

Reassessing Wiggans' Method for Estimating Daily Yields in 3× Milking Systems: What Have We Learned over Four Decades?

X.-L. Wu ¹, G. Wiggans ¹, M. J. Caputo ¹, C. P. Van Tassell ², R. L. Baldwin VI ², S. Sievert ³, J. Mattison ³, J. B. Cole ¹, J. Burchard ¹, and J. Dürr ¹

¹ Council on Dairy Cattle Breeding; ² USDA AGIL; ³ National DHIA



Background: The Wiggans (1986) Model

- Let y_{ij} denote the proportional daily yield of the j -th milking for the i -th cow, and t_{ij} represent the corresponding milking interval (h).

$$y_{ij} = \alpha_j + \beta t_{ij} + e_{ij}$$

where α_j is the intercept for milking j , β is the common regression coefficient across all three milkings, and e_{ij} is the residual effect.

- *A posteriori* adjustment of the intercepts:

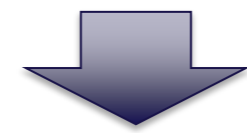
$$\sum_{j=1}^3 \hat{\alpha}_j = 1.$$

Critics of the Wiggans (1986) Model

- Biological impossibility: $y_{ij} = a_j > 0$ when $t_{ij} = 0$.
- The intercept equality constraint, $\sum_{j=1}^3 \hat{\alpha}_j = 1$, imposed after model fitting does not minimize the residual sum of squares.
- Mathematically, $\sum_{j=1}^3 \hat{\alpha}_j = 1$ has no unique solutions, because it has three unknown parameters in a single equation.

Re-parameterized Least Squares (RLS)

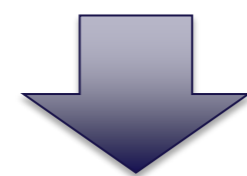
$$\alpha_3 = 1 - \alpha_1 - \alpha_2$$



$$\begin{aligned} y_{ij} &= (\alpha_1 x_{i1} + \alpha_2 x_{i2} + (1 - \alpha_1 - \alpha_2)x_{i3}) + \beta t_{ij} + e_{ij} \\ &= (x_{i1} - x_{i3})\alpha_1 + (x_{i2} - x_{i3})\alpha_2 + x_{i3} + \beta t_{ij} + e_{ij} \end{aligned}$$



$$y_{ij} - x_{i3} = (x_{i1} - x_{i3})\alpha_1 + (x_{i2} - x_{i3})\alpha_2 + \beta t_{ij} + e_{ij}$$



$$y_{ij}^* = x_{i1}^* \alpha_1 + x_{i2}^* \alpha_2 + t_{ij} \beta + e_{ij}$$

Constraint least squares (CLS)

$$y_{ij} = \alpha_j + \beta(t_{ij} - 8) + e_{ij}$$

Subject to:

$$\sum_{j=1}^3 \hat{\alpha}_j = 1$$

Matrix Notation

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e}$$

$$\mathbf{X} = \begin{bmatrix} 1 & 0 & 0 & t_{i1} & -8 \\ 0 & 1 & 0 & t_{i2} & -8 \\ 0 & 0 & 1 & t_{i3} & -8 \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}; \mathbf{b} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \beta \end{bmatrix}$$

Equality Constraint on the Intercepts

$$\mathbf{c}\mathbf{b} = d \quad \longleftarrow \quad \sum_{j=1}^3 \hat{\alpha}_j = 1.$$

where $\mathbf{c} = (1 \quad 1 \quad 1 \quad 0)$ and $d = 1$.

$$\hat{\mathbf{b}} = \hat{\mathbf{b}}_0 - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{c}'(\mathbf{c}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c}')^{-1}(\mathbf{c}\hat{\mathbf{b}}_0 - d)$$

where $\hat{\mathbf{b}}_0 = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ is the unconstrained OLS estimator.

Assume “Perfect” Linearity

$$y(t) = \frac{t}{24}$$

$$a_1 = a_2 = a_3 = 1/3 \approx 0.3333$$

$$\beta = \frac{dy(t)}{dt} = \frac{1}{24} \approx 0.0417$$

3× Milking Data

- Milking data were collected from 2023 to 2025 at four 3× milking Holstein dairy farms (A, B, C, and D) across three U.S. states.
- Milkings were classified by start time:
 - ✓ 1 (02:00 to 09:59 h); 2 (10:00 to 17:59 h); 3 (18:00 to 01:59 h).
- Milkings were recorded weekly for each cow up to 120 days in milk, then monthly until 305 days or the end of lactation.
- Data cleaning removed duplicate, missing, and incomplete records, leaving 15,702 × 3 (A), 13,574 × 3 (B), 8,515 × 3 (C), and 10,821 × 3 (D) records.

Estimated Model Parameters

Method	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\alpha}_3$	$\hat{\beta}$
Milk yield				
RLS	0.334	0.332	0.334	0.0381
CLS	0.334	0.332	0.334	0.0381
WLS1	0.029	0.028	0.029	0.0381
WLS0	0.077	0.068	0.066	0.0329
Fat yield				
RLS	0.336	0.338	0.326	0.0246
CLS	0.336	0.338	0.326	0.0246
WLS1	0.139	0.142	0.129	0.0246
WLS0	0.186	0.186	0.182	0.0186
Protein yield				
RLS	0.332	0.334	0.334	0.0390
CLS	0.332	0.334	0.334	0.0390
WLS1	0.020	0.022	0.021	0.0390
WLS0	0.072	0.065	0.063	0.0333

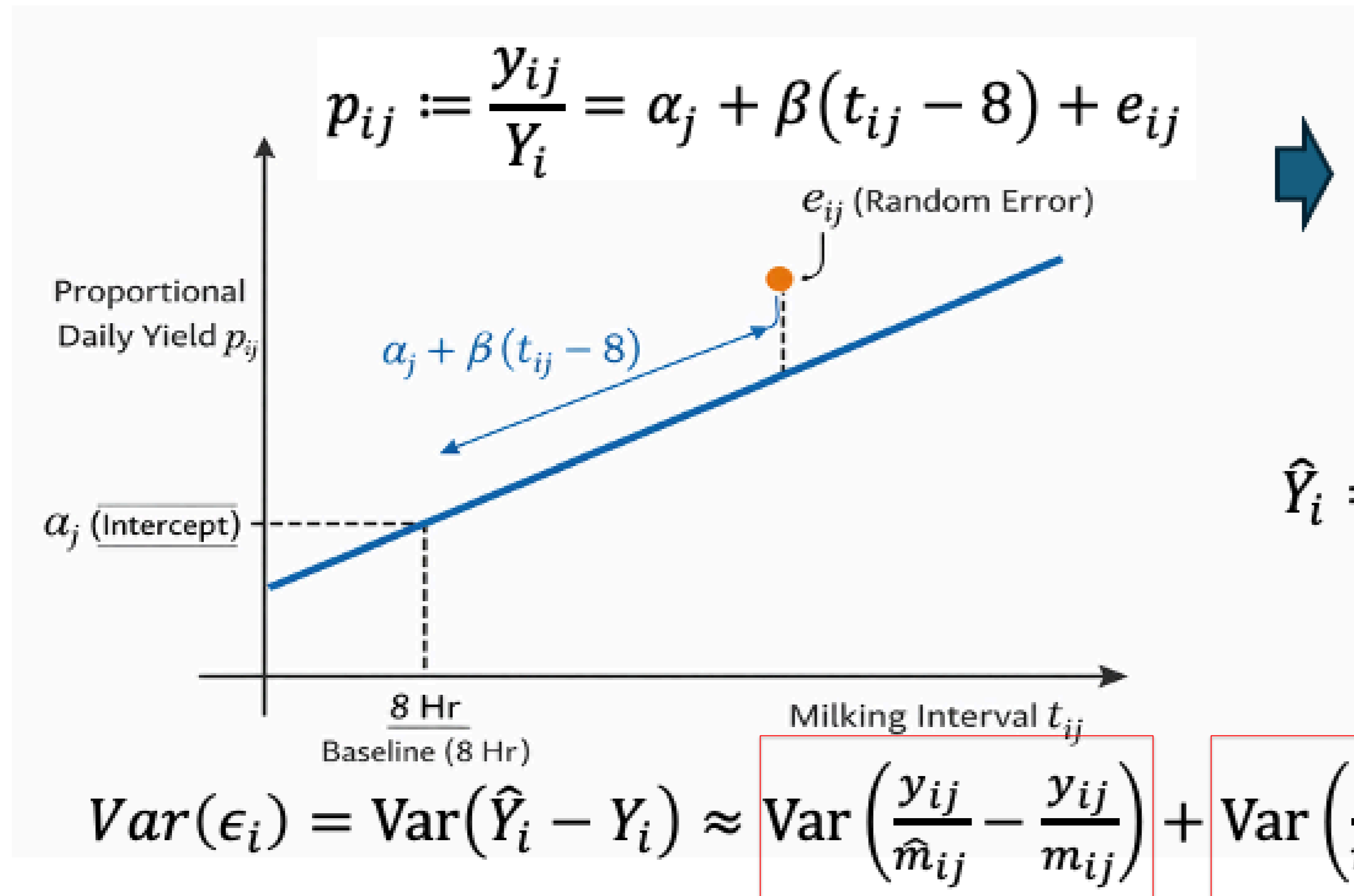
- CLS = constrained ordinary least squares;
- RLS = reparameterized CLS;
- WLS0 = Wiggans (1986) model fitted to historical data;
- WLS1 = Wiggans (1986) model fitted to current data.
- Values in the table are means across ten replicates.

Accuracy of Estimated Daily Yields

Traits	Metrics	RLS	CLS	WLS1	WLS0
MY	COR	0.906	0.906	0.906	0.904
	MAE	3.03	3.03	3.03	3.08
	MSE	21.7	21.7	21.7	22.1
FY	COR	0.833	0.833	0.833	0.831
	MAE	0.183	0.183	0.183	0.184
	MSE	0.071	0.071	0.071	0.072
PY	COR	0.885	0.885	0.885	0.883
	MAE	0.094	0.094	0.094	0.957
	MSE	0.020	0.020	0.020	0.022

- Correlation = correlation between the estimated and actual daily yields;
- MAE = mean absolute error of the estimated yield;
- MSE = mean squared error of the estimated yield.

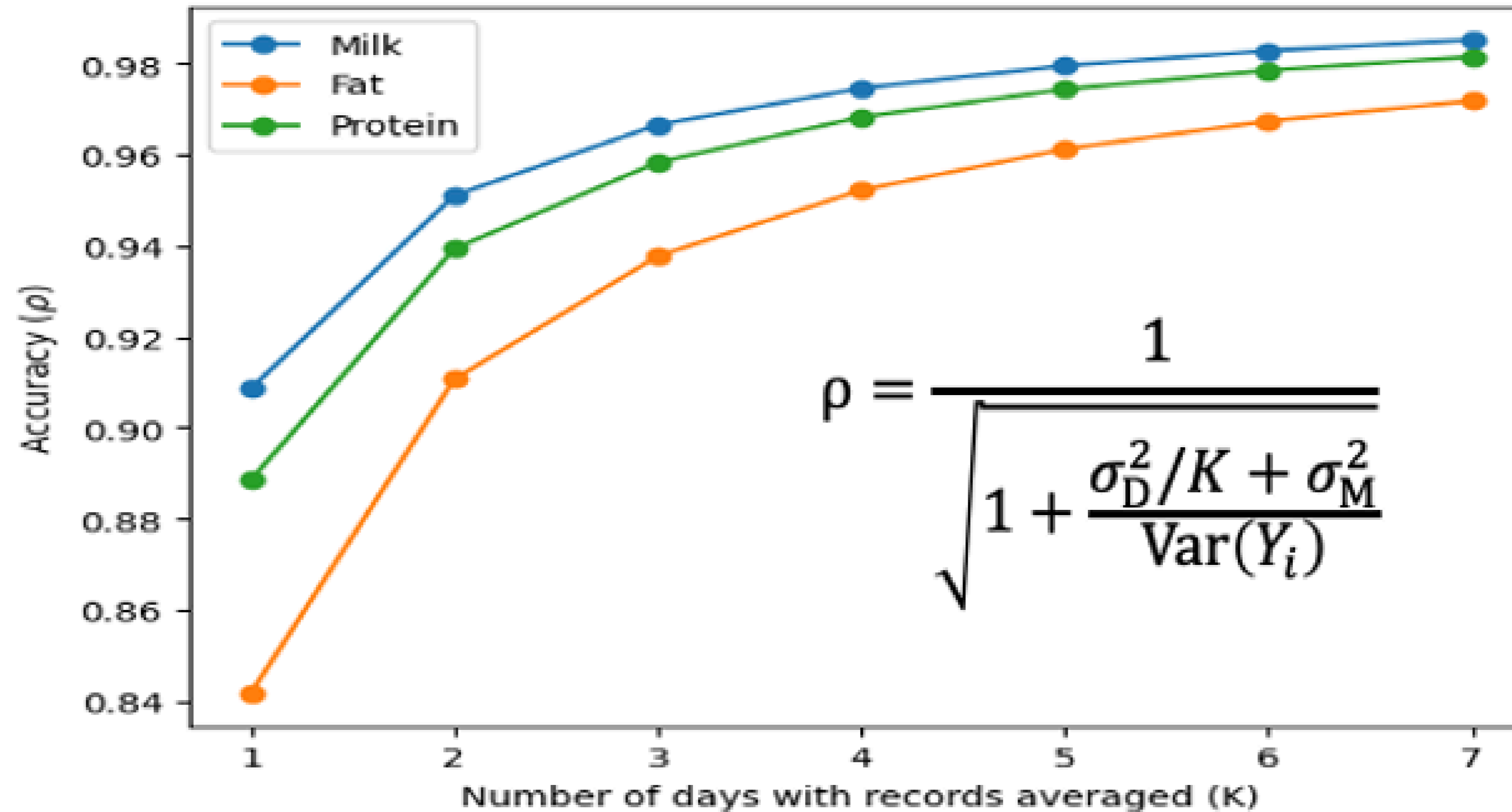
Aftermath



Decomposition of Error Variance

Farm	Var(Y)	Var(\hat{Y})	Var(ϵ)	σ_D^2	σ_M^2	ρ_ϵ	ρ_{MD}	Cor(Y, \hat{Y})
Milk								
A	78.70	94.83	18.05	18.05	0.336	0.902	0.900	0.900
B	98.18	108.7	11.93	11.93	0.261	0.944	0.943	0.944
C	122.2	149.0	33.80	33.79	0.0293	0.885	0.885	0.880
D	113.9	138.3	30.03	30.03	0.324	0.890	0.889	0.885
Fat								
A	0.115	0.175	0.0642	0.0619	3.5E-03	0.801	0.798	0.796
B	0.130	0.172	0.0448	0.0423	1.1E-03	0.862	0.866	0.860
C	0.206	0.281	0.105	0.103	1.9E-04	0.814	0.816	0.793
D	0.177	0.253	0.091	0.0887	1.1E-03	0.813	0.815	0.801
Protein								
A	0.0526	0.0673	0.0162	0.0148	2.8E-04	0.874	0.882	0.872
B	0.0683	0.0790	0.0117	0.0103	3.5E-04	0.924	0.930	0.923
C	0.0997	0.126	0.0344	0.0333	1.6E-05	0.862	0.866	0.853
D	0.0869	0.111	0.0293	0.0283	2.6E-04	0.865	0.868	0.858

Using multiple-day average



Take-Home Messages

- The re-parameterized model is biologically more intuitive and mathematically self-contained.
- Biological variation in proportional yields is the primary constraint on achieving attainable accuracy, whereas parameter estimation-induced variance is negligible.
- The accuracy of estimated daily yields can be effectively improved by utilizing multiple-day pooling or average.