ISSN 0189-0514
National Animal Production Research Institute
Ahmadu Bello University
P.M.B 1096, Shika-Zaria,
Kaduna State, Nigeria.
Email:japr@napri.gov.ngUrl:journals.napri.gov.ng
J. Anim. Prod. Res. (2018) 30(1):153-168

PREDICTION OF ENERGY BALANCE USING PRINCIPAL COMPONENT REGRESSION OF MILK COMPOSITION MEASURES

[*1]Alphonsus C., [2]Barje, P.P.,[2]Nwagu, B.I.,[3]Raji, A. O., 4Yashim, S. M., 4Mukasa, C., and [2]Achi, N.P.

[1]Animal Science Department, Kaduna state University, Kafanchan Campus, Nigeria
[2]National Animal Production Research Institute/Ahmadu Bello University, Shika-Zaria
[3]Department of Animal Science, University of Maiduguri, Borno state
4Department of Animal Science, Ahmadu Bello University, Zaria
*****Correspondent Author: cyprian.alphonsus@kasu.edu.ng

ABSTRACT

The purpose of this study was to investigate whether there is a stable relationship between energy balance (EB) and milk composition measures, and, how accurate could EB be predicted using group mean data. Energy balance was evaluated using data in which milk composition was measured at each milking from thirteen (13) primiparous and 47 multiparous (F_1) Friesian x Bunaji cows. The relationship between EB and milk measures was quantified by Principal Component Regression (PCR). The initial model included 17 combinations of milk composition measures allowing a range of nonlinear effects. These combinations were as follows: DMY, TMY, MFC, MPC, MLC, MFY, MPY, MLY, FPR, FLR, PLR, together with 10 milk composition change variables "diff()". These are the current minus the previous values of the milk measures in question {diff(my), diff(MFC), diff(MPC), diff(MLC), diff(mFy), dff(mPy), diff(mLy), diff(FPR), diff(FLR), diff(PLR)}.Using group mean data, a very high proportion of the variability (Adjusted R^2) in EB (over 99%)was explained by the PCR models. The prediction errors (RMSEP) for the PCR models were low, ranging from 7.54 to 18.3MJ/d. Therefore, using the PCR procedure that based only on milk composition measures, EB was predicted with high degree of reliability both within and across lactation stages. The potentials indicators of EB were the rate of change composite variable. Based on this observations, it is concluded that the mean EB of Friesian x Bunaji cows can be predicted throughout lactation using a common equation base on milk yield, milk fat and milk protein measures.

Key words: energy balance, milk composition, Friesian –Bunaji crosses

INTRODUCTION

Energy balance (EB) is a trait that is receiving more attention in dairy breeding programs as of late, this is probably because in breeding and genetic improvement program, it is desirable to understand the energy profile changes of animals under selection for milk yield, so as to identify animals that are genetically predisposed to retain energy and avoid lengthy intervals due to Negative Energy Balance (NEB) state. Furthermore, modeling and predicting the EB of a cow during its lifetime could become a useful management tool on-farm. This would assist in culling decisions, reproductive management, and other related practices (Banos*et al.*, 2005).

After parturition, almost every cow experiences a period with high energy requirement relative to milk production. Thus, it is normal for the cow to mobilize its body fat to offset the deficit during early lactation. As a consequence, marked changes in milk yield and milk composition are typical of these phases of lactation. Thus, possible relationship between energy balance and milk composition can be

postulated. It is therefore, hypothesized that changes in milk composition and ratios of milk components could help monitor EB status of dairy cows during lactation. One way of validating this hypothesis is to assess the relationship between the milk components ratios and energy balance at different stages of lactation. A clear understanding of this relationship in dairy cows would enable the development of low cost indicator of energy balance status of the cows.

Although the used of milk composition measures as a management tool to monitor the energy balance status of high yielding dairy cows has been established. However, the expected beneficial effect of its use as a selection criterion for EB status in moderate-low yielding dairy cows of the tropics has not been established.

Given the central role of energy balance in the health–fertility-production complex (Friggens, 2003; Ingvartsen*et al.,* 2003), being able to predict energy balance of cows using milk composition measures would be expected to be of great value to the dairy farmer. The purpose of this study was to find out whether there is a stable relationship between EB and milk composition measures, and how accurate EB could be predicted using principal component regression (PCR).

MATERIALS AND METHODS

Experimental site

The study was conducted on the dairy herd of the National Animal Production Research Institute (NAPRI) Shika, Nigeria, located between latitude 11^0 and 12^0N at an altitude of 640 m above sea level, and lies within the Northern Guinea Savannah Zone (Oni *et al.,* 2001). The mean annual rainfall in this zone is 1,100mm, which commenced from May and last till October, with 90% falling between June and September. (Malau–Aduli and Abubakar, 1992).

Animals and managements:

Data were collected from 13 primiparous and 47 multiparous (F1) Friesian x Bunaji cows from 2011 to 2013. The cows were raised during the rainy season on both natural and paddock–sown pasture, while hay or silage supplemented with concentrate of undelinted cotton seed cake, were offered during the dry season. They had access to water and salt lick *ad-libitum*. Unrestricted grazing was allowed under the supervision of herdsmen for about 7 – 9 hours per day. Routine spraying against ticks and other ecto-parasites was observed, while vaccination was carried out against endemic diseases.

Milk composition measures (energy balance indicators)

Cows were milked twice daily (morning and evening) and milk yield was recorded on daily basis. The milk sampled for the determination of fat, protein and lactose percentages was taken once per week starting from 4 days postpartum to the end of each lactation. The milk samples were frozen immediately after collection and stored at -20^{o}C until analyzed. The milk composition analysis was carried out at the Food Technology Laboratory of Institute of Agricultural Research, Ahmadu Bello University, Zaria, Nigeria. The yield values and the ratios were derived from the percentages values (Friggens*et al.,* 2007 Lφvendehl*et al.*,2010). The following milk composition measures were calculated: milk fat content (MFC), milk protein content (MPC),milk lactose content (MLC),Milk yield (MY), milk fat yield (MFY), milk protein yield (MPY), milk lactose yield (MLY),Fat-protein ratio (FPR), Fat-Lactose ratio (FLR), Protein - Lactose ratio (PLR), change in milk yield "diff"(MY), change in milk protein yield "diff"(mPy), change in milk fat yield "diff"(mFy), change in milk lactose yield "diff"(mLy), change in fat protein ratio "diff"(FPR), change in fat lactose ratio "diff"(FLR), change in protein-lactose ratio "diff"(PLR)

Calculation of Energy Balance using Milk composition traits

The energy balance (EB) was calculated weekly within 305 days postpartum from weekly measurements of milk production, milk fat, protein and lactose content and their ratios. Energy balance from milk measures was calculated using the equation described by Lφvendehl*et al*. (2010) as follows:

EBal Milk = 132.769 + 13.0675 x MFC – 140.304 x F/P – 95.1219 x diff (MY) -172.65 x diff (F/P) + 802.306 x diff(mPy)

Where: MFC = Milk fat content, F/P= fat: protein ratio, and MY = milk yield, together with 3 change "diff()" variables. These are the current minus the previous value of the milk measure in question: diff(MY), diff(F/P), and the changes in milk protein yield, diff(mPy).

Statistical Analysis and model development

The models for the estimation of energy balance was developed using Principal Component Regression (PCR) Technique. The Principal component Regression (PCR) was computed in 3 steps; first, principal component analysis (PCA) was run to verify whether the collinear predictors could be combined to form a composite score. Secondly, the component structure was verified by partial least square (PLS). Lastly composite index was used to run the ordinary least square (OLS) regression model.

Incomplete test-day records of less than 100days lactation length were excluded in the analyses. Thus, 36,120 test day milk yield records and 2580 cumulative milk composition records of 60 cows from week 1 to 43 post-partum were used for the analyses. Data were checked for outliers and the normality of the data was tested using ProcUnivariate Procedure of SAS (2000).

The high correlations observed amongst the predictor variables (milk composition measures) (Table1) necessitate the application of principal component analysis (PCA). Principal component analysis is a method for transforming the variables in a multivariate data set X_2, X_2,…….X_n, into new variables, Y_1, Y_2,……..Y_n, which are uncorrelated with each other and account for decreasing proportions of the total variance of the original variables, defined as

$Y_1 = P_{11}X_1 + P_{12}X_2 +\ldots\ldots\ldots\ldots\ldots +P_{1n}X_n$

$Y_2 = P_{21}X_1 + P_{22}X_2 + \ldots\ldots\ldots\ldots\ldots + P_{2n}X_n$

$Y_3 = P_{n1}X_1 + P_{n2}X_2 + \ldots\ldots\ldots\ldots\ldots + P_{nn}X_n$

With the coefficient being chosen so that;

Y_1, Y_2, ……..Y_n account for decreasing proportion of the total variance of the original variables X_1, X_2 …..X_n(Everit*et al.*,2001).

The principal component analysis was run using PROC Factor SAS software (SAS, 2000).

Partial Least Square (PLS) was employed to verify the suggestion made by principal component analysis (PCA). Thus, the PLS was run using PROC PLS procedure of SAS (2000) with cross validation "CV = one" and CV test (stat = PRESS). The proper number of components to be retained was determining by the predicted residual sum of square (PRESS). Taking both of the results yielded from PCA and PLS into account, it was obvious that some variables were measuring the same construct thus, could be combined as a single composite variable. Therefore, 7 principal components (PCs) were chosen and used for the ordinary least square (OLS) regression.

The relationship between the energy balance (EB) and the principal components (PCs) was quantified using Ordinary Least Square (OLS) regression procedure of SAS (2000). The OLS regression model with multiple explanatory variables was used for the prediction of EB. The form of the model used was:

$Y = \alpha + \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \beta_iX_i + e_i$

Where:

Y = dependent variable (energy balance),

α = intercept, β_1, β_2, β_3 and β_i = regression coefficients of explanatory or independent variables (principal components) X_1, X_2, X_3, X_i,

which indicates the average change in Y that is associated with a unit change in X, and e_i is the random error (Hutchesson, 2011).

In all the PCR model run during this study, the X- data (principal components) were mean centered and scaled to have equal variance, and the Y- data (EB) were also mean centered.

The model development was done using Mallows'sCp selection criteria of OLS regression to find models (in each stage of lactation) with 1, 2, 3,….,7 predictors (PCs) that maximizes the Adjusted- R^2 and minimizes the AIC and RMSEP.

Finally, model validation was carried out by cross-validation using 4 segments: the percentages of variance in EB explain by the PCR model. Cross- validation on the same data set was used to indicate the goodness of fit of the PCR model. The RMSEP from cross-validation on the same data set and from validation on the other stages data set was used to assess the predictive ability of the PCR models. The RMSEP can be interpreted as a proxy for the expected standard deviation of the EB, when using the PCR model to predict EB from similar milk composition data sets.

RESULTS AND DISCUSSION

The purpose of using PCA was to remove the observed multicollinearity amongst the predictor variables (milk composition measures) and to reduce the dimensionality without losing much of the information. Since PCs are labeled according to the size of their variances, the first principal component (PC) explains larger amount of variation amongst the variables, while the last PC explain the least. The first few PCs are normally the most informative ones. Therefore, the reduction of dimensionality can be achieved by choosing only the first few PCs to use for the regression analysis (Maitra and Yan 2008). In the present study, the PCA was ran using all the 17 milk composition variables and their derivatives. The first seven PCs which explained over 99% of the variations in the milk composition measures were chosen (Table 2). Since the PCs were labeled according to the size of their variance, the first PC explained the larger amount (38.88%) of the variation amongst the variables, while the 7^{th} PC explained the least (02.41%). The eigenvalues showed the same trend as the percentage variance explained by each principal component. The communality estimates (H) of the 17 variables was very high ranging from 99.40 to 99.96%.

There was a clear grouping of variables, evident by the loading pattern on each of the eigenvector yielded from the PCA of all the independent variables (the best loading of each variable is indicated by the bolded values; Table 2). It was obvious that the change variables (dmy, dmFy, dmPy, dmLy) loaded heavily on the first PC thus, were labeled "diff composite values", also the yield variables (ADMY, MFY, MPY and MLY) loaded heavily on the second PC and were labeled "yield composite values". Other variables like MFC, MLC and FPR loaded heavily on the third PC and were labeled "energy composite values". The change in fat ratios (dFPR, dFLR) loaded heavily on the fourth PC and were labeled "diff fat- ratio". The other variables like MPC, dPLR and FLR loaded each on the 5^{th}, 6^{th} and 7^{th}PCs, respectively and were therefore, treated individually as independent variables.

The Partial Least Square (PLS) was also employed to verify the suggestion made by the PCA. The result of the prediction residual error sum of squares (PRESS) statistics (Table 3) suggested that 11 components could be retained contrary to the 7 PCs suggested by PCA. To locate the optimal point, the number of principal components were plotted against the percent of variance explained, (Figure 1). Going from the first PC to the third PC resulted in a substantive gain, but as more and more components were added, the gain becomes less and less. When the number increases from 7 to 8, the two points form a

plateau to the 11th PC. Therefore, checking the variance explained in the model effect (Table 4) a seven-variable solution seems to be optimal hence, the 7 PC as suggested by the result of the PCA were chosen and used in the development of the PCR models. The PCR extracts factors to explain as much predictor sample variation as possible. The PCR, apart from minimizing sample response prediction error, seeks linear function of the predictors that explain as much variation in each response as possible. In this case, the 7 PCs that were chosen as the predictor variables were uncorrelated with each other, thus certifying the assumption of the independence of the predictor variables.

Table 5 showed the evaluation criteria and prediction variables (PCs) of the models fit for the prediction of EB. Generally, using the PCs of the group mean data, the percentage of variance in EB explained by the PCR models was very high (about 99%) for all the 7 PCR models developed for both within and across lactation stages. Likewise, the prediction errors of the models were relatively low (7.54 – 18.3MJ/d) this was within the results reported by Friggen*et al*. (2007). The high percentage of variance explained and the corresponding low prediction error obtained using the PCs clearly suggests the presence of a stable biological based relationship between EB and milk composition measures.

The single most informative PC for the prediction of EB was the rate of change composite variables ("diff" variables; model 1). This suggested that the rate of change variables might provide suitable indicators of EB in dairy cows.

The adjusted – R^2 values and the prediction error (RMSEP) across all lactation periods for model 1 was 99.01% and 16.01MJ/d, respectively, while within the lactation stages S_1, S_2 and S_3, the R^2adj and RMSEP were 99.07%, 99.33% and 98.77%, and 16.09, 12.78 and 18.38MJ/d, respectively.

The PCR models derived in one stage had relatively low or high prediction ability when applied to other lactation stage. For example, when model 1 derived in S_1 was applied to data on S_2 the AIC and RMSEP decreased; 180.27 and 12.78MJ/d, respectively, and the R^2adj increased (99.33%), which showed an improvement in the model. However, when the same model was applied to S_3 the AIC and RMSEP increased; 199.92 and 18.40Mj/d, respectively, while the R^2adj decreased 98.84%.

In contrast, the PCR model 1 when applied to the whole data set across lactation stages had a high predictive ability deviating from within stages RMSEP values by an average of 0.86MJ/d (range - 2.37 to + 3.23). Similar trends were observed in all the 7 models developed for all the lactation stages.

Table 1: Correlations amongst milk yield and milk composition variables used for prediction energy balance (EB)

[i]Milk measures	ADY	MFC	MPC	MLC	MFY	MPY	MLY	FPR	FLR	PLR	dmy	dmFy	dmPy	dmLy	dFPR	dFLR
ADMY	-															
MFC	-0.264	-														
MPC	-0.195	0.452	-													
MLC	-0.321	0.853	0.305	-												
MFY	0.956	0.025	-0.054	-0.078	-											
MPY	0.986	-0.189	-0.029	-0.275	0.966	-										
MLY	0.939	0.019	-0.089	0.014	0.988	0.943	-									
FPR	-0.177	0.853	-0.079	0.773	0.063	0.191	0.078	-								
FLR	0.183	0.044	0.169	-0.484	0.203	0.218	0.016	-0.048	-							
PLR	0.240	-0.669	0.162	-0.889	0.056	0.272	-0.057	-0.841	0.579	-						
Dmy	-0.669	0.037	0.232	-0.058	-0.681	-0.645	-0.728	-0.095	0.161	0.171	-					
DmFy	-0.674	-0.056	0.120	-0.085	-0.714	-0.668	-0.742	-0.133	0.056	0.144	0.980	-				
DmPy	-0.671	0.000	0.117	-0.068	-0.695	-0.666	-0.734	-0.118	0.118	0.123	0.985	0.989	-			
DmLy	-0.653	0.021	0.187	-0.084	-0.671	-0.634	-0.723	0.183	0.183	0.176	0.992	0.983	0.985	-		
dFPR	-0.182	-0.388	-0.002	-0.154	-0.279	-0.183	-0.220	-0.352	-0.352	-0.152	0.240	0.345	0.212	0.246	-	
Dflr	-0.129	-0.433	-0.336	0.017	-0.254	-0.196	-0.126	-0.284	-0.691	0.171	-0.021	0.142	0.063	-0.036	0.605	-
Dplr	0.070	-0.061	-0.391	0.147	0.038	-0.005	0.115	0.165	-0.385	-0.363	-0.297	-0.23	0.459	-0.321	-0.427	0.459

[i]milk composition variables indicated by the following: average daily milk yield (ADMY), milk fat content (MFC), milk protein content (MPC), milk lactose content (MLC), milk fat yield (MFY), milk protein yield (MPY), milk lactose yield(MLY), fat protein ratio (FPR), fat lactose ratio (FLR), protein lactose ratio (PLR). Variables abbreviations starting with "diff" are the current minus the previous values of milk measures in question. Yield values are in kilogram per day (kg/day), content values are in percentages (%) and ratios are unitless. The measures used were group mean averages.
[2]cummulative percentages of variation explained with increasing number of PC indicated

Table 2: The relationships amongst milk composition measures[1] expressed as loadings in a principal component analysis.

Items	Principal Components							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**H**
Variable explained[2]	38.88	60.01	75.00	85.30	92.38	97.46	99.86	-
ADMY	-0.34	**0.93**	-0.02	0.00	-0.00	-0.05	-0.00	99.81
MFC	0.02	0.13	**0.85**	-0.23	0.42	0.05	0.15	99.96
MPC	-0.04	0.15	0.05	0.07	**0.98**	-0.09	-0.03	99.96
MLC	0.05	0.13	**0.82**	0.08	0.37	0.08	-0.38	99.89
MFY	-0.33	**0.92**	0.17	-0.05	0.08	-0.03	0.03	99.88
MPY	-0.34	**0.93**	-0.01	-0.01	0.14	-0.06	-0.01	99.90
MLY	-0.34	**0.91**	0.19	0.01	0.07	-0.02	0.10	99.83
FPR	0.05	0.04	**0.90**	-0.27	-0.25	0.13	0.19	99.97
FLR	0.10	-0.04	-0.15	-0.48	-0.05	-0.07	**0.86**	99.97
PLR	0.03	-0.06	-0.86	-0.09	0.17	-0.18	0.43	99.86
diff'my	**0.94**	-0.33	0.01	-0.06	0.01	0.02	0.03	99.40
diff'mFY	**0.94**	-0.32	-0.02	0.06	-0.03	-0.02	0.02	99.77
diff'mPy	**0.94**	-0.32	0.02	-0.06	-0.04	0.07	0.04	88.81
diff'mLy	**0.95**	-0.31	-0.01	-0.07	-0.02	-0.07	0.05	99.77
diff'FPR	-0.04	0.03	-0.25	**0.81**	0.04	-0.49	-0.15	99.94
diff'FLR	-0.04	-0.05	-0.08	**0.92**	-0.03	0.29	-0.24	99.95
diff'PLR	-0.01	-0.09	0.22	0.02	-0.09	**0.96**	-0.08	99.96
% variance[3]	38.88	21.20	14.92	10.30	07.08	0.508	02.41	-
Eigenvalues	6.610	3.604	2.536	1.751	1.204	0.863	0.409	-

[1]milk composition variables indicated by the following: average daily milk yield (ADMY), milk fat content (MFC), milk protein content (MPC), milk lactose content (MLC), milk fat yield (MFY), milk protein yield (MPY), milk lactose yield(MLY), fat protein ratio (FPR), fat lactose ratio (FLR), protein lactose ratio (PLR). Variables abbreviations starting with "diff" are the current minus the previous values of milk measures in question. Yield values are in kilogram per day (kg/day), content values are in percentages (%) and ratios are unitless. The measures used were group mean averages.

[2]cummulative percentages of variation explained with increasing number of PC indicated

3 percentage variance explained by each principal components

H= communalities estimates is a variance in observed variables accounted for by a common factor

Table 3: Prediction residual sum of squares (PRESS) statistics

Number of extracted factors	Root mean PRESS	Probability > PRESS
0	1.0357	<0.0001
1	1.0145	<0.0001
2	1.0920	<0.0001
3	0.9080	<0.0001
4	0.7746	<0.0001
5	0.6369	<0.0001
6	0.4139	0.0820
7	0.0769	0.0820
8	0.0893	0.0100
9	0.0605	0.0400
10	0.0573	0.0510
11	0.0523	0.1950
12	0.0522	0.1690
13	0.0521	0.2470
14	0.0584	0.0180
15	0.0467	1.0000
Minimum root mean PRESS		0.0467
Minimizing number of factors		15
Smallest number of factor with P>0.1		11

PRESS = prediction residual error sum of squares

Table 4: percentage variation accounted for by principal components

Number of extracted factors	Model effects		Dependent variables	
	Current	Total	Current	Total
1	35.1086	35.1086	9.1798	9.1798
2	20.9954	56.1039	6.4167	15.5965
3	17.8714	73.9753	29.6695	45.2660
4	11.2851	85.2604	20.6239	65.8899
5	7.0656	92.3259	11.5431	77.4330
6	4.7109	97.0368	13.9746	91.4076
7	2.0268	99.0636	8.3683	99.7759
8	0.8926	99.9562	0.0020	99.7780
9	0.0212	99.9774	0.1201	99.8981
10	0.0097	99.9871	0.0191	99.9172
11	0.0064	99.9935	0.0142	99.9314

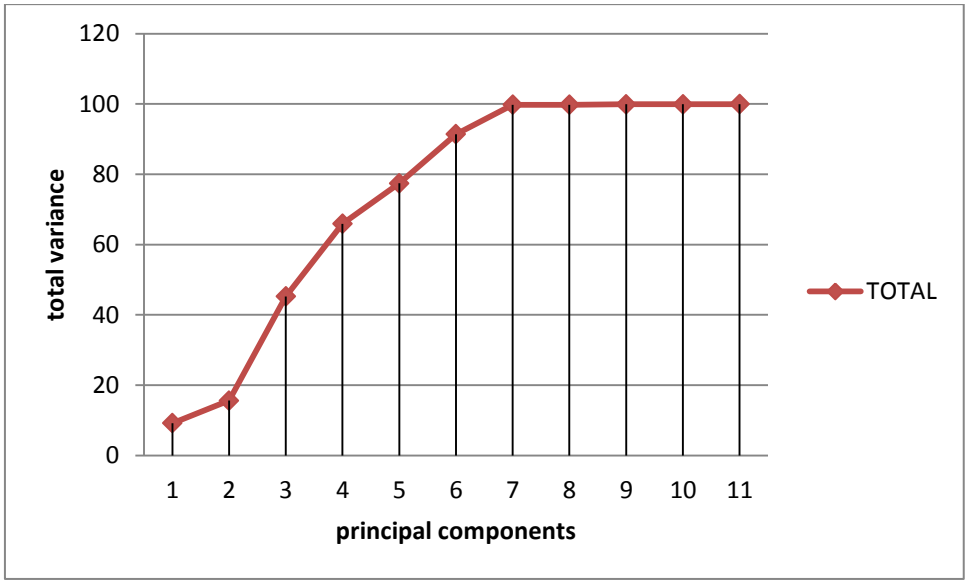


Figure 1. The number of principal components against the percent of variance explained

Table 5: Evaluation criteria and prediction variables for the prediction of EB using principal component regression (PCR)

Model	Evaluation criteria					Predictor variables (principal components)
	CP	R^2	Ra^2	AIC	RMSEP	
All stages						
1	11.53	99.02	99.01	578.76	16.01	diff" values
2	53.71	99.31	99.30	544.36	13.50	diff" values, diff" FR,
3	15.34	99.50	99.48	514.19	11.63	diff" values, diff"FR, diff" CP,
4	5.08	99.55	99.53	504.06	11.02	diff" values, diff"FR, diff" CP, energy
5	4.52	99.56	99.54	503.33	10.93	diff" values, diff"FR, diff" CP, energy, yield
6	6.15	99.56	99.54	504.94	10.97	diff"values, diff"FR, diff"CP, energy, yield, MPC
7	8.00	99.57	99.53	506.77	11.02	diff" values, diff"FR,diff" CP, energy, yield, MPC FLR
1st stage						
1	66.77	99.10	99.07	196.44	16.09	diff" values
2	65.34	99.13	99.07	197.19	16.05	diff" values, diff" FR,
3	19.99	99.57	99.52	174.78	11.51	diff" values, diff"FR, diff" PLR,
4	12.18	99.66	99.61	168.61	10.41	diff" values, diff"FR, diff"PLR, energy
5	4.71	99.74	99.70	160.32	9.14	diff" values, diff"FR, diff"PLR, energy, yield
6	6.27	99.75	99.69	161.75	9.23	diff"values, diff"FR, diff"PLR, energy, yield, MPC
7	8.00	99.75	99.69	163.41	9.35	diff" values, diff"FR,diff"PLR, energy, yield, MPC FLR
2nd stage						
1	63.66	99.35	99.33	180.27	12.78	diff" values
2	32.29	99.58	99.55	167.05	10.44	diff" values, diff" FR,
3	8.31	99.7	99.73	149.76	8.05	diff" values, diff"FR, diff" PLR,

		6				
4	8.28	99.77	99.74	149.69	7.95	diff' values, diff'FR, diff'PLR, energy
5	8.49	99.78	99.74	149.76	7.86	diff' values, diff'FR, diff'PLR, energy, yield
6	10.46	99.78	99.74	151.71	7.99	diff'values, diff'FR, diff'PLR, energy, yield, MPC
7	8.00	99.81	99.76	148.36	7.54	diff' values, diff'FR,diff'PLR, energy, yield, MPC FLR
3rd stage						
1	54.82	98.84	98.77	199.97	18.38	diff' values
2	4.29	99.55	99.52	169.14	11.53	diff' values, diff' FR,
3	3.15	99.59	99.55	167.66	11.14	diff' values, diff'FR, diff' PLR,
4	2.64	99.63	99.57	166.59	10.83	diff' values, diff'FR, diff'PLR, energy
5	4.01	99.63	99.57	167.77	10.89	diff' values, diff'FR, diff'PLR, energy, yield
6	6.00	99.63	99.55	169.77	11.09	diff'values, diff'FR, diff'PLR, energy, yield, MPC
7	8.00	99.63	99.54	171.77	11.30	diff' values, diff'FR,diff'PLR, energy, yield, MPC FLR

R^2 = coefficient of determination, R_a^2 = adjusted R^2, CP = Mallows's values, AIC = Akaike information criteria (small value is better), RMSEP = root mean square error of prediction. Variables starting with "diff" are the current minus the previous values of milk measures in question. FR= fat ratios, PLR = protein lactose ratio, FLR = fat lactose ratio, MPC = milk protein content

Table 6 : The regression coefficients and p-values of models for the prediction of energy balance

Models	Intercept	Regression Coefficient							P-value
		diff" values	diff"FR	diff" PLR	mixed	yield	MPC	FLR	
All stages									
1	48.377	-54.117							<0.0001
2	48.278	-54.407	-71.876						<0.0001
3	48.139	-54.480	-66.917	120.066					<0.0001
4	107.846	-54.584	-74.301	135.594	-6,483				<0.0001
5	104.119	-54.104	-73.581	140.955	-6.935	0.971			<0.0001
6	97.962	-54.104	-74.013	143.005	-7.284	0.960	2.256		0.0002
7	122.362	-54.075	-79.906	140.376	-7.906	0.963	2.632	-19.872	0.0088
1st Stage (P1)									
1	44.515	-54.984							<0.0001
2	44.439	-55.104	-29.297						<0.0001
3	43.508	-54.975	-33.909	218.818					<0.0001
4	118.399	-54.946	-50.987	236.445	-7.801				0.0001
5	135.173	-55.345	-53.922	234.840	-8.805	-0.794			0.0014
6	259.505	-55.078	-60.813	216.594	-5.502	-0.754	-37.664		<0.0001
7	221.049	-55.084	-52.044	226.382	4.557	-0.682	-38.344	31.260	0.0216
2nd stage (P2)									
1	51.37	-53.121							<0.0001
2	51.539	-53.132	-83.348						<0.0001
3	51.588	-53.278	-74.455	108.490					<0.0001

4	126.246	-53.402	-81.836	128.350	-8.139				0.0294
5	126.287	-52.825	-80.004	136.659	-9.254	1.207			0.0280
6	123.092	-52.817	-80.920	140.963	-10.331	1.243	2.952		0.0428
7	330.218	-52.657	-123.07	159.074	-24.148	1.488	9.587	-107.54	0.0067
3rd stage (P3)									
1	49.142	-54.161							<0.0001
2	49.061	-54.837	-88.167						<0.0001
3	49.066	-54.908	-85.139	56.795					<0.0001
4	58.917	-54.952	-86.019	59.409	-1.114				0.0933
5	58.921	-54.018	-85.093	72.961	-2.674	1.990			0.0831
6	68.315	-54.056	-81.443	60.319	0.482	1.998	-9.090		0.0977
7	71.718	-54.058	-82.161	60.019	0.429	1.991	-9.156	-2.561	0.3557

Variables starting with "diff" are the current minus the previous values of milk measures in question. FR= fat ratios, PLR = protein lactose ratio, FLR = fat lactose ratio, MPC = milk protein content

Table 7: Prediction error and goodness of fit for regression model relating energy balance (EB) to principal components (PCs) using group mean data

Model	Prediction error (RMSEP)				Explained variance (Adjusted -R^2)			
	All	S_1	S_2	S_3	All	S_1	S_2	S_3
1	16.01	16.09	12.78	18.39	99.01	99.07	99.33	98.77
2	13.50	11.87	10.44	11.53	99.30	99.07	99.55	99.52
3	11.63	11.51	8.05	11.14	99.48	99.52	99.73	99.55
4	11.02	10.41	7.95	10.83	99.55	99.61	99.74	99.57
5	10.93	10.52	7.86	10.89	99.54	99.60	99.74	99.57
6	10.97	9.23	7.99	11.09	99.54	99.69	99.74	99.55
7	11.02	9.35	7.54	11.30	99.53	99.69	99.76	99.54

S_1, S_2, S_3 = three stages of lactation

The regression coefficient and goodness of fit for the 7 energy balance models developed in each of the lactation stages is shown in Table 6 and 7. For the models developed using the whole data set across lactation stages, the first 5 models showed high prediction ability but beyond the 5th model the prediction ability decreases as more predictor variables (PCs) were added to the model. However, when these models that were developed from the whole data set were tested in each of the lactation stages, only the first 3 models in all the 3 lactation stages showed high prediction ability. Although the R^2adj increases as the CPs were been added up to the 4th model, but beyond the 4th model the R^2adj decreases as more PCs were been included in the model. Likewise, the prediction error (RMSEP) decreases with the addition of PCs up to the 4th model and then increases as more PCs were added.

The stability and reliability of the PCR model was not in doubt considering the evaluation criteria such as the P- values, R^2adj and the RMSEP. Although the entire 7 model developed had high R^2 and R^2adj values, only the first 3 models were highly significant ($P< 0.0001$). Also the F-statistics (not shown) indicates that the error sums of squares values were relatively smaller compared to the model sum of square. This suggests that the variation explain by the PCR models was not due to chance. Therefore, using the PCR procedure that based only on milk composition measures, EB was predicted with high degree of reliability both within and across lactation stages.

Based on the foregoing, it is concluded that the mean EB of Friesian x Bunaji cows can be predicted throughout lactation using a common equation based on milk yield, milk fat and milk protein measures.

REFERENCES

Banos, G. and Coffey, M. P. (2010) Genetic association between direct and indirect measures of body energy in dairy cows.*Proceedings of the 126th Annual Meeting of the British Society of Animal Science.* Belfast. April 12 to 14, 2010. p 3

Banos, G Coffey, M. P. and Brotherstone S. (2005) Modeling Daily Energy Balance of Dairy Cows in the First Three Lactations *J. Dairy Sci*. 88:2226–2237

Cammell, S. B., D. E. Beever, J. D. Sutton, J. France, G. Alderman, and D. J. Humphries. 2000. An examination of energy utilization in lactating dairy cows receiving a total mixed ration based on maize silage. Anim. Sci. 71:585–596.

Clark, C. E. F., Fulkerson, W. J., Nandra, K. S., Barchia, I. & Mac Millan,K. L. (2005) The use of indicators to assess the degree of mobilization of body reservesin dairy cows in early lactation on a pasture-based diet. *Livestock Production Science* 94,199-211

De Vries, M. J., and Veerkamp. R. F. (2000). Energy balance of dairy cattle in relation to milk production variables and fertility.*J. Dairy Sci*. 83:62–69.

Ellis, J. L., F. Qiao, and Cant, J. P (2006). Evaluation of net energy expenditures of dairy cows according to body weight changes over a full lactation. *Journal Dairy Science*, 89:1546–1557.

Everitt, B.S., Landau, S. Leese, M.: Cluster Analysis. 4th edn., Arnold Publisher, London. 2001.

Fox, D. G., Van Amburgh, M. E. &Tylutki, t. P. (1999) Predicting requirements for growth, maturity, and body reserves in dairy cattle. *Journal of Dairy Science*82, 1968-1977.

Friggens, N. C. (2003). Body lipid reserves and the reproductive cycle: Towards a better understanding. *Livestock Production Science.* 83:219–226.

Friggens N.C., Ridder C and Løvendahl, P (2007).On the use of milk composition measures to predict the energy balance of dairy cows.*Journal of Dairy Science*. 90:5453-5467

Huer C, Van Straalen W.M., Schukken Y.H., Dirkzwager A., Noordhuizen J.P.T.M (2000). Prediction of energy balance in a high yielding dairy herd in early lactation: model development and precision, *Livestock Production Science,* 65, 91-105.

Heuer, C., W. M. Van Straalen, Y. H. Schukken, A. Dirkzwager, and J. P. T. M. Noordhuizen.(2001). Prediction of energy balance in high yielding dairy cows with test-day information.*Journal Dairy Science*, 84:471–481

Hutcheson, G. D. (2011). Ordinary Least-Squares Regression. In L. Moutinho and G. D.Hutcheson, The SAGE *Dictionary of Quantitative Management Research*. P.224-228.

Ingvartsen, K. L., Dewhurst, R. J andFriggens, N. C. (2003). On the relationship between lactational performance and health: Is it yield or metabolic imbalance that cause production diseases in dairy cattle? A position paper.*Livestock Production Science*. 83:277–308.

Komaragiri MVS, Erdman R.A(1997).Factors affecting body tissue mobilization in early lactation dairy cows. 1. Effect of dietary protein on mobilization of body fat and protein. *Journal of Dairy Scienc,* 80:929–937

Konigsson. K, Savoini, G Govoni, N. Invernizzi, G Prandi, A. Kindahl, H andM. C. Veronesi (2008) Energy balance, leptin, NEFA and IGF-I plasma concentrations andresumption of postpartums ovarian activity in swedish red and white breed

cows. *ACTA VeterinariaScandinavica*50:(3) 1-7

Løvendahl, P., Ridder, C and Friggens N. C (2010). Limits to prediction of energy balance from milk composition measures at individual cow level.*J. Dairy Sci*. 93:1998–2006

Lucy, M. C. (2001) Reproductive loss in high-producing dairy cattle: where will it end? *Journal of Dairy Science* 84, 1277-1293

MalauAduli, A. E. O. and B. Y. Abubakar (1992). Estimation of 305 – day yield from milk yield in Bunaji and Friesian – Bunaji Crosses. *Nigeria Journal of Animal production* 19(182): 141-143.

Maitra.S. and Yan.J. (2008).Principle Component Analysis and Partial Least Squares:Two Dimension Reduction Techniques for Regression.*Principal Component Analysis and Partial Least Squares: Two Dimension Reduction Techniques for Regression*

Oni, O. O, Adeyinka, I. A., Afolayan, R. A., Nwagu, B. I., Malau-Aduli, A .E. O., Alawa, C. B .I and Lamidi, O. S (2001). Relationships between milk yield, post-partum body weight and reproductive performance in Friesian x Bunaji Cattle. *Asian–Australian Journal*, 14 (11):1516 – 1519.

Reist, M., Erdin, D., VonEuw. D., Tschuemperlin K., Leuenberger, H., Chilliard, Y., Hammon, H. M., Morel. C., Philipona. C., Zbinden Y., Kuenzi N., and Blum .J. W (2002).Estimation of Energy Balance at the Individual and Herd Level Using Blood and Milk Traits in High-Yielding Dairy Cows. *J. Dairy Sci.* 85:3314–3327

SAS (2000) SAS User's Guide Version 8.1. Statistical Analysis system institute Inc, Cary, Nc, USA.

Staples, C. R., W. W. Thatcher, and J. H. Clark.(1990). Relationship between ovarian activity and energy status during the early postpartum period of high producing dairy cows.*Journal Dairy Sci*ence, 73:938–947.