Available online at website: https://jurnal.iaii.or.id/index.php/RESTI

JURNAL RESTI
(Rekayasa Sistem dan Teknologi Informasi)
Vol. 9 No. 5 (2025) 1078 - 1090 e-ISSN: 2580-0760

A Data-Driven Comparison of Linear Mixed Model and Mixed Effects Regression Tree Approaches for Dairy Productivity Analysis

Achmad Fauzan[1,2], Fatkhurokhman Fauzi[3,4], Rhendy K P Widiyanto[5,6]
Khairil Anwar Notodiputro[7], Bagus Sartono[8]
[1,3,5,7,8]Study Program of Statistics and Data Science, School of Data Science, Mathematics and Informatics
IPB University, Bogor, Indonesia
[2]Department of Statistics, Universitas Islam Indonesia, Yogyakarta, Indonesia
[4]Department of Statistics, Universitas Muhammadiyah Semarang, Semarang, Indonesia
[6]Center for Agricultural Data and Information System, Ministry of Agriculture, Jakarta, Indonesia

[1]fauzanachmad@apps.ipb.ac.id, [2]achmadfauzan@uii.ac.id, [3]fauzifatkhurokhman@apps.ipb.ac.id
[4]fatkhurokhmanf@unimus.ac.id, [5]stat.rhendy@apps.ipb.ac.id, [6]rhendy@pertanian.go.id
[7]khairil@apps.ipb.ac.id, [8]bagusco@apps.ipb.ac.id

Abstract

Dairy productivity studies often involve hierarchical and longitudinal data structures that violate the assumptions of linear regression. This study compares two modeling approaches, Linear Mixed Model (LMM) and Mixed Effects Regression Tree (MERT), in predicting dairy productivity based on the 2024 National Dairy Productivity Survey data. Predictive performance was evaluated using MSEP, RMSEP, MAD, and MAPE, with MERT consistently outperforming LMM in accuracy and robustness. Permutational Multivariate Analysis of Variance (PERMANOVA) test results reinforced this finding, yielding a pseudo-F value of 224.7 and a p-value of 0.001, indicating statistically significant differences in model performance. Key predictors of MERT model included farm altitude, the previous week's milk production, and the amounts of concentrate feed given, which are part of significant predictor variables in LMM. This finding underscores MERT's superiority in modeling complex agricultural datasets while providing interpretable insights through data-driven segmentation. The study advocates policy focus in sustainable milk production as well as the availability of high quality of feed and altitude-based dairy farms location to improve milk productivity. Should these focuses implemented by the industry, combined with the MBG Program, Indonesia would be progressing better towards achievement of SDGs Goal 2 and 3.

Keywords: hierarchical data; indonesian dairy cow milk productivity survey 2024; linear mixed model (LMM); mixed effects regression tree (MERT); SDGs: goal 2 and 3

How to Cite: Achmad Fauzan, Fatkhurokhman Fauzi, Rhendy K P Widiyanto, Khairil Anwar Notodiputro, and Bagus Sartono, "A Data-Driven Comparison of Linear Mixed Model and Mixed Effects Regression Tree Approaches for Dairy Productivity Analysis", *J. RESTI (Rekayasa Sist. Teknol. Inf.)* , vol. 9, no. 5, pp. 1078 - 1090, Oct. 2025.
Permalink/DOI: https://doi.org/10.29207/resti.v9i5.6751

Received: June 9, 2025
Accepted: September 20, 2025
Available Online: October 13, 2025

This is an open-access article under the CC BY 4.0 License
Published by Ikatan Ahli Informatika Indonesia

1. Introduction

In the analysis of longitudinal data and hierarchical data, the linear regression model (LM) is often inadequate because the assumption of independence between observations is not met. Such data typically has a nested (clustered) or repeated measuresstructure, requiring a modeling approach that can capture both the variability among observations (variability due to fixed effects) and the variability within observations due to cluster differences (variability due to random effects) [1]. One widely used approach is the Linear Mixed Model (LMM). LMM is designed to combine fixed and random effects, allowing for a better explanation of the variation structure in such complex data [2].

The main advantage of LMM over LM is its ability to handle correlation between observations that arise due to cluster differences, utilizing certain random effects component. Furthermore, LMM offers flexibility in modeling data with varying numbers of observations across clusters (unbalanced data). Additionally, by capturing inter-cluster variability, LMM is known to provide parameter estimates with lower bias compared to LM [3]. Moreover, the LMM model is also known to be more effective in capturing longitudinal variation

compared to conventional marginal models [4]. Here, the LMM yielded models with smaller residual errors and a more realistic covariance structure in public health behavior studies. Another study found that LMM remained a competitive baseline model, achieving up to 82% predictive accuracy in hierarchical INVALSI data simulations [5]. Recently, by employing a two-level model, the multilevel regression approach based on LMM is shown to be more effective in explaining the variation in National Examination scores across schools and districts. The model was also able to identify significant variables at both levels thus the resulted model was the best-fitting model with the smallest residual variance [6].

Nevertheless, LMM has certain limitations. The model assumes a linear relationship between predictor and response variables, as well as the normal distribution of residuals and random effects components. When these assumptions are not met, i.e. in scenarios involving complex non-linearity or high-level interactions among predictors, LMM may produce inadequate model and biased interpretations [7]. Additionally, specifying the covariance structure in LMM is often nontrivial and can significantly influence estimation results.

Non-parametric and semi-parametric approaches have been developed to address the limitations of LMM, one of which is the Mixed Effects Regression Tree (MERT). MERT combines the strengths of regression tree models in capturing nonlinear and flexible relationships with the random effects structure of mixed effect linear models to accommodate correlations within clusters [8] - [10]. Unlike parametric approaches (e.g., LMM), MERT does not assume a specific functional form between predictor and response variables, making it highly useful in situations where the relationship structure is unknown or difficult to model explicitly. The MERT approach also significantly outperformed conventional mixed linear models in longitudinal data with subgroup heterogeneity and able to reduce the Mean Squared Error (MSE) by up to 15% [11]. Thus reinforce the conclusion that MERT provides better predictive performance when the relationships between variables are difficult to model parametrically [12].

The development of MERT in recent years has shown great potential in longitudinal data and panel data analysis. Various variants have been created, including integration with ensemble algorithms such as Random Forest and Boosting to improve predictive accuracy and model stability [13], [14]. Nonetheless, both methods are considered as black box approaches which require more analysis to interpret the model. Further improvements are the integration of MERT with Bayesian frameworks [15] and regularization models (penalized likelihood) to enhance interpretability and enable variable selection within the generated regression tree structure [16]. Recently, the 3Trees-EvTree and 3Trees-CTree approaches were developed to overcome the limitations of the CART algorithm, which tends to be biased in split selection and prone to overfitting. These methods have proven to be more accurate and stable, particularly in the context of semi linear models and complex hierarchical data structures [17].

Based on the above, the choice between LMM and MERT highly depends on the characteristics of the data and the objectives of the analysis. We restrained our study to only employ the LMM and MERT as the previous mentioned approaches were required advanced algorithms which might be impractical in some settings. Here, LMM remains the primary choice in cases where linear model assumptions are acceptable, in contrast, MERT offers a more flexible alternative for handling complex data structures and nonlinear relationships that are difficult to model explicitly. Thus, a study to model data structures with clusters and fixed effects should carefully choose which approach is more appropriate. This is due to the unknown nature of the exact structure of the relationship between predictor and response variables.

Dairy milk is one of the most important food sources for a healthy life [18]. Especially in rural areas, farmers often sell fresh milk as a source of daily income [19], [20]. The dairy industry is known to provide employment and has a positive impact on improving GDP [21]. In Indonesia, recently, the dairy sector has become more pivotal due to realization of Free Nutritious Meal Program (Makan Bergizi Gratis/MBG) of current President of Indonesia which include milk as part of the menu. The MBG in Indonesia is aimed to improve the wellbeing of children and reduce the prevalence of malnutrition problems [22] which directly improve the attainment of Goal 3: Good Health and Well-Being, in United Nation's Sustainable Development Goals (SDGs) [23]. In the case that the meal is given to children and pregnant woman from low socio-economic status, the program will also directly help Indonesia progressing in Goal 2: Zero Hunger in SDGs. Nevertheless, due to limited domestic dairy production capacity [24], more insight from multi-disciplinary studies is needed, including approaches from machine learning models utilizing recently available datasets.

The majority of previous studies in the field of animal husbandry are still limited to simple linear models or descriptive analysis, and have not utilized the potential of the LMM and MERT approaches in handling hierarchical data structures. This research specifically compares two modeling approaches, LMM and MERT, in analyzing the relationship between dairy cow productivity and various influencing factors within the context of dairy farming in Indonesia. The dataset used was based on the 2024 National Dairy Productivity Survey conducted by the Ministry of Agriculture, Indonesia (MoA). This study is novel as there has been no prior research comparing the application of these two models in dairy cow productivity data in Indonesia. It is because the number of research done based on national survey data with aim to understand key factors for

improving Indonesia's dairy productivity is still limited. Moreover, the available data do not provide sufficient evidence to ascertain whether the relationship between the response variables and the predictors is linear. Therefore, a nonlinear model, namely the (MERT), was employed for further investigation. Let me know if you'd like it tailored to a specific field (e.g., psychometrics, statistics, education). In addition, the data cannot be ascertained whether the relationship between the response variable and the predictor variable is linear, so it is explained using a non-linear model, namely MERT. Therefore, this study is a new contribution to the development of applied statistical methods in the livestock sector in Indonesia.Thus, the use of LMM and MERT models on national survey data of dairy cows' productivity in this research is expected to provide insights into analytical tools that can be utilized for agricultural research, particularly in Indonesia. Findings of this study could also serve as basis for policy recommendations for improving dairy production in Indonesia.

2. Methods

2.1 Data and Research Methods

As mentioned previously, the data used in this study were based on a survey concerning dairy cow milk productivity done by the Ministry of Agriculture, Republic of Indonesia, in 2024. The survey aimed specifically to estimate the average of dairy productivity in Indonesia. The sampling unit was dairy farms recorded in IDENTIK PKH, a database of cattle and dairy farms managed by Directorate General of Livestock and Animal Health, MoA. The survey was conducted in November 2024 with reference date October 31, 2024. Here, farmer was asked regarding condition of their herd and milk production as if it was on reference date. A total of 1020 dairy farm samples were planned to be collected under systematic random sampling design with 20-25% sampling error. Nevertheless, due to changes in farms in the time of survey, e.g. farms could not be found, farms have been closed, or enumerators could not have access to farms, the observations collected were only 1003 dairy farms. Samples were collected from farms in Jawa Barat, Jawa Tengah, DI Yogyakarta, and Jawa Timur province. In addition to the volume of milk production on every farm, the survey also collected information assumed to be related to the amount of milk production. This assumption was the basis motivation of this study in which we are seeking to model and provide interpretation of the relationship. Details of the variables used in this study can be seen in Table 1.

The first step taken in this study is the data preprocessing stage. Data preprocessing aims to clean data, prepare data in a more appropriate format, and improve model performance and accuracy. Data preprocessing includes deleting missing values and reformatting the naming of factors in each variable to be consistent [25]. If the response variable contains outliers, then the Box-Cox transformation [26]. Box-Cox transformation is also needed to meet the normality assumption required for the LMM model approach [27]. However, in the MERT model, the transformation process is not carried out on the response variable because it is not very sensitive to parametric assumptions [9], [10].

Table 1. Research Variables

Variables	Notes
Daily Milk Production (Y)	Volume of milk production per cow per day (liter/cow/day)
Farm location altitude (X_1)	Three-category classification based on altitude of farm: Highlands (above 700 meters above sea level) (H), Midlands (400 to 700 meters above sea level) (M), or Lowlands (below 400 meters above sea level) (L) *
Bull population (X_2)	Heads, as of October 31, 2024
Dairy heifer population (< 2-year-old) (X_3)	Heads, as of October 31, 2024
Lactating dairy cow population (X_4)	Heads, as of October 31, 2024
Dry cow population (X_5)	Heads, as of October 31, 2024
Cull cow population (X_6)	Heads, as of October 31, 2024
Population of lactating cows in previous week (X_7)	Heads
Milk production in previous week (on average) (X_8)	Liter
Lactating period (on average) of a cow in a year (X_9)	Days
Dry period (on average) of a cow in a year (X_{10})	Days
Management system (X_{11})	Two categories: Intensive (I)* or Semi-Intensive (SI).
Cattle housing type (X_{12})	Five categories: 1. Closed pen (CP), 2. Free stall (FS), 3. Tethered (T), 4. Combination of tethered and closed pen (CTCP). 5. Combination of tethered and free stall (CTFS)*.
Ownership of Forage Land (X_{13})	Two categories: Yes or No*
Amount of Forage Feed per Cow per Day (X_{14})	Kilograms
Amount of Concentrate Feed per Cow per Day (X_{15})	Kilograms
Amount of Agricultural By-Products Feed per Cow per Day (X_{16})	Kilograms
Regular Vaccination Status (X_{17})	Two categories: Regularly and Occasionally (RO) or Never (N)*.
Type of Milking Technique Used (X_{18})	Three categories: Manual by hand (M), using Machine (MC), manual and mechanical (MM)*.
Did Cow have Reproductive Disease this past year (X_{19})	Two categories: Yes* or No.
Did Cow have Udder Inflammation this past year (X_{20})	Two categories: Yes* or No.

*defined as reference category.

The flow diagram of works done in this study is presented in Figure 1.

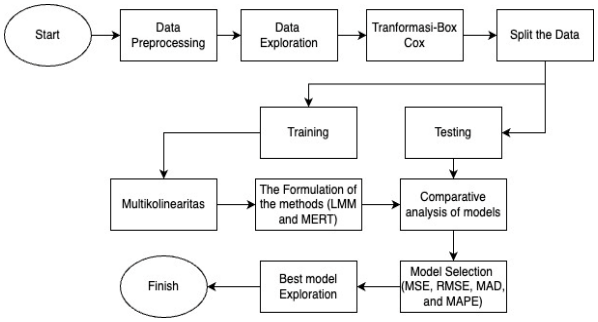


Figure 1. Research Flowchart

The second step is to divide the data into 80% for training data and 20% for testing data, according to the hold-out method approach [28]. In splitting the training and testing data, we pay attention to the representation of observations at the sub-district level, i.e. at least two observations from each sub-district are in training data. This is done so that the model can be trained well to predict testing data. The third step is checking multicollinearity using the Variance Inflation Factor (VIF). The formula of the VIF value is presented in Formula 1 [29].

$$VIF = \frac{1}{1-R_j^2}, j = 1,2,\cdots,k \quad (1)$$

R_j^2 is the coefficient of determination of the j^{th} predictor variable against all other predictor variables. If a predictor variable is not correlated with the others, then the R_j^2 value will approach zero, and the VIF value will be 1. However, if the VIF value is more than 1, it will indicate a relationship between the variables and the others. For example, if the R_j^2 = 0.95 (very strong relationship), then the VIF will be 20 [30]. To supplement multicollinearity checking in predictor variables, a correlation matrix is provided. A correlation matrix is a table that visually summarizes the pairwise relationships among multiple variables [31], commonly constructed based on Pearson correlation coefficients (ρ), which quantify the linear association between variable pairs [32]. The value of ρ = 1 indicates a perfect positive linear correlation, while ρ = −1 signifies a perfect negative linear correlation. When ρ = 0, it suggests no linear correlation, although the variables may still be nonlinearly related [33]. Variables with VIFs above 5 or a Pearson correlation close to 0.8 are considered to exhibit potential collinearity [34] - [37]. To address this, those variables will be removed from the model, except for the one that has the strongest correlation with the response variable. After that, the VIF and correlation matrix are re-examined to ensure that multicollinearity has been resolved. Utilizing the variables that do not provide multicollinearity conditions, the fourth step is continued, namely modeling using LMM and MERT. Subsequent to the modelling, a comparison of the performance of the two models is conducted through the utilization of evaluation metrics.

The evaluation of the predictive model's performance is imperative for the determination of its capacity to represent the actual data. The evaluation of models employs a range of metrics which are done to the testing data, including Mean Squared Error of Prediction (MSEP), Root Mean Squared Error of Prediction (RMSEP), Mean Absolute Deviation (MAD), and Mean Absolute Percentage Error (MAPE). MSEP quantifies the proximity of data points to the regression line, with lower values indicating enhanced accuracy [38]. RMSEP is a measure of the average of the squared differences between the actual and predicted values; the smaller the value, the higher the model accuracy [39]. MAPE has been shown to calculate the average percentage error between the predicted and actual values, thereby providing a relative measure of predictive accuracy that can be used to compare model performance across datasets [40]. The calculations for each of these measures are presented in Formulas 2 to 5 [41].

$$MSEP = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (2)$$

$$RMSEP = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (3)$$

$$MAD = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (4)$$

$$MAPE = \sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{\hat{y}_i}\right| \times 100\% \quad (5)$$

y_i is the actual value and $\hat{y}_i$ is the predicted value of the model used. We repeated the LMM and MERT modeling with randomization on different training and testing data and quantified them. This is to provide a view of the consistency of the two methods in modeling dairy productivity. The Permutational Multivariate Analysis of Variance (PERMANOVA) test was conducted to test whether the LMM and MERT performances were statistically significantly different or not [42].

2.2 Linear Mixed Models (LMM)

Linear Mixed Models (LMM) are an extension of the linear regression model that allows the combination of fixed effects and random effects in one analysis framework [43]. The use of random effects in this model is intended as an effort to capture the diversity in the response variable that can be explained by a certain level or cluster. Thus, this model is very suitable for data that has a hierarchical structure or repeated measurements. An example of the use of random effects is when there is an assumption that different livestock maintenance ecosystems will have an impact on the phenotype of the animal. Another example is when giving different diets to livestock is thought to have an impact on the production of the animal. The addition of random effects in the model allows for increased predictive ability in the model; this is because the

variation in the response variable can be better explained than if it is only explained through the residuals. The LMM model can be written with Formula 6.

$$y_{ij} = \beta_0 + \beta_1 x_{1,ij} + \cdots + \beta_k x_{k,ij} + b_j + \varepsilon_{ij} \quad (6)$$

y_{ij}: response of the i^{th} individual in the j^{th} group, β_0: fixed intercept, $\beta_1, \cdots, \beta_k$: fixed coefficients for each predictor variable, $x_{1,ij}, \cdots, x_{k,ij}$: values of the 1st to kth predictor variables in the i^{th} individual in the j^{th} group, $b_j \sim N(0, \sigma_b^2)$: random effect of the j^{th} group, and $\varepsilon_{ij} \sim N(0, \sigma^2)$: residual errors that are independent of b_j.

In this case, the y_i values from a particular cluster will have a different model compared to other clusters. This condition is true if the variables in the same trait have moderate to high correlation [44]. Parameter estimation in LMM is often done using the Maximum Likelihood (ML) or Restricted Maximum Likelihood (REML) method [43]. LMM allows predictions using more diverse predictor variables (a mixture of numeric and categorical variables). Recent research on LMM highlights the importance of decomposing the variation explained by fixed and random effects separately in one model to understand how much each covariate contributes to data variability [45]. With this flexibility, LMM has become a very important tool in various fields, from genetics to social sciences and engineering.

2.3 Mixed Effect Regression Tree

The model has been developed to accommodate data characterized by a hierarchical or grouped structure, incorporating non-linear relationships and intricate interactions between variables. The MERT model is presented in Formula 7 [8].

$$y_{ij} = f(x_{ij}) + b_j + \varepsilon_{ij} \quad (7)$$

y_{ij} is the response of the i^{th} observation in the j^{th} group, $f(x_{ij})$ is a tree-based non-parametric regression function to capture the non-linear relationship in the fixed effects for the covariate x_{ij}, b_j is the random effect for the j^{th} group, and ε_{ij} denotes the residual of the i-th observation in the j^{th} group, $b_j \sim N(0, \sigma_u^2)$, $\varepsilon_{ij} \sim N(0, \sigma^2)$. In summary, in Formula 7, when compared to Formula 6, the linear structure for $\beta_0 + \beta_1 x_{1,ij} + \cdots + \beta_k x_{k,ij}$ is replaced by $f(x_{ij})$ which is a form of tree-based non-parametric regression. Parameter estimation in MERT focuses on the separation between fixed and random effects components. The fixed effect component ($f(x_{ij})$) is estimated using the regression tree algorithm. MERT utilizes the iterative Expectation-Maximization (EM) algorithm.

To implement the Linear Mixed Models (LMM), we employed the lme4 package in R , specifically using the lmer() function to estimate fixed and random effects efficiently [46]. This package is widely used in statistical modeling due to its robust handling of hierarchical data structures. For the Mixed Effects Regression Tree (MERT), we utilized the glmertree package in R, which implements the Generalized Linear Mixed Model Tree algorithm as proposed by Fokkema and Zeileis [10]. This method integrates tree-based partitioning with mixed-effects modeling, allowing for the identification of subgroups with distinct fixed effects while accounting for random effects in the data. All computational analyses were conducted in R (version 4.4.2), and scripts were executed on a standard computing environment (Intel Core i7, 64GB RAM) to ensure reproducibility and scalability for larger datasets.

3. Results and Discussions

3.1 Exploratory Data Analysis and Preprocessing Procedures

Table 1 outlines the predictor variables used in the analysis, which include a combination of numerical and categorical types. To better understand how the categorical predictors relate to the response variable, their distributions are visualized using box plots. Meanwhile, Figure 2 highlights the relationship between the altitude variable (X_1) and the response variable (Y), revealing potential patterns that may inform further modeling.

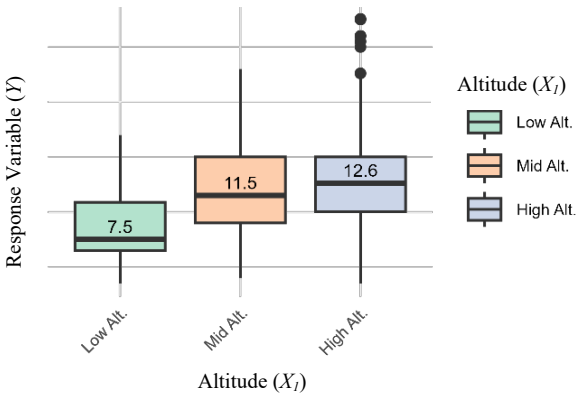


Figure 2. Box-plot the relationship between the altitude variable (X_1) and the response variable (Y)

As shown in Figure 2, dairy cows raised in lowland areas (< 400 m) exhibit lower and more similar milk productivity. Productivity improves in mid-altitude regions (400–700 m), showing both higher yields and greater variability. High-altitude areas (> 700 m) demonstrate the highest potential, though with substantial variation and a few exceptionally high-producing cows. This pattern indicates that milk productivity tends to increase with farm elevation but also becomes more dispersed.

The variable X_8, which represents Milk production in the previous week (on average), exhibits a high degree of variability in its distribution. According to the descriptive statistics, X_8 ranges from 0 to 576.14 liters, with a mean of 46.21 liters and a standard deviation of 55.99 liters, indicating a widespread and the potential presence of outliers. The scatterplot between X_8 and Daily Milk Production reveals a distribution pattern that

tends to decline as X_8 increases, although no clear linear relationship is observed (see Figure 3).

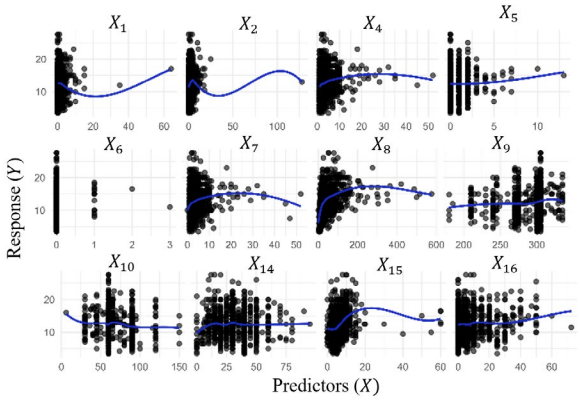


Figure 3. Scatter plots show the relationship between numerical predictor variables and the response variable.

This finding is consistent with previous study [47], which stated that milk production in the previous week can only partially explain the current production due to short-term influences such as thermal stress, cow health conditions, and fluctuations in daily feed intake. In contrast, the variable X_{15}, which measures the Amount of Concentrate Feed per Cow per Day, ranges from 0 to 60 kg, with a mean of 6.90 kg and a standard deviation of 5.17 kg. The median value of 6.00 kg and the upper quartile of 9.00 kg indicate that most cows receive a moderate amount of concentrate feed.

The scatterplot between X_{15} and Daily Milk Production suggests that increasing the amount of concentrate feed per cow per day has a more consistent impact on increasing average milk production (see Figure 3). This observation aligns with the findings of Molle et al. [48] and Huhtanen et al. [49], who emphasized the importance of concentrate feeding in enhancing lactation efficiency and milk output. Concentrate feed provides the additional energy and protein needed to support milk synthesis, especially when nutritional requirements cannot be met through forage alone.

Therefore, X_{15} can be considered a more reliable variable in explaining the variation in Daily Milk Production. In addition, analysis of variable X_{11} reveals that cows managed under semi-intensive systems generally produce more milk with higher variability compared to those in intensive systems, highlighting the influence of management practices on productivity outcomes. Regarding the type of housing system (X_{12}), free-stall barns are associated with higher milk productivity. Tethered or mixed housing types yield moderate productivity with considerable variation, while close pens exhibit the lowest and most uniform productivity levels.

Animal health factors such as routine vaccination (X_{17}), history of reproductive disorders (X_{19}), and mastitis incidence (X_{20}) also show potential associations with productivity variation. Cows that are regularly vaccinated and free from health issues tend to achieve higher and more symmetrically distributed milk yields compared to other groups. The relationship between numerical predictor variables and the response variable is visualized using scatter plots with trend lines, as presented in Figure 3.

As shown in Figure 3, most relationships between the numerical predictor variables (X) and the response variable (Y) do not follow a strict linear pattern. The blue trend lines, generated using loess smoothing, reveal fluctuations or directional shifts in several variables, such as X_3, X_5, X_7, and X_{15}. These patterns suggest that simple linear models may be insufficient to capture the underlying complexity of the data. Moreover, the dataset exhibits a hierarchical structure, with observations nested within subdistricts. In this context, both Linear Mixed Models (LMM) and Mixed Effects Regression Trees (MERT) are appropriate modeling approaches, as they account for data dependence within groups through random effects while also incorporating fixed-effect predictors.

Following the exploration of the relationships between predictor variables and the response variable, it is also important to examine the distribution of the response variable. Figure 4 presents this distribution using both a histogram and a box plot.

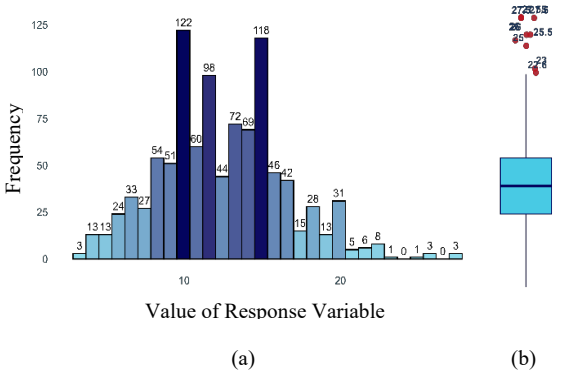


Figure 4. Distribution of the response variable visualized through (a) a histogram and (b) a box plot

As shown in Figure 4(a), the histogram indicates a bimodal distribution, with two prominent peaks around values of 10 and 15, suggesting the presence of two dominant groups within the data. This pattern is further supported by the box plot in Figure 4(b), which shows that the data are spread between approximately 5 and 25, with a few outliers above 25 marked by red dots. The median falls below 15, indicating a slight right-skewness in the distribution. Box-Cox transformation was applied to the response variable due to the presence of outliers and an asymmetric distribution. The transformed response variable is denoted as Y^*, using a lambda value of $\lambda = 0.5$. Formulas 8 and 9 represent the transformation and back-transformation processes, respectively.

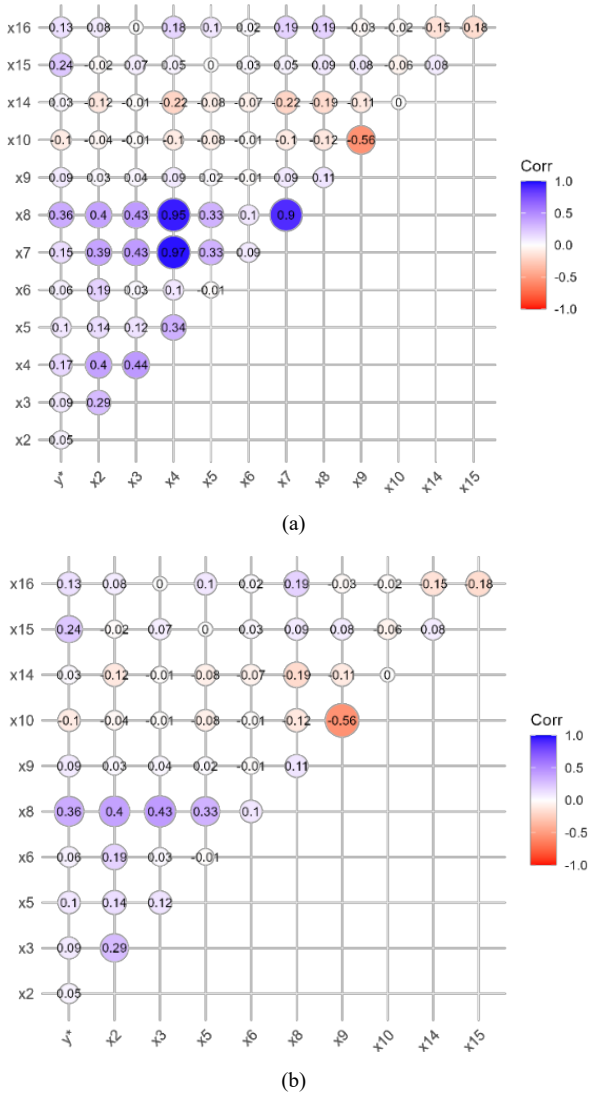
$$Y^* = Y^{0.5} = \sqrt{Y} \quad (8)$$

$$Y = (Y^*)^2 \quad (9)$$

Following the transformation, the next step involved splitting the dataset, as described in the methodology section. Within each subset, a multicollinearity diagnostic was performed to assess potential collinearity issues among the numerical predictor variables using the Variance Inflation Factor (VIF). The VIF values are presented in Table 2.

Table 2. Variance Inflation Factor (VIF) values for numerical predictor variables.

Variable	Correlation between X and Y	VIF (before elimination)	VIF (after elimination)
X_2	0.046	1.269	1.265
X_3	0.091	1.294	1.274
X_4	0.169	29.103	-
X_5	0.099	1.143	1.137
X_6	0.057	1.046	1.046
X_7	0.151	16.303	-
X_8	0.358	10.100	1.620
X_9	0.093	1.505	1.502
X_{10}	-0.102	1.487	1.487
X_{14}	0.033	1.106	1.091
X_{15}	0.239	1.083	1.071
X_{16}	0.132	1.108	1.105



(a)

(b)

Figure 5. Correlation matrix, (a) before elimination, (b) after elimination.

Meanwhile, the correlation matrices before and after the elimination process are presented in Figure 5. As shown in Table 2, strong correlations were observed among variables X_4, X_7, and X_8. Variable X_8 was retained due to its stronger correlation with the response variable compared to the other two.

This selection process resulted in 18 predictor variables to be used in the subsequent LMM modeling.

3.2 Linear Mixed Model (LMM)

Here, we fit the 18 predictor variables free from multicollinearity to dairy productivity with LMM. We used subdistricts as random effects; different subdistricts will have different effects on the response variable. Table 3 presents the fixed effects estimates from the Linear Mixed Model (LMM), which evaluates the influence of various predictors on dairy cow productivity. Several variables are found to be statistically significant. Farm location in highland areas (X_1(H)) shows a positive coefficient of 0.287 ($p-value = 0.049$), indicating that cows raised at higher altitudes produce more milk compared to those in lowlands (reference category), while the other variables remain unchanged. This finding aligns with West et al. [50], who reported that cooler microclimates in highland areas reduce thermal stress and improve milk yield. Likewise, the semi-intensive management system (X_{11}(SI)) has a significant positive effect ($\beta_{11} = 0.155, p-value = 0.020$), supporting Brandt et al. [51], who emphasized the benefits of combining controlled feeding and free movement in enhancing physiological performance in dairy cows.

Table 3. Estimated Effects of Fixed Factors in the LMM

Variable	Est.	Std.error	t value	p-value
Intercept	1.696	0.782	2.167	0.031*
X_1 (H)	0.287	0.145	1.981	0.049*
X_1 (M)	0.228	0.144	1.589	0.113
X_2	-0.013	0.006	-2.059	0.040*
X_3	-0.005	0.004	-1.254	0.210
X_5	0.018	0.015	1.180	0.238
X_6	0.191	0.122	1.567	0.117
X_8	0.003	0.000	8.671	0.000*
X_9	0.000	0.001	0.229	0.819
X_{10}	0.001	0.001	0.402	0.688
X_{11} (SI)	0.155	0.066	2.342	0.020*
X_{12} (CP)	0.625	0.499	1.251	0.211
X_{12} (FS)	0.301	0.558	0.540	0.589
X_{12} (T)	0.597	0.456	1.310	0.191
X_{12} (CTCP)	0.479	0.650	0.737	0.461
X_{13} (Yes)	-0.014	0.073	-0.189	0.850
X_{14}	0.001	0.002	0.579	0.563
X_{15}	0.013	0.003	3.931	0.000*
X_{16}	0.006	0.002	3.367	0.001*
X_{17} (Yes)	0.125	0.136	0.923	0.356
X_{18} (M)	0.409	0.474	0.862	0.389
X_{18} (MC)	0.473	0.481	0.984	0.325
X_{19} (No)	0.021	0.040	0.527	0.598
X_{20} (No)	-0.040	0.040	-1.005	0.315

*significant at $\alpha = 5\%$.

Furthermore, continuous variables such as the previous week's milk production (X_8), the amount of concentrate feed (X_{15}), and agricultural by-products (X_{16}) also show

strong positive effects with highly significant coefficient. The effect of past production (X_8) confirms that historical productivity is a reliable indicator of future performance, as discussed by Zang et al. [52] in their persistence-of-lactation model. Nutritional factors such as concentrate and agricultural by-product intake highlight the role of feed quality in sustaining high productivity, consistent with Abebe et al. [53], who stressed the importance of ration formulation in optimizing milk output. Conversely, the bull population (X_2) negatively impacts productivity ($\beta_2 = -0.013, p - value = 0.040$), which may be attributed to the dilution of productive female-to-total population ratio, as observed by Rehman et al [54].

As shown in Table 3, although some variables do not exhibit statistically significant effects, they are retained in the model due to their substantive relevance in representing the livestock population structure, which may influence production patterns and farm management practices. Subsequently, the variance of the random effects is estimated to capture the unobserved heterogeneity across groups that are not explained by the fixed effects. This variance reflects the extent to which average effects differ among subdistricts. Table 4 presents the estimated variance components of the random effects.

Table 4. Estimated Variance Components of the Random Effects

Variance Component	Estimated Variance	Std.dev
Subdistrict	0.054	0.232
Residual	0.201	0.449

The estimated variance components indicate that the intercept variance across subdistricts is 0.054, while the residual variance at the individual level is 0.201. These results suggest the presence of heterogeneity in baseline values (intercepts) between subdistricts. Based on these values, the proportion of total variance attributable to differences between subdistricts is calculated using the Intra-Class Correlation (ICC). The ICC measures the degree of similarity or homogeneity in response values within a group or cluster [55] and is computed using Foemula 10.

$$ICC = \frac{\sigma_\alpha^2}{\sigma_\alpha^2+\sigma_\varepsilon^2} = \frac{0.054}{0.054+0.201} = 0.212 \quad (10)$$

σ_α^2: the random effect variance term and σ_ε^2: the error variance term from LMM (fitting a random intercept only). With an ICC value of 0.212 (greater than 0.05) [56], the results indicate that a substantial proportion of the variation in the response variable can be attributed to differences between subdistricts. This supports the use of a multilevel modeling approach over a standard linear regression, as it accounts for the hierarchical structure of the data and the dependency among observations within subdistricts.

As shown in Table 3, farms located in highland and mid-altitude subdistricts tend to exhibit higher milk productivity compared to those in lowland regions. Furthermore, semi-intensive management systems appear to be more effective than intensive systems. Other contributing factors include average milk yield in the previous week, the amount of concentrate feed, and the use of agricultural by-products (e.g., tofu waste or fermented straw), all of which play a significant role in enhancing productivity.

In the prediction process, each coefficient is combined with the value of b_j, representing the random intercept for subdistrict j, where $j = 1,2, \ldots ,49$, corresponding to the number of subdistricts. The distribution of the random effect estimates (b_j) from the LMM is visualized in Figure 6.

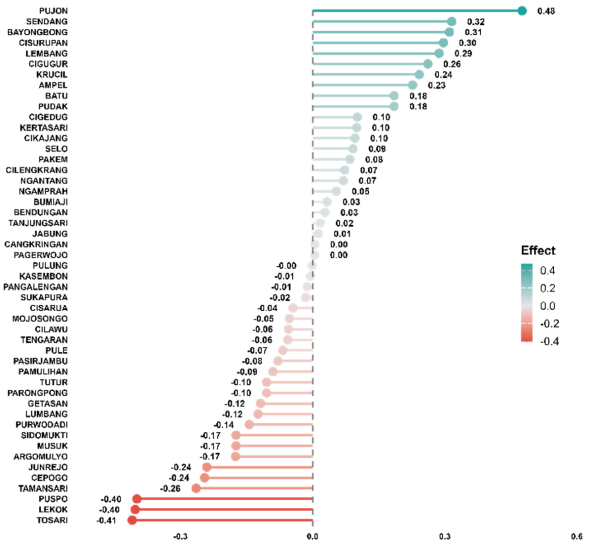


Figure 6. Random intercept estimates from the LMM.

Figure 6 displays the random effects estimated from the LMM. Red indicates negative random effects, while green represents positive values. A positive random effect suggests that a subdistrict contributes additional gains to milk productivity beyond what is explained by fixed factors such as feed or altitude. This implies the presence of localized advantages not captured by the model variables, such as superior farming practices or strong institutional support. Conversely, negative random effects indicate lower-than-expected productivity, suggesting the presence of local challenges. To ensure the appropriateness of the LMM, a normality test was conducted on the model residuals. The normality of residuals was assessed using Quantile–Quantile Plot (Q–Q plot), as shown in Figure 7.

The red dashed line represents the reference line for a theoretical normal distribution. Most of the points closely follow this line, with only slight deviations at the upper-right tail, suggesting that the residuals are approximately normally distributed [57] . This visual assessment is supported by the Lilliefors test, which yielded a p-value of 0.18—greater than the significance level of α = 0.05—indicating no sufficient evidence to reject the null hypothesis of normality. Thus, there is insufficient evidence to reject the null hypothesis that the residuals follow a normal distribution [58]. This

result supports the fulfillment of the normality assumption, which is a key requirement for the application of LMM.

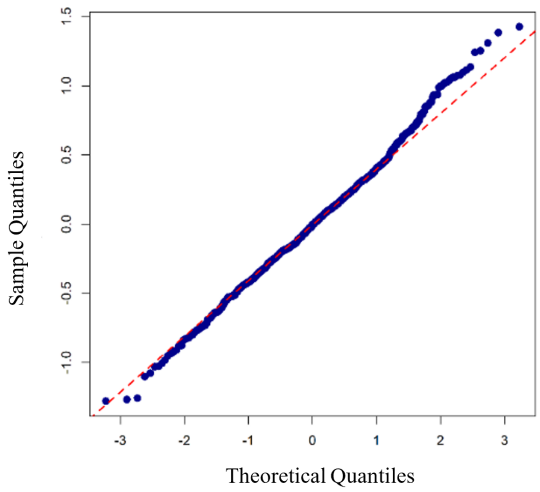


Figure 7. The Normal Distribution Q-Q Plot of residuals.

3.3 Mixed Effect Regression Trees (MERT)

The application of LMM to dairy cow productivity data is limited by its assumption of linearity in the fixed effect structure. When such structure is absent or unknown in the data, the LMM may struggle to provide accurate predictions of the response variable. In contrast, the MERT model adopts a non-parametric approach to estimate fixed effects using tree-based regression. This method relaxes the linearity assumption, thereby enhancing the model's predictive performance while still making it easy to interpret the model built. Unlike the LMM, which produces a linear regression equation, the MERT model is presented in the form of a nonparametric decision tree structure. This approach enables the partitioning of data into homogeneous groups based on predictor variables. The tree structure obtained from the MERT model, using the same set of predictor variables as in the previous LMM (Table 3) and subdistricts as a random effect, is presented in Figure 8.

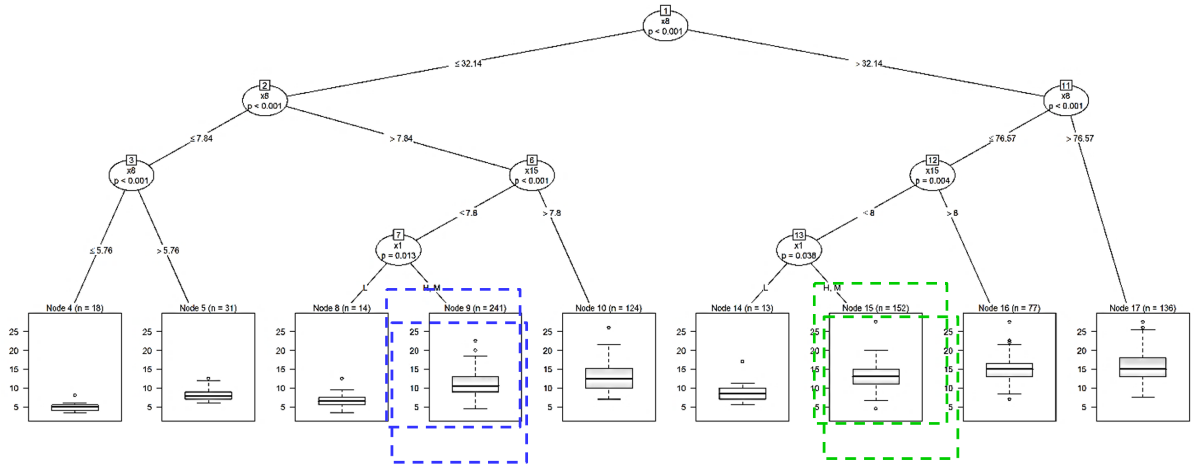


Figure 8. MERT Tree Structure Using All Predictor Variables in the Dataset

As shown in Formula 7, the fixed effect values derived from Figure 8 begin from the leftmost node (Node 4) and proceed sequentially to the right (Node 17). These values are presented in Table 5.

Table 5. Fixed Effect Estimates from the MERT Model

Node	$f(x)$	Node	$f(x)$	Node	$f(x)$
4	5.710	9	11.199	15	13.109
5	8.777	10	12.882	16	14.680
8	8.256	14	10.352	17	15.485

As shown in Figure 8, the resulting regression tree identifies the average milk yield over the past week (X_8), the amount of concentrate feed (X_{15}), and farm altitude (X_1) as the primary splitting variables. These variables are consistent with the findings from the LMM, which also identified them as significant predictors (see Table 3). Variable X_8 appears as the primary divider, indicating that historical milk production serves as an initial indicator of current productivity. However, the effect of X_8 does not act in isolation—as illustrated in the blue box, when past production is low ($X_8 \leq 32.14$), variables X_1 and X_{15} play a critical role in determining milk yield. For instance, farms with low historical production but sufficient concentrate feeding ($X_{15}>$ 7.8) can still achieve competitive productivity levels.

Conversely, when the concentrate feed is limited, geographical characteristics such as highland or mid-altitude areas (X_1) can partially compensate for this constraint, in contrast to lowland areas, which tend to exhibit the lowest productivity. On the other hand, within the group with high historical production ($X_8 >$ 32.14), the effect of concentrate feed remains significant, as illustrated in the green box. Farms with high X_{15} values ($X_{15} >$ 8) demonstrate optimal productivity, while those with lower X_{15} values show a greater dependency on geographic factors, where highland and mid-altitude farms outperform those in the lowlands. The emergence of variable X_1 underscores the role of altitude in influencing milk productivity, potentially due to differences in microclimate, types of

forage available, and levels of thermal stress experienced by the animals. Meanwhile, X_{15} reaffirms the importance of feed management, particularly the provision of concentrate feed as a key nutritional component directly affecting lactation performance.

The prominence of X_8, representing the average milk yield over the past week at the top of the regression tree underscores the critical role of lactation continuity in determining overall milk productivity. This tree-based model highlights the need for more context-specific policy interventions in certain regions, for instance, by promoting adaptive management systems tailored to local agroecological conditions and improving access to high-quality concentrate feed to ensure consistent milk production.

These findings emphasize that improving milk productivity cannot be attributed to a single factor but rather to the interplay of various management and environmental variables. Location-specific interventions are, therefore, essential. Farms located in lowland areas with persistently low productivity should be prioritized for targeted support, particularly by enhancing their access to quality concentrate feed. Moreover, efforts should also be made targeted to make cows able to produce milk as often as it could be. Consistent recording and monitoring of production history should also be strengthened as a predictive tool and early warning system for productivity declines. Farmer empowerment strategies should be aligned with local characteristics, integrating feed management, environmental conditions, and regular performance evaluation. Such an approach would offer a more adaptive and targeted solution for boosting national milk production.

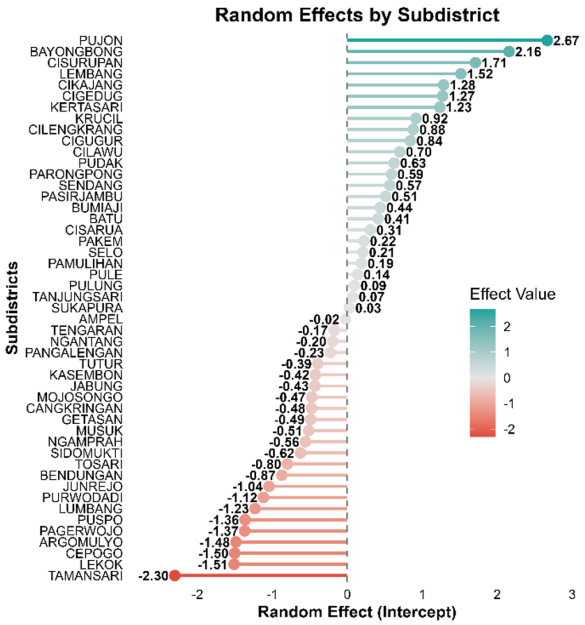


Figure 9. Random Intercept Estimates from the MERT Model

Subsequently, as described in Formula 7, Figure 9 presents the estimated random intercept coefficients for each subdistrict under the MERT model. Figure 9 shows that several subdistricts, such as Pujon, Bayongbong, and Cisurupan, have positive estimated random intercepts, indicating that dairy cows in these areas tend to produce more milk than the overall average. In contrast, subdistricts with negative random effects exhibit lower-than-average productivity.

Nevertheless, with insight from Figure 9, the productivity from these subdistricts could be improved should cows in the regions are able to consistently produce milk and feed with adequate concentrate feed. Following these results with insights from MERT model, efforts on increasing domestic dairy production can be better guided. Should the availability of milk be adequate, the MBG program could then regularly provide milk in the menu which eventually helps the improvement the nutritious diet and reduce malnutrition problems among children. This can be considered as a realization of Goal 3 of SDGs: improving people's health and wellbeing.

3.4 Model Comparison

The comparison between the LMM and MERT models is based on evaluation metrics, including MSEP, RMSEP, MAD, and MAPE, in accordance with Formulas (2) to (5). The evaluation results for MAPE and MSEP are presented in Figure 10, while a summary of all evaluation metrics is provided in Table 6.

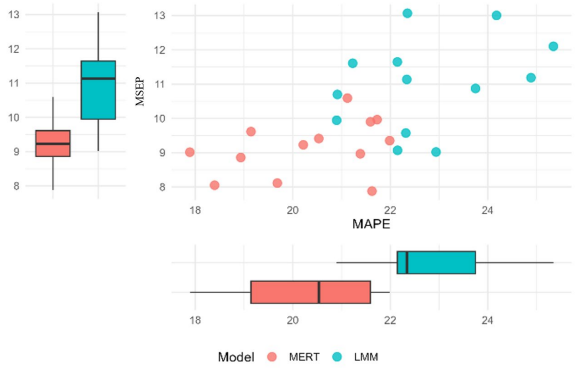


Figure 10. Distribution of MAPE and MSEP Values Across Models

These metrics were calculated using testing data. Repetitions were conducted using the *set.seed()* function during the data-splitting process to assess the consistency of each model's performance.

As shown in Figure 10, the MERT model demonstrates superior predictive performance in terms of MAPE, with a lower median value compared to the LMM model. This indicates that the percentage prediction error of MERT is the smallest among the two models. A similar pattern is observed for MSEP, where MERT yields the lowest mean squared error. This consistency is further supported by the RMSEP and MAD values, both of which show that MERT produces the lowest average squared error, as well as the smallest median and dispersion of prediction errors.

Results presented in Table 6 further suggest that MERT is not only more accurate but also more stable, with

fewer extreme deviations from actual values. In contrast, the LMM model consistently exhibits the poorest performance across all four evaluation metrics. The higher median error and wider spread of residuals indicate lower accuracy and greater vulnerability to outliers or extreme prediction errors.

Table 6. Summary of Evaluation Metrics for Model Comparison

No	Evaluation Metrics		LMM	MERT
1.	MSEP	Min	9.021	7.880
2.		Average	10.995	9.151
3.		Median	11.137	9.229
4.		Max	13.066	10.593
5.	RMSEP	Min	3.003	2.807
6.		Average	3.310	3.022
7.		Median	3.337	3.038
8.		Max	3.615	3.255
9.	MAD	Min	2.369	2.160
10.		Average	2.547	2.312
11.		Median	2.542	2.299
12.		Max	2.744	2.500
13.	MAPE	Min	20.900	17.889
14.		Average	22.726	20.326
15.		Median	22.338	20.533
16.		Max	25.344	21.987

Overall, the evaluation results indicate that the MERT model is the most optimal to model dairy productivity data in terms of predictive accuracy and consistency based on the metrics used. This conclusion is further supported by the results of the PERMANOVA test, which was conducted to assess differences in predictive performance between MERT and LMM statistically. The test yielded a pseudo-F value of 224.700 and a p-value of 0.001, which is below the significance threshold of α=0.05. These results confirm that there is a statistically significant difference in predictive performance between the two models. Furthermore, compared to Table 3, in which one should carefully consider all 18 predictor variables to determine key factors, result from MERT model is simpler, easier and ready to use (see Figure 8). This conclusion was also substantiated by Hajjem [59] in which the MERT proven to be superior compared to LMM in prediction power.

4. Conclusions

This study compares two modeling approaches, Linear Mixed Model (LMM) and Mixed Effects Regression Tree (MERT), in analyzing dairy productivity based on the 2024 National Dairy Survey data. Here, LMM is effective in explaining the influence of fixed variables and capturing variations between subdistricts, making it suitable for inferential analysis. In contrast, MERT is more flexible in identifying non-linear patterns and complex interactions, demonstrating better predictive performance based on MSEP, RMSEP, MAD, and MAPE, while still convenient to interpret. Key factors affecting dairy productivity are the altitude of dairy farms, farming systems, amount of concentrate feed given, and production history. This finding highlights the importance of policy intervention aimed at improving milk production sustainability and the supply of cow's feed particularly the concentrate feed. Should a new dairy farm be built, then high altitude regions must be a priority due to its determining contribution to milk productivity. Combining these insights with the Government's plan to import cows, it will then increase the domestic supply of milk which in turn increases the nutritious value of MBG menu. The domino effect will then eventually reach the improvement of nation's capacity in fulfilling many SDGs goals, especially the Goal 3 and Goal 2. A follow-up study could explore the utilization of hybrid models combining LMM's inferential strength and MERT's predictive flexibility, as well as expansion into spatial models and ensemble approaches better to capture subdistrict-level patterns, uncertainties, and complex interactions. Bayesian tree based semi-parametric and black-box machine-learning approaches would potentially improve predictive ability of the model. This approach is expected to support more adaptive and data-driven policymaking.

Acknowledgements

The authors would like to acknowledge the Directorate General of Livestock and Animal Health, Ministry of Agriculture, Republic of Indonesia for granting access to the data essential to this study. We are grateful for their assistance and for enabling this research. The authors declare no conflict of interest. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

[1] J. F. Hair Jr. and L. P. Fávero, "Multilevel modeling for longitudinal data: concepts and applications," *RAUSP Management Journal*, vol. 54, no. 4, pp. 459–489, Oct. 2019, doi: 10.1108/RAUSP-04-2019-0059.

[2] F. Rahmawati, K. A. Notodiputro, and L. O. A. Rahman, "A comparison of fixed effect and mixed effect models in analyzing telecommunication products," in *International Conference on Statistics and Data Science 2021*, 2022, p. 020013. doi: 10.1063/5.0111015.

[3] S. Hilbert, M. Stadler, A. Lindl, F. Naumann, and M. Bühner, "Analyzing longitudinal intervention studies with linear mixed models," *TPM Test Psychom Methodol Appl Psychol*, vol. 26, no. 1, pp. 101–119, Mar. 2019, doi: 10.4473/TPM26.1.6.

[4] I. Funatogawa and T. Funatogawa, "Longitudinal Data and Linear Mixed Effects Models," in *Longitudinal Data Analysis: Autoregressive Linear Mixed Effects Models*, I. Funatogawa and T. Funatogawa, Eds., Singapore: Springer Singapore, 2018, pp. 1–26. doi: 10.1007/978-981-10-0077-5_1.

[5] G. Vannucci, A. Gottard, L. Grilli, and C. Rampichini, "Random effects regression trees for the analysis of INVALSI data," in *ASA 2021 Statistics and Information Systems for Policy Evaluation*, Firenze University Press, 2021, pp. 29–34. doi: 10.36253/978-88-5518-304-8.07.

[6] K. Nurfadilah, M. N. Aidi, K. A. Notodiputro, and B. Susetyo, "Multilevel Regressions for Modeling Mean Scores of National Examinations," *BAREKENG: Jurnal Ilmu Matematika dan Terapan*, vol. 18, no. 1, pp. 0323–0332, Mar. 2024, doi: 10.30598/barekengvol18iss1pp0323-0332.

[7] L. Fahrmeir and G. Tutz, *Multivariate Statistical Modelling Based on Generalized Linear Models*. New York, NY: Springer New York, 2001. doi: 10.1007/978-1-4757-3454-6.

[8] A. Hajjem, F. Bellavance, and D. Larocque, "Mixed effects regression trees for clustered data," *Stat Probab Lett*, vol. 81, no. 4, pp. 451–459, Apr. 2011, doi: 10.1016/j.spl.2010.12.003.

[9] A. Gottard, G. Vannucci, L. Grilli, and C. Rampichini, "Mixed-effect models with trees," *Adv Data Anal Classif*, vol. 17, no. 2, pp. 431–461, Jun. 2023, doi: 10.1007/s11634-022-00509-3.

[10] M. Fokkema, J. Edbrooke-Childs, and M. Wolpert, "Generalized linear mixed-model (GLMM) trees: A flexible decision-tree method for multilevel and longitudinal data," *Psychotherapy Research*, vol. 31, no. 3, pp. 329–341, Apr. 2021, doi: 10.1080/10503307.2020.1785037.

[11] M. G. Kundu and J. Harezlak, "Regression trees for longitudinal data with baseline covariates," *Biostat Epidemiol*, vol. 3, no. 1, pp. 1–22, 2019, doi: 10.1080/24709360.2018.1557797.

[12] W. Y. Loh and W. Zheng, "Regression trees for longitudinal and multiresponse data," *Annals of Applied Statistics*, vol. 7, no. 1, pp. 495–522, Mar. 2013, doi: 10.1214/12-AOAS596.

[13] W. Fu and J. S. Simonoff, "Unbiased regression trees for longitudinal and clustered data," *Comput Stat Data Anal*, vol. 88, pp. 53–74, Aug. 2015, doi: 10.1016/j.csda.2015.02.004.

[14] R. J. Sela and J. S. Simonoff, "RE-EM trees: a data mining approach for longitudinal and clustered data," *Mach Learn*, vol. 86, no. 2, pp. 169–207, Feb. 2012, doi: 10.1007/s10994-011-5258-3.

[15] J. L. Speiser, B. J. Wolf, D. Chung, C. J. Karvellas, D. G. Koch, and V. L. Durkalski, "BiMM forest: A random forest method for modeling clustered and longitudinal binary outcomes," *Chemometrics and Intelligent Laboratory Systems*, vol. 185, pp. 122–134, Feb. 2019, doi: 10.1016/j.chemolab.2019.01.002.

[16] S. Nestler and S. Humberg, "A Lasso and a Regression Tree Mixed-Effect Model with Random Effects for the Level, the Residual Variance, and the Autocorrelation," *Psychometrika*, vol. 87, no. 2, pp. 506–532, 2022, doi: 10.1007/s11336-021-09787-w.

[17] Asrirawan, K. A. Notodiputro, B. Susetyo, and S. D. Oktarina, "Modified tree-based selection in hierarchical mixed-effect models with trees: A simulation study and real-data application," *MethodsX*, vol. 14, p. 103312, Jun. 2025, doi: 10.1016/j.mex.2025.103312.

[18] M. Akram, M. Sami, O. Ahmed, P. F. Onyekere, and C. Egbuna, "Health Benefits of Milk and Milk Products," in *Functional Foods and Nutraceuticals*, Cham: Springer International Publishing, 2020, pp. 211–217. doi: 10.1007/978-3-030-42319-3_12.

[19] Food and Agriculture Organization of the United Nations, "The Global Dairy Sector: Facts," Oct. 2016.

[20] John. Moran, *Tropical dairy farming : feeding management for small holder dairy farmers in the humid tropics*. Landlinks Press, 2005.

[21] M. C. R. Oliveros, "The Dairy Industry in Southeast Asia: Perspective, Challenges and Opportunities," in *IOP Conference Series: Earth and Environmental Science*, IOP Publisher, Nov. 2019, p. 012068. doi: 10.1088/1755-1315/372/1/012068.

[22] Soma Rathma Inna, Azhar Aidil, and Uchiyama Tomohiro, "Food preferences in Indonesian schoolchildren and the parents' perspectives on the upcoming nutritious free meal program," in *E3S Web Conf.*, 2024, p. 2004. doi: 10.1051/e3sconf/202457702004.

[23] D. Lopez de Romaña, A. Greig, A. Thompson, and M. Arabi, "Successful delivery of nutrition programs and the sustainable development goals," *Curr Opin Biotechnol*, vol. 70, pp. 97–107, 2021, doi: https://doi.org/10.1016/j.copbio.2021.03.004.

[24] S. Jahroh, J. Atmakusuma, A. Fadillah, and others, "Comparative analysis of dairy farming management and business model between East Java and West Java, Indonesia," *Jurnal Manajemen & Agribisnis*, vol. 17, no. 1, p. 96, 2020.

[25] A. P. Joshi and B. V. Patel, "Data Preprocessing: The Techniques for Preparing Clean and Quality Data for Data Analytics Process," *Oriental journal of computer science and technology*, vol. 13, no. 0203, pp. 78–81, Jan. 2021, doi: 10.13005/ojcst13.0203.03.

[26] M. Kutner, C. Nachtsheim, and J. Neter, *Applied Linear Regression Models*. McGraw Hill, 2004.

[27] A. Soberón and W. Stute, "Assessing skewness, kurtosis and normality in linear mixed models," *J Multivar Anal*, vol. 161, pp. 123–140, Sep. 2017, doi: 10.1016/j.jmva.2017.07.010.

[28] H. Farhood, I. Joudah, A. Beheshti, and S. Muller, "Evaluating and Enhancing Artificial Intelligence Models for Predicting Student Learning Outcomes," *Informatics*, vol. 11, no. 3, p. 46, Jul. 2024, doi: 10.3390/informatics11030046.

[29] N. Shrestha, "Detecting Multicollinearity in Regression Analysis," *Am J Appl Math Stat*, vol. 8, no. 2, pp. 39–42, Jun. 2020, doi: 10.12691/ajams-8-2-1.

[30] S. G. Singh and S. V. Kumar, "Dealing with Multicollinearity Problem in Analysis of Side Friction Characteristics Under Urban Heterogeneous Traffic Conditions," *Arab J Sci Eng*, vol. 46, no. 11, pp. 10739–10755, Nov. 2021, doi: 10.1007/s13369-020-05213-y.

[31] N. B. A. Warif *et al.*, "Copy-move forgery detection: Survey, challenges and future directions," *Journal of Network and Computer Applications*, vol. 75, pp. 259–278, Nov. 2016, doi: 10.1016/j.jnca.2016.09.008.

[32] M. Bolic, "Concepts in performance evaluation and uncertainty analysis," in *Pervasive Cardiovascular and Respiratory Monitoring Devices*, Elsevier, 2023, pp. 1–35. doi: 10.1016/B978-0-12-820947-9.00008-8.

[33] P. K. Choudhary and H. N. Nagaraja, *Measuring Agreement*. Wiley, 2017. doi: 10.1002/9781118553282.

[34] D. S. Young, *Handbook of Regression Methods*. CRC Press, 2017.

[35] T. Kyriazos and M. Poga, "Dealing with Multicollinearity in Factor Analysis: The Problem, Detections, and Solutions," *Open J Stat*, vol. 13, no. 03, pp. 404–424, 2023, doi: 10.4236/ojs.2023.133020.

[36] D. M. Levine, D. F. Stephan, and K. A. Szabat, *Statistics for managers using Microsoft Excel*. Pearson, 2016.

[37] D. Szabo *et al.*, "Admission lactate level and the GRACE 2.0 score are independent and additive predictors of 30-day mortality of STEMI patients treated with primary PCI—Results of a real-world registry," *PLoS One*, vol. 17, no. 11, p. e0277785, Nov. 2022, doi: 10.1371/journal.pone.0277785.

[38] T. O. Hodson, T. M. Over, and S. S. Foks, "Mean Squared Error, Deconstructed," *J Adv Model Earth Syst*, vol. 13, no. 12, Dec. 2021, doi: 10.1029/2021MS002681.

[39] A. Gogineni, I. K. Panday, P. Kumar, and R. Kr. Paswan, "Predicting compressive strength of concrete with fly ash and admixture using XGBoost: a comparative study of machine learning algorithms," *Asian Journal of Civil Engineering*, vol. 25, no. 1, pp. 685–698, Jan. 2024, doi: 10.1007/s42107-023-00804-0.

[40] N.-H. Nguyen, K. T. Tong, S. Lee, A. Karamanli, and T. P. Vo, "Prediction compressive strength of cement-based mortar containing metakaolin using explainable Categorical Gradient Boosting model," *Eng Struct*, vol. 269, p. 114768, Oct. 2022, doi: 10.1016/j.engstruct.2022.114768.

[41] P. Ruba, S. Aarthi, and P. Bhuvaneshwari, "Prediction of ultimate strength for high strength concrete (HSC) using machine learning approaches - optimized by PSO technique," *Discover Computing*, vol. 28, no. 1, p. 17, Mar. 2025, doi: 10.1007/s10791-025-09517-1.

[42] Y. Yue and Y.-J. Hu, "Extension of PERMANOVA to Testing the Mediation Effect of the Microbiome," *Genes (Basel)*, vol. 13, no. 6, 2022, doi: 10.3390/genes13060940.

[43] O. A. Montesinos López, A. Montesinos López, and J. Crossa, "Linear Mixed Models," in *Multivariate Statistical Machine Learning Methods for Genomic Prediction*, Cham: Springer International Publishing, 2022, pp. 141–170. doi: 10.1007/978-3-030-89010-0_5.

[44] G. M. Fitzmaurice and N. M. Laird, "Linear Mixed Models," in *International Encyclopedia of the Social & Behavioral Sciences*, Elsevier, 2015, pp. 162–168. doi: 10.1016/B978-0-08-097086-8.42016-7.

[45] N. Schreck and M. Wiesenfarth, "Decomposition of Explained Variation in the Linear Mixed Model," 2023. [Online]. Available: https://arxiv.org/abs/2205.15090

[46] D. Bates, M. Mächler, B. M. Bolker, and S. C. Walker, "Fitting linear mixed-effects models using lme4," *J Stat Softw*, vol. 67, no. 1, Oct. 2015, doi: 10.18637/jss.v067.i01.

[47] A. Akıllı, H. Atıl, and Ç. Takma, "A Decision Tree Model to Determine Some Environmental Factors Affecting 305-Day Milk Yield in Simmental Cows," *Journal of Agricultural Faculty of Gaziosmanpaşa University*, vol. 39, no. 3, pp. 191–198, 2022, doi: 10.55507/gopzfd.1175502.

[48] G. Molle, M. Decandia, A. Cabiddu, S. Y. Landau, and A. Cannas, "An update on the nutrition of dairy sheep grazing Mediterranean pastures," *Small Ruminant Research*, vol. 77, no. 2, pp. 93–112, 2008, doi: https://doi.org/10.1016/j.smallrumres.2008.03.003.

[49] P. Huhtanen, M. Rinne, and J. Nousiainen, "Evaluation of the factors affecting silage intake of dairy cows: a revision of the relative silage dry-matter intake index," *Animal*, vol. 1, no. 5, pp. 758–770, 2007, doi: https://doi.org/10.1017/S175173110773673X.

[50] S. V Rama Rao *et al.*, "Performance, carcass variables and immune responses in commercial broiler chicks fed graded concentrations of threonine in diet containing sub-optimal levels of protein," *Anim Feed Sci Technol*, vol. 169, no. 3, pp. 218–223, 2011, doi: https://doi.org/10.1016/j.anifeedsci.2011.06.013.

[51] P. Brandt, T. Rousing, M. S. Herskin, E. V Olsen, and M. D. Aaslyng, "Development of an index for the assessment of welfare of finishing pigs from farm to slaughter based on expert opinion," *Livest Sci*, vol. 198, pp. 65–71, 2017, doi: https://doi.org/10.1016/j.livsci.2017.02.008.

[52] H. Zhang *et al.*, "Genetic parameters and genome-wide association analyses for lifetime productivity in Chinese Holstein cattle," *J Dairy Sci*, vol. 107, no. 11, pp. 9638–9655, Nov. 2024, doi: 10.1016/j.jods.2024.10.001.

[53] K. Abebe, B. Tamir, G. Goshu, T. Demissie, and F. Regassa, "Effect of Late Gestation and Early Postpartum Concentrate Supplementation on Milk Production, Milk Composition and Reproductive Performances of HF × Boran Dairy Cows," *Advances in Life Science and Technology*, vol. 61, pp. 28–37, 2017, [Online]. Available: www.iiste.org

[54] H. U. Rehman *et al.*, "Sustainable Strategies for the Development of Livestock Production Sector: Moving Towards Zero Hunger and Food Security," *Journal of Population Therapeutics and Clinical Pharmacology*, vol. 31, no. 7, 2024, doi: 10.53555/jptcp.v31i7.4394.

[55] B. T. West, K. B. Welch, and A. T. Gatecki, *Linear Mixed Models: A Practical Guide Using Statistical Software*, 3rd ed. Taylor & Francis Group, LLC, 2022.

[56] P. Bressoux, "Using multilevel models is not just a matter of statistical adjustment. Illustrations in the educational field," 2020.

[57] C. Feng, L. Li, and A. Sadeghpour, "A comparison of residual diagnosis tools for diagnosing regression models for count data," *BMC Med Res Methodol*, vol. 20, no. 1, p. 175, Dec. 2020, doi: 10.1186/s12874-020-01055-2.

[58] N. Khatun, "Applications of Normality Test in Statistical Analysis," *Open J Stat*, vol. 11, no. 01, pp. 113–122, 2021, doi: 10.4236/ojs.2021.111006.

[59] A. Hajjem, F. Bellavance, and D. Larocque, "Mixed-effects random forest for clustered data," *J Stat Comput Simul*, vol. 84, no. 6, pp. 1313–1328, Jun. 2014, doi: 10.1080/00949655.2012.741599.