Hamadani, A., Hamadani, H., Ganai, N., Khan, H., Ali, R. (2026) Effects of various factors on reproductive efficiency in dairy cattle, *Proceedings 13th World Congress on Genetics Applied to Livestock Production* 23918: https://doi.org/10.31274/wcgalp.23918
©2026 Hamadani, et al. CC BY-NC-ND

Effects of various factors on reproductive efficiency in dairy cattle

A. Hamadani*, H. Hamadani, N.A. Ganai, H.M. Khan, and R. Ali

*Sher-e-Kashmir University of Agricultural Sciences and Technology of Kashmir, Srinagar, Jammu and Kashmir, India; *Corresponding Author: qambrineh@gmail.com*

Introduction

One of the factors that reproductive management in dairy herds depends on is understanding how parity and breed composition affect key fertility traits (Crowe, Hostens and Opsomer, 2018). Understanding these relationships is particularly important for making informed management decisions. Reproductive performance, as indicated by factors like inter-calving period (ICP), age at first calving (AFC), and gestation length, directly impacts farm profitability and animal welfare.

Genetic background, represented through sire effects, introduces clustering in reproductive data. However, small herds often have limited sire-level replication. Due to the limited sample sizes, genetic diversity may be constrained. This is challenging for traditional mixed-effects modeling approaches. This necessitates the use of alternative methods such as generalized estimating equations (GEE) and cluster-robust regression that can handle sparse clustering structures.

Recent advances in statistical methodology offer opportunities to improve inference: robust regression methods are less sensitive to outliers than standard ordinary least squares (OLS) (Brossart, Parker and Castillo, 2011). Quantile regression allows examination of heterogeneous effects across the outcome distribution (Khadka *et al.*, 2025). In this study, we apply mixed-effects models, generalized estimating equations, and cluster-robust ordinary least squares as univariate approaches, alongside robust regression and quantile regression, to quantify the effects of parity, breed, and other factors on fertility outcomes. We further use multivariate analysis to examine relationships among outcomes, assess the consistency of results across statistical methods, and identify the most important predictors of reproductive efficiency in this small herd context.

Materials & Methods

Study Population and Data Collection. The reproduction records of 192 dairy cows maintained at the Mountain Livestock Research Institute, SKUAST-K herd, representing Jersey and Crossbreds from 2008 to 2020, were used in the study. Data collection included average inter-calving period (ICP, days), age at first calving (AFC, days), average gestation length (days), parity (lactation number), birth weight (kg), and sire identification. The number of sires in the data was 30. Multiple Imputation by Chained Equations (MICE) was used to handle missing data. All analysis was done in Python. Cluster-robust Ordinary Least Squares (OLS), Generalized Estimating Equations (GEE), and mixed-effects models were treated as primary analyses, while robust regression, quantile regression, and random forest were used as complementary sensitivity analyses.

Model Specifications. ICP models: Fixed effects included parity (continuous), birthweight_(kg), AFC_(days), and breed. Gestation length models: Fixed effects included parity, birthweight (kg), and breed. AFC models: Fixed effects included parity, birthweight, and breed. Parity and birth weight were modeled as continuous fixed-effect covariates, whereas breed was modeled as a categorical fixed-effect factor. Model fit was compared using Akaike and Bayesian Information Criteria (AIC/BIC) (Kuha, 2004). Lower values indicate a better fit; differences <2 were considered negligible, and model interpretation was guided by the consistency and robustness of effects across methods. AIC and BIC were used only for comparisons among closely related linear models for a given outcome and were not used to compare fundamentally different modeling frameworks.
Generalized Estimating Equations (GEE): GEE models were fitted to account for clustering of daughters within sires. **Cluster-Robust Ordinary Least Squares (OLS):** Standard OLS regression with standard errors adjusted for sire-level clustering using the cluster-robust variance estimator. Cluster-robust OLS was used for: parity effects on ICP and gestation length, and breed effects on AFC. **Mixed-Effects Linear Models (MixedLM):** Mixed-effects models with sire as a random intercept were fitted using restricted maximum likelihood (REML) estimation with the L-BFGS optimization method. MixedLM was used for parity effects on ICP and gestation length. **Robust Regression:** Robust regression using Huber's T M-estimator was applied to reduce sensitivity to outliers. Robust regression was used for parity effects on ICP and gestation length. **Quantile Regression:** Quantile regression models were fitted at the 25th, 50th, and 75th percentiles to examine heterogeneous effects of predictors across different quantiles of the outcome distribution. **Random Forest Regression:** Random Forest regression with conservative parameters (n_estimators=100, max_depth=3, min_samples_split=10, random_state=42) was used to assess non-parametric feature importance. This ensemble method provides a measure of variable importance that captures non-linear relationships and interactions without assuming a specific functional form. Random Forest was used to predict ICP and identify the relative importance of predictors (parity, birth weight, AFC, gestation average, gestation length). Gestation average was modelled as an outcome in the gestation models and as a continuous predictor (covariate) only in the Random Forest analysis of ICP feature importance. Gestation average was included only as a predictor in the random forest analysis for ICP.

Results

GEE models with an exchangeable sire-level correlation structure converged successfully for all outcomes. Parity showed a strong and consistent association with reproductive traits. For ICP, each unit increase in parity was associated with a reduction of approximately 8.9 days ($p = 0.0003$). For gestation length, parity was associated with an increase of 0.65 days per parity ($p = 0.0008$). Significant breed differences were observed for AFC.
Cluster-robust OLS models produced nearly identical results, confirming the robustness of these findings. Parity effects remained strongly negative for ICP (−9.0 days per parity, $p = 0.0005$) and positive for gestation length (+0.70 days per parity, $p = 0.0006$), while breed effects on AFC remained significant.
Mixed-effects models with sire as a random intercept also converged, although with reduced precision. The direction of effects remained consistent, with parity associated with reduced ICP (−10.7 days per parity, $p = 0.049$) and increased gestation length (+0.56 days

per parity, p = 0.073). The AFC model also converged successfully. Overall, inclusion of sire-level random effects did not materially alter the estimated fixed effects.

Sensitivity analyses supported these findings. Robust regression confirmed significant associations, though with smaller effect sizes (ICP: −3.52 days per parity; gestation: +0.27 days per parity; both $p < 0.001$), suggesting some influence of outliers. Quantile regression revealed heterogeneity in ICP effects, with stronger reductions at lower quantiles (−10.4 days at the 25th percentile) and negligible effects at higher quantiles, indicating that parity effects are most pronounced among more efficient cows. In contrast, gestation length showed consistent positive parity effects across quantiles.

Random forest analysis highlighted the relative importance of predictors for ICP. Age at first calving (AFC) emerged as the most important variable (43.8%), followed by gestation average (26.9%), birth weight (18.5%), and parity (10.8%). This suggests that early-life factors are more predictive of reproductive performance than parity alone.

Within-family model comparisons showed minimal differences in fit between GEE and cluster-robust OLS (ICP: AIC = 2056.77; BIC = 2082.2; gestation: ΔAIC = 0.13; ΔBIC = 0.13), indicating no meaningful separation between these closely related models.

Across all approaches, the direction and general magnitude of key effects—particularly parity for ICP and gestation, and breed for AFC—were consistent, supporting the robustness of the main findings.

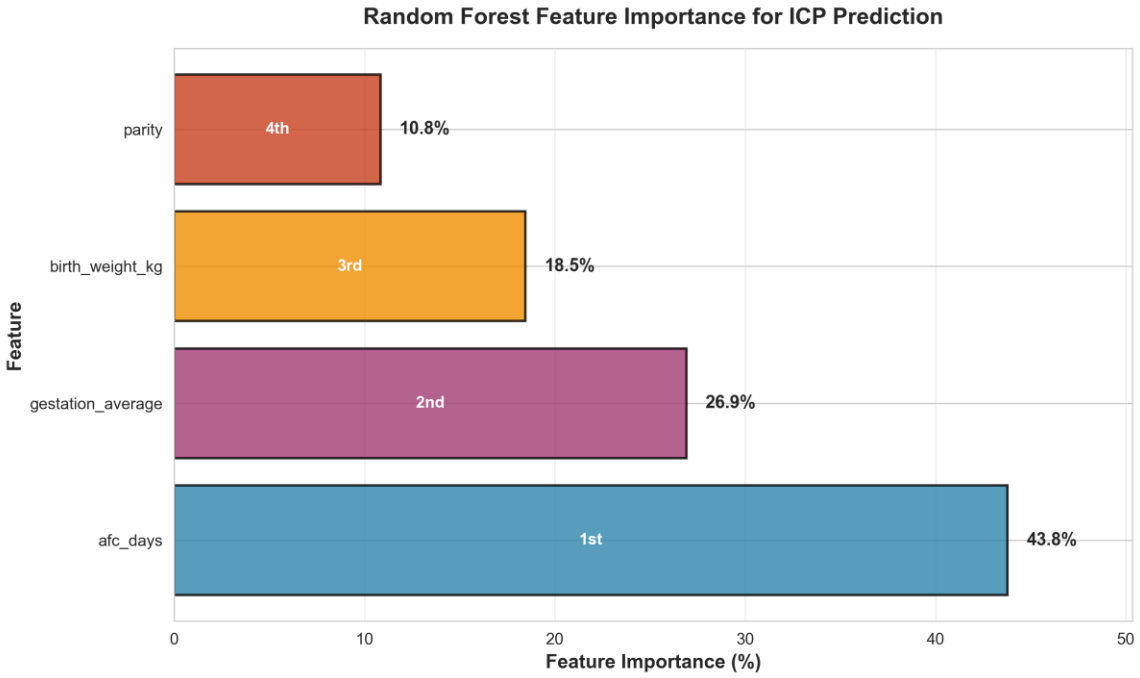

Figure 1: Feature Importance as per random forests for ICP Prediction

Discussion

Across multiple modeling approaches, parity showed a consistent negative association with ICP and a small positive association with gestation length, while breed and early-life factors were more influential for AFC. The highly significant negative association between parity and inter-calving period (ICP) and positive association with gestation length were confirmed across all methods, demonstrating that higher-parity cows tend to return to calving faster while experiencing slightly prolonged gestations (Atashi and Asaadi, 2019).

Quantile regression revealed that parity effects on ICP are strongest for efficient cows but minimal for problematic cows. This indicates that parity benefits are most pronounced in already-efficient animals. There is an importance of heifer development and early-life management in determining lifetime reproductive efficiency (Dallago, Warner and Vasseur, 2025).

These results have practical implications for dairy herd management, which includes investment in heifer development programs, as early-life factors are more predictive of later reproductive performance than parity, and breed selection for reproductive efficiency should consider breed-specific patterns, particularly for heifer development. The non-parametric analysis confirms that early-life factors (AFC, gestation) are substantially more predictive of later fertility performance than parity, supporting the importance of heifer development programs and management in determining reproductive efficiency (Hamadani *et al*., 2020). Our findings on parity effects align with previous reports that reproductive performance improves with increasing parity (Twomey and Cromie, 2023). Age at first calving is widely recognized as a key determinant of lifetime reproductive and production performance in dairy cattle (Shekure *et al.*, 2024). Breed differences in age at first calving observed here also mirror broader patterns in cross-bred dairy populations, where genetic background influences maturity and timing of reproductive onset (Allan *et al.*, 2024). Some authors suggest that gestation is more strongly influenced by fetal and environmental factors in many populations (Nogalski and Piwczyński, 2011). Improving AFC and heifer development likely yields more gains than relying on parity progression.

Conclusion

Our results indicate that parity was associated with shorter inter-calving period and slightly longer gestation. Age at first calving was the most influential predictor of inter-calving period. We suggest that strategies aimed at improving heifer development and reducing delayed AFC will yield greater reproductive benefits than parity-based management alone.

References

Allan, F. K. *et al.* (2024) Tropical Animal Health and Production, 56(7), p. 218. https://doi.org/10.1007/s11250-024-04074-z

Atashi, H. and Asaadi, A. (2019) Animal Reproduction*, 16*(4), pp. 846–852. https://doi.org/10.21451/1984-3143-AR2019-0005

Brossart, D. F., Parker, R. I. and Castillo, L. G. (2011) Behavior Research Methods*,* 43(3), pp. 710–719. https://doi.org/10.3758/s13428-011-0079-7

Crowe, M. A., Hostens, M. and Opsomer, G. (2018) Irish Veterinary Journal, 71(1), p. 1. https://doi.org/10.1186/s13620-017-0112-y

Dallago, G. M., Warner, D. and Vasseur, E. (2025) Journal of Dairy Science, 108(8), pp. 8594–8605. https://doi.org/10.3168/jds.2024-26191

Hamadani, H. *et al.* (2020) Asian Journal of Dairy and Food Research, 39(2), pp. 135–138. https://doi.org/10.18805/ajdfr.DR-1688

Khadka, A. *et al.* (2025) American Journal of Epidemiology, 194(7), pp. 2075–2084. https://doi.org/10.1093/aje/kwae246

Kuha, J. (2004) Sociological Methods & Research, 33(2), pp. 188–229. https://doi.org/10.1177/0049124103262065

Nogalski, Z. and Piwczyński, D. (2011) Asian-Australasian Journal of Animal Sciences, 25(1), pp. 22–27. https://doi.org/10.5713/ajas.2011.11084

Shekure, T. *et al.* (2024) Scientific Reports*,* 14(1), p. 30703. https://doi.org/10.1038/s41598-024-79626-2

Twomey, A. J. and Cromie, A. R. (2023*)* Journal of Animal Science, 101, p. skad008. https://doi.org/10.1093/jas/skad008