THE GLOBAL STANDARD
FOR LIVESTOCK DATA

Procedure 2 of Section 2 of ICAR Guidelines – Computing of Accumulated Lactation Yield

Computing Lactation Yield

Version January, 2020

Network. Guidelines. Certification.

Table of Contents

Tables

Equations

Figures

Change Summary

Date of Change	Nature of Change
July 17	Reformatted using new template.
August 17	Headings changed from Guideline A to Procedure 2.
August 17	Table numbers added with appropriate captions.
August 17	Equation numbers added to selected equations.
August 17	Table and Equation index added to Table of Contents.
August 17	Stopped Track change sand accepted all previous changes.
August 17	Moved the file to the new template (v2017_08_29).
August 17	Version updated to August 2017.
August 17	Equations not showing corrected.
October 2017	All cross-references were reconstructed since lost in the previous version.
January 2020	Edits at request of DCMR WG.

1 The Test Interval Method (TIM) (Sargent, 1968)

Test Interval Method is the reference method for calculating accumulated yields. Another adaptation of the method is the Centering Date Method where the yields from the preceding recording are used until the mid point of the recording interval and then substituted by the yields from the following recording.

The following equations are used to compute the lactation record for milk yield (MY), for fat (and protein) yield (FY), and for fat (and protein) percent (FP).

Equation 1. Cumulative yield calculation (ISLC).

$$MY = I_0M_1 + I_1 * \frac{(M_1 + M_2)}{2} + I_2 * \frac{(M_2 + M_3)}{2} + I_{n-1} * \frac{(M_{n-1} + M_n)}{2} + I_nM_n$$

$$FY = I_0F_1 + I_1 * \frac{(F_1 + F_2)}{2} + I_2 * \frac{(F_2 + F_3)}{2} + I_{n-1} * \frac{(F_{n-1} + F_n)}{2} + I_nF_n$$

$$FP = \frac{FY}{MY} * 100$$

Where:

M_1, M_2, M_n are the weights in kilograms, given to one decimal place, of the milk yielded in the 24 hours of the recording day.

F_1, F_2, F_n are the fat yields estimated by multiplying the milk yield and the fat percent (given to at least two decimal places) collected on the recording day.

I_1, I_2, I_{n-1} are the intervals, in days, between recording dates.

I_0 is the interval, in days, between the lactation period start date and the first recording date.

I_n is the interval, in days, between the last recording date and the end of the lactation period.

The equation applied for fat yield and percentage must be applied for any other milk components such as protein and lactose.

Details of how to apply the formulae are shown in Table 3 using the example data in Table 1, below.

Table 1. Raw data used in example (TIM).

Data:
Calving March 25

	Date of recording	Number of days	Quantity of milk weighed in kg	Fat percentage	Fat in grams
April	8	14	28.2	3.65	1 029
May	6	28	24.8	3.45	856
June	5	30	26.6	3.40	904
July	7	32	23.2	3.55	824
August	2	26	20.2	3.85	778
August	30	28	17.8	4.05	721
September	25	26	13.2	4.45	587
October	27	32	9.6	4.65	446
November	22	26	5.8	4.95	287
December	20	28	4.4	5.25	231

Table 2. Lactation period summary (TIM).

Beginning of lactation:	March 26
End of lactation:	January 3
Duration of lactation period:	284 days
Number of testings (weighings):	10

Table 3. Computations using Test Interval Method.

Interval both days included			Days	Daily production		Sum	
				Kg milk	Grams of fat	Kg milk	Kg fat
Mar 26	-	Apr 8	14	28.2	1 029	395	14.410
Apr 9	-	May 6	28	(28.2+24.8)/2	(1 029+856) /2	742	26.389
May 7	-	June 5	30	(24.8+26.6) /2	(856+904) /2	771	26.400
June 6	-	July 7	32	(26.6+23.2) /2	(904+824) /2	797	27.648
July 8	-	Aug. 2	26	(23.2+20.2) /2	(824+778) /2	564	20.817
Aug. 3	-	Aug 30	28	(20.2+17.8) /2	(778+721) /2	532	20.980
Aug 31	-	Sept. 25	26	(17.8+13.2) /2	(721+587) /2	403	17.008
Sept. 26	-	Oct. 27	32	(13.2+9.6) /2	(587+446) /2	365	16.541
Oct. 28	-	Nov. 22	26	(9.6+5.8) /2	(446+287) /2	200	9.536
Nov. 23	-	Dec. 20	28	(5.8+4.4) /2	(287+231) /2	143	7.253
Dec. 21	-	Jan. 3	14	4.4	231	62	3.234
			284			4973	190.216

Total quantity of milk: 4 973. kg
Total quantity of fat: 190 kg
Average fat percentage (190.216 / 4973) x 100 = 3.82%

2 Interpolation using Standard Lactation Curves (ISLC) (Wilmink, 1987)

With the method 'Interpolation using Standard Lactation Curves' missing test day yields and 305 day projections are predicted. The method makes use of separate standard lactation curves representing the expected course of the lactation, for a certain herd production level, age at calving and season of calving and yield trait. By interpolation using standard lactation curves, the fact that after calving milk yield generally increases and subsequently decreases is taken into account. The daily yields are predicted for fixed days of the lactation: day 0, 10, 30, 50 etc.

The cumulative yield is calculated as follows in :

Equation 2. Cumulative yield calculation (ISLC).

$$\sum_{i=1}^{n}\left[(INT_i - 1) * y_i + (INT_i + 1) * y_{i+1}\right]/2$$

where:

y_i = the i-th daily yield;

INT_i = the interval in days between the daily yields y_i and y_{i+1};

n = total number of daily yields (measured daily yields and predicted daily yields).

The next example illustrates the calculation of a record in progress. The cow was tested at day 35 and day 65 of the lactation. To determine the lactation yield, daily milk yields are determined for day 0, 10, 30 and 50 of the lactation, by means of the standard lactation curves. The daily yields are in Table 4.

Table 4. Measured and derived daily yields, used to calculate the record in progress in the example (ISLC).

Day of lactation	Milk (kg)	Note
0	25.9	Predicted
10	27.8	Predicted
30	31.7	Predicted
35	31.8	Measured
50	32.9	Interpolated using standard lactation curve
65	33.0	Measured

Next, the record in progress can be calculated by means of the formula for a cumulative yield as follows:

[(10 - 1) * 25.9 + (10+1) * 27.8] / 2 +

[(20 - 1) * 27.8 + (20+1) * 31.7] / 2 +

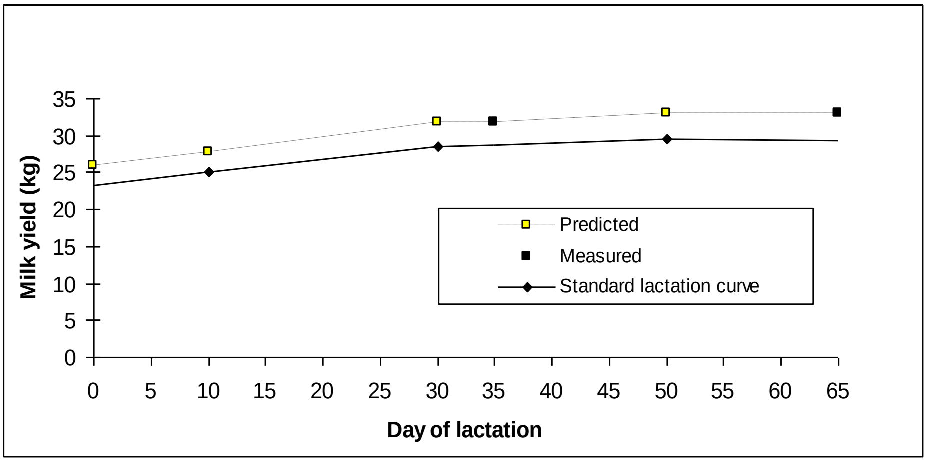
[(5 - 1) * 31.7 + (5+1) * 31.8] / 2 +

[(15 - 1) * 31.8 + (15+1) * 32.9] / 2 +

[(15 - 1) * 32.9 + (15+1) * 33.0] / 2 = 2005.3 kg.

This corresponds to the surface below the line through the predicted and measured daily yields (see Figure 1).

Figure 1. Example of calculation of record in progress..



3 Best prediction (BP) (VanRaden, 1997)

Recorded milk weights are combined into a lactation record using standard selection index methods. Let vector y contain M_1, M_2, to M_n and let E(**y**) contain corresponding the expected values for each recorded day. The E(y) are obtained from standard lactation curves for the population or for the herd and should account for the cow's age and other environmental factors such as season, milking frequency, etc. The yields in **y** covary as a function of the recording interval between them (I). Diagonal elements in Var(y) are the population or herd variance for that recording day and off diagonals are obtained from autoregressive or similar functions such as Corr(M_1, M_2)=0.995^I for first lactations or 0.992^I for later lactations. Covariances of one observation with the lactation yield, for example Cov(M_1, MY), are the sum of 305 individual covariances. E(MY) is the sum of 305 daily expected values. Lactation milk yield is then predicted as Equation 3:

Equation 3. Lactation yield prediction (BP).

$$MY = E(MY) + Cov(y, MY)' \; Var(y)^{-1} \, [\boldsymbol{y} - E(\boldsymbol{y})]$$

With best prediction, predicted milk yields have less variance than true milk yields. With TIM, estimated yields have more variance than true yields. The reason is that predicted yields are regressed toward the mean unless all 305 daily yields are observed. With best prediction, the predicted MY for a lactation without any observed yields is E(MY) which is the population or herd mean for a cow of that age and season. With TIM, the estimated MY is undefined if no daily yields are recorded.

Milk, fat, and protein yields can be processed separately using single-trait best prediction or jointly using multi-trait best prediction. Replacement of M_1, M_2, to M_n with F_1, F_2, to F_n or P_1, P_2, to P_n gives the single-trait predictions for fat or for protein. Multi-trait predictions require larger vectors and matrices but similar algebra. Products of trait correlations and autoregressive correlations, for example, may provide the needed covariances.

4 Multiple-Trait Procedure (MTP) (Schaeffer and Jamrozik, 1996)

The Multiple-Trait Procedure predicts 305-d lactation yields for milk, fat, protein and SCS, incorporating information about standard lactation curves and covariances between milk, fat, and protein yields and SCS. Test day yields are weighted by their relative variances, and standard lactation curves of cows of similar breed, region, lactation number, age, and season of calving are used in the estimation of lactation curve parameters for each cow. The multiple-trait procedure can handle long intervals between test days, test days with milk only recorded, and can make 305-d predictions on the basis of just one test day record per cow. The procedure also lends itself to the calculation of peak yield, day of peak yield, yield persistency, and expected test-day yields, which could be useful management tools for a producer on a milk recording program.

The MTP method is based upon Wilmink's model in conjunction with an approach incorporating standard curve parameters for cows with the same production characteristics. Wilmink's function for one trait is given by Equation 4.

Equation 4. Wilmink function for one trait (MTP).

$$y = A + Bt \pm C exp\,(-0.05t) + e$$

where y is yield on day t of lactation, A, B, and C are related to the shape of the lactation curve.

The parameters A, B, and C need to be estimated for each yield trait. The yield traits have high phenotypic correlations, and MTP would incorporate these correlations. Use of MTP would allow for the prediction of yields even if data were not available on each test day for a cow.

The vector of parameters to be estimated for one cow are designated:

$$\hat{c} = \begin{vmatrix} A_M \\ B_M \\ C_M \\ A_F \\ B_F \\ C_F \\ A_P \\ B_P \\ C_P \\ A_S \\ B_S \\ C_S \end{vmatrix}$$

where M, F, and P represent milk, fat, and protein, respectively, and S represents somatic cell score. The vector **c** is to be estimated from the available test-day records. Let $\mathbf{c_o}$ represent the corresponding parameters estimated across all cows with the same production characteristics as the cow in question.

Let

$$y_k = \begin{vmatrix} M_k \\ F_k \\ P_k \\ S_k \end{vmatrix}$$

be the vector of yield traits and somatic cell scores on test k at day t of the lactation.

The incidence matrix, X_k, is constructed as follows:

$$X'_k = \begin{vmatrix} 1 & 0 & 0 & 0 \\ t & 0 & 0 & 0 \\ \exp(-0.05t) & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & t & 0 & 0 \\ 0 & \exp(-0.05t) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & t & 0 \\ 0 & 0 & \exp(-0.05t) & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & t \\ 0 & 0 & 0 & \exp(-0.05t) \end{vmatrix}$$

The MTP equations are:

Equation 5. MTP equations.

$$\mathbf{(X'R^{-1}X + G^{-1})\,\hat{c} = X'R^{-1}y + G^{-1}c_o}$$

where

$$X'R^{-1}X = \sum_{k=1}^{n} X_k^{'} R_k^{-1} X_k$$

and

$$X'R^{-1}y = \sum_{k=1}^{n} X_k^{'} R_k^{-1} y_k$$

and n is the number of tests for that cow. R_k is a matrix of order 4 that contains the variances and covariances among the yields on k^{th} test at day t of lactation. The elements of this matrix were derived from regression formulas based on fitting phenotypic variances and covariances of yields to models with t and t^2 as covariables. Thus, element i_j of R_k would be determined by

$$r_{ij}(t) = ß_{oij} + ß_{1ij}\,(t) + ß_{2ij}\,(t^2)$$

G is a 12 x 12 matrix containing variances and covariances among the parameters in $\hat{\mathbf{c}}$ and represents the cow to cow variation in these parameters, which includes genetic and permanent environmental effects, but ignores genetic covariances between cows. The parameters for $\boldsymbol{G}$ and $\boldsymbol{R_k}$ vary depending on the breed, but must be known. Initially, these matrices were allowed to vary by region of Canada in addition to breed, but this meant that there could exist two cows with identical production records on the same days in milk, but because one cow was in one region and the other cow was in another region, then the accuracy of their predictions would be different. This was considered to be too confusing for dairy producers, so that regional differences in variance-covariance matrices were ignored and one set of parameters would be used for all regions for a particular breed. Estimation of G is described later.

If a cow has a test, but only milk yield is reported, then

$$y'_k(Mk \quad 0 \quad 0 \quad 0)$$

and

$$R_k = \begin{vmatrix} r_{MM}^{(t)} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{vmatrix}$$

The inverse of $\boldsymbol{R_k}$ is the regular inverse of the nonzero submatrix within $\boldsymbol{R_k}$, ignoring the zero rows and columns. Thus, missing yields can be accommodated in MTP.

Accuracy of predicted 305-d lactation totals depends on the number of test-day records during the lactation and DIM associated with each test. Thus, any prediction procedure will require reliability figures to be reported with all predictions, especially if fewer tests at very irregular intervals are going to be frequent in milk recording. At the moment, an approximate procedure is applied that uses the inverse elements of $(\mathbf{X'R^{-1}X + G^{-1}})^{-1}$.

4.1 Example calculations

Four test day records on a 25 month old, Holstein cow calving in June from Ontario are given in the Table 5 below.

Table 5. Example test day data for a cow (MTP).

Test no.	DIM=t	Exp(-0.05t)	Milk (kg)	Fat (kg)	Protein (kg)	SCS
1	15	0.47237	28.8			3.130
2	54	0.06721	29.2	1.12	0.87	2.463
3	188	0.000083	23.7	0.97	0.78	2.157
4	250	0.0000037	20.8			2.619

Notice that two tests do not have fat and protein yields, and that intervals between tests are irregular and large. The vector of standard curve parameters based on all available comparable cow, is

$$\hat{\mathbf{c}} = \begin{vmatrix} 27.533957 \\ -0.024306 \\ -2.996587 \\ 0.874776 \\ -0.000044 \\ 0.172253 \\ 0.801297 \\ -0.000208 \\ -0.109917 \\ 2.042824 \\ 0.001917 \\ 0.997263 \end{vmatrix}$$

The R^(-1)_k matrices for each test day need to be constructed. These matrices are derived from regression equations. The equations for Holsteins were:

$$r_{MM}(t) = 71.0752 - 0.281201t + 0.0004977t^2$$
$$r_{MF}(t) = 2.4365 - 0.013274t + 0.0000302t^2$$
$$r_{MP}(t) = 2.0504 - 0.008286t + 0.0000163t^2$$
$$r_{MS}(t) = -1.7993 + 0.013209t - 0.000056t^2$$
$$r_{FF}(t) = 0.1312 - 0.000725t + 0.000001586t^2$$
$$r_{FP}(t) = 0.0739 - 0.000386t + 0.000000926t^2$$
$$r_{FS}(t) = - 0.0386 + 0.000292t - 0.000001796t^2$$
$$r_{PP}(t) = 0.066 - 0.000267t + 0.0000005636t^2$$
$$r_{PS}(t) = - 0.0404 + 0.000369t - 0.000001743t^2$$
$$r_{SS}(t) = 3.0404 - 0.000083t - 0.000006105t^2$$

The inverses of the residual variance-covariance matrices for yields for the four test days are as follows:

$$\mathbf{R\text{^}(-1)\_1} = \begin{vmatrix} 0.0151259 & 0 & 0 & 0.0080354 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0.0080354 & 0 & 0 & 0.3334553 \end{vmatrix}$$

$$\mathbf{R\text{^}(-1)\_2} = \begin{vmatrix} 0.1685584 & 0.345947 & -4.851935 & 0.0254775 \\ -0.345947 & 26.830915 & -17.40281 & -0.041445 \\ -4.851935 & -17.40281 & 187.18579 & -0.584885 \\ 0.0254775 & -0.041445 & -0.584885 & 0.3365425 \end{vmatrix}$$

$$\mathbf{R\text{^}(-1)\_3} = \begin{vmatrix} 0.2620161 & 0.1479068 & -7.943903 & 0.0316069 \\ 0.1479068 & 54.446977 & -56.01333 & 0.3306741 \\ -7.943903 & -56.01333 & 317.9609 & -0.92601 \\ 0.0316069 & 0.3306741 & -0.92601 & 0.3654369 \end{vmatrix}$$

$$\mathbf{R\text{^}(-1)\_4} = \begin{vmatrix} 0.0329465 & 0 & 0 & 0.0251039 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0.0251039 & 0 & 0 & 0.3981981 \end{vmatrix}$$

Inverse matrix G^(-1) of order 12 is the same for all cows of the same breed:

Upper Left 6 x 6

0.1071767	0	0	-0.136926	0	0
	7715.8655	0	0	-17488.1	0
		0.0081987	0	0	-0.01643
			23.298605	0	0
				1758757	0
					1.235102

Upper Right 6 x 6

-3.277253	0	0	0.0159958	0	0
0	-216515.9	0	0	2036.696	0
0	0	-0.2011	0	0	0.002507
-18.22036	0	0	0.1014891	0	0
0	-1261220	0	0	8337.366	0
0	0	-0.585594	0	0	-0.00712

and the **Lower Right** 6 x 6

135.02387	0	0	-0.425106	0	0
	8761798.3	0	0	-29796.5	0
		7.648804	0	0	-0.04183
			0.2667083	0	0
				18672.44	0
					0.021171

Note that many covariances between different parameters of the lactation curves have been set to zero. When all covariances were included, the prediction errors for individual cows were very large, possibly because the covariances were highly correlated to each other within and between traits. Including only covariances between the same parameter among traits gave much smaller prediction errors.

The elements of the MTP equations of order 12 for this cow are shown in partitioned format also:

X'R⁻¹X =

Upper Left 6 x 6

0.4786468	66.824686	0.0184957	-0.19804	9.125325	-0.023239
	11814.771	0.7230498	9.125325	4218.8356	-1.253252
		0.0041365	-0.023239	-1.253252	-0.001563
			81.277893	11684.901	1.8078249
				2002612.9	98.228104
					0.1212006

Upper Right 6 x 6

-12.79584	-1755.458	-0.326758	0.0902237	13.714392	0.0055107
-1755.458	-294917.5	-17.73328	13.714392	2762.2093	0.1499179
-0.326758	-17.73328	-0.021917	0.0055107	0.1499179	0.001908
-73.41614	-11470.26	-1.174292	0.2892287	59.928675	-0.002758
-11470.26	-2030482	-64.03474	59.928675	11566.49	-0.14526
-1.174292	-64.03474	-0.078612	-0.002758	-0.14526	-0.000187

And the **Lower Right** 6 x 6

505.14668	69884.681	12.607147	-1.510895	-205.6737	-0.039387
	11783844	684.32233	-205.6737	-34434.43	-2.137195
		0.8455549	-0.039387	-2.137195	-0.002642
			1.4336329	191.42681	0.1801651
				38859.772	3.5902121
					0.0759252

$$
\mathbf{X'R^{-1}y} = \begin{vmatrix} 1.813004 \\ 257.30912 \\ 0.24295 \\ 18.048269 \\ 2762.4114 \\ 0.3174273 \\ 3.6520902 \\ 653.97454 \\ 0.0166432 \\ 4.9935515 \\ 678.86668 \\ 0.6708264 \end{vmatrix}
$$

and

$$
\mathbf{G^{-1}c_o} = \begin{vmatrix} 0.2378446 \\ -137.8976 \\ -0.002795 \\ 2.2183526 \\ 624.94513 \\ 0.3192604 \\ 1.1512642 \\ 3441.8785 \\ -0.380704 \\ 0.7334103 \\ -7.895393 \\ 0.0169737 \end{vmatrix}
$$

The solution vector for this cow is

$$
\hat{\mathbf{c}} = \begin{vmatrix} 28.875659 \\ -0.028768 \\ -0.454583 \\ 0.9842104 \\ -0.000124 \\ 0.3339813 \\ 0.8375506 \\ -0.00034 \\ -0.038198 \\ 2.084599 \\ 0.0017539 \\ 1.9446955 \end{vmatrix}
$$

To predict 305-day yields, Y_{305}

Equation 6. Equation for 305-day yield prediction.

$$
Y_{305} = \sum_{t=1}^{305} (\hat{A} + Bt + C\exp(-.05t))
$$

$$
= 305(A) + 46665(B) + 19.504162(C),
$$

is used separately for each trait (milk, fat, protein, and SCS). The results for this cow were 7456 kg milk, 301 kg fat, and 239 kg protein. The result for SCS is divided by 305 to give an average daily SCS of 2.477.

5 References

1) Sargent, F. D., V. H. Lyton, and o. G. Wall, J r . 1968. Test interval method of calculating Dairy Herd Improvement Association records. J. Dairy Sci. 51:170.

2) Schaeffer, L.R., and J. Jamrozik. 1996. Multiple-trait prediction of lactation yields for dairy cows. J. Dairy Sci. 79:2044-2055.

3) VanRaden, P.M. 1997. Lactation yields and accuracies computed from test day yields and (co)variances by best prediction. J. Dairy Sci. 80:3015-3022.

4) Wilmink, J.B.M. 1987. Comparison of different methods of predicting 305-day milk yield using means calculated from within-herd lactation curves. Livest. Prod. Sci. 17:l.