces and urine excretions were measured daily using a total collection method along with DM and nutrient intake, milk yield and milk composition measurements. Feed and faecal samples dried at 65°C were used for determination of ether extract (EE), NDF, ADF and permanganate lignin using the procedure described by Goering and Van Soest (1970). Feed, faecal and urinary N and C were determined respectively using the Kjeldahl procedure and by combustion in an atmosphere of oxygen with volumetric determination of CO_2 produced (Smith *et al.* 1965). All moisture determinations were made by drying to a constant weight in a forced draft oven at 100°C. Milk samples collected daily over two consecutive milkings were analysed for true milk protein and fat percentages by infrared analysis (Moe *et al.* 1972, 1973; Tyrrell and Moe 1974; Tyrrell *et al.* 1988; Sechen *et al.* 1989; Andrew *et al.* 1991). A list of measured variables and their summary statistics are given in Table 1. The data came from 47 energy balance studies using 315 lactating Holstein cows, of which, 265 cows provided multiple observations ranging from 2 to 24 observations per cow. Data from 31 studies (742 observations from 218 cows) were randomly selected for model construction and the rest (n = 364) were allocated to model evaluation.

Variable and model selection

A set of linear mixed-effects models was constructed to predict separately faecal DM (F_{DM}), water (F_{Water}), C (F_C), N (F_N), NDF (F_{NDF}), ADF (F_{ADF}) and lignin (F_{Lignin}) output, and total urine (U_E), urinary C (U_C) and urinary N (U_N) output (all in kg/day). Selection of variables to develop the models began with a primary pool including DMI (kg/day), dietary DM percentage, dietary CP, NDF, ADF, lignin, EE and total ash contents (% of DM) and N, NDF and ADF intake (kg/day), milk yield (kg/day) and milk protein and fat percentages, BW (kg/cow), age (years) and DIM (Fig. 1). Two separate variable selection schemes were initiated with and without DMI data for each response variable. Variables for which the absolute value of the Pearson's correlation coefficient $|r| \geq 0.5$ were not included simultaneously in order to minimise multi-collinearity issues such as inaccurate model parameterisation, decreased statistical power and exclusion of significant predictor variables during model construction (Graham 2003). Consequently, variable selection that included DMI data was proceeded by two subschemes; one using DMI and corresponding dietary nutrient content (e.g. DMI and CP) and the other using related nutrient intake (e.g. N intake) because of notable linear correlations among them (Table 2). Moreover, within each selection scheme, milk yield and DMI (r = 0.78), dietary ADF and NDF contents (r = 0.85), dietary ADF and lignin contents (r = 0.63), dietary CP and ash contents (r = 0.50), milk yield and DIM (r = –0.61), and milk yield and milk protein percentage (r = –0.53) were not included together. Therefore, several subpools of variables had to be formed within each selection scheme (Fig. 1). For instance, the scheme including DMI and dietary nutrient content but not nutrient intake had four subpools of predictor variables (Table 3), whereas the scheme excluding DMI data had eight subpools of predictor variables (Table 3).

All possible combinations of variables in each subool were regressed separately against the response variable in question. For example, one of the subpools using DMI (Subpool 1 in Table 3) had 11 independent variables (P = 11), which led to 2048 potential regression models ($2^P = 2^{11} = 2048$). Each regression was carried out accounting for random animal and study effects as shown in following linear mixed-effects model:

$$y_{ijk} = \mathbf{x}_{ijk}^{\mathrm{T}}\boldsymbol{\beta} + \alpha_i + \gamma_j + \varepsilon_{ijk},$$

Table 1. Summary statistics for the data (n = 1106)

F_{Water}, F_{DM}, F_C, F_N, F_{NDF}, F_{ADF}, F_{HC} *and* F_{CL} = faecal water, dry matter (DM), carbon, nitrogen, neutral detergent fibre (NDF), acid detergent fibre (ADF), hemicellulose and cellulose outputs, respectively. U_E, U_C and U_N = total urine output, urinary carbon and nitrogen outputs, respectively. T_E, T_C, T_N, $R_{C:N}$ and C_{DM} = total fresh manure output, total carbon and nitrogen outputs, carbon to nitrogen ratio in fresh manure and dry matter concentration in fresh manure, respectively

Variable	Mean	s.d.	CV%	Minimum	Maximum
Independent or predictor variables					
Diet composition					
DM (% of diet)	68.0	20.0	29	30.2	93.8
CP (% of DM)	16.1	2.40	15	10.3	21.9
NDF (% of DM)	33.8	7.13	21	16.1	57.2
ADF (% of DM)	19.6	4.30	22	8.97	31.4
Lignin (% of DM)	4.33	1.48	34	1.26	8.44
Ether extract (% of DM)	2.56	0.75	29	0.52	4.95
Ash (% of DM)	6.31	1.11	18	3.54	9.99
Intake (kg/day)					
DM	15.6	4.08	26	6.40	28.7
N	0.41	0.13	32	0.14	0.93
NDF	5.31	1.81	34	1.15	12.0
ADF	3.08	1.06	35	0.70	6.82
Lignin	0.69	0.31	45	0.12	1.84
Production and other characteristics					
Milk yield (kg/day)	21.6	9.80	45	1.04	49.1
Milk fat (%)	3.50	0.76	22	1.30	7.60
Milk protein (%)	3.27	0.41	13	2.30	5.75
Age (years)	5.77	2.33	40	2.00	15.4
Bodyweight (kg/cow)	603	78.3	13	351	854
Days in milk	175	90.0	51	0.00	488
Response variables					
Faecal excretions (F_x, kg/day) and faecal C content (C_C, kg/kg of DM)					
F_{Water}	25.0	9.80	39	4.03	59.8
F_{DM}	5.20	1.77	34	1.18	10.7
F_C	2.40	0.80	33	0.54	4.76
F_N	0.13	0.04	31	0.05	0.25
F_{NDF}	3.06	1.04	34	0.54	7.21
F_{ADF}	1.91	0.65	34	0.34	4.24
F_{HC}	1.17	0.49	42	0.10	3.22
F_{CL}	1.16	0.42	36	0.21	2.76
C_C	0.46	0.02	04	0.38	0.52
Urinary excretions (U_x, kg/day)					
U_E	16.6	6.60	40	4.38	34.9
U_C	0.22	0.07	32	0.07	0.43
U_N	0.16	0.08	48	0.03	0.40
Total output (T_x, kg/day), C : N ratio ($R_{C:N}$), and DM content (C_{DM}, w/w) of fresh manure					
T_E	46.7	14.0	30	16.9	98.5
T_C	2.59	0.80	31	0.68	5.09
T_N	0.29	0.10	34	0.09	0.66
$R_{C:N}$	9.50	2.69	28	4.24	19.6
C_{DM}	0.11	0.02	17	0.05	0.19

where y_{ijk} is the kth measurement of the response variable of ith animal in jth study (k = 1, . . ., n_{ij}), $\mathbf{x}_{ijk}$ is the p-dimensional vector of independent variables, $\boldsymbol{\beta}$ is the vector of regression coefficients, α_i is the random effect associated with the ith animal (i = 1,., 218), γ_j is the random effect associated with the jth experiment (j = 1, . . ., 31) and ε_{ijk} is the error. It is assumed that random effects and errors are mutually independent and normally distributed. All the mixed-models in each subpool were fitted first to data using the *lme4* package in R (version 2.12.2, R Foundation for Statistical Computing, Vienna, Austria) and then ranked by descending Bayesian information criterion (BIC; Schwarz 1978) values. The BIC was calculated as:

$$\mathrm{BIC} = -2 \times \ln L_{\max} + K \times \ln N,$$

where $L_{\max}$ is the maximum likelihood achievable by the model, K is the number of parameters of the model, and N is the total number of observations used in the fit. The absolute BIC value of a model has no interpretation but can be compared with the values of other models. Lower BIC values imply proper balance between model complexity and model fit (Myung 2000). Another criterion widely used in model selection is the Akaike information criterion (AIC; Akaike 1974). Because BIC leans more towards lower-dimensional models (Schwarz 1978) compared with AIC, BIC was used to select the simplest model that best predicts the response of interest.

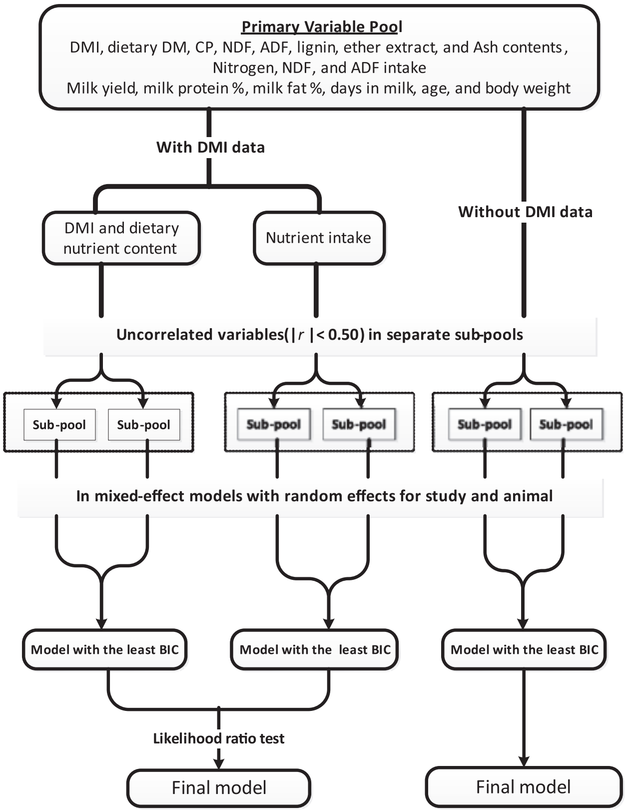

Fig. 1. Schematic diagram illustrating variable and model selection schemes.

The least BIC-associated model was chosen from the analysis of each subpool within each scheme. For example, the schemes with and without DMI had four and eight models because they had four and eight subpools, respectively (Table 3). These models within each scheme were ranked by descending BIC values and the model with the least BIC value was chosen as the best prediction model within each scheme (Fig. 1). However, the two schemes with DMI; one using DMI and the dietary nutrient content separately and the other using them for calculating the nutrient intake (Fig. 1), were associated with two best models for predicting each of F_{NDF}, F_{ADF}, F_N and U_N. Only one model was ultimately chosen as the final prediction model using the likelihood ratio test (Fig. 1). The models including DMI and the corresponding dietary nutrient content provided a significant improvement in model fitting ($P < 0.001$) and had smaller BIC compared with models including nutrient intake (Table 4). So those models were chosen as the final prediction models that require DMI data. Variance inflation factors (VIF) and the determinant of the correlation matrix ($det\{\mathrm{R}\}$) of predictor variables included in the final prediction models were calculated to verify degree of multi-collinearity. Numbered (Eqns 1–18) final prediction models are presented along with corresponding maximum VIF and $det\{\mathrm{R}\}$ in Table 5.

Secondary predictions

Faecal C concentration was fairly constant within the data (CV = 4%, Table 1). Consistently, faecal DM output (F_{DM}, kg/day) had a strong linear relationship with faecal C (F_C, kg/day) output ($r = 0.98$, data not shown). Hence an additional simple linear mixed-effects model (Eqn 19) was developed while accounting for random study and animal effects in order to determine F_C using only F_{DM} estimated with (Eqn 1) or without DMI (Eqn 10).

Table 2. Correlations (*r*) among candidate predictor variables

DM = dietary dry matter (% of diet), CP, ADF, NDF, LIG, EE and Ash = dietary crude protein, acid detergent fibre, neutral detergent fibre, lignin, crude fat and dietary ash contents, respectively (% of DM); Milk = milk yield (kg/day); mPrt = milk protein percentage; mFat = milk fat percentage; DIM = days in milk; BW = bodyweight (kg/cow); Age = age of the cows (years); iN, iNDF and iADF = nitrogen, NDF and ADF intake (kg/day)

	DM	CP	NDF	ADF	LIG	EE	Ash	Milk	mPrt	mFat	DIM	BW	Age	iN	iNDF	iADF
DMI	–0.34	0.22	0.05	0.11	0.27	0.37	0.14	0.78	–0.17	0.04	–0.37	0.33	–0.15	0.91	0.8	0.84
DM	–	0.13	–0.46	–0.41	–0.28	–0.48	0.06	–0.32	0.15	–0.32	0.09	0.03	–0.15	–0.25	–0.53	–0.50
CP	–	–	–0.33	–0.17	–0.06	–0.01	0.50	0.31	–0.03	–0.09	–0.23	0.15	–0.14	0.58	-0.01	0.10
NDF	–	–	–	0.85	0.47	0.24	0.23	–0.06	0.05	0.41	–0.08	–0.26	0.02	–0.08	0.61	0.49
ADF	–	–	–	–	0.63	0.30	0.44	–0.01	0.04	0.45	–0.08	–0.25	–0.03	0.04	0.56	0.62
LIG	–	–	–	–	–	0.28	0.33	0.05	0.13	0.43	0.1	0.01	–0.06	0.22	0.44	0.54
EE	–	–	–	–	–	–	0.14	0.32	–0.04	0.23	–0.07	–0.06	–0.04	0.33	0.41	0.46
Ash	–	–	–	–	–	–	–	0.11	0.05	0.20	–0.18	–0.11	–0.13	0.32	0.21	0.34
Milk	–	–	–	–	–	–	–	–	–0.53	–0.1	–0.61	0.15	–0.12	0.79	0.61	0.63
mPrt	–	–	–	–	–	–	–	–	–	0.31	0.43	0.01	–0.04	–0.14	–0.14	–0.13
mFat	–	–	–	–	–	–	–	–	–	–	–0.03	–0.22	0.03	0.02	0.23	0.24
DIM	–	–	–	–	–	–	–	–	–	–	–	0.11	0.16	–0.39	–0.33	–0.33
BW	–	–	–	–	–	–	–	–	–	–	–	–	0.15	0.30	0.11	0.12
Age	–	–	–	–	–	–	–	–	–	–	–	–	–	–0.18	–0.12	–0.15

Table 3. Predictor variables (|*r*| <0.5) and number of fitted models, including all possible combinations of variables in each subpool when models were selected using schemes with and without dry matter intake (DMI)

DM = dietary dry matter (% of diet); CP, ADF, NDF, LIG, EE and ash = dietary crude protein, acid detergent fibre, neutral detergent fibre, lignin and crude fat content, respectively (% of DM); Milk = milk yield (kg/day); mPrt = milk protein percentage; mFat = milk fat percentage; DIM = days in milk; BW = bodyweight (kg); Age = age of the cows (years)

Subpool	Predictor variables	Number of models
	With DMI	
1	DMI, DM, CP, NDF, LIG, EE, DIM, mPrt, mFat, BW, Age	2048
2	DMI, DM, CP, ADF, EE, DIM, mPrt, mFat, BW, Age	1024
3	DMI, DM, Ash, NDF, LIG, EE, DIM, mPrt, mFat, BW, Age	2048
4	DMI, DM, Ash, ADF, EE, DIM, mPrt, mFat, BW, Age	1024
	Without DMI	
1	Milk, DM, CP, NDF, LIG, EE, mFat, BW, Age	512
2	Milk, DM, CP, ADF, EE, mFat , BW, Age	256
3	Milk, DM, Ash, NDF, LIG, EE, mFat, BW, Age	512
4	Milk, DM, Ash, ADF, EE, mFat, BW, Age	256
5	DIM, mPrt, DM, CP, NDF, LIG, EE, mFat, BW, Age	1024
6	DIM, mPrt, DM, CP, ADF, EE, mFat, BW, Age	512
7	DIM, mPrt, DM, Ash, NDF, LIG, EE, mFat, BW, Age	1024
8	DIM, mPrt, DM, Ash, ADF, EE, mFat, BW, Age	512

Table 4. Bayesian information criteria (BIC), log likelihood value (ℓ) of final models from two schemes, one using both DMI and relevant dietary nutrient concentration and the other using only intake of the corresponding nutrient

F_{NDF}, F_{ADF}, F_N, U_N = faecal neutral detergent fibre (NDF), acid detergent fibre (ADF) and nitrogen, and urinary nitrogen outputs (kg/day), respectively. CP, ADF, NDF and LIG = dietary crude protein, ADF, NDF and lignin contents, respectively (% of DM), mPrt = milk protein percentage, DIM = days in milk, BW = bodyweight (kg/cow), iN, iNDF and iADF = nitrogen, NDF and ADF intake (kg/day). *P*-values are from likelihood-ratio test

Response variable	Predictor variables	BIC	ℓ	*P*-value
F_{NDF}	DMI, CP, NDF	320	–137	<0.001
	iNDF, ADF	540	–264	–
F_{ADF}	DMI, ADF	–305	173	<0.001
	iADF	–247	140	–
F_N	DMI, CP, LIG, BW	–4089	2071	<0.001
	iN, ADF, mPrt	–3823	1935	–
U_N	DMI, CP, DIM, BW	–3354	1703	<0.001
	iN, DIM	–3276	1658	–

$$F_C = 0.461 \pm 0.002 \times F_{DM} \tag{19}$$

Faecal hemicellulose (F_{HC}) output (kg/day) was derived from estimated F_{NDF} (Eqn 3 and 12) and F_{ADF} (Eqns 4, 13) as:

$$F_{HC} = F_{NDF} - F_{ADF}. \tag{20}$$

Furthermore, assuming that lignin is 100% indigestible, faecal cellulose output (F_{CL}) was estimated using the difference between F_{ADF} and lignin intake (I_{Lignin}, kg/day).

$$F_{CL} = F_{ADF} - I_{Lignin}. \tag{21}$$

Total manure output (T_E, kg/day) was determined by summation of F_{Water}, F_{DM} and U_E.

$$T_E = F_{Water} + F_{DM} + U_E. \tag{22}$$

Output of total C (T_C, kg/day) and N (T_N, kg/day) in fresh manure were estimated by adding their faecal and urinary counterparts.

$$T_C = F_C + U_C, \tag{23}$$

$$T_N = F_N + U_N. \tag{24}$$

Total C to total N ratio in fresh manure ($R_{C:N}$) was calculated using T_C and T_N.

$$R_{C:N} = T_C / T_N. \tag{25}$$

Model evaluation

The models were evaluated with data not used in their construction (n = 364). Model adequacy statistics were calculated to determine sources of prediction error. The square root of mean square prediction error (RMSPE) is directly comparable to observed values so that RMSPE% was calculated and expressed as a percentage of the average observed values of the response variables. Mean square prediction error was further decomposed into mean bias, slope bias and non-systematic or random variability of data to give relative estimates of sources of error (Bibby and Toutenburg 1977). All the analyses in the present study were carried out with R (version 2.12.2, R Foundation for Statistical Computing, Vienna, Austria).

Table 5. Prediction equations (standard errors of parameters in parentheses), maximum variance inflation factor (max_VIF) and the determinant of the correlation matrix (*det*{R}) of the selected variables

F_{DM} = faecal dry matter, F_C = faecal carbon, F_{NDF} = faecal neutral detergent fibre (NDF), F_{ADF} = faecal acid detergent fibre (ADF), F_N = faecal nitrogen, F_{Water} = faecal water, U_E = total urine, U_C = urine carbon, and U_N = urine nitrogen outputs (all in kg/day). DMI = dry matter intake (kg/day), *CP*, *ADF*, *NDF* and *LIG* = dietary crude protein, ADF, NDF and lignin content, respectively (% of DM). Milk = milk yield (kg/day), mPrt = milk protein percentage, DIM = days in milk, BW = bodyweight (kg/cow), Age = age of the cows (years)

Equation	Variables and parameter estimates ± standard error	max_VIF	*det*{R}
	With DMI		
(1)	$F_{DM} = -0.576 \pm 0.222 + (0.370 \pm 0.006 \times DMI) + (-0.075 \pm 0.010 \times CP) + (0.059 \pm 0.006 \times ADF)$	1.02	0.90
(2)	$F_C = (0.169 \pm 0.003 \times DMI) + (-0.034 \pm 0.004 \times CP) + (0.027 \pm 0.003 \times ADF) + (-0.075 \pm 0.019 \times m\mathrm{Pr}\,t)$	1.04	0.87
(3)	$F_{NDF} = -0.864 \pm 0.172 + (0.217 \pm 0.004 \times DMI) + (0.035 \pm 0.003 \times NDF) + (-0.039 \pm 0.007 \times CP)$	1.01	0.86
(4)	$F_{ADF} = -1.272 \pm 0.084 + (0.125 \pm 0.003 \times DMI) + (0.061 \pm 0.003 \times ADF)$	1.00	0.99
(5)	$F_N = -0.0368 \pm 0.007 + (0.0096 \pm 0.000 \times DMI) + (0.0022 \pm 0.000 \times CP) + (0.0034 \pm 0.001 \times lignin) + (-0.000043 \pm 0.000010 \times BW)$	1.09	0.80
(6)	$F_{Water} = (1.987 \pm 0.034 \times DMI) + (0.348 \pm 0.032 \times ADF) + (-0.412 \pm 0.052 \times CP) + (-0.074 \pm 0.009 \times DM) + (-0.0057 \pm 0.0012 \times DIM)$	1.16	0.80
(7)	$U_E = -7.742 \pm 2.367 + (0.388 \pm 0.055 \times DMI) + (0.726 \pm 0.096 \times CP) + (2.066 \pm 0.421 \times m\mathrm{Pr}\,t$	1.05	0.94
(8)	$U_C = -0.1601 \pm 0.0169 + (0.0082 \pm 0.0005 \times DMI) + (0.0107 \pm 0.0008 \times CP) + (0.00013 \pm 0.00002 \times BW)$	1.13	0.84
(9)	$U_N = -0.2837 \pm 0.0135 + (0.0068 \pm 0.0004 \times DMI) + (0.0155 \pm 0.0006 \times CP) + (0.00013 \pm 0.00001 \times DIM) + (0.000092 \pm 0.000017 \times BW)$	1.34	0.66
	Without DMI		
(10)	$F_{DM} = 0.846 \pm 0.469 + (0.098 \pm 0.004 \times Milk) + (-0.097 \pm 0.021 \times CP) + (0.080 \pm 0.012 \times ADF) + (0.0038 \pm 0.0005 \times BW)$	1.04	0.80
(11)	$F_C = 0.468 \pm 0.232 + (0.046 \pm 0.002 \times Milk) + (-0.047 \pm 0.010 \times CP) + (0.037 \pm 0.006 \times ADF) + (0.0016 \pm 0.0002 \times BW)$	1.04	0.80
(12)	$F_{NDF} = (0.056 \pm 0.003 \times Milk) + (-0.059 \pm 0.010 \times CP) + (0.0435 \pm 0.0042 \times NDF) + (0.0023 \pm 0.0003 \times BW)$	1.00	0.77
(13)	$F_{ADF} = -0.973 \pm 0.152 + (0.0325 \pm 0.0016 \times Milk) + (0.0675 \pm 0.0043 \times ADF) + (0.0014 \pm 0.0002 \times BW)$	1.00	0.98
(14)	$F_N = (0.00245 \pm 0.00011 \times Milk) + (0.00643 \pm 0.00082 \times \mathrm{LIG}) + (0.000094 \pm 0.000009 \times BW)$	1.00	0.99
(15)	$F_{Water} = (0.559 \pm 0.025 \times Milk) + (0.521 \pm 0.060 \times ADF) + (0.569 \pm 0.100 \times CP) + (0.024 \pm 0.003 \times BW) + (-0.033 \pm 0.012 \times Age)$	1.17	0.66
(16)	$U_E = -0.644 \pm 0.226 + (0.778 \pm 0.099 \times CP) + (1.520 \pm 0.426 \times m\mathrm{Pr}\,t)$	1.00	0.99
(17)	$U_C = -0.1167 \pm 0.0201 + (0.0013 \pm 0.0002 \times Milk) + (0.0106 \pm 0.0009 \times CP) + (0.00024 \pm 0.00002 \times BW)$	1.03	0.87
(18)	$U_N = -0.2578 \pm 0.0183 + (0.0152 \pm 0.0007 \times CP) + (0.0132 \pm 0.0031 \times m\mathrm{Pr}\,t) + (0.00021 \pm 0.00002 \times BW)$	1.01	0.98

Results and discussion

Data

Lactating cows in the EMU database (Table 1) have milk production and composition records representative of an average dairy herd in North America (Wilkerson *et al.* 1997). Dietary characteristics are similar to those recommended by the NRC (Wilkerson *et al.* 1997). Overall the data show a wide range in both the predictor and response variables (Table 1), allowing the models to capture the relationships between them well. DMI and milk yield, the main predictors of fresh manure excretions (Nennich *et al.* 2005) and dietary nutrient composition have a CV ranging from 15% to 45%. Data on the response variables are even more variable as the CV is greater than 32% in all the cases (Table 1).

Total manure excretion (T_E) from the individual lactating Holstein cows was on average 46.7 kg/day and ranged from 16.9 to 98.5 kg/day (Table 1). Dry matter content (w/w) of fresh manure (faeces plus urine) varied from 5.0% to 19.0% with a mean of 11%. An average lactating Holstein cow excreted daily 2.59 kg of C and 0.29 kg of N, which resulted in a mean fresh manure C : N ratio of 9.50 (Table 1). Overall, 92% of total C in fresh manure came from faeces, whereas 55% of the total N was from urine. As rubber mats were used without any bedding material in the EMU calorimeters, the fresh manure output data can be directly compared with ASAE (2005) standards for prediction of faeces and total manure from lactating dairy cows. Such comparisons need to be adjusted for differences in at least DMI and milk yield between the EMU and ASAE (2005) datasets. For example, cows in EMU studies produced on average 21.6 kg/day milk and had a mean DMI of 15.6 kg/day, whereas the mean milk yield and DMI of the ASAE (2005) cows were 40 and 21 kg/day, respectively. As stated previously by Wilkerson *et al.* (1997), comparisons between the EMU data and data from farm trials should address procedural differences in two types of trials. For example, the urinary N excretion and moisture content measures in the EMU data are greater but more accurate than farm trial data because the EMU studies used experimental procedures to minimise N loss (Muck and Richards 1983) and moisture loss during collection (Wilkerson *et al.* 1997).

Faecal excretion predictions

Faecal dry matter

Biological interpretation of the magnitude of partial regression coefficient estimates is challenging (Alexopoulos 2010). Therefore, potential biological mechanisms associated only with the sign of model parameters are discussed. As expected, F_{DM} (kg/day) was strongly related (r = 0.94 and Fig. 2) to DMI. Regardless of the level of DMI, dietary CP and ADF (%

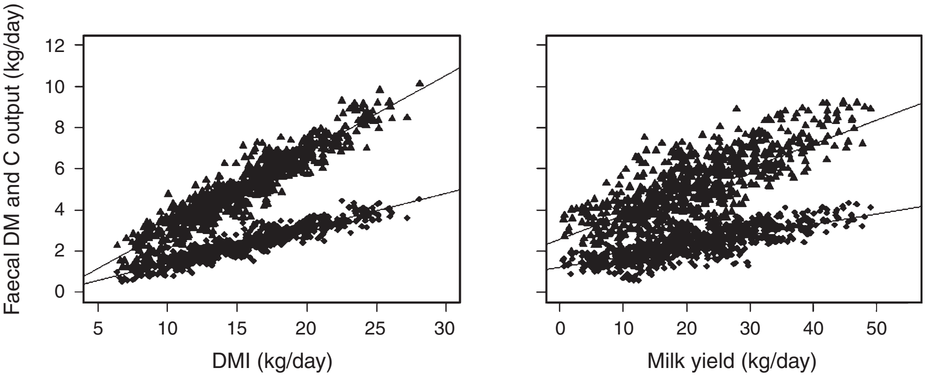

Fig. 2. Relationships of faecal dry matter (DM, ▲) and carbon (C, ◆) output with dry matter intake (DMI) and milk yield in lactating Holstein cows (n = 742).

Table 6. Root mean square prediction error as a percentage of average observed value (RMSPE%) and its decomposition into mean (MB), slope (SB), and random (RB) components, when the models with and without dry matter intake (DMI) were evaluated using data (n = 364) not used for model development

F_{DM} = faecal dry matter, F_C = faecal carbon, F_{NDF} = faecal neutral detergent fibre (NDF), F_{ADF} = faecal acid detergent fibre (ADF), F_N = faecal nitrogen, F_{Water} = faecal water, F_{HC} = faecal hemicellulose, F_{CL} = faecal cellulose, U_E = total urine, U_C = urine carbon, U_N = urine nitrogen, T_E = total fresh manure, T_C = total carbon and T_N = total N outputs (all in kg/day) and $R_{C:N}$ = C : N ratio in fresh manure

Variable	Equation	RMSPE%	MB	SB	RB	Equation	RMSPE%	MB	SB	RB
		With DMI					Without DMI			
				Primary predictions						
F_{DM}	1	11.4	0.2	4	95.8	10	21.2	1.1	2.2	96.7
F_C	2	11.1	0.8	1.5	97.7	11	20.5	0.8	1.1	98.1
F_{NDF}	3	13.7	3.9	0.2	95.9	12	21.5	1.2	1.2	97.6
F_{ADF}	4	13.8	8.3	5.5	86.2	13	20.1	1.9	1.2	96.9
F_N	5	10.8	6.9	0.4	98.7	14	18.3	2.2	0.3	97.5
F_{Water}	6	14.6	0.5	0.1	99.4	15	23.1	1.9	2.9	95.3
U_E	7	28.9	5.5	0.5	94	16	31.2	2.8	0.1	97.1
U_C	8	17.1	3.0	8.4	88.6	17	21.1	1.6	3.6	94.8
U_N	9	17.5	0.7	0.7	98.6	18	24.6	0.1	0.7	99.2
				Secondary predictions						
F_C	19	11.1	0.3	1.6	98.1	19	20.6	0.6	1.2	98.2
F_{HC}	20	19.2	0.9	5.6	93.5	20	29.1	3.5	4.2	92.3
F_{CL}	21	20.7	7.8	5.6	86.5	21	30.3	1.6	9.5	88.9
T_E	22	13	0.5	0.1	99.4	22	20.4	1.8	2.3	95.9
T_C	23	9.7	1.2	2.8	96	23	19.4	1.1	1.7	97.2
T_N	24	9.1	2	1.7	96.3	24	17.9	0	0.9	99.1
$R_{C:N}$	25	14.3	6.4	14.3	79.3	25	15.2	4.1	13	82.9

of DM) had a negative and a positive relationship with F_{DM}, respectively (Eqn 1). The positive association between dietary ADF and F_{DM} is not surprising as increasing dietary lignocellulose contents are generally related to decreasing total-tract DM digestibility in ruminants (Van Soest 1965). The negative association of dietary CP might be related partly to improved fibre digestion with increasing dietary CP levels as demonstrated by Broderick (2003). In the absence of DMI (Eqn 10), milk yield (kg/day) had a significant positive relationship with F_{DM} (Fig. 2). This association appeared to derive from a more basic relationship between milk yield and DMI (r = 0.78, Table 2). Prediction error of the model without DMI was double compared with the one using DMI (RMSPE% = 11.4% vs 21.2%, Table 6). However, in both cases, systematic bias was negligible (mean and slope bias <5% of total bias, Table 6), indicating the absence of noticeable under or over-prediction tendencies over the data range. Dietary CP and ADF contents continued to have the same relationships with F_{DM} even in the absence of DMI. Additionally, BW (kg/cow) had a positive effect on F_{DM} (Eqn 10). The trend was consistent with ASAE (2005) and Nennich *et al.* (2005) predictions, which showed positive effects of BW in the absence of DMI data. Again, this association may be related to more a basic relationship between BW and DMI (r = 0.33, Table 2).

Faecal carbon

In the model using DMI as input, dietary CP and ADF contents had significant impact on faecal C output (Eqn 2). Similarly, a meta-analysis by Nousiainen *et al.* (2009) and Shaver *et al.* (1988)

showed a positive effect of CP content and a negative effect of dietary ADF content on organic matter digestibility, respectively, in lactating dairy cows. Additionally, increasing milk protein percentage is related to decreasing F_C in Eqn 2. Because indigestible non-fibre carbohydrates (NFC) can also contribute significantly to faecal organic matter, this negative relationship between milk protein content and F_C may, in part, be linked to the positive association of milk protein concentration to digestible NFC in lactating dairy cows (Firkins *et al.* 2001; Cabrita *et al.* 2007). In the presence of DMI, F_C predictions were associated with an RMSPE% of 11.1% (Table 6). As in faecal DM output, faecal C output was predicted with RMSPE% of 20.5% even without including DMI. Nonetheless, F_C could be determined with similar RMSPE using the F_{DM} estimates from Eqns 1 and 10 assuming a constant faecal C concentration (0.461 w/w, Eqn 19).

Faecal neutral detergent fibre and acid detergent fibre

Faecal output of a nutrient is primarily a function of its intake and digestibility. Intake of a nutrient is usually estimated using DMI and the dietary nutrient content. However, besides contributing merely to intake, DMI can also regulate digestibility of the nutrient, e.g. through its impact on digesta passage rate. The dietary nutrient content itself may explain its own digestibility. For example, dietary NDF content was associated negatively (r = –0.47) with total-tract NDF digestibility in the EMU trials (data not shown). Hence, DMI and corresponding dietary nutrient content should be able to explain more of the variation in a faecal nutrient output than just the nutrient intake. Consistently, models incorporating DMI and the dietary nutrient content fitted better ($P < 0.001$, Table 4) than those incorporating the corresponding nutrient intake estimates with regard to all faecal nutrient outputs including faecal NDF (F_{NDF}) and ADF (F_{ADF}). Besides DMI and dietary NDF content, the final model for predicting F_{NDF} included a negative effect of dietary CP (Eqn 3). This is in agreement with the findings of Broderick (2003) and Colmenero and Broderick (2006) who demonstrated a positive association between dietary CP content and apparent total-tract NDF digestibility. As expected, the model using DMI was able to explain a greater amount of variability in F_{NDF} than the model not using DMI (RMSPE% = 13.7% vs 21.5%, Table 6). Regardless of DMI in the models, F_{NDF} predictions showed minor systematic bias (mean and slope bias <5% of total bias, Table 6).

Unlike F_{NDF}, F_{ADF} (kg/day) was unrelated to dietary CP content (% of DM), suggesting that increasing dietary N supplementation may primarily enhance hemicellulose degradation in the rumen. However, Colmenero and Broderick (2006) found a significant positive quadratic relationship between dietary CP and apparent total-tract ADF digestibility in lactating Holstein cows, although they were not linearly related. In the present study, polynomial effects of independent variables were not tested because of obvious multi-collinearity issues. To further investigate this observation by Colmenero and Broderick (2006), a quadratic form of dietary CP was included in the model and tested against the data; however, it was not significantly related (P = 0.059, data not shown). F_{ADF} was well predicted with models that included DMI (RMSPE% = 13.8%), although there were mean and slope biases (8.3% and 5.5% of total error, Table 6). As expected, in the absence of DMI, RMSPE% increased to 20.1% of the average observed value.

With known F_{NDF}, F_{ADF} and lignin intake (all in kg/day), faecal hemicellulose (F_{HC}) and cellulose (F_{CL}) outputs could be estimated using Eqns 23 and 24, respectively. Average daily excretion of F_{HC} and F_{CL} by lactating Holstein cows was estimated to be similar (1.16 and 1.17 kg/day respectively, Table 1). Given that Holstein cows make up ~90% of the 9.3 million total dairy cow population in the US (Dungan *et al.* 2012), approximately eight million metric tons of F_{HC} and F_{CL} are produced annually. This represents a large source of carbohydrate that can be broken down to simple sugars such as glucose, xylose, arabinose and galactose through various hydrolytic processes (Liao *et al.* 2005). Relative proportions of hemicellulose, cellulose and lignin in manure significantly affect the rate of hydrolysis of carbohydrates in manure (Liao *et al.* 2005). Therefore, F_{HC} and F_{CL} estimates would help optimise hydrolytic processes. Data in Hashimoto *et al.* (1981), Külling *et al.* (2002) and Hindrichsen *et al.* (2005) indicate that variability in CH_4 production from dairy manure slurry can be better represented when the hemicellulose and cellulose contents in manure are known. When DMI was used, F_{HC} and F_{CL} were predicted with RMSPE% value of 19.2% and 20.7%, respectively. In the absence of DMI as a prediction variable, the prediction error, particularly of F_{CL}, increased to 30.3%. However, over 88% of the errors were random (Table 6).

Faecal nitrogen

Two separate models were developed for determining faecal (F_N) and urinary N (U_N) outputs. Such a distinction may have significant impact on predicting NH_3 volatilisation from dairy barns because urinary urea N is the main source for NH_3 volatilisation, about a half of which can occur before manure reaches the storage facility (Muck and Richards 1983; Hristov *et al.* 2011). Separate estimates of F_N and U_N should also allow more accurate determination of manure N output reaching storage facilities. Dietary CP was positively related with F_N regardless of DMI level (Eqn 5). Dietary lignin content also had a positive relationship with F_N presumably due to its negative impact on forage CP digestibility (Lloyd *et al.* 1961). Bodyweight had a positive linear relationship with F_N (r = 0.22, data not shown). However, when adjusted for DMI, it was negatively related to BW (Eqn 5), indicating that less N was excreted in proportion to intake as BW increased. Consistently, the fraction of N intake excreted in faeces was negatively associated with BW (r = –0.20, data not shown). The equation using DMI predicted F_N with a RMSPE% of 10.8% with 98.7% of the error being random. In the absence of DMI, milk yield appeared to be the predominant predictor of F_N given its relatively strong positive correlation (r = 0.76, data not shown). Interestingly, the equation did not include either dietary CP content or milk protein content. Similarly, a model including only milk yield and a model including milk yield, milk true protein percentage, dietary CP content and BW exhibited similar goodness of fit when regressed against F_N data in ASAE (2005). As discussed above, F_N was positively related to BW when DMI was not taken into consideration. Even though DMI was not used, the F_N predictions from Eqn 14 had a RMSPE% of 18.3% with the majority (97.5%) of the error coming from random sources.

Faecal water output

About 90% of urine and faeces from lactating Holstein cows consists of water (Knowlton *et al.* 2010; Khelil-Arfa *et al.* 2012). Therefore, accuracy of faecal and urinary water estimates considerably affects total fresh manure volume estimates. Fresh manure volume and wash water [e.g. the estimates in Harner *et al.* (2013)] predominantly determine total manure volume, which is a critical factor in designing storage facilities. Manure volume is required to calculate manure nutrient concentrations, which directly affect chemical reactions releasing nutrients into the environment. Therefore, manure volume and associated nutrient concentrations are vital input variables for manure and soil models, e.g. Manure-DNDC (Li *et al.* 2012). Furthermore, manure volume plays a key role in land applications. With the traditional use of manure as fertiliser, for example, manure volume is important in deciding storage and transportation requirements. Furthermore, fresh manure water output estimates assist in quantifying the water footprint of dairy cows (Mekonnen and Hoekstra 2012).

Dry matter intake was strongly related to F_{Water} ($r = 0.95$, data not shown). This was consistent with the fact that DMI drives the two major water inputs to lactating dairy cows, drinking water (Holter and Urban 1992; Murphy 1992) and water in feed (Seo *et al.* 2007). When evaluated with data not used for model development, Eqn 6 predicted F_{Water} using DMI data well (RMSPE% = 14.6% with less than 1% of the error being systematic, Table 6). As expected, prediction error of F_{Water} increased if DMI was excluded as an input (RMSPE% = 23.1%, Table 6). The negative relationship between DIM and F_{Water} (Eqn 6) implies that cows in early lactation may excrete more water in faeces than cows in late lactation. Again, this may be related to increased free water intake in early lactating cows compared with late lactating cows because cows producing more milk tend to drink more water independently of DMI (Murphy *et al.* 1983; Holter and Urban 1992; Khelil-Arfa *et al.* 2012). Dietary DM content had a negative impact on F_{Water}. This is because dry feeds contribute less to water intake via feed, representing 17% of the total water intake by cows (Khelil-Arfa *et al.* 2012). Dietary CP content was also negatively related to F_{Water}. This is perhaps related to increased urinary water output caused by elevated blood urea concentrations as net water transfer from gut to blood responds positively to depletion in extracellular water volume due to diuresis or excessive respiratory and cutaneous water losses (Silanikove and Tadmore 1989). The positive effect of dietary ADF content on F_{Water} (Eqn 6) could be related to positive associations of dietary ADF content with saliva input to the rumen and fractional water passage rate from the rumen (Appuhamy *et al.* 2014).

Urinary excretion predictions

Urine output

The average urine output from lactating Holstein cows in the EMU experiments was 16.6 kg/day, which was less than the average 25 kg/day estimated in ASAE (2005). However, average DMI and milk yield of the ASAE (2005) cows were greater (21 and 40 kg/day, respectively) than those of the EMU cows (15.6 and 21.6 kg/day, respectively). Nonetheless, U_E was highly variable across the EMU trials (CV = 40%, Table 1). Regardless of whether DMI was used or not, the U_E predictions had similar and considerably larger RMSPE% (28.9–31.2%), a significant proportion of which ($\geq$94%) was random. In the absence of DMI, dietary CP content appears to mostly explain U_E variability. According to Colmenero and Broderick (2006), dietary CP might be able to explain U_E variation slightly better than N intake. However, U_E predictions could have been improved if dietary sodium and potassium data had been available because blood sodium and potassium significantly contribute to renal osmolality, which predominantly drives urine volume in dairy cows (Maltz and Silanikove 1996).

Urinary carbon output

Average urinary C excretion by lactating Holstein cows was 0.22 kg/day (Table 1). Urinary C output (U_C) contributed to only 8% of total fresh manure C output, while the rest was from faeces (F_C, Table 1). The majority of the urinary C comes from nitrogenous organic compounds present in urine. When calculated with respect to mean U_N and U_C (Table 1) using the average proportional contributions to U_N in Dijkstra *et al.* (2013), nitrogenous organic compounds in urine contribute ~70% of U_C. Although urea (C:N atomic ratio = 0.5) contributes on average 73.0% of U_N (Dijkstra *et al.* 2013), its contribution to U_C was 23%, whereas hippuric acid (C:N atomic ratio = 9), while contributing only 5.5% of U_N (Dijkstra *et al.* 2013), made up 31% of U_C. Equation 8 predicting U_C using DMI also includes a positive effect of dietary CP content. Consistently, dietary CP content increasing from 13.5% to 19.4% of DM had a significant positive linear relationship with urinary urea output, and tended to be associated with increasing urinary purine derivative output in Colmenero and Broderick (2006). Despite the slope bias (9.1% of total prediction error, Table 6), U_C predictions appear overall to be acceptable (RMSPE% = 17.1%) and the bulk of the prediction error (88.6%) was non-systematic. Prediction error increased by 23% in the absence of DMI (Table 6). This is in line with the proportional difference between simple linear associations of U_C to DMI and milk yield; ($r = 0.71$ and 0.57, respectively, data not shown).

Urinary nitrogen output

Urinary N (U_N, mean = 0.16 kg/day) made up on average 55% of total fresh manure N output (T_N, mean = 0.29 kg/day) and was more variable than F_N (CV = 48% vs 32%, Table 1). Consistently, in Weiss *et al.* (2009), U_N made up 56% of T_N and had a 3.5 times greater variance than that of F_N. Equation 9 using DMI predicted U_N with a RMSPE% of 17.5%, 98.6% of which was random (Table 6). Nitrogen intake, a product of DMI and dietary CP content, is usually associated positively with U_N (Dijkstra *et al.* 2013). However, dietary CP content appeared to contribute predominantly to this relationship as it alone explained more of the variability in U_N than DMI (log-likelihood = –3694 vs –3791, data not shown). Moreover, the fraction of N intake excreted in urine had a significant positive relationship with dietary CP content (Fig. 3), in agreement with Colmenero and

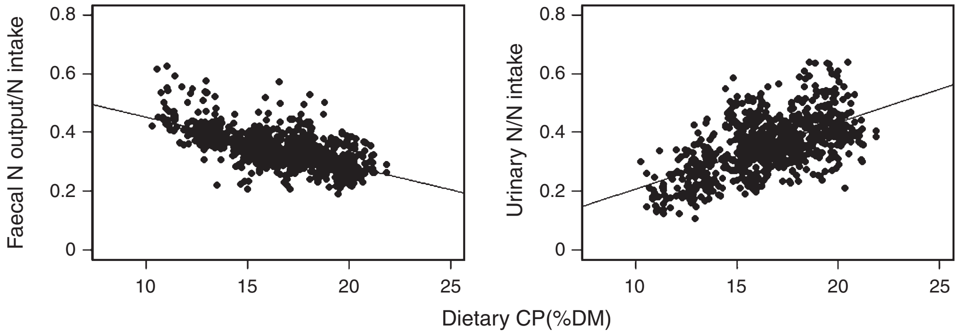

Fig. 3. Fraction of nitrogen intake excreted in faeces and urine (kg/kg) versus dietary crude protein content (% of dry matter) in lactating Holstein cows (n = 742).

Broderick (2006). As observed with F_N (Eqn 14), U_N had a positive linear relationship with BW (r = 0.35 and Eqn 18). Unlike that of F_N (Eqn 5), the relationship of U_N remained positive even when adjusted for DMI (Eqn 9). These differential associations of F_N and U_N with BW adjusted for DMI agree with the inverse relationship between the fraction of N intake excreted in urine and faeces (Fig. 3). The positive association between DIM and U_N apparently contradicts data in Knowlton *et al.* (2002) demonstrating early lactation cows excrete more urinary N than late lactating cows. However, when adjusted for N intake, the late lactating cows [in the control treatment of Knowlton *et al.* (2002)] excreted 11% more urinary N than the early lactating cows. In line with U_C, prediction error of U_N increased notably (RMSPE% = 17.5% vs 24.6%) in the absence of DMI. However, 99.2% of total prediction error was still random. Nonetheless, data allowing for more detailed characterisation of urinary N would yield benefits in terms of NH_3 volatilisation prediction because different urinary nitrogenous compounds alone or through interactions with others appear to be associated with different NH_3 volatilisation rates (Dijkstra *et al.* 2013).

Total fresh manure output and the C : N ratio

Total fresh manure output (T_E) ranged from 16.9 to 98.5 kg/day with a mean of 46.7 kg/day (Table 1), close to the average T_E of 52.9 kg/day reported in farm trials (Bulley and Holbek 1982). Manure volume estimates are required to calculate the nutrient concentrations, which govern primarily the rate of reactions responsible for manure nutrient transformations (Li *et al.* 2012). In the presence of DMI, the model (Eqns 1, 6, 7 and 22, collectively) predicted T_E well (RMSPE% = 13.0%). The mean and slope bias were negligible (0.6% of total prediction error) indicating that the model was able to determine T_E without notable under-prediction or over-prediction throughout the data range. Nonetheless, overall prediction error increased (RMSPE% = 20.4%) when DMI was not used, although the systematic bias remained small (4.1% of the total error). Similarly, total fresh manure C and N output (T_C and T_N, respectively), could be determined well using DMI. The overall prediction error was even less than 10% of the average observed values in both cases (Table 6). The T_C is determined in the Manure-DNDC model by taking the difference between total N intake and milk N output. In calculating the milk N output, 25% of N fed to dairy cows is assumed to be secreted in milk (Li *et al.* 2012). When predicted with respect to the EMU data using such an assumption, T_C were associated with a greater RMSPE (12.4%), 38% of which was systematic (data not shown). Therefore, incorporation of the present models into Manure-DNDC model should not only allow F_N and U_N to be determined distinctly but also allow T_C to be predicted more accurately using DMI data. Prediction error increased as expected when DMI data was not used. However, T_C and T_N were still associated with acceptable errors (RMSPE% = 19.4% and 17.9%, respectively), ≥96% of which was due to random variability of the data (Table 6). The T_C and T_N predictions then allowed determination of the C : N ratio in fresh manure ($R_{C:N}$, Eqn 25). Manure $R_{C:N}$ is important in estimating greenhouse gas emissions from a whole dairy farm system (Amon *et al.* 2006; Dijkstra *et al.* 2011). Moreover, manure $R_{C:N}$ is a major determinant of organic N mineralisation and thereby affect its fertiliser value. For instance, high $R_{C:N}$ has been shown to reduce the value of dairy manure as a N fertiliser (Powell *et al.* 2006).

Given the average milk yield of cows in the EMU trials was relatively low at 21.6 kg/day, applicability of the present models to high producing cows was tested using a subset of data (not used for model development) with daily milk yields >25 kg/day. This evaluation included 121 observations with milk yield ranging from 25.0 to 48.5 kg/day (mean = 30.9 kg/day). In the presence of DMI, F_{DM}, F_N, F_{NDF}, U_N, T_E, T_C and T_N predictions had RMSPE% values of 10.2%, 9.8%, 12.8%, 15.3%, 11.7%, 8.4% and 7.6%, respectively. Mean and slope bias were small as random bias accounted for 92.3–98.8 of total bias in all cases (data not shown). In the absence of DMI as a model input, the predictions had a respective RMSPE% of 19.0%, 17.1%, 19.9%, 21.3%, 16.7%, 16.8% and 15.8%. Only U_N of high producing cows was notably under-predicted (mean bias = 13.6% of total bias), which in turn caused T_N to be under-predicted in the absence of DMI (data not shown). All the other predictions were associated with negligible systematic bias (mean bias + slope bias <7.5% of total bias). Overall, the present models appear to predict manure composition and volume of Holstein cows successfully irrespective of their production level.

The main objective of the present study was to construct a set of equations for predicting important organic matter excretions from lactating dairy cows. Because the equations are predictive, the construction process had to ensure that they possess sound generalisability. The generalisability was established in two ways: (1) by obtaining robust parameter estimates of predictor variables while accounting for random study and animal variability and (2) by reducing model complexity by avoiding multi-collinearity and conducting model selection based not solely on goodness of fit but also on model simplicity (Myung 2000). Multi-collinearity among predictor variables in the final prediction equations was explored using maximum VIF and *det*{R} of the predictor variables (Table 5). Variance inflation factor of a predictor variable estimates how much the variance of a parameter estimate is inflated due to linear correlations of that variable with the other predictor variables. For example, VIF of 1.5 indicates that variance of the coefficient is 50% larger than it would be if the predictor variable in question was completely uncorrelated with the other predictors. However, smaller *det*{R} values indicate the presence of considerable multi-collinearity. Because *det*{R} ranges between 0 and 1, values close to 1 indicate absence of notable multi-collinearity. The maximum VIF was <1.35 and *det*{R} >0.65 in all cases (Table 5) indicating that the selected equations were fairly free of multi-collinearity. As the model parameters were estimated with respect to a considerably variable (CV = 28–48%) large dataset (*n* = 742) and the models use routinely collected data, they can be applied to various production systems (e.g. both confinement and grazing systems).

Conclusions

The models constructed in this study were capable of predicting fresh manure water, DM, C, N, NDF and ADF output from lactating Holstein cows accurately (RMSPE% = 10–21%) using DMI and other routinely collected information on dietary nutrient composition, milk yield and milk composition, BW and DIM. Even when DMI was not used as a regressor variable, the predictions (except for urine output) were satisfactory with RMSPE% ranging from 20% to 25%. The models also allow accurate determination of C : N ratio in fresh manure and hemicellulose and cellulose outputs, particularly when DMI is included (RMSPE% = 9–21%). In all cases, systematic bias of nutrient output predictions was less than 14% of total bias, indicating the predictions were reasonably close to observed values over the data range. The models can be integrated with process-based manure and soil models to assess dairy manure nutrient transformations in various dairy production systems.

Acknowledgements

The project was supported by the Agricultural Greenhouse Gas Program (Canada), the Sesnon Endowed Chairs Program (UC Davis) and Canada Research Chairs Program.

References

Akaike H (1974) A new look at the statistical model identification. *IEEE Transactions on Automatic Control* **19**, 716–723. doi:10.1109/TAC.1974.1100705

Alexopoulos EC (2010) Introduction to multivariate regression analysis. *Hippokratia* **14**, 23–28.

Amon T, Amon B, Kryvoruchk V, Bodiroza V, Potsch E, Zollitsch W (2006) Optimising methane yield from anaerobic digestion of manure: effects of dairy systems and of glycerine supplementation. *International Congress Series* **1293**, 217–220. doi:10.1016/j.ics.2006.03.007

Andrew SM, Tyrrell HF, Reynolds CK, Erdman RA (1991) Net energy for lactation of calcium salts of long-chain fatty acids for cows fed silage-based diets. *Journal of Dairy Science* **74**, 2588–2600. doi:10.3168/jds.S0022-0302(91)78437-3

Appuhamy JA, Wagner-Riddle C, Casper DP, France J, Kebreab E (2014) Quantifying body water kinetics and fecal and urinary water output from lactating holstein dairy cows. *Journal of Dairy Science*, in press. doi:10.3168/jds.2013-7755

ASAE (2005) ‘Manure production and characteristics. ASAE standards D384.2.’ (American Society of Agricultural and Biological Engineers: St Joseph, MI)

Bannink A, Valk H, Van Vuuren AM (1999) Intake and excretion of sodium, potassium, and nitrogen and the effects on urine production by lactating dairy cows. *Journal of Dairy Science* **82**, 1008–1018. doi:10.3168/jds.S0022-0302(99)75321-X

Bibby J, Toutenburg H (1977) ‘Prediction and improved estimation in linear models.’ (John Wiley & Sons: London)

Broderick GA (2003) Effects of varying dietary protein and energy levels on the production of lactating dairy cows. *Journal of Dairy Science* **86**, 1370–1381. doi:10.3168/jds.S0022-0302(03)73721-7

Bulley NR, Holbek N (1982) Nitrogen mass balance for dairy farms from feed to field. *Canadian Agricultural Engineering* **24**, 19–23.

Cabrita AR, Bessa RJ, Alves SP, Dewhurst RJ, Fonseca AJ (2007) Effects of dietary protein and starch on intake, milk production, and milk fatty acid profiles of dairy cows fed corn silage-based diets. *Journal of Dairy Science* **90**, 1429–1439. doi:10.3168/jds.S0022-0302(07)71628-4

Colmenero JJ, Broderick GA (2006) Effect of dietary crude protein concentration on milk production and nitrogen utilization in lactating dairy cows. *Journal of Dairy Science* **89**, 1704–1712. doi:10.3168/jds.S0022-0302(06)72238-X

Dijkstra J, Oenema O, Bannink A (2011) Dietary strategies to reducing n excretion from cattle: implications for methane emissions. *Current Opinion in Environmental Sustainability* **3**, 414–422. doi:10.1016/j.cosust.2011.07.008

Dijkstra J, Oenema O, van Groenigen JW, Spek JW, van Vuuren AM, Bannink A (2013) Diet effects on urine composition of cattle and N2O emissions. *Animal* **7**(Suppl. 2), 292–302. doi:10.1017/S1751731113000578

Dungan RS, Klein M, Leytem AB (2012) Quantification of bacterial indicators and zoonotic pathogens in dairy wastewater ponds. *Applied and Environmental Microbiology* **78**, 8089–8095. doi:10.1128/AEM.02470-12

Firkins JL, Eastridge ML, St-Pierre NR, Noftsger SM (2001) Effects of grain variability and processing on starch utilization by lactating dairy cattle. *Journal of Animal Science* **79**, 218–238.

Goering HK, Van Soest PJ (1970) ‘Forage fiber analysis (apparatus, reagents, procedures, and some applications). Agric. Handbook 379. (ARS-USDA: Washington, DC)

Graham MH (2003) Confronting multicollinearity in ecological multiple regression. *Ecology* **84**, 2809–2815. doi:10.1890/02-3114

Harner JP, Brouk MJ, Potts J, Bradford B, Smith JF (2013) Scientific data for developing water budgets on a dairy. Available at http://www.wdmc.org/2013 [Verified 29 April 2014]

Hashimoto AG, Varel VH, Chen YR (1981) Ultimate methane yield from beef cattle manure: effect of temperature, ration constituents, antibiotics and manure age. *Agricultural Wastes* **3**, 241–256. doi:10.1016/0141-4607(81)90011-1

Hindrichsen IK, Wettstein HR, Machmuller A, Jorg B, Kreuzer M (2005) Effect of the carbohydrate composition of feed concentrates on methane emission from dairy cows and their slurry. *Environmental Monitoring and Assessment* **107**, 329–350.

Holter JB, Urban WE (1992) Water partitioning and intake prediction in dry and lactating Holstein cows. *Journal of Dairy Science* **75**, 1472–1479. doi:10.3168/jds.S0022-0302(92)77904-1

Hristov AN, Hanigan M, Cole A, Todd R, McAllister TA, Ndegwa PM, Rotz A (2011) Review: ammonia emissions from dairy farms and beef feedlots. *Canadian Journal of Animal Science* **91**, 1–35. doi:10.4141/CJAS10034

Jonker JS, Kohn RA, Erdman RA (1998) Using milk urea nitrogen to predict nitrogen excretion and utilization efficiency in lactating dairy cows. *Journal of Dairy Science* **81**, 2681–2692. doi:10.3168/jds.S0022-0302(98)75825-4

Kauffman AJ, St-Pierre NR (2001) The relationship of milk urea nitrogen to urine nitrogen excretion in Holstein and Jersey cows. *Journal of Dairy Science* **84**, 2284–2294. doi:10.3168/jds.S0022-0302(01)74675-9

Khelil-Arfa H, Boudon A, Maxin G, Faverdin P (2012) Prediction of water intake and excretion flows in holstein dairy cows under thermoneutral conditions. *Animal* **6**, 1662–1676. doi:10.1017/S175173111200047X

Knowlton KF, McKinney JM, Cobb C (2002) Effect of a direct-fed fibrolytic enzyme formulation on nutrient intake, partitioning, and excretion in early and late lactation Holstein cows. *Journal of Dairy Science* **85**, 3328–3335. doi:10.3168/jds.S0022-0302(02)74421-4

Knowlton KF, Wilkerson VA, Casper DP, Mertens DR (2010) Manure nutrient excretion by Jersey and Holstein cows. *Journal of Dairy Science* **93**, 407–412. doi:10.3168/jds.2009-2617

Külling DR, Menzi H, Kröber TF, Neftel A, Sutter F, Lischer P, Kreuzer M (2001) Emissions of ammonia, nitrous oxide and methane from different types of dairy manure during storage as affected by dietary protein content. *The Journal of Agricultural Science* **137**, 235–250. doi:10.1017/S0021859601001186

Külling DR, Dohme F, Menzi H, Sutter F, Lischer P, Kreuzer M (2002) Methane emissions of differently fed dairy cows and corresponding methane and nitrogen emissions from their manure during storage. *Environmental Monitoring and Assessment* **79**, 129–150. doi:10.1023/A:1020248700255

Li C, Salas W, Zhang R, Rotz AI, Mitloehner F (2012) Manure-DNDC: a biogeochemical process model for quantifying greenhouse gas and ammonia emissions from livestock manure systems. *Nutrient Cycling in Agroecosystems* **93**, 163–200. doi:10.1007/s10705-012-9507-z

Liao W, Liu Y, Liu C, Chen S (2004) Optimizing dilute acid hydrolysis of hemicellulose in a nitrogen-rich cellulosic material–dairy manure. *Bioresource Technology* **94**, 33–41. doi:10.1016/j.biortech.2003.11.011

Liao W, Wen Z, Hurley S, Liu Y, Liu C, Chen S (2005) Effects of hemicellulose and lignin on enzymatic hydrolysis of cellulose from dairy manure. *Applied Biochemistry and Biotechnology* **124**, 1017–1030. doi:10.1385/ABAB:124:1-3:1017

Lloyd LE, Jeffers HFM, Donefer E, Crampton EW (1961) Effect of four maturity stages of timothy hay on its chemical composition, nutrient digestibility and nutritive value index. *Journal of Animal Science* **20**, 468–473.

Maltz E, Silanikove N (1996) Kidney function and nitrogen balance of high yielding dairy cows at the onset of lactation. *Journal of Dairy Science* **79**, 1621–1626. doi:10.3168/jds.S0022-0302(96)76525-6

Mason CH, Perreault WD (1991) Collinearity, power, and interpretation of multiple regression analysis. *JMR, Journal of Marketing Research* **28**, 268–280. doi:10.2307/3172863

Mekonnen M, Hoekstra A (2012) A global assessment of the water footprint of farm animal products. *Ecosystems* **15**, 401–415. doi:10.1007/s10021-011-9517-8

Moe PW, Flatt WP, Tyrell HF (1972) Net energy value of feeds for lactation. *Journal of Dairy Science* **55**, 945–958. doi:10.3168/jds.S0022-0302(72)85601-7

Moe PW, Tyrell HF, Hooven NW (1973) Energy balance measurements with corn meal and ground oats for lactating cows. *Journal of Dairy Science* **56**, 1149–1153. doi:10.3168/jds.S0022-0302(73)85325-1

Muck RE, Richards BK (1983) Losses of manurial nitrogen in free-stall barns. *Agricultural Wastes* **7**, 65–79. doi:10.1016/0141-4607(83)90005-7

Murphy MR (1992) Water metabolism of dairy cattle. *Journal of Dairy Science* **75**, 326–333. doi:10.3168/jds.S0022-0302(92)77768-6

Murphy MR, Davis CL, McCoy GC (1983) Factors affecting water consumption by Holstein cows in early lactation. *Journal of Dairy Science* **66**, 35–38. doi:10.3168/jds.S0022-0302(83)81750-0

Myung IJ (2000) The importance of complexity in model selection. *Journal of Mathematical Psychology* **44**, 190–204. doi:10.1006/jmps.1999.1283

Nennich TD, Harrison JH, VanWieringen LM, Meyer D, Heinrichs AJ, Weiss WP, St-Pierre NR, Kincaid RL, Davidson DL, Block E (2005) Prediction of manure and nutrient excretion from dairy cattle. *Journal of Dairy Science* **88**, 3721–3733. doi:10.3168/jds.S0022-0302(05)73058-7

Nousiainen J, Rinne M, Huhtanen P (2009) A meta-analysis of feed digestion in dairy cows. 1. The effects of forage and concentrate factors on total diet digestibility. *Journal of Dairy Science* **92**, 5019–5030. doi:10.3168/jds.2008-1833

Powell JM, Wattiaux MA, Broderick GA, Moreira VR, Casler MD (2006) Dairy diet impacts on fecal chemical properties and nitrogen cycling in soils. *Soil Science Society of America Journal* **70**, 786–794. doi:10.2136/sssaj2005.0286

Schievano A, Pognani M, D'Imporzano G, Adani F (2008) Predicting anaerobic biogasification potential of ingestates and digestates of a full-scale biogas plant using chemical and biological parameters. *Bioresource Technology* **99**, 8112–8117. doi:10.1016/j.biortech.2008.03.030

Schwarz G (1978) Estimating the dimension of a model. *Annals of Statistics* **6**, 461–464. doi:10.1214/aos/1176344136

Sechen SJ, Bauman DE, Tyrrell HF, Reynolds PJ (1989) Effect of somatotropin on kinetics of nonesterified fatty acids and partition of energy, carbon, and nitrogen in lactating dairy cows1,2. *Journal of Dairy Science* **72**, 59–67. doi:10.3168/jds.S0022-0302(89)79080-9

Seo S, Lanzas C, Tedeschi LO, Fox DG (2007) Development of a mechanistic model to represent the dynamics of liquid flow out of the rumen and to predict the rate of passage of liquid in dairy cattle. *Journal of Dairy Science* **90**, 840–855. doi:10.3168/jds.S0022-0302(07)71568-0

Shaver RD, Satter LD, Jorgensen NA (1988) Impact of forage fiber content on digestion and digesta passage in lactating dairy cows. *Journal of Dairy Science* **71**, 1556–1565. doi:10.3168/jds.S0022-0302(88)79719-2

Silanikove N, Tadmore A (1989) Rumen volume, saliva flow rate, and systemic fluid homeostasis in dehydrated cattle. *The American Journal of Physiology* **256**, 809–815.

Smith LW, Flatt WP, Barnes KA, Van Soest PJ (1965) A rapid method of carbon analysis for biological materials. *Journal of Association of Official Agricultural Chemists* **48**, 1261.

St-Pierre NR, Thraen CS (1999) Animal grouping strategies, sources of variation, and economic factors affecting nutrient balance on dairy farms. *Journal of Animal Science* **77**(Suppl. 2), 72–83.

Tyrrell HF, Moe PW (1974) Net energy value of a corn and a barley ration for lactation. *Journal of Dairy Science* **57**, 451–458. doi:10.3168/jds.S0022-0302(74)84913-1

Tyrrell HF, Brown AC, Reynolds PJ, Haaland GL, Bauman DE, Peel CJ, Steinhour WD (1988) Effect of bovine somatotropin on metabolism of lactating dairy cows: energy and nitrogen utilization as determined by respiration calorimetry. *The Journal of Nutrition* **118**, 1024–1030.

Van Soest PJ (1965) Symposium on factors influencing the voluntary intake of herbage by ruminants: voluntary intake in relation to chemical composition and digestibility. *Journal of Animal Science* **24**, 834–843.

Weiss WP, Willett LB, St-Pierre NR, Borger DC, McKelvey TR, Wyatt DJ (2009) Varying forage type, metabolizable protein concentration, and carbohydrate source affects manure excretion, manure ammonia, and nitrogen metabolism of dairy cows. *Journal of Dairy Science* **92**, 5607–5619. doi:10.3168/jds.2009-2248

Wilkerson VA, Mertens DR, Casper DP (1997) Prediction of excretion of manure and nitrogen by Holstein dairy cattle. *Journal of Dairy Science* **80**, 3193–3204. doi:10.3168/jds.S0022-0302(97)76292-1