Analysis of Factors Affecting Daily Milk Yields: An Initial Case Study in an Automatic Thrice-Milking Farm

X.-L. Wu[1,2], M. J. Caputo[1], G. R. Wiggans[1], H. D. Norman[1], A. M. Miles[3], C. P. Van Tassell[3], R. L. Baldwin VI[3], S. Sievert[4], J. Mattison[4], J. Burchard[1] and J. Dürr[1]

[1] Council on Dairy Cattle Breeding, Bowie, MD

[2] Department of Animal and Dairy Sciences, University of Wisconsin, Madison, WI

[3] USDA Animal Genomics and Improvement Laboratory, Beltsville, MD

[4] National Dairy Herd Information Association, Verona, WI

ICAR & INTERBULL
MAY 19-24, 2024
SLOVENIA

COUNCIL ON DAIRY CATTLE BREEDING

Background

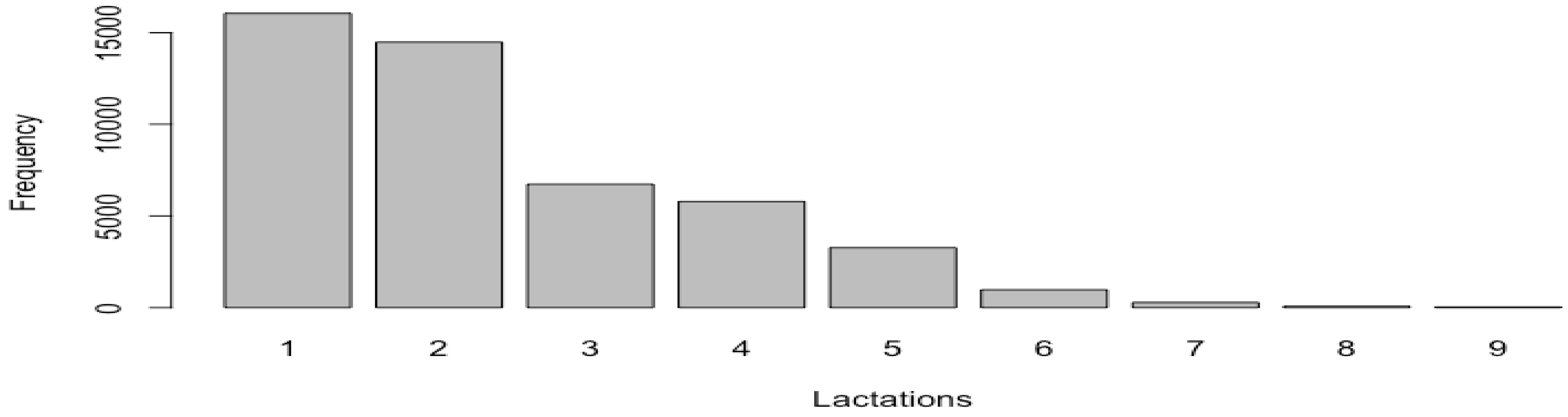
- The methodologies and parameters used for estimating day milk yields (DMY) in the United States were primarily developed from the 1960s through the 1990s. Elaborating these methods and parameters to the current system is now an essential undertaking.
- Recently, an initiative has been established between the Council on Dairy Cattle Breeding, USDA-AGIL, and the National DHIA in the U. S. in the U.S. to collect milking data from dairy farms on a large scale and update these methods and parameters for estimating DMY.
- This study represents an initial case study aiming to evaluate factors influencing daily milk yields and compare the performance of the existing method with a new one for estimating daily milk yields.

Farm 1 milking records

- Farm 1: 63,562 milking from 2,200 Holstein cows in a trice-milking farm (New York, USA) were extracted.
- Trice-milking: 4am-12pm (1st milking), 12pm-8pm (2nd milking), and 8pm-4am (3d milking).
- Data collection: Milk samples were collected and weighed at all 3 milkings for 18 weeks starting May 5 and ending September 1, 2023. After that, three-day monthly milking data collections were carried out. Milking records with prolonged lactation beyond 305d for up to one month are retained.
- Data cleaning: Records removed: (1) DIM greater than 335 days (~ 0.6%); (2) incomplete and missing data. After data cleaning, we retained 47,670 milking records representing 1,869 cows.

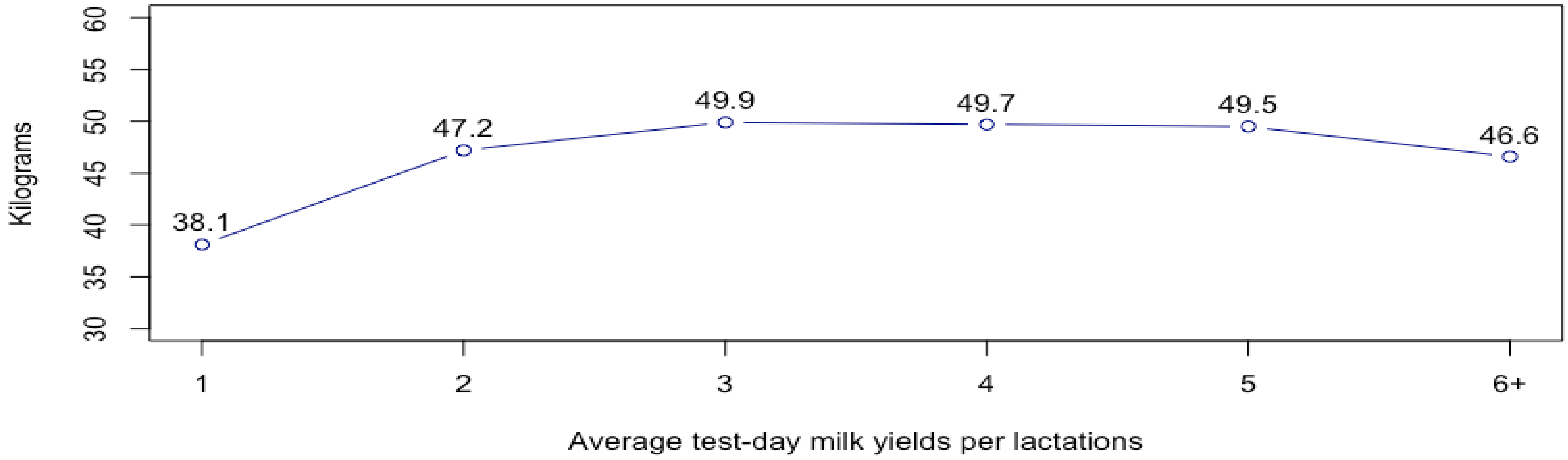
Data summary

- The cleaned data represented up to nine lactations (Figure 1), with 33.7% from the first lactation, 63.5% from lactations 2 to 5, and 2.9% from lactation 6 and beyond.
- Milking data from the sixth lactation and beyond are pooled.



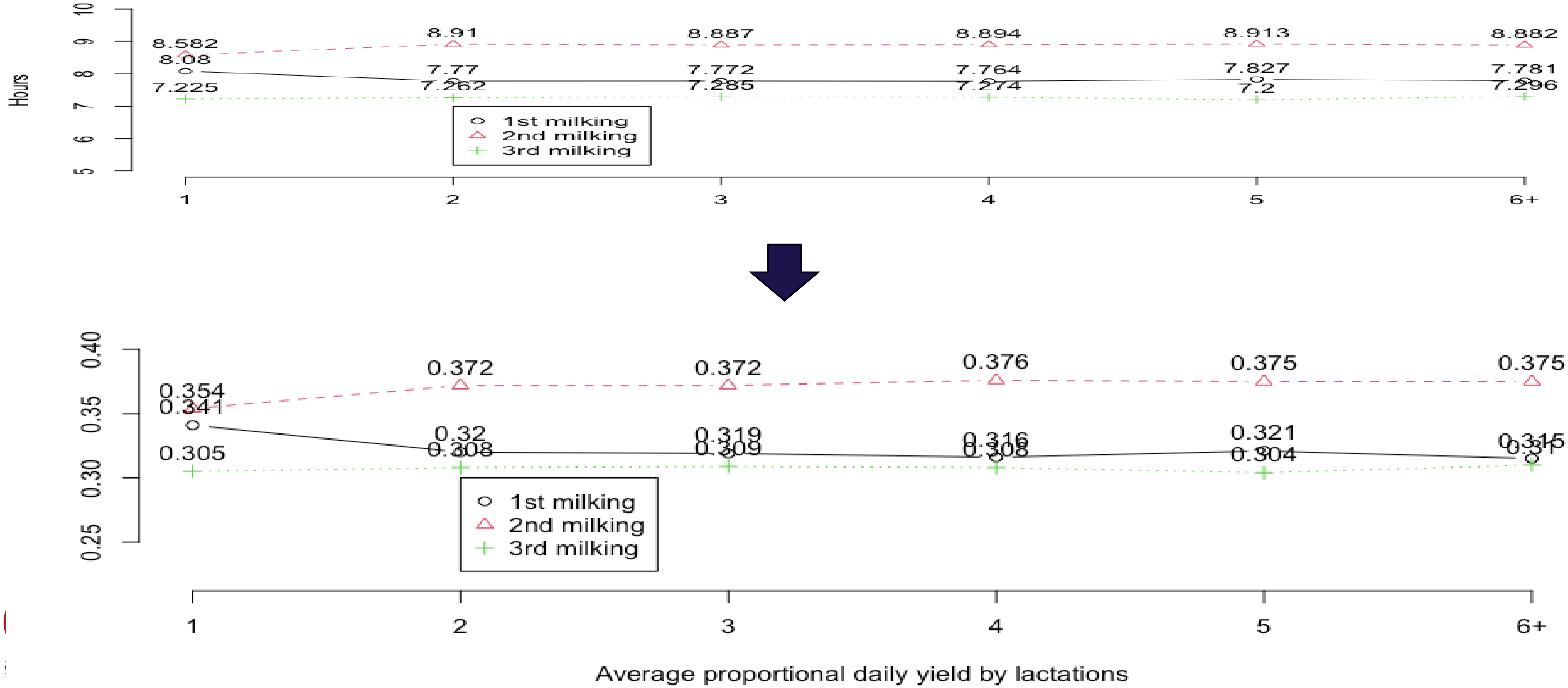
Average daily milk yields by lactations

- Average daily milk yield increased from lactation 1 to 2, remained at similar levels from 3 to 5, and then decreased slightly from lactation 5 to 6.



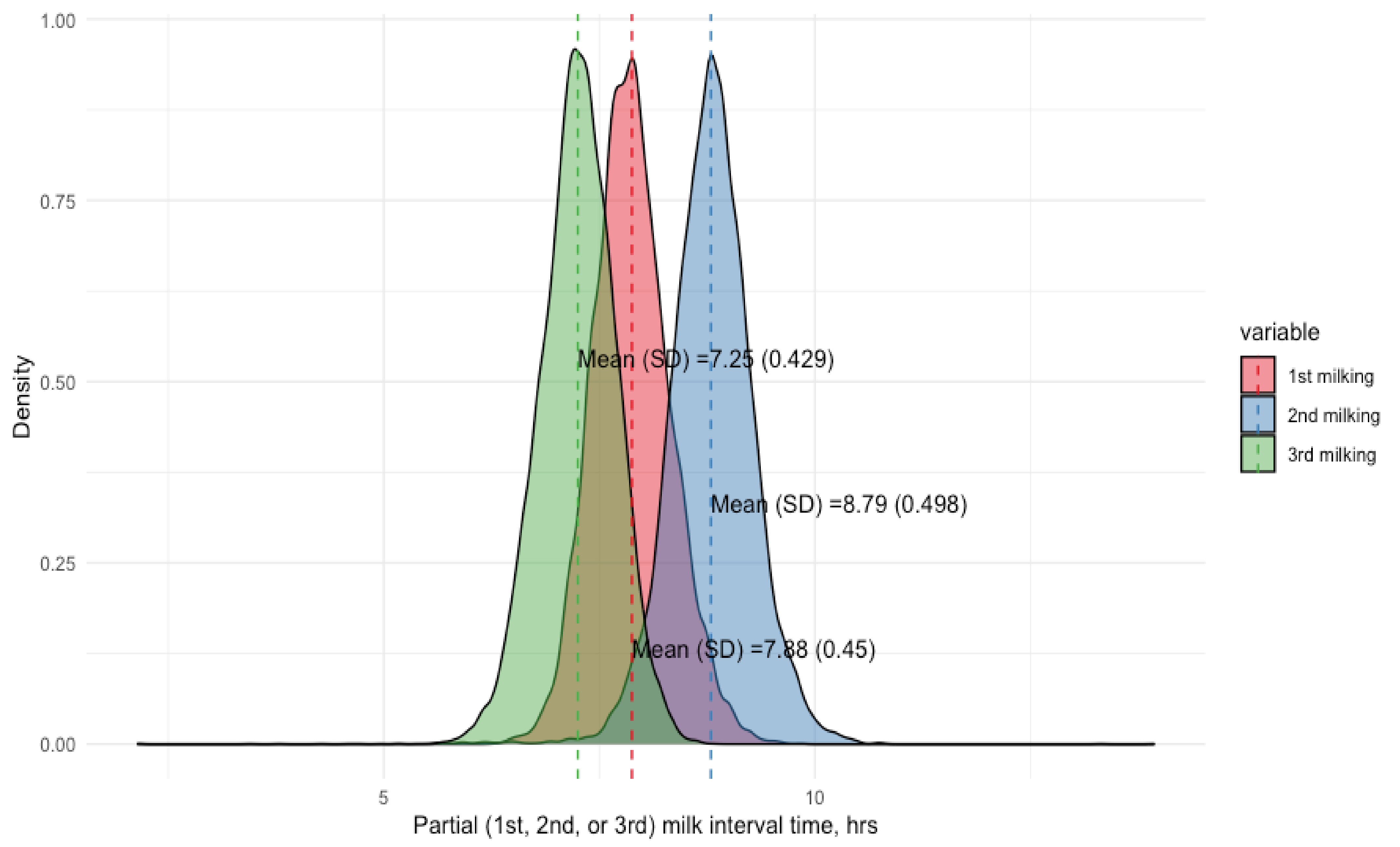
Proportional daily yields by lactations

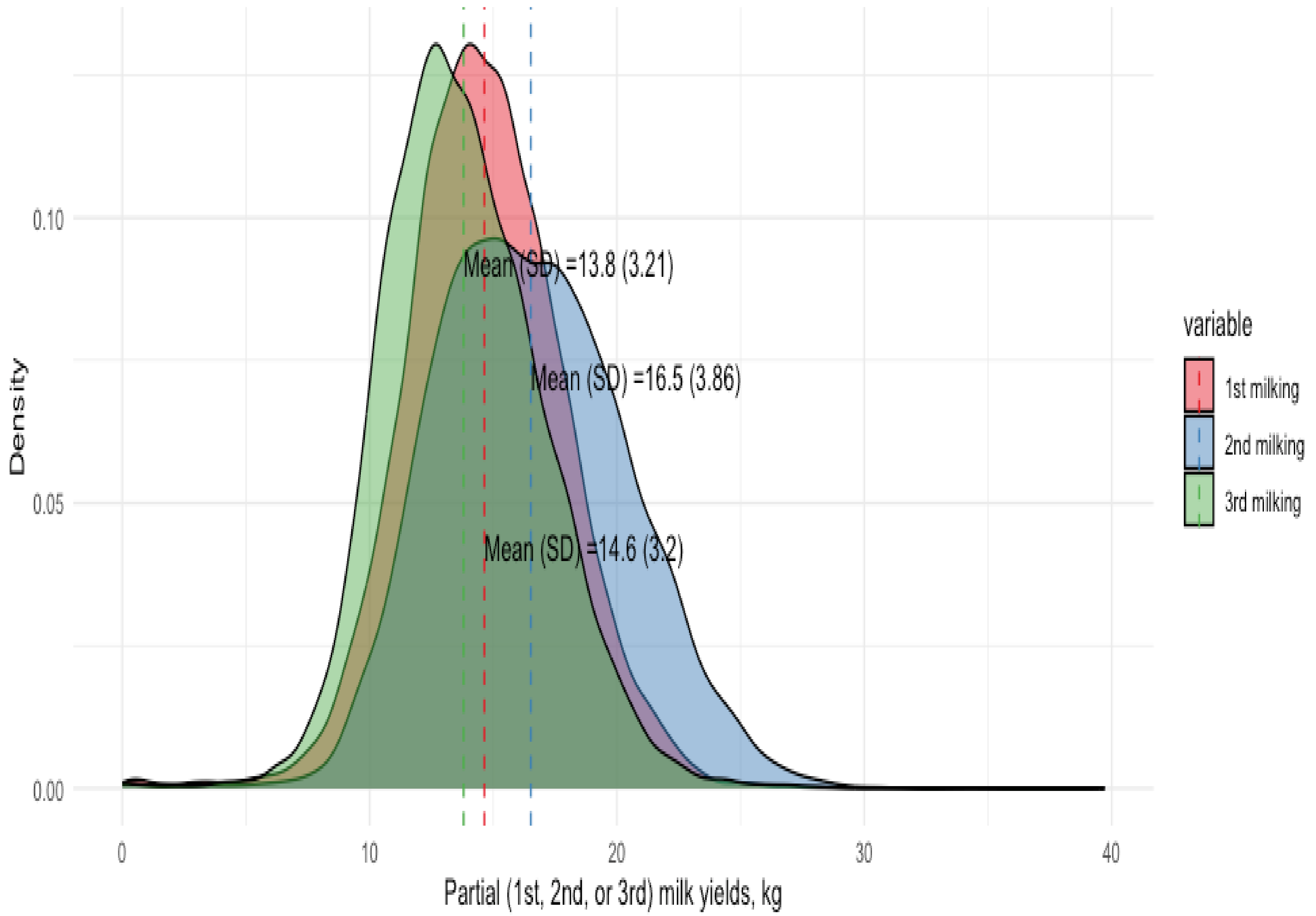
- In contrast, proportional daily milk yields remain relatively stable in concordance with the patterns of milk interval time.



Milking interval time and partial daily yields

- Average milking interval time: 2nd > 1st > 3rd;
- Average partial yields: 2nd > 1st > 3rd.





Relative importance (LMG R2) of predictor variables

- Model 1: GW (Wiggans, 1986)

$$\frac{x_{ijl}}{y_{ijl}} = \alpha + \beta t_{ijl} + m_j + \gamma_l + \varepsilon_{ijl}$$

ICAR & INTERBULL
MAY 19-24, 2024
SLOVENIA

Predictors	1st milking			2nd milking			3rd milking		
	Mean	Q2.5%	Q97.5%	Mean	Q2.5%	Q97.5%	Mean	Q2.5%	Q97.5%
t	0.157	0.145	0.171	0.135	0.121	0.149	0.159	0.146	0.172
m	0.002	0.001	0.004	0.004	0.003	0.007	0.002	0.002	0.005
γ	0.04	0.035	0.046	0.032	0.026	0.037	0.001	0.001	0.003
Sum	0.199			0.17				0.148	

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

y = daily milk yield (DMY); x = partial yield (PY: 1st, 2nd, or 3rd milking);
t = milking interval time (MIT); m = months in milk; γ = lactations.

Relative importance (LMG R2) of predictor variables

- Model 2: PIR (polynomial-interaction-regression)

ICAR & INTERBULL
MAY 19-24, 2024
SLOVENIA

$$y_{ijl} = (b_0 + b_1 t_{ijl} + b_2 t_{ijl}^2) x_{ijl} + m_j + \gamma + \epsilon_{ijl}$$
$$= b_0 x_{ijl} + b_1 (t_{ijl} x_{ijl}) + b_2 (t_{ijl}^2 x_{ijl}) + m_j + \gamma_l + \epsilon_{ijl}$$

Predictors	1st milking			2nd milking			3rd milking		
	Mean	Q2.5%	Q97.5%	Mean	Q2.5%	Q97.5%	Mean	Q2.5%	Q97.5%
x	0.285	0.280	0.290	0.280	0.276	0.284	0.274	0.269	0.279
tx	0.226	0.222	0.230	0.244	0.24	0.247	0.225	0.222	0.229
t^2x	0.158	0.154	0.162	0.199	0.196	0.202	0.172	0.168	0.175
m	0.022	0.02	0.025	0.021	0.019	0.024	0.02	0.018	0.023
γ	0.129	0.124	0.133	0.083	0.080	0.086	0.101	0.096	0.106
SUM	0.82			0.83			0.79		

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

y = daily milk yield (DMY); x = partial yield (PY: 1st, 2nd, or 3rd milking);
t = milking interval time (MIT); m = months in milk; γ = lactations.

R2 Accuracy of estimated daily milking yields

Methods	1st milking			2nd milking			3rd milking		
	Corr	R^2	K	Corr	R^2	K	Corr	R^2	K
Before variance rescaling									
GW1	0.880	0.781	1.237	0.901	0.809	1.253	0.875	0.769	1.285
GW2	0.879	0.791	1.152	0.902	0.801	1.3207	0.875	0.769	1.283
PIR1	0.883	0.800	1.205	0.903	0.815	1.2277	0.877	0.777	1.249
PIR2	0.906	0.847	0.821	0.909	0.852	0.8278	0.889	0.827	0.792
After variance rescaling									
GW1	0.880	0.806	1.000	0.901	0.835	1.000	0.875	0.800	1.000
GW2	0.879	0.806	1.000	0.902	0.836	1.000	0.875	0.800	1.000
PIR1	0.883	0.811	1.000	0.903	0.837	1.000	0.877	0.803	1.000
PIR2	0.906	0.841	1.000	0.909	0.847	1.000	0.889	0.819	1.000

$$R^2\ Accuracy = \frac{Var(y)}{Var(y) + MSE}$$

PIR Model: Estimated parameters (all data)

Model parameters	1st Milking		2nd Milking		3rd Milking	
	Estimate	SE	Estimate	SE	Estimate	SE
M1a: Excluding the effects due to months in milk and locations						
b_0	8.358	0.353	5.185	0.288	7.554	0.535
b_1	-1.003	0.088	-0.326	0.066	-0.832	0.147
b_2	0.042	0.005	0.005	0.004	0.032	0.010
M1b: Including the effects due to months in milk and lactations						
b_0	5.973	0.290	2.781	0.254	5.313	0.457
b_1	-0.754	0.02	-0.014	0.057	-0.605	0.126
b_2	0.034	0.004	-0.008	0.003	0.025	0.009
m_1	10.49	0.172	10.51	0.170	11.75	0.184
……						
m_{11}	9.650	0.316	7.892	0.314	10.70	0.340
γ_2	3.563	0.083	1.438	0.084	2.559	0.089
……						
γ_6	3.897	0.188	1.136	0.186	2.476	0.203

Comparing MCF: REF, GW, and PIR

- Reference (Wiggans, 1986)

$$F_{1st} = \frac{1}{0.077 + 0.033t} \qquad F_{2nd} = \frac{1}{0.068 + 0.033t} \qquad F_{3rd} = \frac{1}{0.066 + 0.033t}$$

- Farm 1 - GW

$$F_{1st} = \frac{1}{0.019 + 0.039t} \qquad F_{2nd} = \frac{1}{0.080 + 0.033t} \qquad F_{3rd} = \frac{1}{0.049 + 0.036t}$$

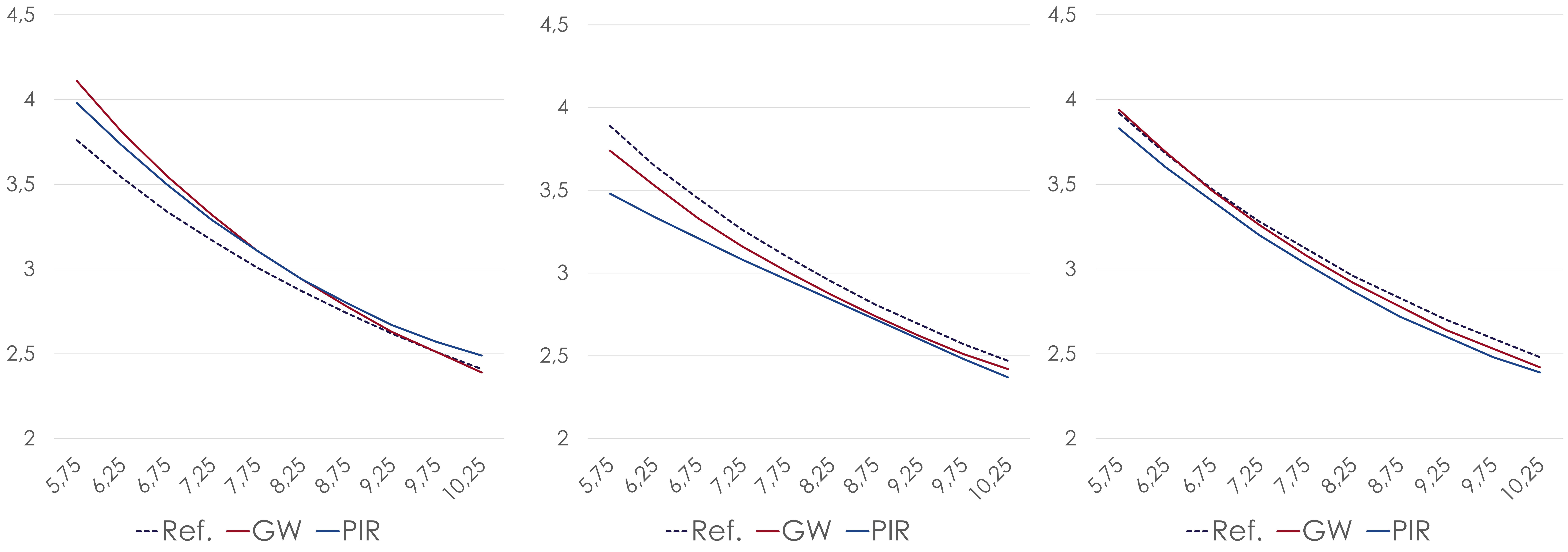
- Farm 1 - PIR

$$F_{1st} = 8.358 - 1.003t + 0.042t^2$$

$$F_{2nd} = 5.185 - 0.326t + 0.005t^2$$

$$F_{3rd} = 7.554 - 0.832t + 0.032t^2$$

Comparing MCF: Ref., GW, and PIR



Take-home messages

- Partial yields and milking interval time are two major variables influencing daily milk yields whereas other variables are secondary.
- Modeling proportional daily milk yields (DMY) as a linear function of milking interval time is a still valid strategy for estimating DMY.
- The polynomial-interaction-regression model offers further improvement (4-6%) in the accuracy of estimated DMY compared to the Wiggans (1986) model.
- Daily yield correction factors showed only minor deviations despite the significant genetic improvement in milk yields in the past decades.
- The current conclusions are preliminary and subject to further large-scale validations.

ICAR & INTERBULL
MAY 19-24, 2024
SLOVENIA

Questions and comments?

COUNCIL ON DAIRY CATTLE BREEDING