# Population dynamics and genetic modelling of performance traits in Vrindavani cattle

# Thesis

**Submitted to the**
**DEEMED UNIVERSITY**
**ICAR-Indian Veterinary Research Institute**
**Izatnagar - 243 122 (U.P.), India**

**Dr. Munish Gangwar**
**Roll No. P-2374**

**IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF**

## Doctor of Philosophy
**(Animal Genetics and Breeding)**

**2025**

भारतीय पशु चिकित्सा अनुसंधान संस्थान
(सम विश्वविद्यालय)
इज्जतनगर -243122, (उ.प्र.), भारत

**DIVISION OF ANIMAL GENETICS**
**INDIAN VETERINARY RESEARCH INSTITUTE**
**(Deemed University)**
**IZATNAGAR - 243 122, U.P., INDIA**

**Dr. G.K. Gaur,**
**Assistant Director General (APB),**
**ICAR, Krishi Bhavan, New Delhi**

**Dated:** 13/02 **2025**

# *Certificate*

*This is to be certified that the research work embodied in this thesis entitled* **"Population dynamics and genetic modelling of performance traits in Vrindavani cattle"** *submitted by* **Dr. Munish Gangwar, Roll No. P-2374**, *for the award of* **Doctor of Philosophy** *Degree in* **Animal Genetics and Breeding** *at ICAR-Indian Veterinary Research Institute, Izatnagar, is the original work carried out by the candidate himself under my supervision and guidance.*

*It is further certified that* **Dr. Munish Gangwar, Roll No. P-2374**, *has worked for more than 30 months in the Institute and has put in more than 300 days attendance under me from the date of registration for the Doctor of Philosophy Degree in this Deemed University, as required under the relevant ordinance.*

**(G.K. Gaur)**
**Chairman**
**Advisory Committee**

# Certificate

*We the undersigned members of Advisory Committee of* **Dr. Munish Gangwar, Roll No. P-2374,** *a candidate for the degree of* **Doctor of Philosophy** *with the major discipline* **Animal Genetics and Breeding,** *agree that the thesis entitled* **"Population dynamics and genetic modelling of performance traits in Vrindavani cattle"** *may be submitted in partial fulfillment of the requirement for the degree.*

*We have gone through the contents of the thesis and are fully satisfied with the work carried out by the candidate, which is being presented for the award of* **Doctor of Philosophy** *Degree of this Institute.*

*It is further certified that the candidate has completed all the prescribed requirements governing the award of* **Doctor of Philosophy** *Degree of the Deemed University, ICAR-Indian Veterinary Research Institute, Izatnagar.*

Signature of External Examiner
Name : Dr. Suresh Malik
Date : 18/04/25

**(G.K. Gaur)**
***Chairman***
***Advisory Committee***
Date : ......................

**MEMBERS OF STUDENT'S ADVISORY COMMITTEE**

**Dr. Pushpendra Kumar**, Principal Scientist & Head
Division of Animal Genetics, ICAR-IVRI, Izatnagar ..........................................

**Dr. Amit Kumar**, Principal Scientist
Division of Animal Genetics, ICAR-IVRI, Izatnagar .......Amit Kumar.......

**Dr. Anuj Chauhan**, Senior Scientist
Livestock Production and Management Section, ICAR-IVRI, Izatnagar .......A Chauhan.......

**Dr. G.V.P.P.S. Ravi Kumar**, Principal Scientist & Head
Animal Biotechnology, NBAGR, Karnal ..........................................

**Dr. Manish Mahawar**, Principal Scientist
Division of Biochemistry, ICAR-CARI, Izatnagar .......Manish.......

# ACKNOWLEDGEMENTS

I would like to express my sincere thanks to almighty who blessed me with the wonderful parents, thanks a lot mummy and papa, for showing faith in me and giving me liberty to choose what I desired. I salute you for the selfless love, care, support and sacrifice you did to shape my life. I would never be able to pay back the love and affection you showered upon me. Also, I would like to express thanks to my brother **Mr Girish Gangwar** and my sister **Mrs Kamla Gangwar** for your unconditional love, empathy and moral support throughout my life. I consider myself the luckiest in the world to have such a supportive family, standing behind me with their love and support. It is his grace bestowed me the patience and courage to successfully complete this dissertation which would have been a tough task.

It is my great pleasure to express deepest gratitude and respect to my advisor, **Dr. G. K. Gaur,** Assistant Director General (APB), ICAR, Krishi Bhavan, New Delhi whose affection, encouragement, guidance, critical appreciation, constructive criticism and scientific temper was always a thrust to usher best out of my capability and to bring laurels to my research carrier. With great pleasure I extend my esteemed sense of gratitude to my advisor for his keen interest and valuable suggestions in planning, friendly cooperation and constant encouragement in executing, sincere and optimistic attitude in analyzing this research work. Thanks, are also due for the real and laudable effort of providing timely arrangement of facilities needed for this research. I feel privileged to have worked under his guidance.

I extend my heartfelt gratitude to **Dr. Sheikh Firdous Ahmad** for his invaluable guidance and support throughout my M.V.Sc and Ph.D. journey. His mentorship, insightful advice, and continuous encouragement have been pivotal to my academic and research growth. I am deeply thankful for his dedication and inspiration, which have greatly influenced my professional development.

I would like to thank members of my advisory committee, **Dr. Pushpendra Kumar**, Principal Scientist and Head of Animal Genetics, **Dr. Amit Kumar**, Principal Scientist, Animal Genetics Division, **Dr. Anuj Chauhan**, Senior Scientist, LPM Section, **Dr. G.V.P.P.S. Ravi Kumar**, Principal Scientist & Head (Animal Biotechnology) NBAGR, Karnal, My reverential thanks to **Director, Joint Director** (Academic) and Scientific Co-ordinator (Academic) of Deemed University IVRI for providing me with necessary facilities during my research work.

I would like to convey my sincere gratitude to the faculty of Animal Genetics Division, ICAR-IVRI, **Dr Subodh Kumar**, Principal Scientist AG Division, **Dr A.K Pandey**, Principal Scientist, **Dr Manjit Panigrahi** Senior Scientist, **Dr. Ayon Tarafdar,** Scientist of LPM Section and other Scientist for their cooperation, insightful comments and help during the course of work.

Words are less to express the help, guidance, cooperation and care of Technical and Nontechnical Staff of the Division of Animal Genetics and LPM Section, ICAR-IVRI. I must express my special gratitude to **Mr. Sonu Kumar**, Technical staff, LPM Section, ICAR-IVRI.

I would like to thanks my seniors and juniors (Drs), **Shiv Kumar Tyagi, Manoj Netra, Tapendra Saini, Rohit Barwar, Abdul Basit Ali, Sonali Nayak, Karan Jain, Roshini Chand, Nitish Gaitri, Anal Bose, Shraddha Dwivedi, Mir Maheroz** and **Mahendra Kumar**.

It is my pleasure to thank my batchmates (Drs) **Amit Kumar, Divya Rajawat, Anuradha Panwar, Chhaya Rani,** and **Sakshi Vaishnav** and for their kind support throughout my research work.

I shall always remember the deep concern, selfless help directory and indirectly by my friends **Vandana Joshi, Vandana Sharma, Chand tanwar, Vishav Jeet Singh** and **Pratibha Yadav** during my research work.

My thesis will be incomplete without saying thanks to **Mr Kuldeep** and **Mr Dharmendra (Chachu)** for their selfless help in setting of this thesis timely.

***Date:*** 13/02/25 ***(Munish Gangwar)***

***Place:*** *ICAR-IVRI, Izatnagar*

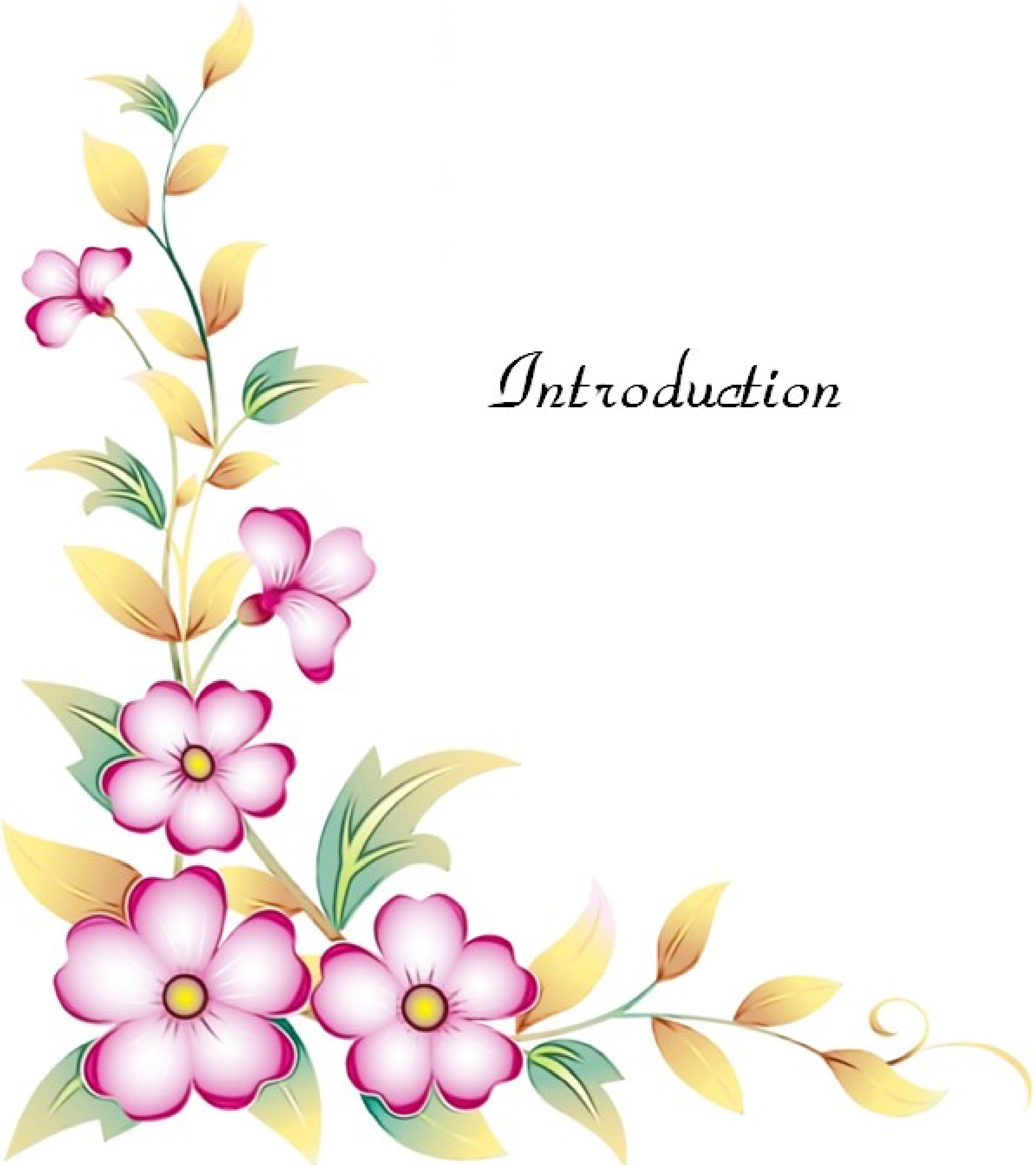
# ABBREVIATIONS

| | |
|---|---|
| AFC | : Age at First Calving |
| ANN | : Artificial Neural Network |
| AR | : Average Relatedness |
| BIC | : Bayesian Information Criterion |
| BLUP | : Best Linear Unbiased Prediction |
| BNN | : Bayesian Neural Network |
| BPN | : Backpropagation Network |
| ICP | : Inter Calving Period |
| CNN | : Convolutional Neural Network |
| DD | : Dry Days |
| EBV | : Estimated Breeding Value |
| GCI | : Genetic Conservation Index |
| GD | : Genetic Diversity |
| GI | : Generation Interval |
| GP | : Gestation Period |
| LL | : Lactation Length |
| LMM | : Linear Mixed Model |
| MAE | : Mean Absolute Error |
| MATLAB | : Matrix Laboratory |
| MLP | : Multilayer Perceptron |
| MSE | : Mean Squared Error |
| Ne | : Effective Population Size |
| POB | : Period of Birth |
| PY | : Peak Yield |
| ReLU | : Rectified Linear Unit |
| REML | : Restricted Maximum Likelihood |
| RNN | : Recurrent Neural Network |
| RMSE | : Root Mean Squared Error |
| SP | : Service Period |
| SOB | : Season of Birth |
| TLMY | : Total Lactation Milk Yield |

# LIST OF TABLES

# LIST OF FIGURES

# CONTENTS



# Introduction

# 1

# INTRODUCTION

India is endowed with a large inventory of farm animal genetic resources (FAnGR) and is one of the 17 biodiversity hotspots of the world, with immense diversity in its flora and fauna resources (Baltenweck *et al.,* 2020). India currently has one of the largest and most diverse livestock populations globally, with 230 registered breeds of livestock, poultry, and dogs (NBAGR, 2025). Over the past several decades, the breeding program for dairy cattle in India has focused mainly on increasing production of milk and milk products to support its ever-growing human population. India is currently the largest producer of milk globally, accounting for around 24.64% contribution to total production worldwide with 239.3 million metric tons (BAHS, 2023). The Indian dairy sector has shown 3.78% annual growth rate, and a per capita availability of around 471 gram milk per day with Uttar Pradesh, Rajasthan, and Madhya Pradesh being the top contributors. Various crossbreeding experiments were started around independence India wherein the adaptation of crosser of various indigenous cattle breed was evaluated after mating with high producing exotic breeds like Holstein – Friesian, Jersy, and Brown-Swiss. On similar lines, the Government of India (GoI) launched a program to cross indigenous Hariana cattle with Holstein-Friesian, Brown Swiss, and Jersey breeds at multiple locations and assess their adaptability and production performance under Indian conditions. This initiative under the All-India Coordinated Research Project (AICRP) led to the development of an elite dairy cattle population named Vrindavani, capable of producing 3,000 kg or more milk per lactation (Singh *et al.,* 2011).

Variation plays a significant role in improving animal germplasm through meticulous selection and mating plans. Understanding population structure and evaluating its variability is crucial for designing effective breed improvement programs (Cervantes *et al.,* 2016). Reliable information on population parameters like inbreeding, effective population size, generation interval, and other similar parameters is also important. Parameters of population structure enable a better understanding of past (recent and distant) events that may have influenced its genetic history and helped in evolution of the current genetic makeup (Ballou *et al.,* 2010).

Scientific management of an animal population considers various genetic parameters, which help predict the probability of genetic erosion in a population. Studying population structure provides insight into significant events and the evolutionary timeline of the population (Špehar *et al.,* 2022). Genetic parameters can help predict the population potential of a population indirectly, with decreased genetic variability and inbreeding resulting in lowered reproduction performance (Poyato-Bonilla *et al.,* 2020). Continuous monitoring of changes in genetic diversity and population structure is necessary to gain insights into population evolution and conservation strategies. However, complete pedigree information is required for efficient and reliable predictions regarding population parameters and diversity estimates. Inaccurate or incomplete information can lead to under- or over-estimation of these parameters, leading to unreliable estimation and spurious inferences. In this context, genomic selection has become a game-changing approach, allowing for more accurate predictions of breeding values and the overall potential of populations. However, its effectiveness heavily depends on having reliable pedigree records and a thorough understanding of population structure. When pedigree information is missing or inaccurate, often the case with animals introduced into the field or those with limited historical records, it becomes difficult to accurately calculate inbreeding levels and other key genetic parameters. This not only reduces the accuracy of breeding value predictions but also puts the genetic diversity and productivity of the population at risk. To fully harness the benefits of genomic selection and support long-term genetic progress in animal populations, it is crucial to maintain detailed pedigree records and continually improve methods for analysing population structure and dynamics (Cervantes *et al.,* 2011).

Milk production traits in dairy cattle are shaped by a complex interaction of quantitative genes with minor effects and environmental factors. Genetic makeup serves as the foundational resource for improving milk production performance, while environmental factors such as nutrition, management practices, and climate significantly impact these traits (Ibrahim *et al.,* 2024). Proper nutrition is vital for achieving optimal milk yield, while poor management and seasonal fluctuations can hinder milk production (Zamoraro *et al.,* 2023). Similarly, growth traits, typically expressed as body weight and growth rate, are influenced by both genetic and various non-genetic factors, including the year and season of birth, sex, age of the dam at calving and sire interaction (Neser *et al.,* 2000). Reproductive traits, such as age at first calving, dry days, intercalving period, and service period, serve as key indicators of female fertility and are important in breeding programs (Gangwar *et al.,* 2023). However, these traits are often constrained by their late expression, low heritability, and negative correlations with productivity, which complicates selection and genetic gain. The reproductive performance of livestock has shown decreasing trend in both tropical countries as well as developed nations (Fekadu *et al.,* 2011). Genetic value for fertility trait is crucial for reducing involuntary culling, improving replacement heifer availability, and accelerating genetic progress. Thus, understanding genetic variability, covariances, and trends in performance traits is crucial for designing effective breed improvement programs (Safari *et al.,* 2007).

Improvement in dairy animal performance could be achieved by either improving the environmental circumstances or the mean breeding values of population, or by a combination of both, measures (Collier *et al.,* 2006). Partitioning total phenotypic variation of economic traits into genetic and non-genetic components is fundamental to understanding and optimizing progress in various selection programs. The genetic evaluation of dairy cattle is influenced by a multitude of factors, including additive genetic effects, maternal genetic influences, and various environmental factors, which collectively shape their performance and overall productivity. Fitting various genetic factors, that incorporate maternal effects, sire effects, and non-genetic factors, improves the accuracy and reliability on covariance components and genetic parameters for performance traits in dairy cows (Mohammadi *et al.,* 2013).

Many computer programs based on restricted maximum likelihood procedures are available for estimating (co)variance components and genetic parameters in animal breeding using univariate linear models, including ASReml (Gilmour *et al.,* 2009), DfReml (Meyer, 1991), and WOMBAT (Meyer, 2007). New methods for inspection and model evaluation can help determine the best statistical model by assessing their suitability and accuracy. The most recent and widely used tools are the log likelihood (Log L), the Schwarz Bayesian information criterion (BIC), and the Akaike information criterion (AIC).

The rapid advancements in modern computing, bio-sensing devices, and cutting-edge information technologies have revolutionized livestock management by enabling large-scale and efficient data collection (Pinna *et al.,* 2023). This progress has made it essential to develop faster and more efficient predictive tools. As animal breeding enters the "big data era", it now utilizes a wide range of data sources, including genomic and phenotypic information. Phenotypic data encompasses traits like growth rates, milk production, and reproductive traits, alongside management data such as feed consumption, housing conditions, and environmental factors. Advanced sensor technologies enable continuous, real-time tracking of animal production, animal behavior, health, and fertility, providing useful insights into how animals interact with their environment. These comprehensive datasets have enormous potential to improve genetic evaluations, refine breeding strategies, and boost the overall productivity of livestock animals.

Traditional predictive models often struggle with the challenges posed by large, complex, and noisy datasets, particularly the 'large P, small N' problem (Chafai *et al.,* 2023). Machine learning (ML) methods, such as artificial neural networks (ANNs) and Bayesian neural networks (BNNs), excel in identifying complex, nonlinear relationships within data. ANN and BNN offer remarkable flexibility and robustness as non parametric models in capturing intricate patterns and interactions (Kufel *et al.,* 2023). They enable the extraction of hidden patterns and similarities, allowing for more accurate and reliable predictions of performance traits. While ANN model delivers high performance, particularly with strong prediction accuracy, they often encounter overfitting issues due to limited generalization, as they fail to incorporate the uncertainties. To mitigate overfitting, ANN typically rely on large datasets and various regularization techniques (Nusrat and Jang, 2018). However, this dependence can pose challenges for applications where data is limited. Moreover, uncertainty estimation is critical in fields like decision making in animal breeding and veterinary medical diagnostics, where machine learning is needed to be augmented with uncertainty (Chakraborty and Ghosh, 2012).

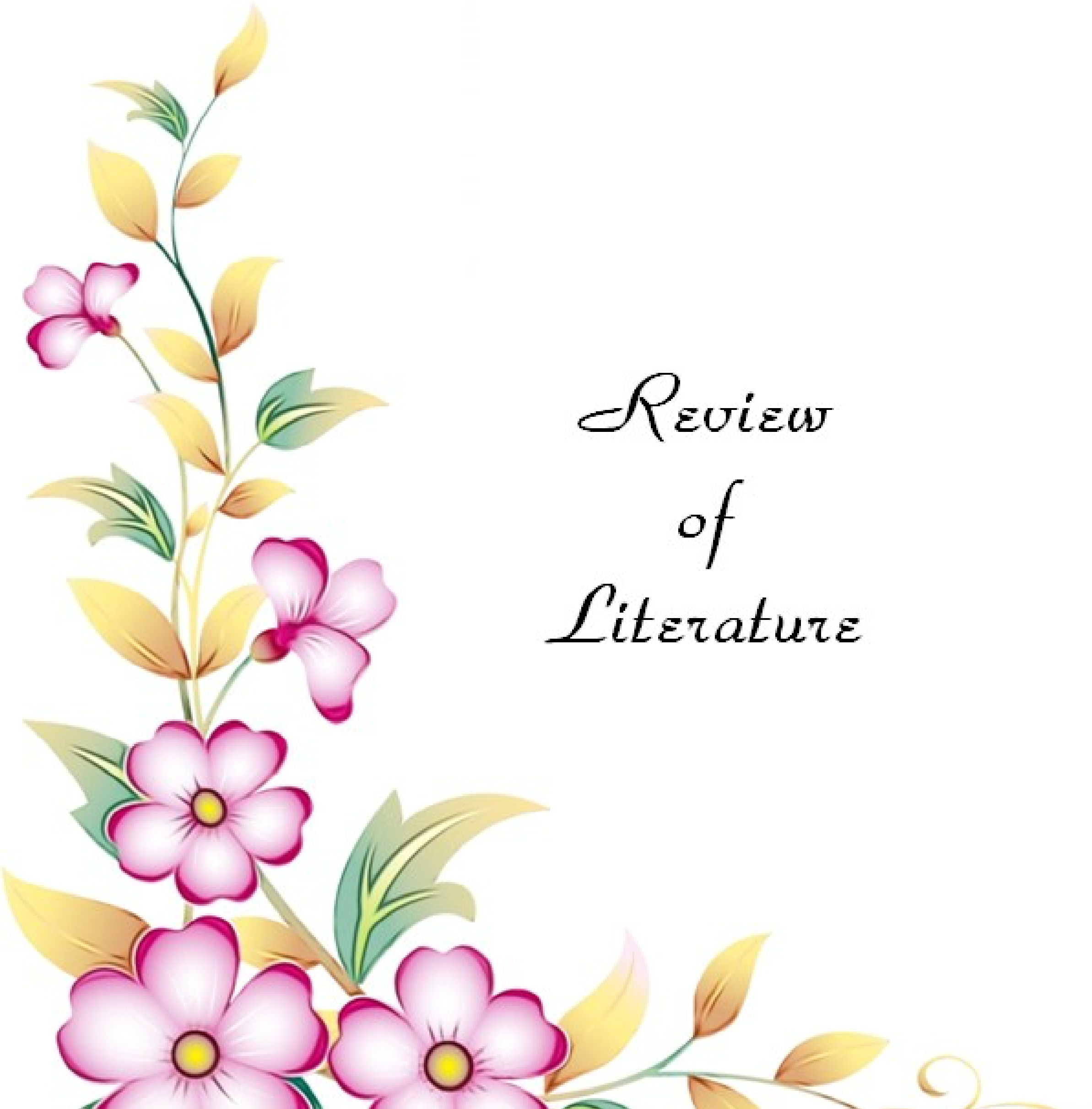
The integration of probabilistic computing approaches with ANN addresses these challenges by enabling regularization and incorporating uncertainties into predictions (Shi *et al.,* 2019). BNN models achieve this by applying Bayes' theorem to the standard neural network framework (Blundell *et al.,* 2025). Unlike traditional ANN, BNN models excel at modeling uncertainty, reducing overfitting, and performing effectively even with smaller datasets. BNN are especially powerful as they operate via ensemble models, effectively combining the capabilities of multiple ANN, with fewer parameters.

ANN and BNN have the potential to transform genetic evaluations and predictions, herd management, and the efficiency of the livestock industry at new levels. These advanced approaches provide the tools to address challenges in animal breeding, outperforming traditional models and enabling informed, and data-driven decisions.

Information on population dynamic parameters, along with comparative evaluation cum modelling of (co)variance components, and the prediction of performance traits using advanced computation techniques are scanty in high yielding Vrindavani cattle. Therefore, the present investigation was taken up in Vrindavani cattle with the following objectives:

- To assess the population structure, demographic pattern, and genetic diversity of Vrindavani cattle herd.
- To determine the efficiency of models for genetic evaluation of growth, production, and reproduction traits in Vrindavani cattle.
- To evaluate the efficiency of advanced computing techniques for predicting performance genetic traits in Vrindavani cattle.

* * *



# Review of Literature

# 2 REVIEW OF LITERATURE

## 2.1 Development of crossbreeding in India

Crossbreeding in India was introduced in the year 1875, although it did not receive much attention from policymakers until 1961. Since 1965, cross-breeding of indigenous breeds with exotic breeds has been done on a large scale to improve milk production. In 1965, the animal husbandry panel recommended the use of Jersey, Holstein Friesian, and Brown Swiss breeds for crossbreeding. Initially crossbreeding using Shorthorn bulls on local cows around Patna began and the *Taylor* breed of cattle was developed. A new breed of cattle called *Jarsindh* was developed by crossing Red Sindhi with Jersey (3/8 to 5/8 inheritance) at Allahabad agricultural institute during 1924-1934. At Bangalore, Tharparkar cows were crossed with Jersey bulls and a new breed of cattle, *Jarthar*, was developed. Brown Swiss and Jersey (5/8) were used on local non- descript cows in Kerala and a new breed of cattle *Sunandini* was developed. Sunandini is a multipurpose breed, which is specially used as a dairy breed with 3.89% fat and 2435 kg milk production. At NDRI, a new dairy breed called *KaranSwiss* was developed by crossing Sahiwal and Red Sindhi with Brown Swiss, in which the inherihance of the Brown Swiss breed was between 50-75%. Milk production of Karan Swiss hybrid females was 5000 to 6000 kg and the fat percentage was 4.78%. In 1971 at NDRI, a new dairy breed named *Karan Frie*s was developed by crossing Tharparkar cows with Holstein Friesian, Brown Swiss and Jersey bulls in which the inheritance of exotic breeds was between 50-62.5%. The milk production of these cows was 3619 kg and the fat percentage was 4.10 - 4.17. Mahatma Phule Krishi Vidyapeeth, Ahmednagar, Maharashtra has developed a crossbred cow named Phule Triveni, which produces 3000 to 3500 litters of milk with 4% fat.

Project Directorate on cattle (PDC), Meerut developed a national crossbreed cow *Frieswal* by crossing Holstein Friesian and Sahiwal breeds, which produced 3600 kg of milk with 4% fat in 300 days of lactation (Wakchaure *et al.,* 2008). In December 2023, Frieswal was officially recognized as first Indian synthetic breed by the National bureau of animal genetic resources (NBAGR).

## 2.2 Development of Vrindavani cattle

The development of the Vrindavani cattle breed was initiated in 1968 at the Indian Veterinary Research Institute (IVRI), Izatnagar, Bareilly, India, under an ad hoc scheme. The objective was to create a high-yielding dairy breed capable of producing an average of 3,000 kg or more milk per lactation. Initially indigenous Hariana cows were used as foundation stock and inseminated with Holstein-Friesian, Jersey, and Brown Swiss semen to produce half-bred progeny. The Holstein-Friesian half-breds were subsequently bred with Jersey and Brown Swiss semen. Similarly, the Jersey and Brown Swiss half-breds were inseminated with Holstein-Friesian semen to develop three-breed crosses. These half-breds and three-breed crosses were interbred and evaluated across seven generations for key traits, including milk production, reproduction performance, heat tolerance, growth, and disease resistance. Performance evaluations showed that genetic combinations of FH (½ Holstein-Friesian + ½ Hariana), FBH (½ Holstein-Friesian + ¼ Brown Swiss + ¼ Hariana), and FJH (½ Holstein-Friesian + ¼ Jersey + ¼ Hariana) had superior production and adaptability compared to other combinations (Dutt and Kumar, 1999). As their performance was nearly equivalent, interbreeding among these groups was conducted irrespective of genotype. Rigorous selection practices were implemented to improve the synthetic strain, with bulls being selected based on the performance of their dams and culled based on the performance of their daughters. The resulting population, consisting of 50–75% exotic inheritance (Holstein-Friesian, Jersey, and Brown Swiss) and 50–25% Hariana inheritance, was officially named **"Vrindavani"** (Singh *et al.,* 2011).

Vrindavani cattle are maintained under a loose housing system with group management. Nutritional needs are met through a balanced combination of dry and green fodder, supplemented with a concentrate mixture. Milk recording begins on the first day of calving and continues until drying. Calves are weaned at birth, and teaser bulls are used for heat detection twice daily.

Routine vaccinations, deworming, and spraying are carried out as per schedule. The Vrindavani breed is well adapted to local agro-climatic conditions and is popular among dairy farmers in Uttar Pradesh, Uttarakhand, Rajasthan, and adjoining regions. Although the exact field population of this breed is unknown, the IVRI cattle and buffalo farm continues to supply Vrindavani semen and live animals through auction and sale, contributing to the breed's propagation and multiplication in these areas (Singh *et al.,* 2011).

## 2.3 Population dynamics

A population is a homogenous group in terms of a specific characteristic. It interacts with other populations or within own population. The interaction leads to an increase or decrease in population size. The population dynamics thus is a study of the changes in the total number of individuals in a population over time, or the study of the impacts of population interaction (Chitra *et al.,*2016). The cattle population is always changing or dynamic. It is influenced by initial strength, sale and purchase of animals and birth and death rate in the population. The population dynamics of an animal is defined by the parameters like natural population growth or a natural increase (NI), ability to produce replacement livestock or net replacement rate (NRR), and ability to cull the remaining replacement stock, either sold or culled out. Population dynamics is computed using population statistics from the previous several years, whereas population output is derived using the number of culled cattle every year and the number of remaining replacement stock (Widyaningrum *et al.,* 2021).

## 2.4 Population dynamics parameters

The different type genealogical parameters are given below:

### 2.4.1 Pedigree completeness index

The pedigree completeness index was assessed by calculating the equivalent number of generations while taking into account the depth and wholeness of each ancestor in the pedigree for several generations back. The completeness of the pedigree up to the best-known generations is very crucial for correct inbreeding estimates. The following parameters were estimated for each individual in addition to assessing the pedigree completeness index (i)

Complete number of generations (CG), detailed as the number of generations delineating the offspring of the farthest generation where the ancestors of second-generation individuals are known. The ancestors with unknown parents were considered founders (generation 0); (ii) the maximum number of generations (MG), determined as the number of generations separating the individual from its ultimate ancestor; (iii) equivalent complete generations (EqG), detailed as the sum over all known ancestors of the terms calculated as the aggregate of (1/2) n, where n is the number of generations separating the individual from each known ancestor (Maignel *et al.* 1996).

Mumtaz *et al.* (2021) observed PDl in Sahiwal cow from first to fifth generation as 92.32, 80.26, 65.22, 49.43 and 32.54%, respectively. Muasya *et al.* (2011) reported a similar trend of pedigree completeness in Sahiwal cattle herd in Kenya. In another study, PDI in Murrah buffaloes for the entire pedigree was noticed: 76.8, 49.2, 27.7 and 12.8% for first, second, third and fourth paternal generations, respectively. Similarly, Ferraz *et al.* (2015) reported 86.6, 44.7 and 18.2% PDI in a closed herd of Jaffarabadi buffaloes for the entire pedigree in first, second and third generation, respectively.

### 2.4.2 Gene origin probabilities

The genetic background in terms of the probability of gene origin was determined by computing the various parameters in the Vrindavani population. The effective number of founders ($f_e$) was calculated as the number of founders who would be expected to contribute equally with genetic material to produce the same genetic diversity as the population under consideration (Lacy, 1989). This was calculated as:

$$f_e = \frac{1}{\sum_{k=1}^{f} q_k^2}$$

Where $f_e$ is the effective number of founders and $q_k$ is the estimated proportional genetic contribution of founder *k*, as determined by the average founder relationship to each animal in the current population.

The effective number of ancestors ($f_a$) refers to the minimum number of individuals (founders or non-founders) that are required to explain the complete genetic diversity of the current population. It was computed to assess the incidence of population bottlenecks in the population (Boichard *et al.*, 1997).

$$f_e = \frac{1}{\sum_{k=1}^{f} q_k^2}$$

Where $f_a$ refers to the total number of ancestors and $q_j$ indicates the marginal contribution of $j^{th}$ ancestor. The marginal contribution refers to the genetic component of an ancestor that could not be explained by other previously selected ancestors (Boichard *et al.*, 1997). The founder genome equivalents ($f_g$) account for genetic variation that may be lost due to random drift in small populations, even if every one of the founders made an equal contribution to the population (Lacy, 1989). The inverse of twice the average coancestry between individuals in the reference population was used to calculate the $f_g$:

$$f_e = \frac{1}{\sum_{k=1}^{f} q_k^2}$$

Where, $f$ indicates the average co-ancestry between participants in the reference population. The $f_g$ is always smaller than both $f_a$ and $f_e$, and it accounts for all factors that influence gene loss during segregation. The non-founder genome equivalents ($f_{ng}$) account for genetic diversity loss due to genetic drift, accumulated over non-founder generations. This was computed as per Caballero and Toro (2002):

$$f_e = \frac{1}{\sum_{k=1}^{f} q_k^2}$$

The genetic bottleneck was determined by calculating the number of ancestors in the population who contributed the majority of 50 percent of genes ($f_a$50) and the $f_e/f_a$ ratio. The $f_a$ is expected to be smaller than the $f_e$ if there is a bottleneck, which can be indicated by the $f_e/f_a$ ratio.

### 2.4.3 Generation interval

Generation interval (GI) is the mean age of parents at the time of birth of their progeny that is kept for perpetuation across generations. Generation intervals were calculated by averaging the four paths of selection models, i.e., sire to son ($L_{ss}$), sire to daughter ($L_{sd}$), dam to son ($L_{ds}$) and dam to daughter ($L_{dd}$). The information on birth dates of retained animals, in addition to their sires and dams were used for these calculations. The average GI was calculated as:

$$f_e = \frac{1}{\sum_{k=1}^{f} q_k^2}$$

For Sahiwal cattle breed, Parveen *et al.* (2016) reported GI (7.2 years). Muasya *et al.* (2011) found a generation interval of 6.9 years in a Sahiwal cattle herd of Kenya. Kamiti *et al.* (2016) and Faria *et al.* (2009) noticed the average GI in Sahiwal cattle population as 6.3 and 8 years, respectively.

### 2.4.4 Inbreeding coefficient and average relatedness

The inbreeding coefficient (F) is the probability that two alleles at a locus, inherited by an individual, are identical by descent (Wright, 1922). The inbreeding coefficient (F) was determined using the algorithm proposed by Meuwissen & Luo (1992). The individual increase in inbreeding for each generation was computed using the basic formula as per González *et al.* (2007) and Gutiérrez *et al.* (2008a 2009a).

$$f_e = \frac{1}{\sum_{k=1}^{f} q_k^2}$$

Where $F_i$ is the individual inbreeding coefficient and t is the equivalent number of complete generations for this individual,

The average relatedness (AR) is considered an alternative to the coefficient of inbreeding to predict the long-term inbreeding of a population. AR refers to the probability that an allele, selected at random from the whole population, belongs to a given individual (Gutiérrez and Goyache 2005). Hence, it was computed as the average of the coefficients in the row; equivalent to the individual in the numerator relationship matrix as described by Dunner *et al.* (1998).

Mumtaz *et al.* (2021) found 2.16% average inbreeding in a herd of Sahiwal. Kamiti *et al.* (2016) noticed average inbreeding of 0.0012% per year in a Sahiwal cattle herd of Kenya over a period of 48 years (1960-2008) which was lower than the recommended limit of 1% (FAO, 1998). Similarly, Muasya *et al.* (2011) observed average inbreeding of 0.06% in a Sahiwal cattle herd of Kenya. In another investigation on Sahiwal cattle, the relationship between inbreeding depression and performance traits was not linear, with greater depression estimate of 15% (Musingi *et al.*,2018). Ferraz *et al.* (2015) reported the average relatedness coefficient of 12.5% in a Jaffarabadi buffalo herd with the highest individual value of 28.95%. Average relatedness coefficient of 3.58% (Santana *et al.*,2011), 2.05% (Malhado *et al.*,2012) and 0.37% (Neto *et al.*,2012) was also observed in previous studies.

### 2.4.5 Effective population size

The effective population size (Ne) refers to the number of breeding individuals that would give rise to a calculated sampling variance or rate of inbreeding if they are bred in the way of an idealized population. Ne was estimated using changes in individuals' inbreeding coefficients (F*i*) as suggested by Gutiérrez *et al.* (2008b) and modified by Gutiérrez *et al.* (2009b). The individual increase in inbreeding coefficients was averaged, and the effective population size was determined as described by Falconer and Mackay (1996):

$$f_e = \frac{1}{\sum_{k=1}^{f} q_k^2}$$

Where Ne = Effective population size; $\Delta F = \text{Inbreeding rate}$

Additionally, Ne was also calculated by individual inbreeding coefficients; regressed on the equivalent generations, complete and maximal generations traced as described by Maignel *et al.* (1996). Santana *et al.* (2011) observed a small effective size of 40 in a Murrah buffalo herd. An effective population size of 500 animals is required to minimise genetic diversity loss due to genetic drift and to maintain a sufficiently flexible population (Frankham, 1995). However, an effective population size of 50 animals is sufficient to avoid inbreeding depression (FAO, 2007).

### 2.4.6 Genetic diversity (GD)

The degree of genetic variation in the reference population in comparison to that existing in the base population is expressed as $N_{ei}$ expected heterozygosity. This was computed as per (Lacy, 1989 1995).

$$GD = 1 - \frac{1}{2 f_g}$$

Genetic diversity in the base population was calculated as:

$$GD^* = 1 - \frac{1}{2f_e}$$

The difference between GD* and GD was computed as per (Caballero and Toro 2002):

$$GD^* - GD = \frac{1}{2f_{ng}}$$

The genetic diversity that has been lost in the population of the founder generation was estimated by 1-GD. The genetic diversity loss by unequal distribution of founder alleles was estimated by 1-GD* as per (Caballero and Toro 2002). Santana *et al.* (2014) revealed 2.30% genetic diversity loss in a Brazilian Gir cattle herd. Mumtaz *et al.* (2021) observed 2.1 and 2.64% genetic diversity loss in Murrah buffalo and Sahiwal cattle herds.

### 2.4.7 Genetic conservation index

Genetic conservation index (GCI) was estimated from genetic contributions of all the identified founders of the reference population as per (Alderson, 1992).

$$GCI_i = \frac{1}{\Sigma p_j^2}$$

Where, $p_j$ is the percentage of genes of $j^{th}$ founder, contributed to the pedigree of $i^{th}$ animal. The estimation of GCI is based on the assumption that breeding program aims to retain complete set of all the alleles from the base population.

## 2.5 Introduction to linear mixed model (LMM)

Mixed models were used significantly to improve plant and animal production over the past 60 years. Charles Henderson and Robin Thompson, along with the development of computing power, introduced the BLUP and REML models for estimating breeding values and variance components respectively. These innovations have led to national and international breeding programs in various livestock species. Karin Meyer and Dorothy Robinson produced software to implement REML methods (in animal breeding). However, analysis was difficult until Robin presented the Average Information method (Gilmour *et al.*, 1995; Johnson and Thompson 1995) underpinning ASReml which become generally available in 1997.

### 2.5.1 Fixed and random effects

A mixed model includes both fixed and random effects (hence 'mixed'). Most statistical publications define a fixed factor as one having few levels, such as age groups, treatments, diets, and years, with variations between the levels to be calculated. If the same experiment was to be replicated; the same ages, treatments, or diets would be predicted to appear again, and their calculated differences would be expected to be consistent between the two samples (Schaeffer, 2019).

A random factor, on the other side, has multiple levels that are large enough to be regarded infinite and large sufficient to be regarded as randomly sampled from an infinite population, such as sires or herd-year-seasons. If the primary research is repeated, sires and herd-year seasons would be completely different, resulting in a new random sample. In animal breeding, experiments are not repeated; instead, new sires and herd-year-seasons are constantly generated, partially depending on the composition of previous samples, and therefore occur only once (Schaeffer, 2019).

### 2.5.2 The equation

A Linear Mixed Model (LMM) combines fixed and random effects, offering a more flexible framework for analyzing data with hierarchical or correlated structures. The model is expressed as:

$$\mathbf{y = Xb + Zu + e}$$

where,

$y$ is an $n \times 1$ vector of n observed records

$b$ is a $p \times 1$ vector of p levels of fixed effects

$u$ is a $q \times 1$ vector of q levels of random effects

$e$ is an $n \times 1$ vector of random, residual terms

X is a known design matrix of order $n \times p$, which relates the records in y to the fixed effects in b

Z is a known design matrix of order n × q, which relates the records in y to the random effects in u

$E(y) = Xb$

$var(u) = G$

$var(e) = R$

$var(y) = ZGZ' + R$

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}Z \\ Z'R^{-1}X & Z'R^{-1}Z + G^{-1} \end{bmatrix} \begin{bmatrix} b \\ u \end{bmatrix} = \begin{bmatrix} X'R^{-1}y \\ Z'R^{-1}y \end{bmatrix}$$

This structure can be expanded in several ways. The vectors b and u could contain more fixed and random effects (e.g., additive genetic, inbreeding, sex, season, year, permanent environmental, etc). The effects in u define the structure of G. If vector y contains additional traits, the vector u will have (breeding) values referring to the various traits (Schaeffer, 2019).

## 2.6 Types of liner mixed model (LLM)

Various types of LMMs are used depending on the objective, data structure, and modeling requirements for animal breeding. Below are the common types of LMMs:

### 2.6.1 Single Trait Animal Model

A single-trait animal model is the simplest mixed model used in animal breeding. It is an ‘animal’ model because we give a breeding value to each animal. The term “single trait” refers to the idea that animals only have information on one character (trait), and that they have only fixed effects and additive genetic effects, with no other random effects like maternal or dominance (Misztal *et al.,*1993).

### 2.6.2 Sire Model

For computational simplicity, this model simply estimates the effect of sires on information on their progeny. The sire EBVs through sire model may be significantly less precise due to lower accuracy and probable bias because dam variances are not accounted. The model generally assumes that each sire’s progeny comes from a different dam, and all dams come from the same homogenous population with the same expected mean. In reality, dams may belong to different breeds, and dams are selected over time, therefore younger dams are likely to surpass older dams (Sun *et al.,* 2009).

### 2.6.3 Repeated records model

This is applied when an animal has several records, such as various milk production records in different lactation stages. The genetic correlation between records is one, while the phenotypic correlation between records is equal to repeatability. The strategy is to establish a permanent environmental effect for every animal, so that when the animal has a second record, not only their breeding value but also a portion of the environmental effects are repeated (Park *et al.,*2012).

An example of a mixed model with repeated records is:

$$\mathbf{y = Xb + Za + Zp + e}$$

where,

y is the vector of the observations, b is the vector of fixed effects,
a is a vector of additive genetic effects,
p is a vector of permanent environmental effects and e is a vector of residual effects.

The matrix X is the incidence matrix for the fixed effects and Z is the incidence matrix relating observations to animals.

### 2.6.4 Maternal effects model

Maternal effects have an impact on several traits, including the survival of new borns and early growth in cattle. The mother has an impact on her offspring's performance through maternal effects. These maternal impacts are mainly environmental for the offspring, although they can include both genetic and environmental components. It is important to take maternal genetic impact when animals are selected from the dam lines. Integrating maternal effects into the model allows for the estimation of maternal effects as well as the adjustment of any bias in the genetic evaluation of the growing animal. It is commonly assumed that maternal effects are inherited; however, some part of it might also be a long-term environmental effect (Kumar *et al.,* 2021).

Maternal Effects Model

$$\mathbf{y = Xb + Z1a + Z2m + e}$$

where,

y is the vector of the observations, b is a vector of fixed effects,

a is a vector of additive genetic effects,

m is a vector of maternal genetic effects and e is a vector of residual effects.

X is the incidence matrix for the fixed effects and Z1 and Z2 are the incidence matrices relating observations to random effects of animal (additive genetic) and dam (maternal genetic), respectively.

Chimonyo *et al.* (2008), Tomiyama *et al.* (2010), Mondal *et al.* (2014) and Alves *et al.* (2018) used different animal models to estimate genetic parameters of body weight in different breeds of animals. Elfadl *et al.* (2016) compared various models and determined genetic parameters in 3464 records of Holstein Friesian cows in 3 successive lactations. To evaluate covariance components and genetic factors, they created 3 multi-trait animal models by including and excluding various fixed and random effects. Their research findings indicated that the range and standard deviations of breeding values varied among models.

Shalaby *et al.* (2016) used linear models to estimate covariance components and genetic parameters in Holstein Friesian dairy cattle using 2846 records. Models were distinguished based on the random effects that were fitted in each model. For all traits, the best-fitting model had the highest Log Likelihood value. This study showed that including maternal effect with direct-maternal genetic covariances in statistical models for genetic evaluations would enhance the existing herd genetically.

Panda *et al.* (2020) used animal models for the genetic evaluation of growth performance in Landlly pigs. Six different models were constructed for growth traits by including or excluding; maternal permanent environmental effect, maternal additive genetic effect and covariance between direct and maternal additive genetic effect. In order to select the most suitable model, they used likelihood ratio test, log likelihood values and Akaike's information criterion. They showed that a model with direct additive genetic and maternal permanent environmental effect fits best for the growth data.

Diop *et al.* (1999) used a multiple-trait animal model to estimate the genetic variances for growth traits in Gobra cattle. They found that a model that includes direct genetic, maternal genetic, and direct-maternal genetic covariances is the most efficient at 18 months of age. It was noticed that fitting models with various fixed and random effects can produce more accurate estimates of genetic parameters for both production and reproduction traits in dairy cows (Miraei *et al.,*2007; Zamani and Mohammadi, 2008; Mohammadi *et al.,*2013). Edriss *et al.* (2006), El-Awady *et al.* (2011) and Hammoud and Salem (2013) demonstrated the significance of direct genetic, maternal genetic and maternal permanent environmental effects on production and reproduction traits in cattle for various types of animal models. Cilek and Kaygisiz (2008) noticed high and positive correlation (0.83) between estimated breeding values and Spearman rank correlations for the different four animal models. Kaya *et al.* (2003) and Seyedsharifi *et al.* (2008) also found similar results.

## 2.7 Artificial neuron network

Artificial neural networks are structured in layers, typically featuring an input layer, one or more hidden layers, and an output layer. Input layers receive data, hidden layers perform computations through weighted connections, and the output layer generates predictions (Fig 2.1 & Table 2.1). ANNs draw inspiration from biological neuron network but employ techniques like backpropagation and gradient descent to adjust weights and optimize performance during training (Chavlis and Poirazi, 2021). This structured approach enables ANNs to handle specific tasks such as classification, regression, and pattern recognition effectively. (A. Pastur-Romay *et al.,* 2016).

**Table 2.1: Relationship between Biological neural network and artificial neural network:**

| Biological Neural Network | Artificial Neural Network |
|---|---|
| Dendrites | Inputs |
| Cell nucleus | Nodes |
| Synapse | Weights |
| Axon | Output |

### 2.7.1 The perceptron

A perceptron is a fundamental model in neural networks used for binary classification. It consists of an input layer where each feature is associated with a weight. The perceptron calculates a weighted sum of these inputs, which is then passed through an activation function. This function typically outputs one value if the weighted sum exceeds a certain threshold and another value if it does not, thus classifying the input into one of two categories. During training, the perceptron adjusts its weights based on the errors it makes, using algorithms like the perceptron learning algorithm to improve its classification accuracy. While a single perceptron can only handle linearly separable problems, combining multiple perceptrons in a network can address more complex tasks. This perceptron model is shown in fig 2.2 in mathematical form.

The mathematical model of the perceptron neuron may write as

$$\mathbf{y = f(\Sigma(Wi * Xi) + b)}$$

**y:** The output of the perceptron.

**f:** The activation function (e.g., sigmoid, ReLU).

**Wi:** The weight associated with the ith input.

**Xi:** The ith input.

**b:** The bias term.

## 2.8 Classification of Artificial Neural Networks

ANN are classified into different types based on their architecture and functionality, each of them with diverse applications in animal breeding. These types of ANNs are described under below:

### 2.8.1 Feedforward neural networks (FNN)

Feedforward neural networks follow the most straight forward ANN architecture, where information moves in one direction—from input nodes, through hidden nodes (if present), and finally to output nodes. FNNs are commonly used in applications like predicting milk production and estimating breeding values due to their ability to generalize well from training data (Hamadani *et al.*, 2023).

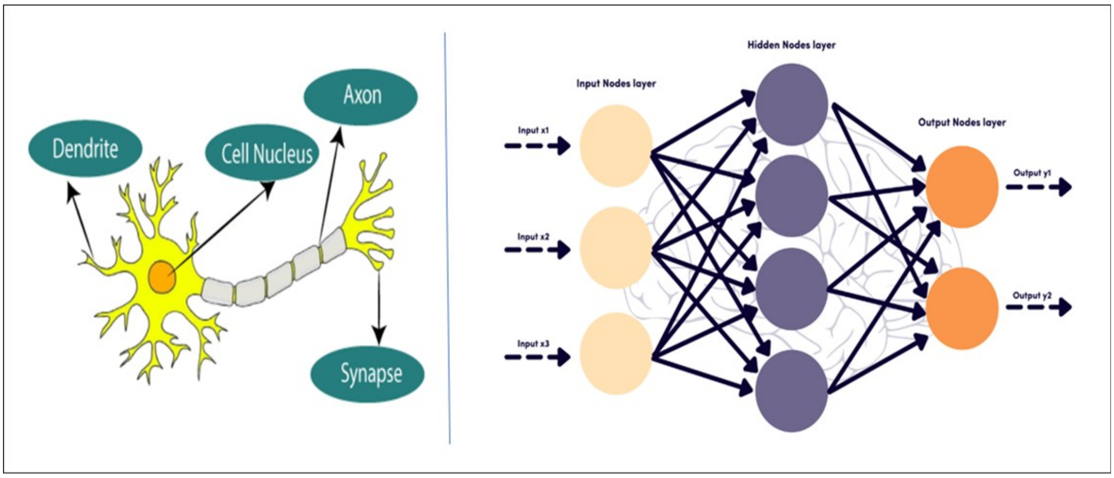


**Fig. 2.1: Biological neuron vs artificial neuron**

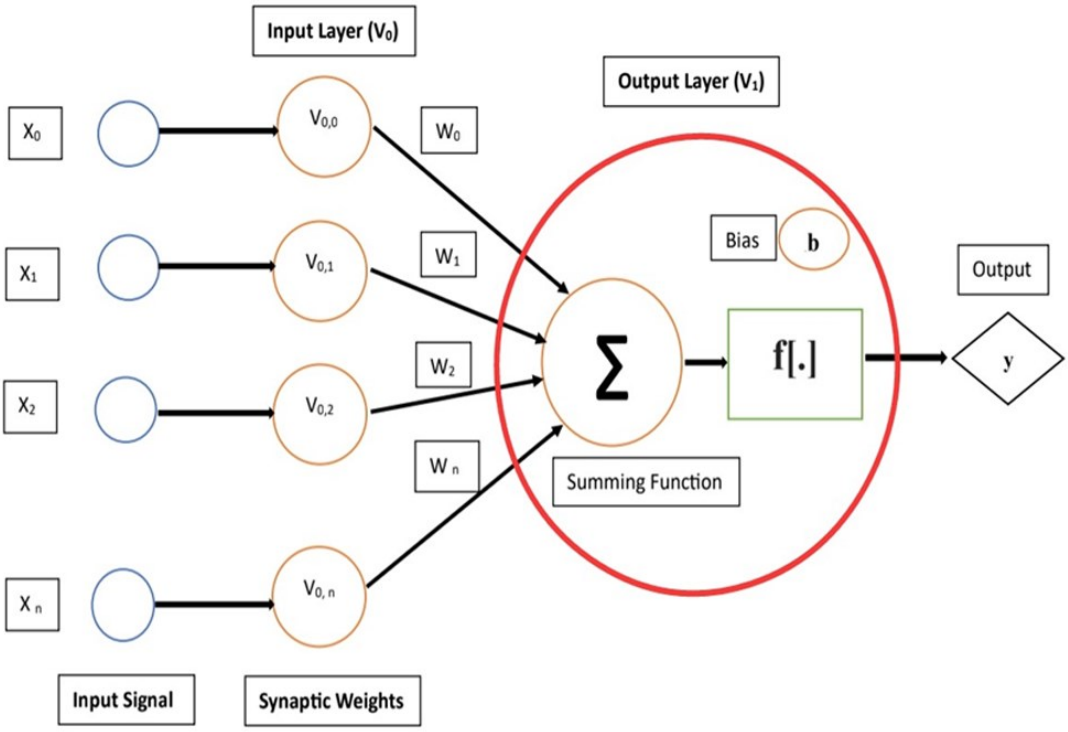


**Fig 2.2: A model perceptron - depicting its various components**

### 2.8.2 Recurrent Neural Networks (RNN)

Recurrent Neural Networks are characterized by connections that loop back on themselves, allowing them to maintain a form of memory. This architecture is useful in scenarios where the time-dependent behavior of data is crucial, such as monitoring livestock health over time. RNNs are particularly adept at handling sequences of data, making them suitable for tasks like predicting disease outbreaks from continuous sensor data (Tealab, 2018).

### 2.8.3 Convolutional neural networks (CNN)

Convolutional Neural Networks are designed to process the data that can be represented as a grid (e.g., images). In the context of animal production, CNNs are applied for image recognition tasks, such as monitoring the physical health of animals through visual inspections. They can also be used in agriculture for tasks like analyzing plant health, which indirectly supports animal production by ensuring feed quality (Li *et al.,* 2021).

### 2.8.4 Multilayer perceptrons (MLP)

Multilayer Perceptrons are a specific type of feedforward neural network with one or more hidden layers between the input and output layers. MLPs are commonly employed in complex predictive tasks in animal production, such as estimating body weight gains in livestock and optimizing feed compositions for better nutrient utilization (Hamadani *et al.,*2023).

### 2.8.5 Need for non-linearity in neural networks

Non-linearity has a diverse role in neural networks. It improves the model's ability to generalize from training data, improves convergence during learning, and allows for complicated predictions that linear models alone cannot generate. This critical component enables neural networks to handle a range of real-world tasks well, making them powerful tools in artificial intelligence (Taye, 2023). Once the weighted sum is calculated in the neural network, it needs to be transformed to introduce non-linearity. This is where activation functions come into play. They apply a non-linear transformation to the output of a neuron, enabling the network to model complex relationships in data. By adding these non-linear activation functions, neural networks can acquire and approximate any continuous function. This capacity is critical for tasks such as complicated data relationships, image recognition, natural language processing, and a wide range of other applications (Shine *et al.,* 2021).

### 2.8.6 Need of activation functions in neural networks

In the absence of an activation function inside a neural network, the resulting output signal would just exhibit a linear action, characterised by a polynomial function of degree one. While linear equations are straightforward and manageable, their capabilities are limited, and they lack the capacity to understand and identify complex connections within datasets. A neural network that lacks activation functions might be regarded as a linear regression model with restricted efficacy and capability across multiple situations (Alzubaidi *et al.,*2021). A neural network should be able to do more than just learn and generate a linear function. Activation functions and artificial neural network techniques, such as deep learning, are employed to effectively analyse complex, high-dimensional, and nonlinear datasets. These approaches are particularly useful when the model consists of multiple hidden layers and a complex architecture, as they facilitate the extraction of valuable information. Ultimately, the objective is to gain a comprehensive understanding of the dataset (Sarker, 2021).

### 2.8.7 Choosing the right activation function

To improve performance and reduce errors, the number of hidden layers, training techniques, hyperparameter adjustment, and activation function are crucial variables to consider. It might be difficult to choose the proper activation. There is no universal rule for choosing an activation function; rather, the decision depends on what task is at hand. Different activation functions offer various pros and cons, depending on the system being designed. An example: A mixture of sigmoid functions helps improve the categorization of variable. Sigmoid and tanh functions are avoided due to the vanishing gradient issue (the gradient becoming zero). The most popular activation function, Rectified Linear Unit (ReLU), works well in most situations. Leaky ReLU can be the used if the network has dead neurons. According to research, the sigmoid and tanh functions are not good for hidden layers because they lose slope as input levels go up or down, which slows the gradient fall. To apply with concealed layers, ReLu is the preferable option since its derivative is 1. Furthermore, leaky ReLU is applicable for zero derivatives. A quicker activation function may be selected for faster training and function approximation (Szanda³a, 2021).

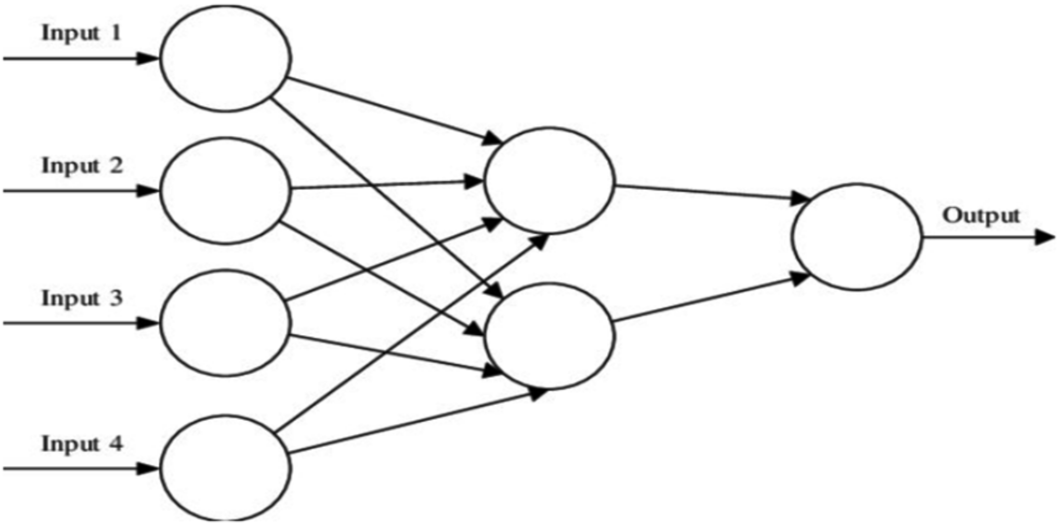


Fig. 2.3: A feedforward neural network model

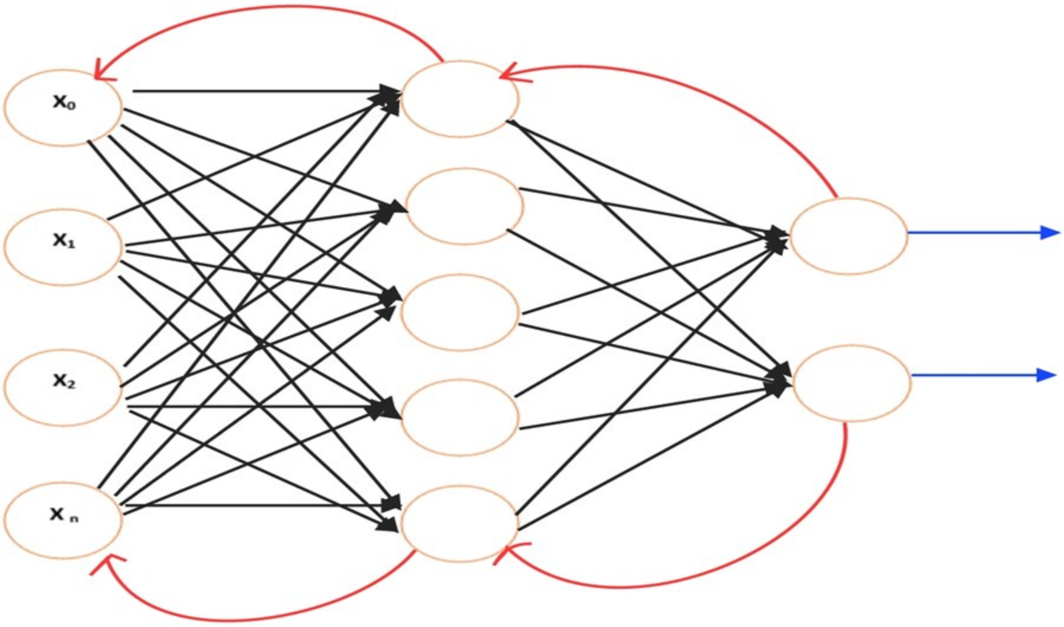


Fig. 2.4: A recurrent neural network model

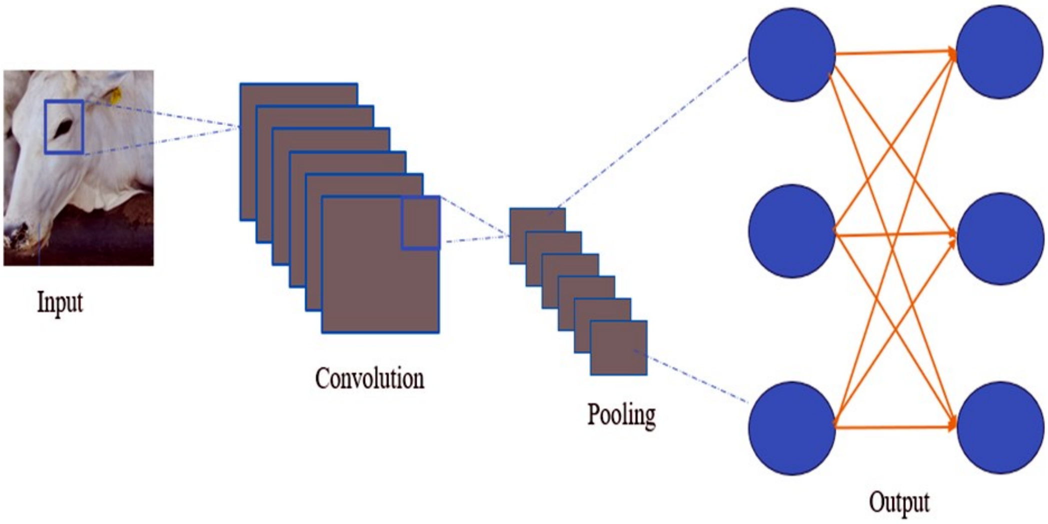


Fig. 2.5: A convolutional neural network model

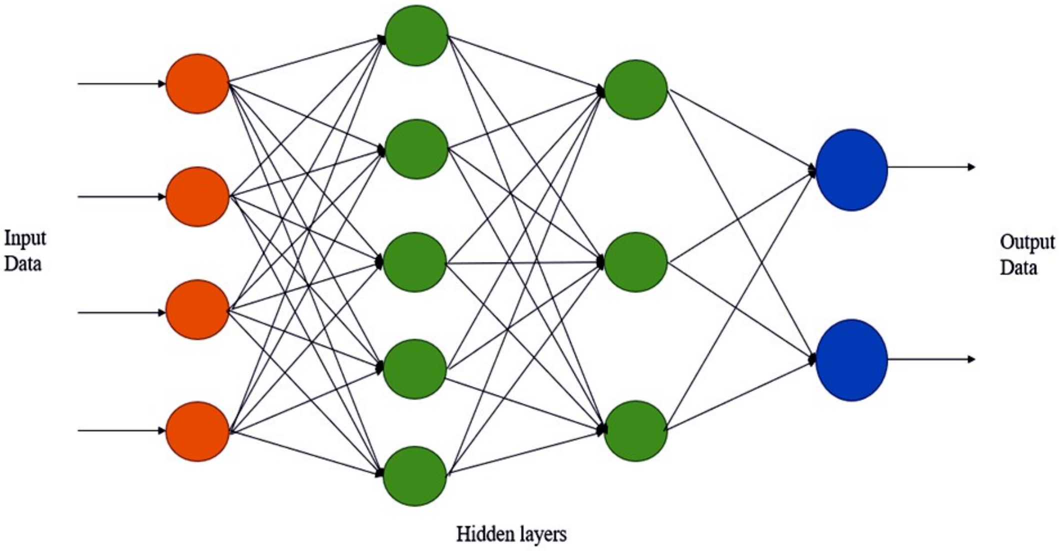


Fig. 2.6: A multilayer perceptron model

## 2.9 Back propagation networks (BPN)

Rumelhart, Hinton, and Williams introduced the concept of BPN in 1986. The term "back propagation network" (BPN) refers to a multi-layer feedforward network in which errors are "back propagated". It is utilised for error detection and correction and makes use of the supervised training process, which is a systematic approach to train the network (Antwi *et al.*,2017). Among supervised learning algorithms for ANN, back-propagation of error is the most popular (Rumelhart *et al.*,1986). When working with large-scale data, it offers good computing attributes, in contrast to other learning methods (like bayesian learning). This model can be seen as either a gradient descent method to identify the most efficient solution (Leven, 1996) or a modification of the delta rule (Riesel, 2007). The algorithm tries to minimise the error function in terms of the weights and usually uses a least-squares solution. This means that the process can be seen as a non-linear regression method.

The BPN is most commonly used for animal genetic and production data because datasets are frequently large and multidimensional, especially when genomic information is included. Because of their ability to learn from several features at once, BPN s are well-suited to dealing with such complexity. For example, estimating milk yield based on hundreds of single nucleotide polymorphisms (SNPs) necessitates the ability to effectively interpret a huge number of input signals. Backpropagation networks may manage this high dimensionality and provide robust predictions, therefore improving outcomes in cattle management (Ehret *et al.*,2015). Another use for backpropagation networks is continuous learning from old data. These networks may update their weights repeatedly, allowing them to adapt when new data is available. In dynamic contexts such as cattle farms, where conditions change frequently having a model like backpropagation networks that can adapt with incoming input is crucial in making prompt sound management decisions (Ehret *et al.*,2015).

## 2.10 Artificial neural network in animal breeding

MLP- and BP-based ANN models are commonly employed in animal studies. Scientific studies on cattle breeding include detection of abnormalities in heifers, analysing the impact of milking on somatic cell content, predicting lactation efficiency, determining insemination problems, and predicting breeding values (Grzesiak *et al.*, 2006, 2010).

### 2.10.1 Prediction of animal traits and breeding values using ANN

Traditional genetic and breeding study approaches sometimes overlook gene-to-gene interactions and non-linearity, leading to inaccurate estimates of animal breeding value. Semi- and non-parametric statistical approaches are increasingly used to estimate animal breeding value, taking into account all possible non-linearity (de les Campos *et al.,*2013). In the current context, machine learning approaches, particularly ANN, have significant potential as a viable approach for simulating complex biological phenomena defined by an undetermined relationship among attributes and hypothesised non-linearities (Calvin and McDowell, 2016). The accurate and reliable ability to predict the phenotypic values of various economic traits in different types of animals, such as cattle (Gianola *et al.,*2011), pigs (Tusell *et al.,*2013), buffaloes (Singh *et al.,*2020), sheep (Abbas *et al.,*2021), and goats (Iqbal *et al.,*2022), using ANN models has previously been established. Additionally, the ANN model has the ability to recognise individual animals through the use of facial recognition in sheep (Song *et al.,*2022), goats (Billah *et al.,*2022), and cattle (Li *et al.,*2022). This leads to the maintenance of precise breeding records, a fundamental requirement for the breeding and genetic enhancement of livestock. In the context of animal breeding programmes, the use of genomic selection has proven to be a valuable tool for accurately estimating the breeding values of animals at an early stage. This approach has demonstrated the potential to accelerate genetic improvement by reducing the generation interval (van der Werf, 2013).

The two most frequently used methods for performing whole genome prediction or genome-wide association studies (GWAS) are the Bayesian regression model and genomic best linear unbiased prediction (Meuwissen *et al.,*2001; Hayes *et al.,*2009). However, these models assume a linear relationship between genotype and phenotype, which may impact GWAS performance for complex traits (Nelson *et al.,*2013). Using a neural network model may improve GWAS performance by detecting high-order epistatic interactions, improving polygenic scoring, and increasing statistical power. This can help emphasise gene variations following analysis (Leem *et al.,*2014; Mieth *et al.,*2016; Paré *et al.,*2017; Raj and Sreeja, 2018; Vitsios and Petrovski, 2019). ANN models may predict complex polygenic traits, like milk composition, genomically at a lower computing cost than traditional models (Ehret *et*

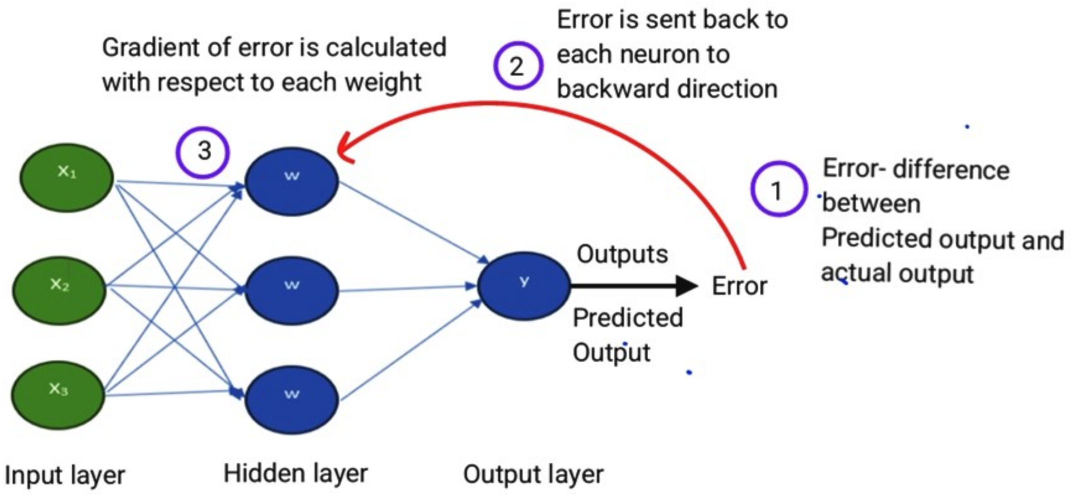


**Fig. 2.7: Back propagation algorithms**

*al.,*2015). Neural network: the open-source tool "JWAS" uses the Bayesian alphabate model (NN Bayes) to compute GWAS with optimal prediction ability, including non-linear phenotypic relationships without affecting genetic interpretability (Zhao *et al.,*2021). Thus, a neural network model that enables predictions and genomic selection could be useful as a dependable method for accelerating the genetic enhancement of livestock (Mcdowell, 2016).

### 2.10.2 ANN for analysing genetic diversity and population improvement

The use of artificial neural networks for the purpose of inferring demographic history and recombination rates within a population has gained popularity as a novel application in the area of population genetics (Schrider and Kern, 2018). Accurate estimates of demographic variables within a population are often hindered by the presence of selection biases, which are in turn influenced by demographic events (Simonsen *et al.,*1995; Ewing and Jensen, 2016). To address this issue, a deep neural network called evoNET (a three-epoch model) can estimate demography and selection all at once, detecting population size fluctuations, regions, and hard and soft selection sweeps under balanced selection (Sheehan and Song, 2016). In the discipline of population genetics, the precise estimation of recombination rates over the whole genome in populations has significant significance. This is mostly due to the fact that linkage patterns have a profound impact on several aspects, including genetic mapping and comprehension of evolutionary history (Schrider *et al.,*2015). It is well known that sequence alignments using convolutional neural networks can be used to draw conclusions about population genetics. These methods have already reached a similar level of computational efficiency (Blum and François, 2010; Flagel *et al.,*2019; Torada *et al.,*2019). ReLERNN is a recurrent neural network model-based software that accurately estimates genome-wide recombination maps. It is accomplished this by fixing problems like demographic model misspecification and genome inaccessibility, which makes it a possible tool for analysing low-quality genomic data (Adrion *et al.,* 2020).

### 2.10.3 ANN for prediction of performance traits

Predicting milk production performance at early ages allows for the selection of high-producing cows and the culling of low-producing animals, enhancing profitability in dairy farming

(Kliœ *et al.,*2021). In order to forecast lactation milk yield, it is necessary to utilise non-linear functions (Dongre and Gandhi, 2016). Traditional milk production forecasting methods underestimate the non-linearity and interdependencies of independent factors. ANN models may accurately forecast milk output in dairy animals, addressing this problem (Chaturvedi *et al.,*2013). An ANN model accurately predicts first lactation production and lifetime milk production in cattle and buffalo at early stages, surpassing traditional methods (Gandhi *et al.,*2009; Singh *et al.,*2022). Therefore, it is a viable tool for optimising breeding programmes and economic models in dairy herds by culling low-yielding animals earlier (Adamczyk *et al.,*2021). The ANN model may be utilised to forecast the weekly milk yield of dairy goats by taking into account the amount of milk that has been produced in the past, the interval between kiddings, and the first milk control record (Fernández *et al.,*2007). also predicts sheep milk yield from udder measurement (Angeles-Hernandez *et al.,*2022). Hence, the utilisation of ANN models may be a viable and dependable approach to enhance milk production in dairy animals through the use of early prediction. The morphometric evaluation and early estimation of body weights play a crucial role in the breeding and selection of meat animals (Bureš and Bartoò, 2012). ANN have considerable potential in the domain of forecasting breeding values for growth and carcass traits (Ghotbaldini *et al.,*2019). This approach has advantages in beef cattle production for selecting elite cattle based on body condition scoring and carcass quality evaluation by grading score predictions at slaughter (Goyache *et al.,*2001; Èandek-Potokar *et al.,*2015). The Visual geometry group (VGG), a convolutional neural network, effectively predicts beef carcass yield, partitions carcasses, classifies marbling beef, and scores muscle (McPhee *et al.,*2017; Kwon *et al.,*2020; Gonçalves *et al.,*2021; Matthews *et al.,*2022). These methods will provide an accurate evaluation of beef quality without the need for animal slaughter while also enabling the calculation of meat shelf life (Prevolnik *et al.,*2011). ANN can help make image-based automated carcass grading (ACG) systems that can differentiate between pigs of various ages and figure out the quality of the pork by estimating the thickness of the back fat, the amount of fat inside the muscles, and the percentage of meat in the carcass. The ANN model can assess meat and carcass characteristics of live sheep and goats, including body weight, fat depot, and carcass composition. This predictive ability can aid in breeding value estimation and optimal meat production management strategies (Huma and Iqbal, 2019; Bautista-Díaz *et al.*,2020).

### 2.11 Bayesian Neural Networks (BNNs)

A BNN is a type of neural network that incorporates Bayesian inference to quantify uncertainty in its predictions. Unlike standard ANN, which provide point estimates of parameters, BNN model used weights and biases as probability distributions rather than fixed values. This allows BNNs to capture uncertainty in both the parameters and the predictions (Jospin *et al.,* 2022).

### 2.12 Key features of Bayesian neural networks

- **Probabilistic modeling**:
- In BNNs, each weight and bias in the network is treated as a random variable with an associated probability distribution.
- The network predicts a distribution over the outputs rather than a single deterministic value.
- **Bayesian inference**:
- Bayesian methods are used to estimate the posterior distributions of the weights, given the prior distributions and observed data.
- The posterior distribution is computed using Bayes' theorem:

$$\frac{P(A \cap B)}{P(B)} = \frac{P(A) \cdot P(B|A)}{P(B)}$$

where:

P(A) = The probability of A occurring

P(B) = The probability of B occurring

P(A|B) = The probability of A given B

P(B|A) = The probability of B given A

P(A∩B) = The probability of both A and B occurring

- **Uncertainty quantification**:
- BNNs provide epistemic uncertainty (due to model uncertainty) and aleatoric uncertainty (due to inherent noise in the data).
- This makes BNNs particularly useful in applications where knowing the confidence in predictions is crucial, such as medicine, autonomous driving, and finance.
- ***Regularization:***
- By incorporating prior distributions over weights, BNNs inherently regularize the model, reducing the risk of overfitting.

### 2.12.1 Advantages of BNNs

- **Uncertainty-aware predictions**:
- Provides not only predictions but also the associated confidence levels.
- **Robustness**:
- Better handles out-of-distribution data compared to standard ANNs.
- **Flexibility in model specification**:
- Prior distributions can be tailored to include domain-specific knowledge.

### 2.12.2 Challenges of BNNs

- **Computational complexity**:
- Exact inference is often computationally infeasible for large networks due to the high-dimensional integrals involved.
- **Approximation methods**:
- Practical implementations rely on approximation techniques like:
- **Variational inference (VI)**: Approximate the posterior with a simpler distribution.
- **Monte Carlo (MC) sampling**: Use stochastic sampling methods (e.g., Markov Chain Monte Carlo).

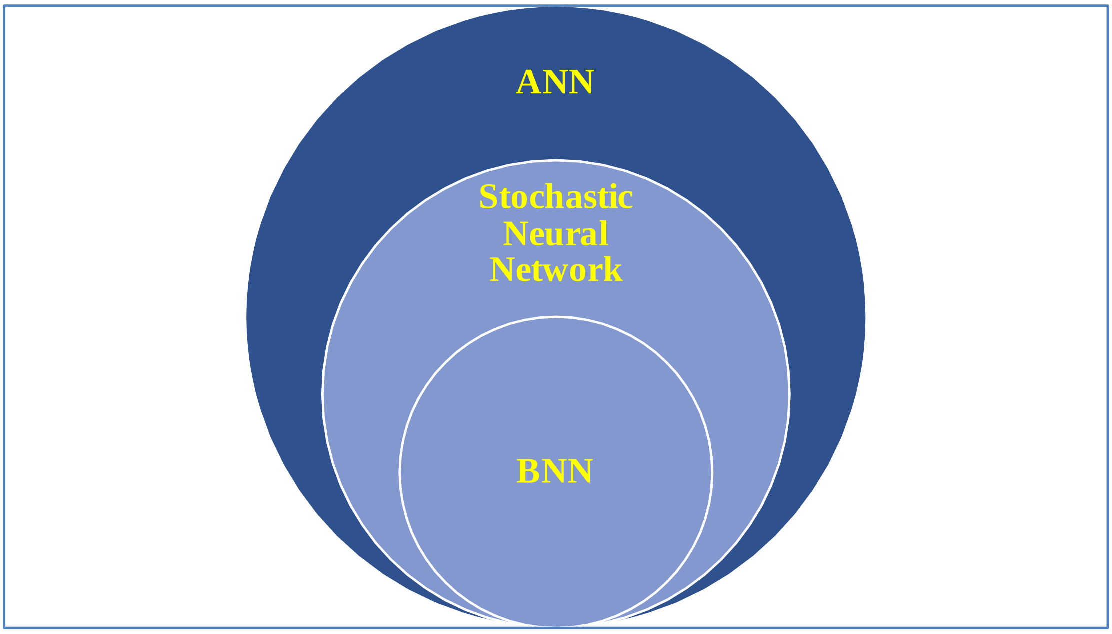


**Fig. 2.8a:** BNNs are a special type of stochastic neural networks, which in turn are specific types of ANNs

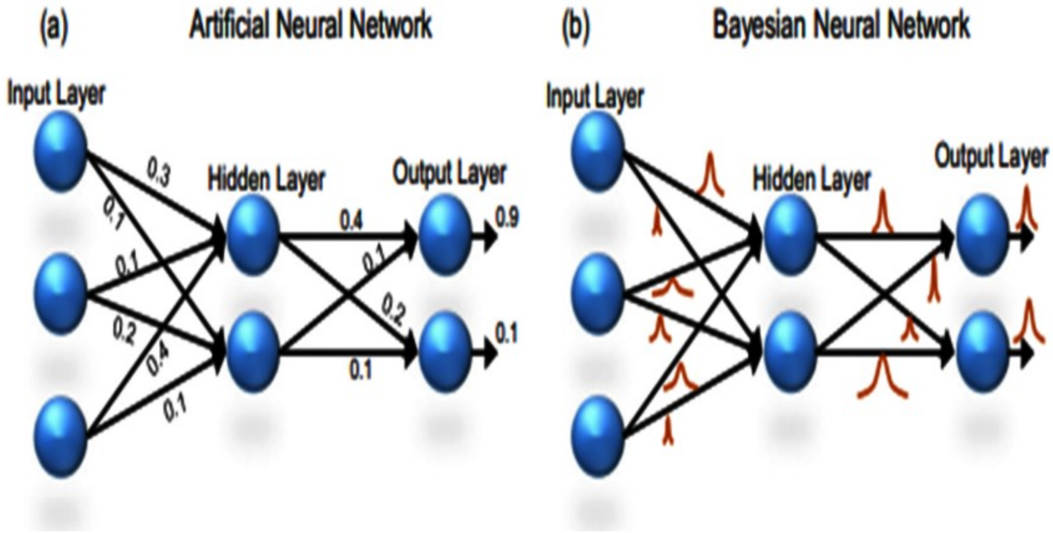


**Fig. 2.8b:** Comparison of an ANN and BNN. (a) an ANN and (b) a BNN. The synapses of ANN are represented by single-valued weights while the synapses of a BNN is represented by probability distributions

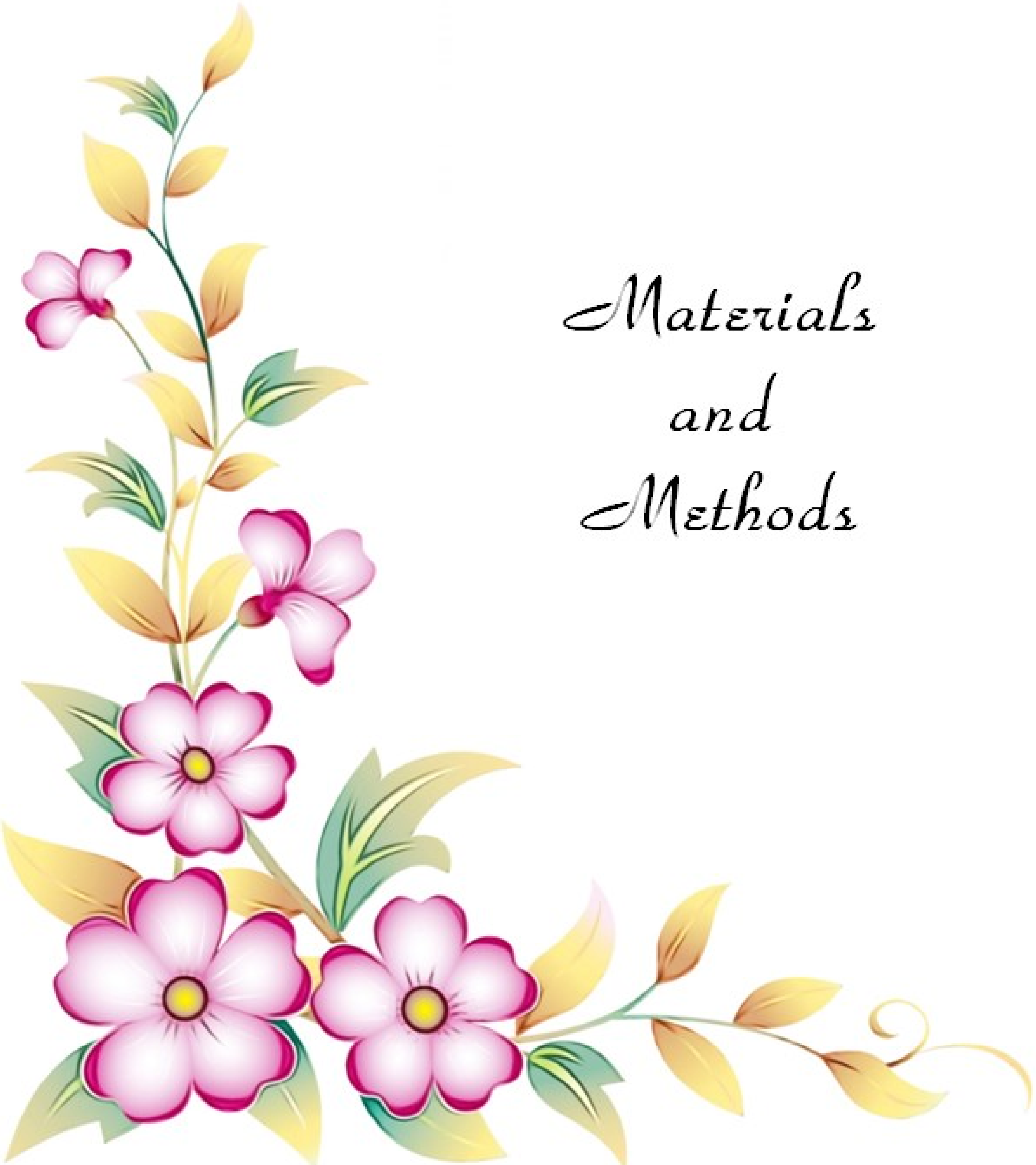
- **Dropout as approximate bayesian inference**: Dropout is applied at both training and inference to approximate uncertainty.
- **Scalability**:
- Training BNNs on large datasets or networks can be resource-intensive.

## 2.13 Study of BNN model

Van Bergen *et al.* (2020) investigated the use of BNN with variable selection for predicting the genetic components of complex traits in animal breeding. The study found that BNNs with variable selection are effective tools for genetic prediction, providing flexibility and the capability to manage complex genetic architectures. A study conducted by Sebastian *et al.* (2022) successfully demonstrated the implementation of a BNN using memtransistors fabricated from two-dimensional materials like MoS2 and WSe2. The BNN achieved a test accuracy of 80.85% on the PIMA Indian diabetes dataset while effectively quantifying uncertainties through entropy decomposition.

* * *

# Materials and Methods



# 3

# MATERIALS AND METHODS

## 3.1. Location of the farm and climate

The study was conducted at the *Cattle and Buffalo Farm* of the Indian Veterinary Research Institute (ICAR -IVRI) at Izatnagar, Bareilly, Uttar Pradesh. The farm is located in the upper gangetic plain with an altitude of 169.2 meters above mean sea level, positioned between 28°22' north latitude and 79°24' east longitude. The average yearly temperature at this area remains around 21°C, with the coldest temperatures around 5°C and the warmest reaching 40°C. The average annual rainfall ranges from 760 to 960 mm, with the highest precipitation occurring in July and August. The relative humidity ranges from 41% to 85% throughout the year (Singh *et al.*, 2011). Therefore, animals at the ICAR-IVRI farm are exposed to varied weather conditions throughout the year.

The farm maintains animals under loose housing and group management, providing a balanced diet of dry and green fodder with concentrate mixture supplementation support. Milk recording starts from calving until drying, and calves are weaned immediately after birth. Teaser bulls detect cows in heat twice daily, while vaccination, deworming, and spraying are done regularly according to standard recommendation.

## 3.2. Data collection and its properties

The following data related to pedigree and performance traits of Vrindavani cattle were collected from the records maintained at the *Cattle and Buffalo Farm*, Livestock Production and Management Section, ICAR-IVRI, Izatnagar, UP, India over a period of 53 years from 1970 to 2023:

1. **Individual information**
   a) Animal number
   b) Sex of the animal
   c) Date of birth
2. **Pedigree information**
   a) Maternal ancestor
   b) Paternal ancestor
3. **Growth traits**
   a) Body weight at birth (BW0)
   b) Body weight at 3 months (BW3)
   c) Body weight at 6 months (BW6)
   d) Body weight at 12 months (BW12)
   e) Body weight at 18 months (BW18)
   f) Body weight at 24 months (BW24)
4. **Production performance traits**
   a) 305 days milk yield (305-DMY)
   b) Total lactation milk yield (TLMY)
   c) Peak yield (PY)
   d) Lactation length (LL)
   e) Milk production in calving month (M1)
   f) Milk production in second (M2) and third month (M3)
5. **Reproduction performance traits**
   a) Age at first calving (AFC)
   b) Inter calving period (ICP)
   c) Dry days (DD)
   d) Service period (SP)

## 3.3 Classification of data

**Table 3.1a: Classification of data for season**

| Season | Code |
|---|---|
| Summer (March to May) | I |
| Monsoon (June to September) | II |
| Autumn (October to November) | III |
| Winter (December to February) | IV |

**Table 3.1b: Classification of data for period**

| Period of birth | Code |
|---|---|
| 1970 to 1979 | I |
| 1980 to 1989 | II |
| 1990 to 1999 | III |
| 2000 to 2009 | IV |
| 2010 to 2023 | V |

**Table 3.1c: Classification of data for parity**

| Parity | Code |
|---|---|
| Parity 1 | I |
| Parity 2 | II |
| Parity 3 | III |
| Parity 4 | IV |
| Parity 5 | V |
| Parity 6 & above | VI |

**Table 3.1d: classification of animal for sex**

| Sex | Code |
|---|---|
| Male | I |
| Female | II |

## 3.4 Population structure, Demographic pattern and Genetics Diversity of Vrindavani cattle herd

Pedigree data from 12,718 Vrindavani animals were compiled over a 53-year period, spanning from 1970 to 2023. The data underwent numerical transformation using R Studio and Pedigree viewer to ensure consistency and accuracy. Subsequently, the processed pedigree data were analyzed using the ENDOG v4.8 program to elucidate population structure, dynamics, and genetic diversity parameters (Pablo Gutiérrez *et al*., 2010). The programme helps in the estimation of population parameters based on the probability of gene origin. The pedigrees of the animals were traced as far back as possible in the birth record registers maintained at the farm. The pedigree data was originally available for 53-years between 1970 and 2023 (n = 12,718); the pedigree information of the cohort of calves born between 2010 and 2023 (n = 2066) were used as a reference dataset for Endog-based analysis. The different genealogical parameters elucidated in the Vrindavani population are given below:

### 3.4.1 *Pedigree completeness index*

The pedigree completeness index was assessed by calculating the equivalent number of generations while taking into account the depth and wholeness of each ancestor in the pedigree for several generations back. The completeness of the pedigree up to the best-known generations is very crucial for correct inbreeding estimates. The following parameters were estimated for each individual in addition to assessing the pedigree completeness index (i) Complete number of generations (CG), detailed as the number of generations delineating the offspring of the farthest generation where the ancestors of second-generation individuals are known. The ancestors with unknown parents were considered founders (generation 0); (ii) the maximum number of generations (MG), determined as the number of generations separating the individual from its ultimate ancestor; (iii) equivalent complete generations (EqG), detailed as the sum over all known ancestors of the terms calculated as the aggregate of (1/2) n, where n is the number of generations separating the individual from each known ancestor (Maignel *et al*., 1996).

### 3.4.2 *Generation interval*

Generation interval (GI) is the mean age of parents at the time of birth of their progeny that is kept for perpetuation across generations. Generation intervals were calculated by averaging the four paths of selection models, i.e., sire to son ($L_{ss}$), sire to daughter ($L_{sd}$), dam to son ($L_{ds}$) and dam to daughter ($L_{dd}$). The information on birth dates of retained animals, in addition to their sires and dams were used for these calculations. The average GI was calculated as:

$$GI = \frac{L_{ss} + L_{sd} + L_{ds} + L_{dd}}{4}$$

### 3.4.3 *Inbreeding coefficient and average relatedness*

The inbreeding coefficient (F) is the probability that two alleles at a locus, inherited by an individual, are identical by descent (Wright, 1922). The inbreeding coefficient (F) was determined using the algorithm proposed by Meuwissen and Luo (1992). The individual increase in inbreeding for each generation was computed using the basic formula as per González *et al.* (2007) and Gutiérrez *et al.* (2008a 2009a).

$$\Delta F_i = 1 - \sqrt[t-1]{1 - F_i}$$

Where $F_i$ is the individual inbreeding coefficient and t is the equivalent number of complete generations for this individual,

The average relatedness (AR) is considered an alternative to the coefficient of inbreeding to predict the long-term inbreeding of a population. AR refers to the probability that an allele, selected at random from the whole population, belongs to a given individual (Gutiérrez and Goyache 2005). Hence, it was computed as the average of the coefficients in the row; equivalent to the individual in the numerator relationship matrix as described by Dunner *et al.* (1998).

### 3.4.4 *Effective population size*

The effective population size (Ne) refers to the number of breeding individuals that would give rise to a calculated sampling variance or rate of inbreeding if they are bred in the way of an idealized population. Ne was estimated using changes in individuals' inbreeding coefficients (F*i*) as suggested by Gutiérrez *et al.* (2008b) and modified by Gutiérrez *et al.* (2009b). The individual increase in inbreeding coefficients was averaged, and the effective population size was determined as described by Falconer and Mackay (1996):

$$\underline{N}_e = \frac{1}{2\Delta \underline{F}}$$

Where Ne = Effective population size; ΔF = Inbreeding rate

Additionally, Ne was also calculated by individual inbreeding coefficients; regressed on the equivalent generations, complete and maximal generations traced as described by Maignel *et al.* (1996).

### 3.4.5 Gene origin probabilities

The genetic background in terms of the probability of gene origin was determined by computing the various parameters in the Vrindavani population. The effective number of founders ($f_e$) was calculated as the number of founders who would be expected to contribute equally with genetic material to produce the same genetic diversity as the population under consideration (Lacy, 1989). This was calculated as:

$$f_e = \frac{1}{\sum_{k=1}^{f} q_k^2}$$

Where $f_e$ is the effective number of founders and $q_k$ is the estimated proportional genetic contribution of founder *k*, as determined by the average founder relationship to each animal in the current population.

The effective number of ancestors ($f_a$) refers to the minimum number of individuals (founders or non-founders) that are required to explain the complete genetic diversity of the current population. It was computed to assess the incidence of population bottlenecks in the population (Boichard *et al.*, 1997).

$$f_a = \frac{1}{\sum_{j=1}^{f} q_j^2}$$

Where $f_a$ refers to the total number of ancestors and $q_j$ indicates the marginal contribution of *j*th ancestor. The marginal contribution refers to the genetic component of an ancestor that could not be explained by other previously selected ancestors (Boichard *et al.*, 1997). The founder genome equivalents ($f_g$) account for genetic variation that may be lost due to random drift in small populations, even if every one of the founders made an equal contribution to the population (Lacy, 1989). The inverse of twice the average coancestry between individuals in the reference population was used to calculate the $f_g$:

$$f_g = \frac{1}{2}\bar{f}$$

Where, $f$ indicates the average co-ancestry between participants in the reference population. The $f_g$ is always smaller than both $f_a$ and $f_e$, and it accounts for all factors that influence gene loss during segregation. The non-founder genome equivalents ($f_{ng}$) account for genetic diversity loss due to genetic drift, accumulated over non-founder generations. This was computed as per Caballero and Toro (2002):

$$\frac{1}{f_{ng}} = \frac{1}{f_g} + \frac{1}{f_e}$$

The genetic bottleneck was determined by calculating the number of ancestors in the population who contributed the majority of 50 percent of genes ($f_a$50) and the $f_e/f_a$ ratio. The $f_a$ is expected to be smaller than the $f_e$ if there is a bottleneck, which can be indicated by the $f_e/f_a$ ratio.

### 3.4.6 *Genetic diversity (GD)*

The degree of genetic variation in the reference population in comparison to that existing in the base population is expressed as $N_{ei}$ expected heterozygosity. This was computed as per (Lacy, 1989 1995).

$$GD = 1 - \frac{1}{2f_g}$$

Genetic diversity in the base population was calculated as:

$$GD^* = 1 - \frac{1}{2f_e}$$

The difference between GD* and GD was computed as per (Caballero and Toro 2002):

$$GD^* - GD = \frac{1}{2f_{ng}}$$

The genetic diversity that has been lost in the population of the founder generation was estimated by 1-GD. The genetic diversity loss by unequal distribution of founder alleles was estimated by 1-GD* as per (Caballero and Toro 2002).

### 3.4.7 *Genetic conservation index*

Genetic conservation index (GCI) was estimated from genetic contributions of all the identified founders of the reference population as per (Alderson, 1992).

$$GCI_i = \frac{1}{\sum p_j^2}$$

Where, $p_j$ is the percentage of genes of $j^{th}$ founder, contributed to the pedigree of $i^{th}$ animal. The estimation of GCI is based on the assumption that breeding program aims to retain complete set of all the alleles from the base population.

## 3.5 Genetic evaluation of growth, production, and reproduction traits in Vrindavani cattle

### 3.5.1 Framework of data for growth traits

The present study used 36,353 body weight records on 9587, Vrindavani animals, that had descended from 253 sire across different generations. The data was sourced from pedigree cum history sheets, which provided details about the animals' dates of birth, sexes, seasons of birth, period of birth and other important parameters. Body weight traits were recorded at birth, three-month, six-month, 12-month, 18 months, and 24-month, Body weight was recorded using traditional and digital scales at different ages. The study also included four non genetic factor i.e, season of birth (December to February, March to May, June to August, and September to November), period of birth (5 levels with 10-year durations each), two levels for sex (male and female), and lactation number of birth (six level). The structure of the dataset has been described in Table 3.5.1. The study aimed to examine how non-genetic factors affected body weight at different life stages.

**Table 3.2: Structure of dataset for growth traits**

| Traits | W0 | W3 | W6 | W12 | W18 | W24 |
|---|---|---|---|---|---|---|
| No of animals | 9587 | 6517 | 5183 | 4099 | 3610 | 3275 |
| No of Sires | 253 | 242 | 234 | 222 | 218 | 216 |
| No of Dams | 3146 | 2718 | 2420 | 2144 | 1994 | 1878 |
| Minimum weight (kg) | 15.00 | 35.00 | 65.00 | 89.00 | 115.00 | 175.00 |
| Maximum weight (kg) | 44.00 | 110.0 | 195.00 | 255.00 | 340.00 | 420.00 |
| Average weight (kg) | 23.15 | 57.46 | 93.71 | 156.04 | 228.78 | 292.91 |
| Standard deviation (kg) | 4.10 | 11.89 | 16.98 | 25.42 | 30.57 | 38.52 |

### 3.5.2 Framework of data for production and reproduction traits

The study utilized production and reproduction data from Vrindavani cattle maintained at the Cattle and Buffalo farm of ICAR-IVRI, spanning multiple generations. A total of 3,074 animals contributed 26,957 production data points across four distinct production traits (Table 3.4), while 2,979 female cattle provided 26,515 reproduction records across five traits (Tables 3.4). These animals were traced their lineage to 147 sires and 1,949 dams for production data

**Table 3.3: Studied productive and production traits in Vrindavani cattle**

| Traits | Definition |
|---|---|
| Total lactation milk yield (TLMY) | Total amount of milk produced by a cow over the entire lactation period |
| 305- Day milk yield (305-DMY) | Total amount of milk produced by a cow in first 305 day of lactation period |
| Lactation length (LL) | Time period for which a cow produces milk following parturition |
| Peak yield (PY) | Highest amount of milk produced by a cow in a single day during her lactation period. |
| Age at first calving (AFC) | The age at which a dairy cow gives birth for the first time |
| Dry days (DD) | Stage of lactation cycle where milk production ceases prior to calving |
| Inter-calving period (ICL) | The period between two successive calving |
| Gestation period (GP) | The time from conception to birth |
| Service period (SP) | The time between the date of calving and the date of next successful conception |

and 149 sires and 1,446 dams for reproduction data. The data, sourced from pedigree and history sheets, included details such as identification numbers of each animal, their sire and dam IDs, dates of birth, periods of calving, and parity information. Fixed effects included in the study were categorized by season of calving (1–4), period of calving (1–5, each covering a ten-year span), and parity (1–6), facilitating a comprehensive analysis of production and reproduction traits across generations.

**Table 3.4: Structure of dataset for production and reproduction traits**

| Records | TLMY | 305-DMY | LL | PY | AFC | ICT | DD | GP | SP |
|---|---|---|---|---|---|---|---|---|---|
| No of animals | 3074 | 2374 | 3074 | 2713 | 2979 | 2558 | 2429 | 2500 | 2614 |
| No. of animals without offspring | 1572 | 1136 | 1562 | 1293 | 1740 | 1261 | 1182 | 1225 | 1276 |
| No. of animals with offspring | 1502 | 1238 | 1512 | 1420 | 1595 | 1297 | 1247 | 1275 | 1338 |
| No. of sires with progeny records | 147 | 136 | 147 | 112 | 149 | 141 | 140 | 139 | 140 |
| No. of dams with progeny records | 1949 | 1102 | 1949 | 1308 | 1446 | 1156 | 1107 | 1136 | 1198 |
| Progeny per sire | 20.91 | 17.45 | 20.91 | 24.22 | 20 | 18.14 | 17.35 | 17.98 | 18.67 |
| Progeny per dam | 1.57 | 2.15 | 1.57 | 2.07 | 2.06 | 2.21 | 2.19 | 2.20 | 2.18 |
| No. of animals with unknown sire | 440 | 353 | 437 | 297 | 458 | 406 | 387 | 398 | 393 |
| No. of animals with unknown dam | 783 | 661 | 785 | 529 | 810 | 721 | 698 | 709 | 700 |
| No. of animals with maternal grandsire | 1783 | 1371 | 1799 | 1844 | 1962 | 1400 | 1319 | 1362 | 1496 |
| No. of animals with maternal granddam | 1624 | 1213 | 1635 | 1648 | 1789 | 1261 | 1181 | 1232 | 1362 |

### 3.5.3 Data analysis using statistical methods

The analysis of body weight, production and reproduction traits in the Vrindavani crossbred population was conducted using the WOMBAT program, which employs the REML algorithm to determine (co)variance structures (Meyer, 2007). Eight distinct univariate animal models were fitted to body weight data (BW0, BW3, BW6, BW12, BW18, and BW24), and eight distinct univariate repeated animal models were fitted to production data (TLMY, 305-DMY, LL, and PY) and reproduction traits (AFC, ICP, DD, GP, and SP) incorporating random effects such as animal additive genetic effect, maternal additive genetic effect, sire effect, permanent environment effect, and the covariance between additive genetic and maternal genetic effects. Fixed effects included season of calving (winter, autumn, summer, rainy), period of calving (1–5), sex of the animal (only for body weight traits), and the number of lactations in which the animal was born (1–6), while the inbreeding coefficient of animal and AFC of the dam were included as fixed covariables.

The models were compared based on the inclusion or exclusion of random effects and the animal inbreeding coefficient as a fixed covariable. Model I incorporated the animal additive genetic effect as the sole random variable, along with residuals, while Model II excluded the inbreeding coefficient as a fixed covariable. Model III added the maternal additive genetic effect as a random effect, and Model IV excluded the inbreeding coefficient. Model V introduced a correlation between animal and maternal additive genetic effects, with Model VI excluding the inbreeding coefficient. Model VII extended Model III by incorporating the sire effect, and Model VIII excluded the inbreeding coefficient from Model VII. Table 2 highlights the fundamental differences in these models, while Table 3.5.3 outlines their key variations among the models.

**Table 3.5: Summary of models used for fitting onto the production and reproduction data of Vrindavani cattle population**

| **Fixed and random effect** | **Models** | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | **I** | **II** | **III** | **IV** | **V** | **VI** | **VII** | **VIII** |
| Fixed effect (season and period of calving, and parity) | Y | Y | Y | Y | Y | Y | Y | Y |
| Animal genetic effect | A | A | A | A | A | A | A | A |
| Maternal genetic effect | - | - | M | M | M | M | M | M |
| Permanent environment effect* | Pe | Pe | Pe | Pe | Pe | Pe | Pe | Pe |
| Animal and Maternal genetic correlated | - | - | - | - | Y | Y | - | - |
| Sire genetic effect | - | - | | | | | S | S |
| Inbreeding (Covariable) | Fx | - | Fx | - | Fx | - | Fx | - |
| AFC of dam (Covariable) | Y | Y | Y | Y | Y | Y | Y | Y |

Y refers inclusion while dash '– 'refers not exclusion; A indicates use of an animal genetic effect; S indicates sire effect as random effect; M indicates use of a maternal genetic; Pe is permanent environmental effects; and Fx indicate Inbreeding covariable. *Not included in growth traits.

The tested models, illustrated in matrix form, are:

$Y = X\beta + Z_1 a + Z_2 Pe + e$ (I)

$Y = X\beta + Z_1 a + Z_2 Pe + e$ (Additive genetic effect without Fx) (II)

$Y = X\beta + Z_1 a + Z_2 Pe + Z_3 m + e$ (III)

$Y = X\beta + Z_1 a + Z_2 Pe + Z_3 m + e$ (Maternal additive genetic effect without Fx) (IV)

$Y = X\beta + Z_1 a + Z_2 Pe + Z_3 m + e$, Cov (a,m) = Aóam (V)

$Y = X\beta + Z_1 a + Z_2 Pe + Z_3 m + e$, Cov (a,m) = Aóam (Maternal additive genetic effect with correlated structure without Fx) (VI)

$Y = X\beta + Z_1 a + Z_2 Pe + Z_3 m + Z_4 s + e$ (VII)

$Y = X\beta + Z_1 a + Z_2 Pe + Z_3 m + Z_4 s + e$ (Sire random effect without Fx: besides other factors) (VIII)

Here, Y represents the vector of observations of a growth, production, and reproduction trait. The vector β encompasses the fixed effects, including the season of calving, period of calving, parity, and fixed covariables like the inbreeding coefficient and the AFC of the dam. The vector *a* denotes the additive genetic effect, *m* represents the maternal additive genetic effect, *Pe* is the permanent environment effect, and *s* captures the sire effects. The vector *e* contains the random residuals.

$X, Z_1, Z_2, Z_3$, and $Z_4$ were corresponding incidence matrices.

It was assumed that the random effects followed a normal distribution with a mean of zero and a non-zero variance.

$$\mathrm{Var}\begin{Bmatrix} a \\ m \\ pe \\ s \\ e \end{Bmatrix} = \begin{Bmatrix} A\sigma^2_a & A\sigma_{am} & 0 & 0 & 0 \\ A\sigma_{am} & A\sigma^2_m & 0 & 0 & 0 \\ 0 & 0 & I\sigma^2_{pe} & 0 & 0 \\ 0 & 0 & 0 & I\sigma^2_s & 0 \\ 0 & 0 & 0 & 0 & I\sigma^2_e \end{Bmatrix}$$

Where

$\sigma^2_a$ is the additive genetic variance,

$\sigma^2_m$ is the maternal additive genetic variance,

$\sigma^2_{pe}$ is the permanent environment effect variance

$\sigma_{am}$ refers to the covariance between additive and maternal genetic effects,

$\sigma^2_s$ is the sire additive genetic variance

$\sigma^2_e$ is the residual variance.

A is the numerator relationship matrix and

I is the identity matrix.

Note: For growth traits, the permanent environment effect due to repeated records was not included, while the sex of the animal was included as a fixed effect. All other components of the models remained consistent with those described above.

### 3.5.4 Comparative evaluation of fitting of different models

The identification of the most suitable (best-fit) univariate model for each trait was undertaken using the log-likelihood ratio test (LRT) as outlined by Meyer (1991). Statistical significance was evaluated at a threshold level of $p<0.05$ by analyzing twice the differences in log-likelihoods, represented as -2logL, against critical values from a chi-squared distribution. The degrees of freedom for this comparison were dictated by the variances associated with the two models being evaluated.

$$\lambda^2 = -2\,(\text{Log L}_{\text{reduced model}} - \text{Log L}_{\text{full model}})$$

Heritability ($h^2$) was obtained in all models; however, maternal heritability ($m^2$) was estimated in models III, IV, V, VI, VII, and VIII. Sire variance $s^2$ was estimated in VII models and VIII, whereas correlations between additive and maternal additive genetic ($r_{am}$) components were obtained under models V and VI.

Assumption of the model were

$$V(a) = A\,\sigma^2_{a,}\ V(m) = V\,\sigma^2_{m,}\ V(s) = V\,\sigma^2_{s,} \text{ and } V(e) = I\sigma^2_e$$

Those estimated (co) variance components were used to obtain,

$(h^2 = \sigma^2_a/\sigma^2p)$ = Animal's own additive effect

$(m^2 = \sigma^2_m/\acute{o}^2p)$ = Maternal additive effect

$(s^2 = \sigma^2s/\sigma^2p)$ = Sire additive effect

$[r_{am} = (\sigma_{am}/\sqrt{\sigma^2_a \times \sigma^2_m})]$ = Additive maternal genetic correlation

## 3.6. Artificial neural network (ANN) analysis

ANN were employed to predict three performance traits: TLMY, 305 DMY, and PY. The analysis was conducted using MATLAB software (version R2024b), utilizing the neural network toolbox.

### *3.6.1 Data structure and description*

The input variables for the study included both categorical and continuous data. The categorical inputs, which were one-hot encoded (convert categorical input into a numerical format), comprised the following: SOB and POB). The continuous inputs, which were normalized, included BW0, BW3, BW6, average milk production during the calving month (M1), milk production during the second month (M2), and milk production during the third month (M3).

Note: Milk production during the second month (M2), and milk production during the third month (M3) was not included as input in prediction of PY

### *3.6.2 Data Preprocessing*

Data preprocessing was performed in Rstudio prior to ANN analysis. Categorical input variables were hot encoded, converting them into binary format suitable for neural network input. Continuous variables were normalized to a range of 0 to 1 to ensure uniformity and enhance the efficiency of model training.

### *3.6.3 Dataset and Training*

The pre-processed dataset was partitioned into three subsets: training (70%), validation (15%), and testing (15%) using MATLAB default splitting mechanism. The training subset was used to optimize network weights, the validation subset monitored generalization and overfitting, and the testing subset evaluated final model performance.

### 3.6.4 *Neural Network Architecture*

ANN used in this study was configured as a feedforward backpropagation multilayer network to model the relationships between input variables and performance traits. The network architecture included an input layer, one hidden layer, and an output layer. The input layer consisted of nodes corresponding to the pre-processed input variables, with categorical variables one-hot encoded and continuous variables normalized to a range of 0 to 1. The hidden layers consisted of neurons that processed the input data through nonlinear transformations. The number of neurons in hidden layer was optimized based on the performance of the model each performance trait.

The activation functions used in the hidden layers were the sigmoid. Sigmoid was used to introduce non-linearity while providing an efficient and fast convergence by avoiding the vanishing gradient problem. The output layer consisted of a linear activation function to undertake continuous predictions for each performance trait (TLMY, 305 DMY, and PY). The network was trained using the Levenberg-Marquardt algorithm, which is well-suited for medium-sized networks and efficiently optimizes the weights and biases to minimize the mean squared error (MSE). The training process was performed for 1000 epochs, with early stopping applied if the validation performance degraded or convergence was achieved.

In addition to the Levenberg-Marquardt algorithm, two alternative training algorithms, including Bayesian regularization (trainbr), which helps prevent overfitting, and scaled conjugate gradient (traincgs), suitable for optimizing larger networks efficiently were also used in study.

## 3.7 Bayesian neural network (BNN) analysis

BNN approaches were employed to predict three performance traits: TLMY, 305 DMY, and PY. The analysis was conducted in Python 4.0, utilizing libraries such as *pandas, numpy*, *scikit-learn*, and *Keras*.

### 3.7.1 *Data structure and description*

The input and output variables used in this study follow the same data structure as those utilized in the ANN model.

### 3.7.2 *Data preprocessing*

Missing values in the dataset were imputed using the mean imputation strategy implemented with the *SimpleImputer* class from the *sklearn library*. Categorical variables (SOB and POB) were encoded into numerical format using *LabelEncoder* for compatibility with machine learning algorithms. To capture potential non-linear relationships among the variables, polynomial features up to the second degree were generated using the *Polynomial* features class from *sklearn*.

### 3.7.2 *Dataset partitioning*

The dataset was split into training and test sets, with 85% allocated for training and 15% for testing. During model training, the training data was further divided into training and validation subsets in an 85:15 ratio to monitor performance and mitigate overfitting. using the *train_test_split* function with a random seed of 42 for reproducibility.

### 3.7.3 *Feature Scaling*

Numerical variables were standardized using the *StandardScaler* function to ensure uniform scaling and optimize model performance during training.

### 3.7.4 *Bayesian neural network architecture*

A BNN model was developed using the *Sequential* model from the *Keras library*. The architecture consisted of:

1. An input layer with 28 neurons and ReLU activation to process the transformed feature set.
2. First hidden layers, each with 56 neurons and ReLU activation, interspersed with a dropout layer (30% dropout rate) for regularization.
3. Second final hidden layer with 28 neurons and ReLU activation.
4. An output layer with a single neuron for predicting continuous target traits.

The model was compiled with the *Adam optimizer* and a mean squared error loss function to minimize prediction error

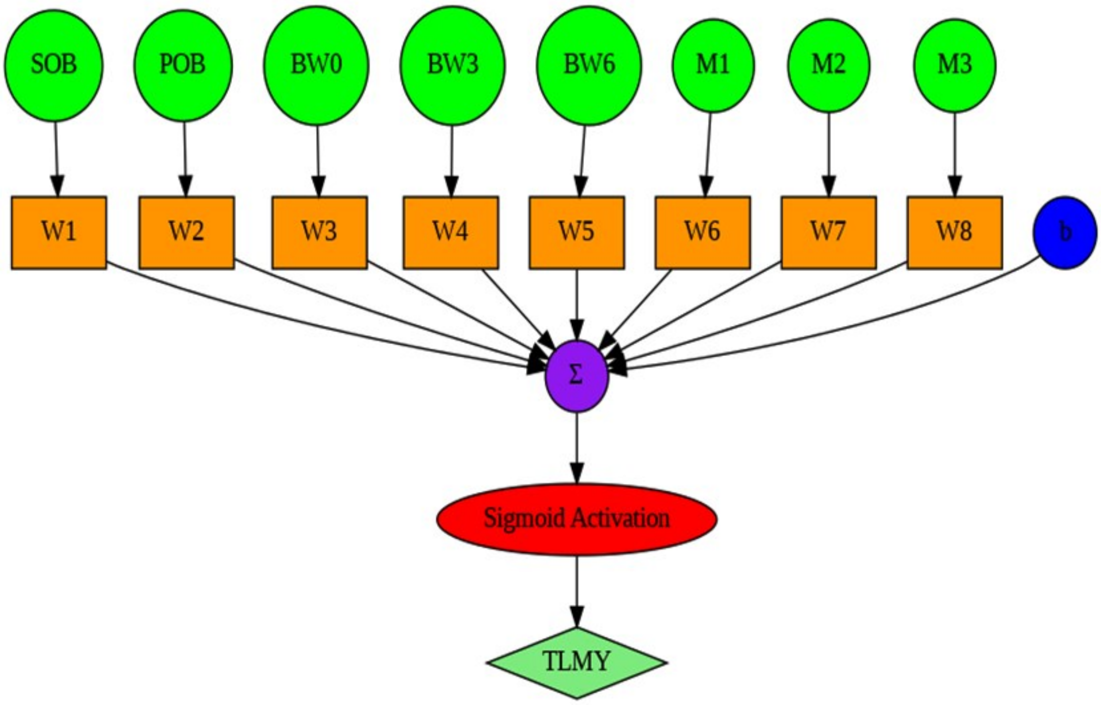


Fig 3.1a: Architecture of ANN for TLMY

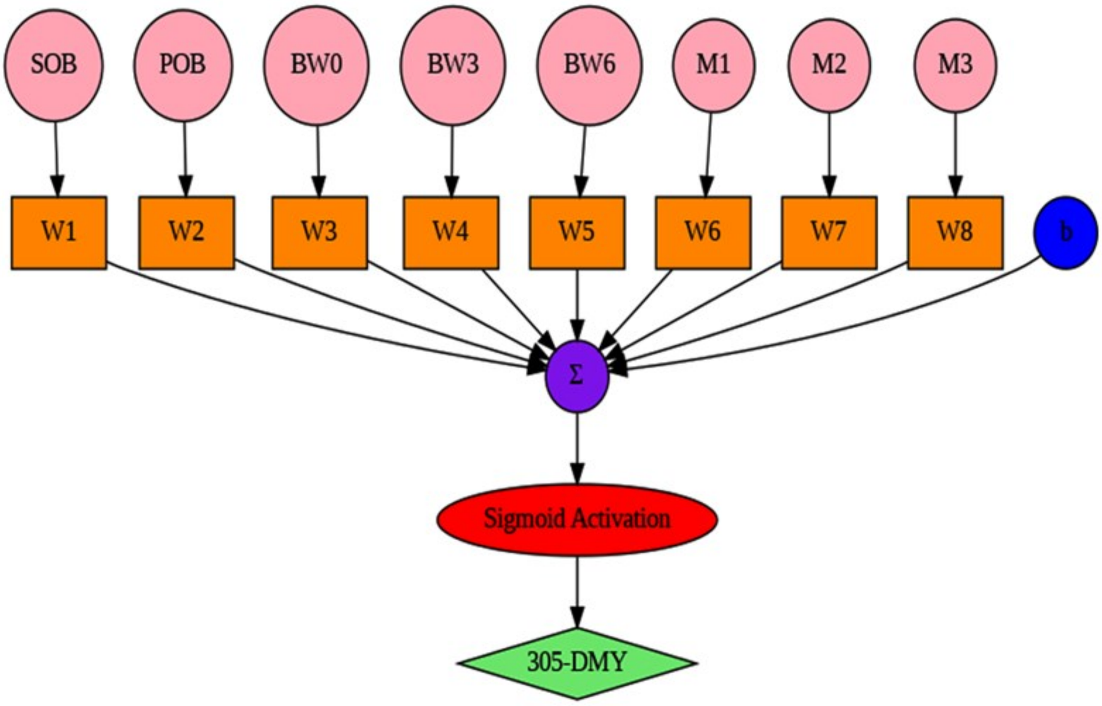


Fig 3.1b: Architecture of ANN for 305-DMY

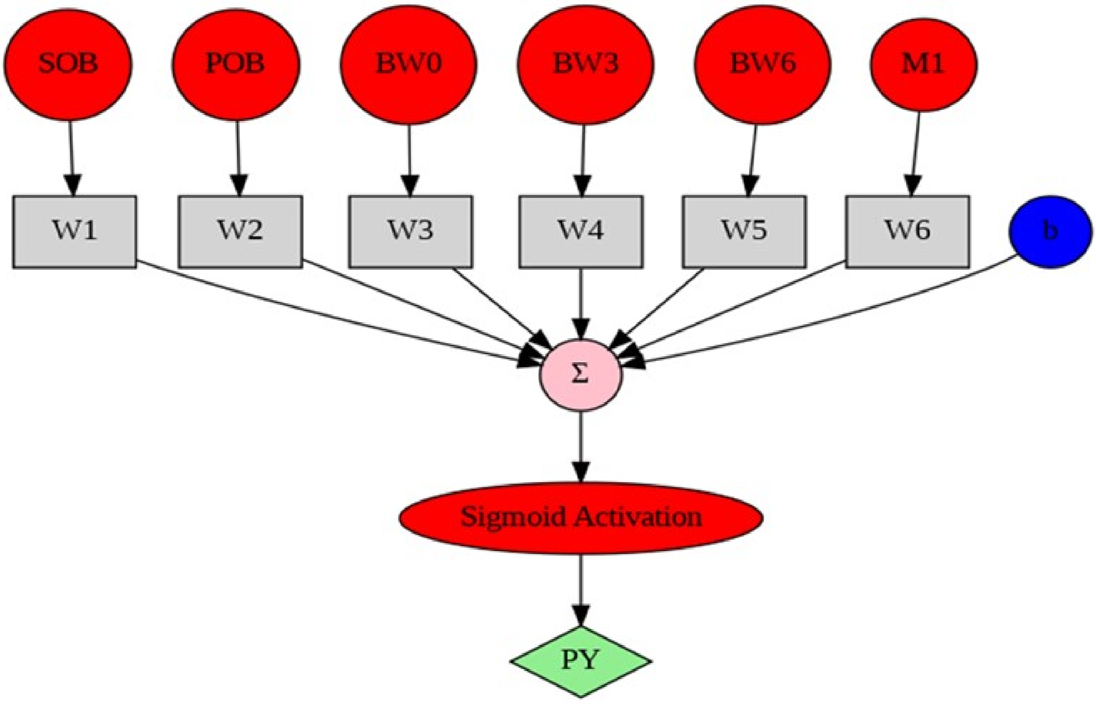


**Fig 3.1c: Architecture of ANN for PY**

### 3.7.5 *Training and Optimization*

The model was trained using a batch size of 52 over a maximum of 1000 epochs. Early stopping with a patience parameter of 40 epochs was applied to monitor validation loss and prevent overfitting. A 15% validation split from the training data was utilized for model validation during training.

### 3.8 *Model assessment and comparison*

The performance of ANN and BNN models was evaluated and compared using standard regression metrics, including mean absolute error (MAE), mean squared error (MSE), root mean squared error (RMSE), R-squared ($R^2$), and Pearson's correlation coefficient (r). These metrics provided a robust framework for assessing prediction accuracy and model reliability.

- MAE: This metric quantifies the average difference between observed and predicted values by averaging the absolute differences across the dataset. Lower MAE values indicate better model accuracy.
- MSE: MSE evaluates the average squared difference between observed and predicted values, penalizing larger errors more significantly, making it suitable for detecting substantial deviations.
- RMSE: RMSE, as the square root of MSE, provides an error measure on the same scale as the original data, offering an intuitive representation of prediction error.
- $R^2$: Also known as the coefficient of determination, $R^2$ measures the proportion of variance in the observed data explained by the model. Ranging from 0 to 1, higher values signify better model performance and closer alignment between predicted and actual values.
- r: This metric evaluates the linear relationship between observed and predicted values. The correlation coefficient (r) ranges from -1 to 1, where values closer to 1 indicate a strong positive correlation, reflecting high predictive accuracy.

$$MAE = \frac{1}{N}\sum_{i=1}^{N} |y_i - \hat{y}|$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N} (y_i - \hat{y})^2$$

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (y_i - \hat{y})^2}$$

$$R^2 = 1 - \frac{\sum (y_i - \hat{y})^2}{\sum (y_i - \bar{y})^2}$$

Where,
$\hat{y}$ − predicted value of $y$
$\bar{y}$ − mean value of $y$

The model that exhibited the highest $R^2$ and r values, along with the lowest MSE, MAE, and RMSE values, was regarded as the most efficient

* * *

# Results

# 4

# RESULTS

Vrindavani is a crossbred population, developed by crossing Hariana cows with sires of exotic inheritance (Holstein-Friesian, Jersey, and Brown Swiss) in various combinations, with subsequent *interse* mating during later generations. Constant assessment of population structure, genetic architecture, and other genealogical parameters, along with their dynamics is essential for the planning and implementation of an efficient breeding programme.

The main demographic characteristics, as derived from genealogical data of the Vrindavani population, have been summarized in Table 4.1. The number of founders with unknown pedigrees (both parents unknown) was 2620, out of the total 12718 animals present in the data, thus accounting for ~20.60% of the total population. The pedigree completeness and number of breeding animals of the Vrindavani population has been summarized in Fig. 4.1; the completeness of ancestral generations decreased with an increase in pedigree depth. The percentage of animals with complete known pedigree in the immediately preceding generation was 80.87%, revealing a near-optimal degree of pedigree completeness, which could be suitable for analysis of the population structure and genetic architecture of this population. The pedigree completeness level was 52.56, 32.87, and 15.97% in the second, third, and fourth generations of the whole pedigree respectively. The pedigree of recent generations was more detailed as compared to older generations. The integrity of pedigree information in a population may be assessed using the equivalent number of complete generations. In the present study, the maximum number of generations (MG) in Vrindavani cattle based on pedigree data was 13 (Fig 4.2), while the number of completed (CG) and equivalent generations (EqG) were 3.23 and 1.95, respectively. The average values of mean maximum generations, mean complete generations, and mean equivalent generations were 3.91, 1.44, and 2.17, respectively.

**Table 4.1: Summary of pedigree analysis of Vrindavani cattle**

| Items | Whole population |
|---|---|
| Total number of animals in whole population | 12718 |
| Total number of effective sires | 478 |
| Total number of effective dams | 4663 |
| Maximum parental family size | 2244 |
| Maximum maternal family size | 2654 |
| Number of matings between half sibs | 146 |
| Number of matings between parent and offspring | 272 |
| Number of inbred animals | 3116 |
| Number of non-inbred animals | 9502 |
| Number of animals with both parent unknown (founder) | 2602 |
| Number of animals with one parent unknown (half founders) | 2435 |
| Number of animals with both known parents | 10098 |
| Number of animals without progeny | 7695 |
| Number of animals with progeny | 5023 |
| Mean maximum generations | 3.91 |
| Mean complete generations | 1.44 |
| Mean equivalent generations | 2.17 |
| Mean inbreeding coefficient (% F) for whole population | 1.11% |
| Mean inbreeding coefficient (% F) for reference population | 3.44% |
| Genetic conservation index (GCI) for the whole population | 4.11 |
| Total number of animals in reference population (year 2010-2023) | 4746 |

## 4.1 Probability of gene origin

The number of founders in the whole population was 2620, while the effective number of founders was 118, representing around 4.48% of the founders. Similarly, the total number of founders in the reference population was 235. On the other hand, the effective number of founders was 16, representing 6.80% of the founder animals in the cohort. The number of ancestors describing 50, 75, and 100 percent of the gene pool in the whole population was 58, 149, and 2194, respectively. The main parameters characterizing probability of gene origin in Vrindavani population summarized in Table 4.2. These losses in population were inevitable for genetic improvement because selected sires were utilized in mating.

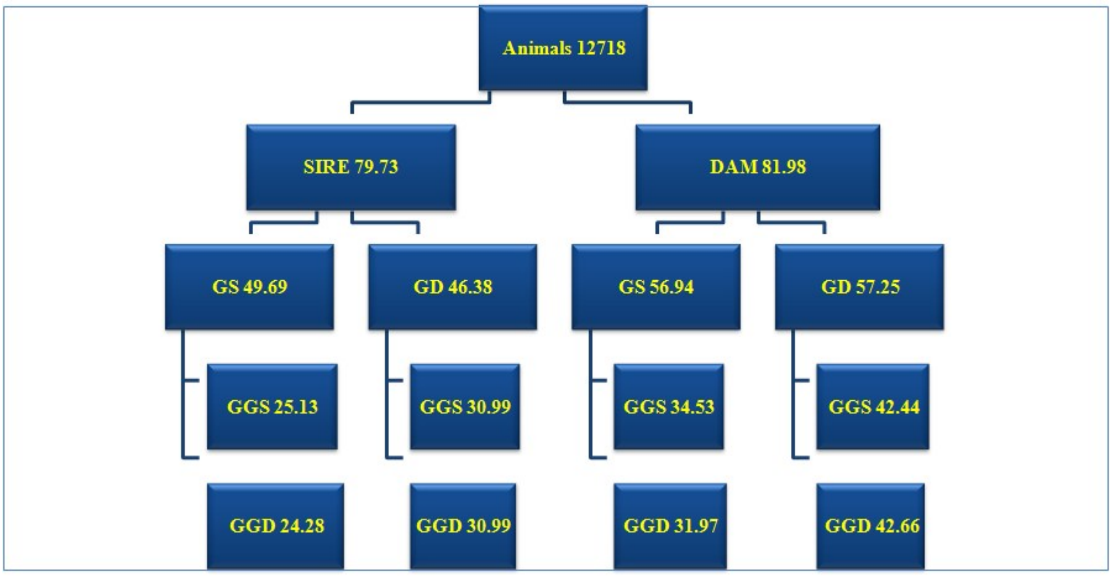


Fig. 4.1: Overview of pedigree complete index (%)

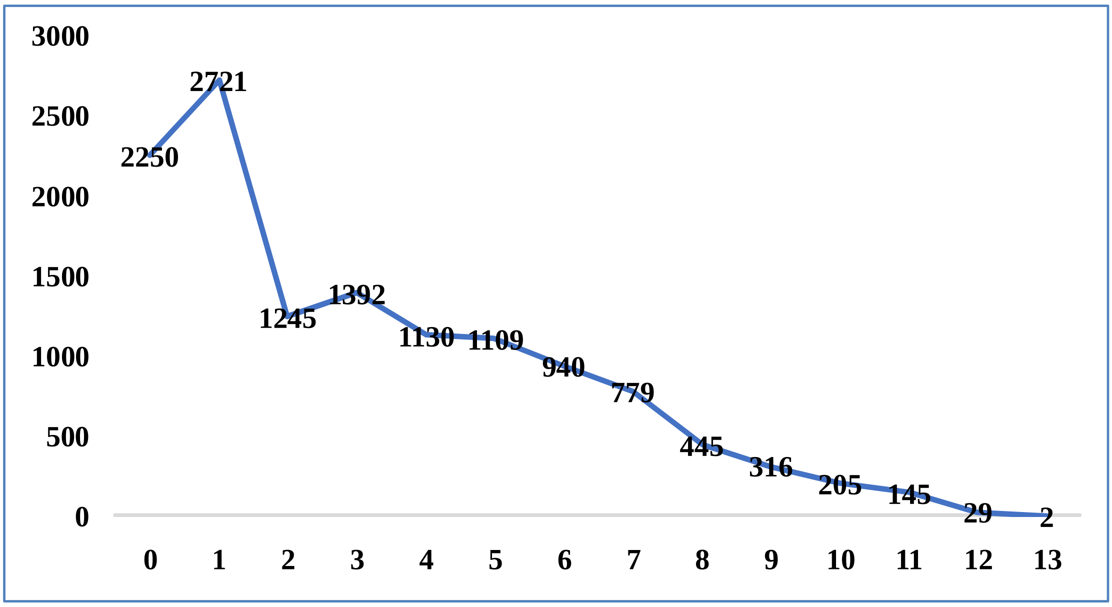


Fig. 4.2: Average known ancestors per generations in Vrindavani cattle

## Table 4.2 Parameters characterizing probability of gene origin in Vrindavani cattle herd

| Parameters characterizing probability of gene origin | Value |
|---|---|
| Total number of animals in population | 12718 |
| **Whole Population** | |
| Number of founders contributing to whole population | 2620 |
| Effective population size of founder | 174.92 |
| Effective number of founders ($f_e$) for whole population | 115 |
| Number of ancestors contributing to whole population | 2194 |
| Effective number of ancestors ($f_a$) in whole population | 78 |
| Effective size obtained from regression on the birth date | 77.40 |
| Effective size obtained from Log regression on the birth date | 71.24 |
| **Reference population** | |
| Number of animals in reference population (2010-2023) | 4746 |
| Number of founders contributing to reference population | 235 |
| Effective number of founders ($f_e$) for reference population | 16 |
| Number of ancestors contributing to reference population | 223 |
| Effective number of ancestors ($f_a$) in reference population | 15 |
| Effective number of Founder genome equivalent ($f_g$) | 9.44 |
| Effective number of non-founder genomes ($f_{ng}$) | 0.17 |
| **Comparison between parameters** | |
| $f_e/f_a$ for whole population | 1.19 |
| $f_e/f_a$ for reference population | 1.20 |
| $1/2f_a$ for reference population (%) | 0.47 |
| $f_e/f_g$ for reference population | 1.69 |
| Expected inbreeding (%) due to unequal founder contributions | 0.29 % |
| Number of ancestors describing 50 %, 75 % and 100 % of the gene pool in whole population | 58, 149, 2194 |
| Number of ancestors explaining 50 %, 75 % and 100 % of the gene pool in reference population | 6, 14, 204 |
| Mean average relatedness (% AR) for whole population | 1.16 % |
| Mean average relatedness (% AR) for reference population | 3.49 |

## 4.2 Generation interval

Generation interval (GI) of a population is an important parameter that is directly related to selection response and yearly trends thereof. Table 4.3 depicts the descriptive details pertaining to the average generation interval of the population and estimates across four possible selection pathways. The mean GI of the whole Vrindavani population was 6.90 years.

**Table 4.3: Generation intervals in (years) for four pathways of Vrindavani cattle for whole population**

| Pathway | Whole population | |
|---|---|---|
| | No. | GI ± SE (years) |
| Sire-Son | 103 | 8.44 ± 0.43 |
| Sire-Daughter | 2763 | 8.49 ± 0.07 |
| Dam-Son | 103 | 6.96 ± 0.58 |
| Dam-Daughter | 2781 | 5.29 ± 0.05 |
| Total | 5750 | 6.90 ± 0.05 |

GI = Generation intervals; SE = Standard error

## 4.3 Inbreeding and average relatedness

The mean coefficient of inbreeding (%F) for the whole population was 1.1%, while the estimate was 3.44% for the reference population. The average increase in the coefficient of inbreeding over maximum, equivalent, and complete generations was 0.43, 1.09, and 0.89%, respectively (Table 4.4). The percentage of inbred animals in the whole population was 25.23%. In the whole population, 25.23% of the animals were inbred, out of which 48.57% were males and the rest were females. However, males (1.19%) had a slightly higher inbreeding coefficient than females (1.05%). This shows an corporately intense selection of males over females. The herd was under constant selection pressure, and accordingly, males were stringently selected based on the dams' yield. Due to the closed flock structure with no inheritance from outside gene pool, the average inbreeding coefficient and percentage of inbred animals increased with increasing generation number.

The mean average relatedness (%AR) for the whole population was 1.16%, while the estimate was 5.49% for the reference cohort. The interpretation of average relatedness and inbreeding parameters should take into contribution the nature of the flock and pedigree depth. AR was greater than the mean inbreeding coefficient in the Vrindavani population. The increase in inbreeding estimates with maximum, complete, and equivalent generations was 0.43, 0.89, and 1.09 %, respectively. The mean inbreeding coefficient, percentage of inbred animals, and average relatedness among animals increased across the completed generations. A higher average relatedness combined with a lower inbreeding coefficient indicates a high degree of relatedness among all individuals in the pedigree dataset. This could lead to difficulties while trying to avoid mating between unrelated or distantly related individuals. About 5.52 % of total matings were highly inbred. Out of which, 1.15% (n=146) were half-sib and 2.14% (n=272) were parent-offspring mating. The expected contribution to inbreeding due to the unbalanced number of founders was 0.29% in the Vrindavani population.

Table 4.5 depicts the estimates of average inbreeding coefficient (F), percentage of inbred individuals (POR), average F for inbred animals (FP), and average relatedness coefficient (AR) across the maximum and completed generations. The mean inbreeding coefficient increased irregularly from the second generation (0.29) onwards during subsequent generations. The estimate was the highest in the 13th generation (6.36%).

**Table 4.4: Estimates of the increase in inbreeding and effective population size in the Vrindavani cattle**

| Parameter | Method of estimation | Value |
|---|---|---|
| Inbreeding increase (%) | Maximum generations | 0.43% |
| | Equivalent generations | 1.09% |
| | Complete generations | 0.89% |
| Effective population size | Maximum generations | 115.76 |
| | Equivalent generations | 40.58 |
| | Complete generations | 46.04 |
| | Individual increase in inbreeding | 56.42 |
| | Regression on equivalent generations | 39.59 |
| | Log regression on equivalent generations | 50.34 |
| | Regression on birth date | 77.40 |
| | Log regression on birth date | 71.24 |

**Table 4.5: Generations (G), number of animals (N), average inbreeding coefficient (F), percentage of inbred individuals (POR), average F for inbred animals (FP), and average relatedness coefficient (AR) per generation for Vrindavani cattle using Maximum number of generations (MG) and complete number of generations (CG) traced**

| Generation | No of animals | Mean F | % Inbred | Average F for inbred | AR |
|---|---|---|---|---|---|
| | **Maximum generation (MG)** | | | | |
| 0 | 2250 | 0.00% | 0.00% | 0.00% | 0.04% |
| 1 | 2721 | 0.00% | 0.00% | 0.00% | 0.16% |
| 2 | 1245 | 0.29% | 1.53% | 19.08% | 0.67% |
| 3 | 1392 | 0.96% | 5.24% | 18.32% | 0.97% |
| 4 | 1130 | 1.42% | 9.56% | 14.81% | 1.69% |
| 5 | 1109 | 1.71% | 12.89% | 13.24% | 2.03% |
| 6 | 940 | 2.51% | 25.43% | 9.88% | 2.29% |
| 7 | 779 | 2.84% | 45.19% | 6.29% | 2.73% |
| 8 | 455 | 4.27% | 62.20% | 6.86% | 3.06% |
| 9 | 316 | 2.89% | 55.70% | 5.19% | 3.08% |
| 10 | 205 | 3.02% | 62.44% | 4.84% | 3.23% |
| 11 | 145 | 4.47% | 83.45% | 5.36% | 3.29% |
| 12 | 29 | 7.05% | 89.66% | 7.86% | 3.39% |
| 13 | 2 | 6.36% | 50.00% | 12.725 | 2.43% |
| | | **Complete generations (CG)** | | | |
| 0 | 2620 | 0.00% | 0.00% | 0.00% | 0.08% |
| 1 | 4688 | 0.77% | 4.61% | 16.78% | 0.73% |
| 2 | 3516 | 1.05% | 8.25% | 12.70% | 1.54% |
| 3 | 1607 | 2.63% | 55.63% | 4.73% | 2.86% |

## 4.4 Effective population size

The estimates of effective population size ($N_e$) across the years in the Vrindavani population have been depicted in Fig 4.3. The average Ne estimates from maximum, equivalent, and complete generations were 115.56, 56.42, and 46.02, respectively. The effective population size based on regression and log regression on birth date was 77.40 and 71.24, respectively. The effective population size of the founders was 174.92

## 4.5 Genetic Conservation Index (GCI) and genetic diversity (GD)

Individuals would typically get equal contributions from all the founding descendants of a population, leading to better values for animal conservation. An animal's conservation value increases with its GCI score. The average estimate of GCI in all animals was 4.11, and the males recorded higher GCI than the females. The mean values of GCI for all individuals in the population improved continuously over time and reached a maximum value of 15.03, which was estimated for one individual. The proportion of animals with a GCI greater than 10 was 3.22%.

The estimated value of GD in the reference population relative to the base population or Nei expected heterozygosity was 0.947, indicating that 5.3 % of the genetic diversity in the base population was lost during the study period. However, the genetic diversity estimated for the base population ($GD^*$) was 0.969. In the founder generation, losses in heterozygosity due to genetic drift and bottleneck effects (1-GD) and uneven contributions of founder alleles (1-GD*) were calculated as 0.053 and 0.031, respectively.

## 4.6 Descriptive statistics of growth, production, and reproduction traits

Table 4.6 presents the descriptive statistics for the growth, production, and reproduction traits of Vrindavani crossbred cows, providing an overview of the number of records, mean, standard deviation (SD), coefficient of variation (COV), and range (minimum to maximum) for each trait. The means for growth traits were 23.15±4.10 kg (BW0), 47.46±10.89 kg (BW3), 84.71±14.98 kg (BW6), 156.05±20.42 kg (BW12), 228.78±25.57 kg (BW18), and 285.03±32.87 kg (BW24). The means for production traits were 3,113.24±885.84 kg (TLMY), 2,977.5±853.05 kg (305-DMY), 298.35±59.8 days (LL), and 14.54±4.80 kg (PY). The means for reproduction traits were 979±180 days (AFC), 435±112.4 days (ICL), 104±27.78 days (DD), 278.6±11.61 days (GP), and 140.58±46.89 days (SP). These results indicate varying levels of diversity across the growth, production, and reproduction traits in Vrindavani crossbred cows.

**Table 4.6: Data structure and summary statistics for growth, production, and reproduction traits of Vrindavani crossbred cattle**

| Traits | No of records | Mean | SD | COV | Minimum | Maximum |
|---|---|---|---|---|---|---|
| BW0 (kg) | 8841 | 23.15 | 4.10 | 17.71% | 15 | 44 |
| BW3 (kg) | 6517 | 47.46 | 10.89 | 22.94% | 35 | 110 |
| BW6 (kg) | 5183 | 84.72 | 14.98 | 17.68% | 40 | 195 |
| BW12 (kg) | 4099 | 156.05 | 20.42 | 13.09% | 89 | 255 |
| BW18 (kg) | 3610 | 228.78 | 25.57 | 11.77% | 115 | 340 |
| BW24 (kg) | 3275 | 285.03 | 32.87 | 11.53% | 175 | 420 |
| TLMY (kg) | 7977 | 3113.24 | 885.84 | 28.54% | 1000 | 8243.4 |
| 305-DMY (kg) | 2374 | 2977.5 | 853.05 | 28.64% | 1275 | 7186.5 |
| LL (days) | 7977 | 298.35 | 59.80 | 20.04% | 84 | 610 |
| PY (kg) | 6833 | 14.54 | 4.80 | 33.01% | 2.5 | 33.01 |
| AFC (days) | 2979 | 979 | 180 | 18.38 % | 703 | 2332 |
| ICL (days) | 6265 | 435 | 112.40 | 25.83 % | 312 | 1671 |
| DD (days) | 5399 | 104 | 27.78 | 26.71 % | 20 | 295 |
| GP (days) | 5873 | 278.60 | 11.61 | 4.17 % | 235 | 433 |
| SP (days) | 5999 | 140.58 | 46.89 | 33.35 % | 43 | 387 |

## 4.7 Covariance components and genetic parameters for growth traits

Table 4.7 provides a comprehensive summary of the (co)variance components and genetic parameters for various body weight traits (BW0, BW3, BW6, BW12, BW18, and BW24) in crossbred cattle across eight models, offering insights into genetic and environmental influences. For BW0, the additive genetic variance ($\sigma^2_a$) ranged from 2.331 in model 8 to 3.428 in model V, reflecting variability in genetic contributions. Maternal genetic variance ($\sigma^2_m$) appears in models III to VIII, spanning from 0.130 to 1.252, while antagonistic maternal-additive covariance ($\sigma^2_{am}$) was evident in models V and VI with negative values (-2.15 and -2.19). Residual variance ($\sigma^2_e$) was stable across models, between 13.632 and 13.772, with phenotypic variance ($\sigma^2_p$) varying slightly from 16.105 to 17.3159. Direct heritability ($h^2$) estimates ranged from 0.134 in model VIII to 0.225 in Model V, while maternal heritability ($m^2$) spanned 0.007 to 0.082 in models III to VIII. Sire variance ($\sigma^2_s$) was significant in models VII and VIII, with values of 0.0636 and 0.0624. A negative genetic correlation ($r_m$) between additive and maternal effects was observed in models V and VI.

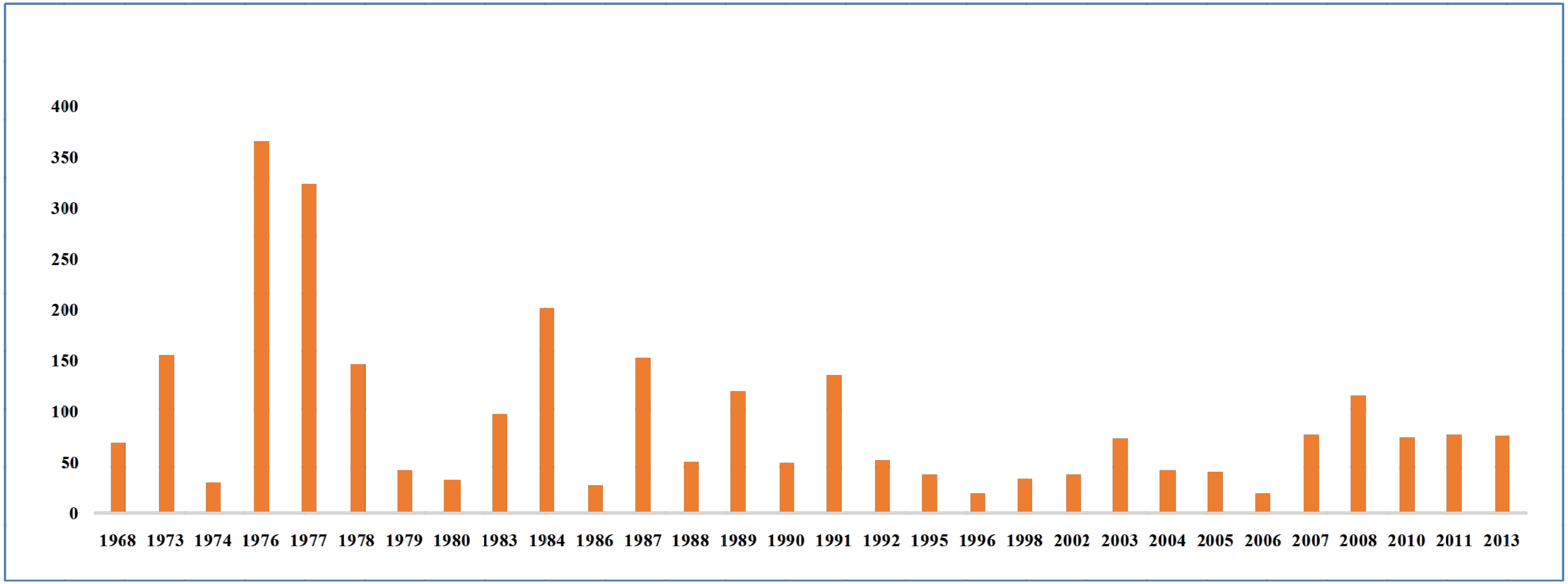


Fig. 4.3: Effective population size (year wise) of Vrindavani cattle

For BW3, $\sigma^2_a$ ranged from 21.390 in model IV to 27.474 in model VI, with $\sigma^2_m$ present in models II to VIII, ranging from 0.10 in models VII and VIII to 16.856 in model VI. Negative $\sigma^2_{am}$ was scored in models V and VI (-12.149 and -11.158), indicating antagonistic interactions. The $\sigma^2_e$ is stable, ranging from 56.309 in model V to 74.084 in model VIII, while $\sigma^2_p$ varied between 88.232 in model V and 97.655 in model VIII. Heritability estimates range from 0.229 in model VIII to 0.3091 in model V, with maternal heritability ranging from 0.019 to 0.188 in models III to V. The $\sigma^2_s$ in Models VII and VIII is 1.194 and 1.168, respectively.

For BW6, $\sigma^2_a$ fluctuated between 25.843 in model IV and 33.354 in model VI, with $\sigma^2_m$ present in models II to VIII, spanning from 0.1 in models VII and VIII to 20.956 in model VI. Negative $\sigma^2_{am}$ was observed in models V and VI (-18.354 and -12.350). The $\sigma^2_e$ ranged from 181.259 in model VI to 199.236 in model VIII, while $\sigma^2_p$ peaked at 229.95 in model VIII and is lowest at 219.43 in model IV. Heritability estimates ranged from 0.118 in model I to 0.149 in model VI, with $m^2$ ranging from 0.007 in models III and IV to 0.093 in model VI. The $\sigma^2_s$ estimates in models VII and VIII were similar (1.322 and 1.331).

For BW12, $\sigma^2_a$ ranged from 15.510 in model III to 18.397 in model VI, and $\sigma^2_m$ in models II to VIII, ranging from 0.10 in models VII and VIII to 4.546 in model VI. The $\sigma^2_e$ was relatively stable, varying from 128.81 in model IV to 134.02 in model VIII, while $\sigma^2_p$ spanned 142.988 in model VI to 230.06 in model VII. Heritability estimates ranged from 0.105 in model III to 0.128 in models V and VI, with $m^2$ ranging between 0.018 and 0.031 in models III to VI. The $\sigma^2_s$ estimates in models VII and VIII were 1.094 and 1.096, respectively.

For BW18, $\sigma^2_a$ and $\sigma^2_m$ trends mirrored those in BW12, with notable negative $\sigma^2_{am}$ in models IV and V (-4.10 and -4.30). The $\sigma^2_e$ ranges from 128.81 in model IV to 134.02 in model VIII, with $\sigma^2_p$ between 142.99 in model VI and 230.06 in model VII. Heritability ranged from 0.105 to 0.128 across the models, and $m^2$ spanned from 0.018 to 0.031 in Models III to VI.

For BW24, $\sigma^2_a$ varied from 8.27 in model VII to 11.90 in model VI, while the $\sigma^2_m$ ranges from 0.10 in models VII and VIII to 2.83 in model VI. A notable negative $\sigma^2_{am}$ was observed in models V and VI (-4.02 and -3.91, respectively). The $\sigma^2_e$ spanned from 67.59 in

**Table 4.7: (Co) Variance structure and genetic parameters for different models applied on body weight traits in Vrindavani while employing various animal models with inclusion or exclusion of different fixed and random effects**

| TRAIT | MODEL | $\sigma^2_a$ | $\sigma^2_m$ | $\sigma^2_{am}$ | $\sigma^2_s$ | $\sigma^2_e$ | $\sigma^2_p$ | $h^2$ | $m^2$ | $s^2$ | $r_{am}$ | Log L |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| BW0 | Model I | 2.39 | - | - | - | 13.74 | 16.13 | 0.15 | - | - | - | -16388.08 |
| BW0 | Model II | 2.39 | - | - | - | 13.74 | 16.13 | 0.15 | - | - | - | -16383.76 |
| BW0 | Model III | 2.35 | 0.15 | - | - | 13.63 | 16.13 | 0.15 | 0.01 | - | - | -16385.36 |
| BW0 | Model IV | 2.35 | 0.15 | - | - | 13.63 | 16.13 | 0.15 | 0.01 | - | - | -16381.06 |
| BW0 | Model V | 3.43 | 1.25 | -2.15 | - | 12.72 | 15.25 | 0.22 | 0.08 | - | -0.96 | -16347.72 |
| BW0 | Model VI | 3.33 | 1.05 | -2.19 | - | 12.82 | 15.01 | 0.22 | 0.07 | - | -0.95 | -16350.07 |
| BW0 | Model VII | 2.35 | 0.13 | - | 1.10 | 13.73 | 17.32 | 0.14 | 0.01 | 0.06 | - | -16346.86 |
| BW0 | Model VIII | 2.33 | 0.13 | - | 1.08 | 13.77 | 17.31 | 0.13 | 0.01 | 0.06 | - | -16350.86 |
| BW3 | Model I | 21.82 | - | - | - | 66.38 | 88.20 | 0.25 | - | - | - | -17663.19 |
| BW3 | Model II | 21.79 | - | - | - | 66.41 | 88.20 | 0.25 | - | - | - | -17663.09 |
| BW3 | Model III | 21.44 | 1.68 | - | - | 65.25 | 88.36 | 0.24 | 0.02 | - | - | -17659.49 |
| BW3 | Model IV | 21.39 | 1.65 | - | - | 65.32 | 88.35 | 0.24 | 0.02 | - | - | -17659.47 |
| BW3 | Model V | 27.28 | 16.79 | -12.15 | - | 56.31 | 88.23 | 0.31 | 0.19 | - | -0.97 | -17578.43 |
| BW3 | Model VI | 27.47 | 16.86 | -11.16 | - | 56.30 | 89.47 | 0.31 | 0.19 | - | -0.95 | -17560.08 |
| BW3 | Model VII | 22.50 | 0.10 | - | 1.19 | 74.02 | 97.72 | 0.23 | 0.00 | 0.01 | - | -17560.78 |
| BW3 | Model VIII | 22.40 | 0.10 | - | 1.17 | 74.08 | 97.65 | 0.23 | 0.00 | 0.01 | - | -17561.25 |
| BW6 | Model I | 26.00 | - | - | - | 193.43 | 219.42 | 0.12 | - | - | - | -16496.54 |
| BW6 | Model II | 26.07 | - | - | - | 193.26 | 219.33 | 0.12 | - | - | - | -16496.71 |
| BW6 | Model III | 25.78 | 1.58 | - | - | 192.17 | 219.52 | 0.12 | 0.01 | - | - | -16495.93 |
| BW6 | Model IV | 25.84 | 1.60 | - | - | 191.99 | 219.43 | 0.12 | 0.01 | - | - | -16496.08 |
| BW6 | Model V | 31.29 | 19.67 | -18.35 | - | 188.00 | 220.61 | 0.14 | 0.09 | - | -0.95 | -16452.72 |
| BW6 | Model VI | 33.35 | 20.96 | -12.35 | - | 181.26 | 223.21 | 0.15 | 0.09 | - | -0.94 | -16452.71 |
| BW6 | Model VII | 29.38 | 0.10 | - | 1.32 | 199.36 | 230.06 | 0.13 | 0.00 | 0.01 | - | -16448.17 |
| BW6 | Model VIII | 29.38 | 0.10 | - | 1.33 | 199.23 | 229.95 | 0.13 | 0.00 | 0.01 | - | -16448.19 |
| BW12 | Model I | 15.52 | - | - | - | 131.39 | 146.91 | 0.11 | - | - | - | -12234.25 |
| BW12 | Model II | 15.67 | - | - | - | 131.26 | 146.93 | 0.11 | - | - | - | -12235.36 |

**Table 4.7 Contd...**

| TRAIT | MODEL | $\sigma^2_a$ | $\sigma^2_m$ | $\sigma^2_{am}$ | $\sigma^2_s$ | $\sigma^2_e$ | $\sigma^2_p$ | $h^2$ | $m^2$ | $s^2$ | $r_{am}$ | Log L |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| BW12 | Model III | 15.51 | 2.59 | - | - | 128.96 | 147.06 | 0.11 | 0.02 | - | - | -12231.97 |
| BW12 | Model IV | 15.65 | 2.62 | - | - | 128.81 | 147.08 | 0.11 | 0.02 | - | - | -12233.06 |
| BW12 | Model V | 18.34 | 4.53 | -9.49 | - | 129.96 | 143.34 | 0.13 | 0.03 | - | -0.99 | -12190.32 |
| BW12 | Model VI | 18.40 | 4.55 | -9.07 | - | 129.11 | 142.99 | 0.13 | 0.03 | - | -0.93 | -12191.36 |
| BW12 | Model VII | 16.58 | 0.10 | - | 1.09 | 134.02 | 230.06 | 0.11 | 0.00 | 0.01 | - | -12193.97 |
| BW12 | Model VIII | 16.56 | 0.10 | - | 1.10 | 134.01 | 229.95 | 0.11 | 0.00 | 0.01 | - | -12195.01 |
| BW18 | Model I | 8.89 | - | - | - | 82.37 | 91.25 | 0.10 | - | - | - | -9924.15 |
| BW18 | Model II | 9.41 | - | - | - | 82.12 | 91.53 | 0.10 | - | - | - | -9928.44 |
| BW18 | Model III | 8.97 | 2.65 | - | - | 79.72 | 91.34 | 0.10 | 0.03 | - | - | -9918.98 |
| BW18 | Model IV | 9.48 | 2.79 | - | - | 79.36 | 91.62 | 0.10 | 0.03 | - | - | -9922.84 |
| BW18 | Model V | 11.36 | 3.52 | -4.42 | - | 78.06 | 89.10 | 0.13 | 0.04 | - | -0.98 | -9872.81 |
| BW18 | Model VI | 11.59 | 3.53 | -4.30 | - | 78.69 | 89.51 | 0.13 | 0.04 | - | -0.97 | -9875.01 |
| BW18 | Model VII | 8.75 | 0.10 | - | 1.24 | 82.91 | 92.91 | 0.09 | 0.00 | 0.01 | - | -9883.08 |
| BW18 | Model VIII | 9.07 | 0.10 | - | 1.28 | 82.86 | 93.22 | 0.10 | 0.00 | 0.01 | - | -9885.43 |
| BW24 | Model I | 8.96 | - | - | - | 70.93 | 79.90 | 0.11 | - | - | - | -8782.17 |
| BW24 | Model II | 9.62 | - | - | - | 70.67 | 80.29 | 0.12 | - | - | - | -8788.31 |
| BW24 | Model III | 9.55 | 2.51 | - | - | 67.99 | 80.05 | 0.12 | 0.03 | - | - | -8776.40 |
| BW24 | Model IV | 10.19 | 2.67 | - | - | 67.59 | 80.45 | 0.13 | 0.03 | - | - | -8781.94 |
| BW24 | Model V | 11.32 | 2.72 | -4.02 | - | 69.16 | 79.19 | 0.14 | 0.03 | - | -0.96 | -8733.37 |
| BW24 | Model VI | 11.90 | 2.83 | -3.91 | - | 69.04 | 80.18 | 0.15 | 0.04 | - | -0.97 | -8736.90 |
| BW24 | Model VII | 8.27 | 0.10 | - | 0.71 | 72.22 | 81.19 | 0.10 | 0.00 | 0.01 | - | -8739.27 |
| BW24 | Model VIII | 8.57 | 0.10 | - | 0.73 | 72.24 | 81.55 | 0.11 | 0.00 | 0.01 | - | -8742.87 |

model IV to 72.24 in model VIII, with $\sigma^2_p$ ranging between 79.90 in model I and 81.55 in model VIII. Heritability was lowest in model I at 0.11 and peaked at 0.15 in model VI. Maternal variance ranged between 0.01 in models VII and VIII and 0.04 in models III and IV.

Across all traits, the findings emphasized the intricate interaction among the additive genetic effects, maternal influences, and environmental factors in determining phenotypic variance in crossbred cattle population. The genetic parameters for various models applied to body weight traits have been presented in Fig. 4.4.

### 4.7.1 *Effect of inbreeding on growth traits across different models*

The effect of inbreeding on various models and growth traits of Vrindavani cattle revealed significant differences when Fx was included or excluded. For body weight traits, models incorporating Fx (Models I, III, V, VII) consistently demonstrated lower ó$^2$a compared to their counterparts without Fx (Models II, IV, VI, VIII), indicating that inbreeding reduces genetic variance. This pattern was observed across all growth traits, including BW0, BW3, BW6, BW12, BW18, and BW24. Additionally, models accounting for Fx were revealed to exhibit slightly reduced heritability estimates, further highlighting the adverse effects of inbreeding. Overall, the inclusion of inbreeding effects led to a reduction in both genetic variance and heritability estimates, underscoring the detrimental impact of inbreeding on the genetic potential and performance of Vrindavani cattle. Specifically, the percentage decrease in ó$^2$a due to inbreeding was: 3.75% for BW0, 4.76% for BW3, 10.7% for BW6, to 4.92% for BW12, 5.52% for BW18, and 12.3% for BW24. These findings highlighted the importance of managing inbreeding in order to preserve the genetic diversity and maintain the productivity of Vrindavani cattle.

## 4.8 Covariance components and genetic parameters for production traits

Table 4.8 provides a comprehensive summary of (co)variance components and genetic parameters for TLMY, 305-DMY, LL, and PY in crossbred Vrindavani cattle across eight models, highlighting the contributions of genetic and environmental factors to the observed phenotypic variance.

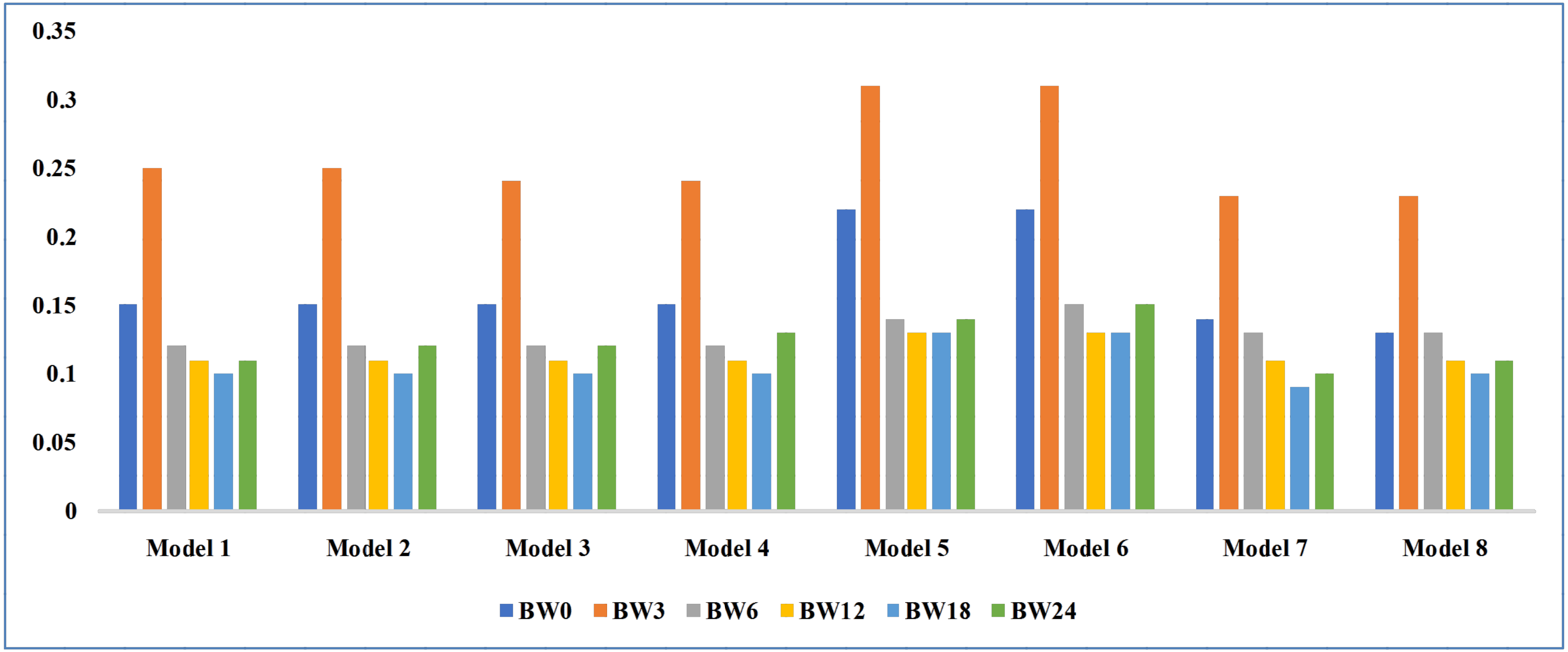


Fig.4.4: Heritability estimates for six growth traits across different models

For TLMY, the $\sigma^{2}$ ranged from 9.2645 in model VII to 17.162 in model VI. The $\sigma^{2}$˜ was present in models III through VIII, with the highest value observed in model VI (1.371) and the lowest non-zero values in models III and IV (0.10). models V and VI also included $\sigma^{2}$˜ , showing negative values (-4.012 and -4.164). the $\sigma^{2}$› was incorporated in models VII and VIII, with values of 2.092 and 2.021, respectively, highlighting sire contributions to phenotypic variance. Permanent environmental variance ($\sigma^{2}$š ‘ ) was consistently present across all models, ranging from 8.896 in models V and VI to 11.719 in models I and III. Heritability estimates varied from 0.131 in model VII to 0.247 in model VI, with the latter indicating a greater proportion of phenotypic variance attributable to additive genetic factors.

In the case of 305-DMY, $\sigma^{2}$ ranged from 3.032 in model VII to 4.241 in model VI, showing an increase of approximately 37% in models V and VI compared to model I. The $\sigma^{2}$˜ was highest in models V and VI. The $\sigma^{2}$› was present in models VII and VIII, showing a slight increase of 2.4% in model VIII compared to model VII. The $\sigma^{2}$š ‘ peaked in model V, with a 12.5% increase relative to model I, reflecting its substantial impact. The $\sigma^{2}$‘ remained stable across models, while heritability estimates ranged from 0.145 in models VII and VIII to 0.195 in model VI, indicating a 31% higher genetic contribution to phenotypic variance in model VI compared to model I.

For LL, the $\sigma^{2}$ showed notable increases in models V and VI compared to model I. The ó$^{2}$˜ was included in models III through VIII, with the highest values observed in models VII and VIII. The covariance between $\sigma^{2}$ ˜ exhibited negative values, emphasizing antagonistic interactions. The $\sigma^{2}$› increased marginally in model VIII relative to model VII, while $\sigma^{2}$š ‘ peaked in model VII, highlighting its influence on lactation length. The $\sigma^{2}$‘ remained consistent, with heritability estimates varying from 0.036 in models VII and VIII to 0.085 in model VI, representing a 39% increase in genetic contributions in model VI.

For PY, $\sigma^{2}$ ranged from 1.850 in model VIII to 3.909 in model V, reflecting variations in additive genetic variance. The $\sigma^{2}$˜ was present in models III through VIII, with values ranging from 0.100 to 0.206. The $\sigma^{2}$‘ remained stable at approximately 11.7 across models. Heritability estimates ranged from 0.11 in model VIII to 0.23 in model V, with maternal and

**Table 4.8: (Co)Variance structure and genetic parameters for different models applied on production traits in Vrindavani while employing various animal models with inclusion or exclusion of different fixed and random effects**

| Traits | MODEL | $\sigma^2_a$ | $\sigma^2_m$ | $\sigma^2_{am}$ | $\sigma^2_s$ | $\sigma^2_{pe}$ | $\sigma^2_e$ | $\sigma^2_p$ | $h^2$ | $m^2$ | $s^2$ | $pe^2$ | $r_{am}$ | Log L |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| TLMY | Model I | 10.96 | - | - | - | 11.72 | 46.25 | 67.11 | 0.16 | - | - | 0.17 | - | -20327.4 |
| TLMY | Model II | 11.11 | - | - | - | 11.56 | 46.26 | 68.93 | 0.16 | - | - | 0.17 | - | -20326.4 |
| TLMY | Model III | 10.97 | 0.10 | - | - | 11.72 | 46.25 | 68.93 | 0.16 | 0.01 | - | 0.17 | - | -20327.4 |
| TLMY | Model IV | 11.12 | 0.10 | - | - | 11.56 | 46.26 | 68.94 | 0.16 | 0.01 | - | 0.17 | - | -20326.4 |
| TLMY | Model V | 16.87 | 1.26 | -4.01 | - | 9.13 | 46.19 | 69.45 | 0.24 | 0.02 | - | 0.13 | -0.87 | -20323.9 |
| TLMY | Model VI | 17.16 | 1.37 | -4.16 | - | 8.90 | 46.20 | 69.47 | 0.25 | 0.02 | - | 0.13 | -0.86 | -20322.6 |
| TLMY | Model VII | 9.26 | 0.10 | - | 2.09 | 13.20 | 46.10 | 70.65 | 0.13 | 0.01 | 0.03 | 0.19 | - | -20324.4 |
| TLMY | Model VIII | 9.78 | 0.10 | - | 2.02 | 13.10 | 46.10 | 71.00 | 0.14 | 0.01 | 0.03 | 0.19 | - | -20323.5 |
| 305-DMY | Model I | 3.11 | - | - | - | 3.26 | 14.56 | 20.93 | 0.15 | - | - | 0.16 | - | -8250.57 |
| 305-DMY | Model II | 3.12 |  |  |  | 3.25 | 14.55 | 20.92 | 0.15 |  |  | 0.16 |  | -8248.54 |
| 305-DMY | Model III | 3.12 | 0.10 | - | - | 3.26 | 14.56 | 20.94 | 0.15 | 0.01 | - | 0.16 | - | -8250.57 |
| 305-DMY | Model IV | 3.13 | 0.10 |  |  | 3.26 | 14.55 | 20.93 | 0.15 | 0.01 |  | 0.16 |  | -8248.55 |
| 305-DMY | Model V | 4.22 | 0.77 | -0.50 | - | 3.67 | 14.55 | 20.67 | 0.19 | 0.01 | - | 0.18 | -0.97 | -8251.33 |
| 305-DMY | Model VI | 4.24 | 0.78 | -0.51 |  | 3.36 | 14.54 |  | 0.20 | 0.01 |  | 0.16 | -0.88 | -8249.33 |
| 305-DMY | Model VII | 3.03 | 0.10 | - | 0.46 | 3.31 | 14.55 | 20.94 | 0.15 | 0.01 | 0.00 | 0.16 | - | -8250.55 |
| 305-DMY | Model VIII | 3.04 | 0.10 | - | 0.47 | 3.30 | 14.55 | 20.94 | 0.15 | 0.01 | 0.00 | 0.16 | - | -8248.53 |
| LL | Model I | 1.75 | - | - | - | 4.24 | 22.45 | 28.43 | 0.06 | - | - | 0.15 | - | -17171.8 |
| LL | Model II | 1.77 | - | - | - | 4.18 | 22.46 | 28.41 | 0.06 | - | - | 0.15 | - | -17168.9 |
| LL | Model III | 1.73 | 0.23 | - | - | 4.23 | 22.45 | 28.43 | 0.06 | 0.01 | - | 0.15 | - | -17171.8 |
| LL | Model IV | 1.75 | 0.36 |  |  | 4.17 | 22.46 | 28.41 | 0.06 | 0.01 |  | 0.15 |  | -17168.9 |
| LL | Model V | 2.39 | 0.37 | -0.68 | - | 3.94 | 22.43 | 28.45 | 0.08 | 0.01 | - | 0.14 | -0.72 | -17170.3 |
| LL | Model VI | 2.42 | 0.38 | -0.69 | - | 3.88 | 22.44 | 28.44 | 0.09 | 0.01 | - | 0.14 | -0.72 | -17167.4 |
| LL | Model VII | 1.02 | 0.34 | - | 0.35 | 4.62 | 22.43 | 28.45 | 0.04 | 0.01 | 0.01 | 0.16 | - | -17170.3 |
| LL | Model VIII | 1.03 | 0.42 | - | 0.35 | 4.57 | 22.44 | 28.44 | 0.04 | 0.01 | 0.01 | 0.16 | - | -17167.4 |

**Table 4.8: Contd...**

| Traits | MODEL | $\sigma^2_a$ | $\sigma^2_m$ | $\sigma^2_{am}$ | $\sigma^2_s$ | $\sigma^2_{pe}$ | $\sigma^2_e$ | $\sigma^2_p$ | $h^2$ | $m^2$ | $s^2$ | $pe^2$ | $r_{am}$ | Log L |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| PY | Model I | 2.32 | - | - | - | 2.79 | 11.71 | 16.82 | 0.14 | - | - | 0.17 | - | -12672 |
| PY | Model II | 2.30 | - | - | - | 2.81 | 11.71 | 16.82 | 0.14 | - |  | 0.17 | - | -12670.2 |
| PY | Model III | 2.32 | 0.10 | - | - | 2.79 | 11.71 | 16.82 | 0.14 | 0.01 | - | 0.17 | - | -12672 |
| PY | Model IV | 2.31 | 0.10 | - | - | 2.81 | 11.71 | 16.83 | 0.14 | 0.01 | - | 0.17 | - | -12670.2 |
| PY | Model V | 3.91 | 0.21 | -0.90 | - | 2.09 | 11.70 | 17.01 | 0.23 | 0.01 | - | 0.12 | -1.00 | -12668.8 |
| PY | Model VI | 3.87 | 0.21 | -0.89 |  | 2.12 | 11.70 | 17.01 | 0.23 | 0.01 |  | 0.13 | -1.00 | -12666.9 |
| PY | Model VII | 1.87 | 0.10 | - | 0.32 | 2.99 | 11.71 | 16.89 | 0.11 | 0.01 | 0.02 | 0.18 | - | -12670.5 |
| PY | Model VIII | 1.85 | 0.10 | - | 0.31 | 0.18 | 11.71 | 16.89 | 0.11 | 0.01 | 0.02 | 0.21 | - | -12668.7 |

sire component of variance remaining low at 0.0001 to 0.012 and consistently at 0.019, respectively, indicating minor but measurable contributions of these effects.

These results highlighted the complex interplay of genetic and environmental factors influencing TLMY, 305-DMY, LL, and PY in crossbred cattle, underscoring the importance of model selection in accurately estimating genetic parameters for these economically significant traits. The genetic parameters for various models applied to production traits are presented in Fig. 4.5.

### 4.8.1 *Effect of inbreeding on production traits across different models*

The effect of inbreeding on various models and traits of crossbred cattle shows notable differences when Fx is included or excluded. For TLMY, models including Fx (models I, III, V, VII) generally showed lower additive genetic variance compared to their counterparts without Fx (Models II, IV, VI, VIII), indicating that inbreeding reduced the genetic variance. Similar trends are observed in the 305-DMY, LL, and PY traits, where models accounting for Fx slightly reduced genetic variances and resulted in lower heritability estimates. Overall, inbreeding effects led to a reduction in genetic variance and heritability estimates across all production traits, highlighting the detrimental impact of inbreeding on the genetic potential and performance of crossbred cattle. Specifically, the percentage decrease in additive genetic variance due to inbreeding in different animal models ranged from 1.32% to 5.25% for TLMY, 0.23% to 0.61% for 305-DMY, 0.29% to 1.30% for LL, and 0.69% to 1.19%.

## 4.9 Covariance components and genetic parameters for reproduction traits

Table 4.9 provides a comprehensive summary of estimates for (co)variance components, genetic parameters, and likelihood values for AFC, ICL, DD, GP, and SP across eight distinct statistical models. The results highlighted the varying contributions of genetic and environmental factors to these traits, as well as the effectiveness of each model in capturing their underlying genetic architecture.

For AFC, $\sigma^2$ ranged from 38.035 in model VIII to 49.465 in model V, reflecting varying levels of genetic contribution to trait variability. The $\sigma^{2\sim}$ was present in models II, III,

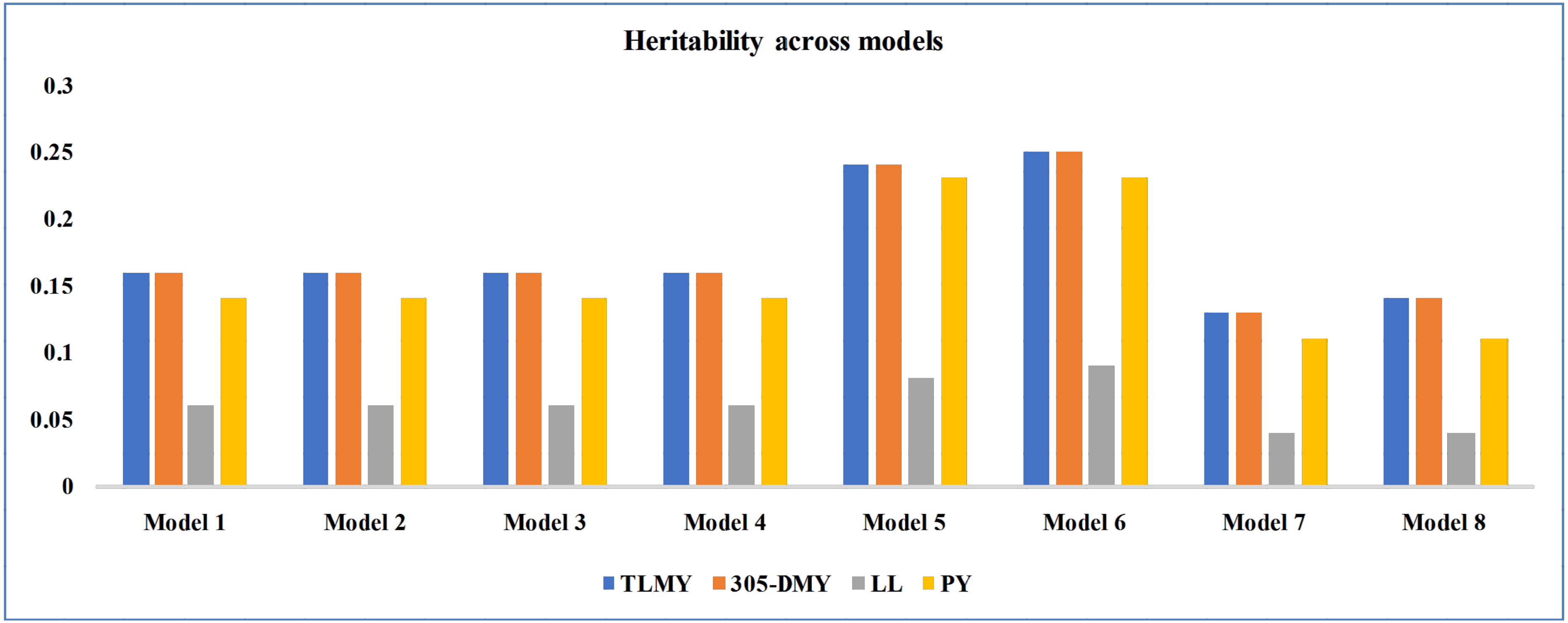


Fig. 4.5: Heritability estimates for production traits across different models

IV V, VI, VII, and VIII, with values ranging from 0.0001 to 5.262, underscoring the role of maternal effects. Notably, models V and VI exhibited negative $\sigma^{2}$ ˜, indicating antagonistic interactions. The $\sigma^{2}$› was significant in models VII and VIII, with values of 7.091 and 7.014, respectively, accompanied by minimal additive genetic and direct heritability estimates. The $\sigma^{2}$‘ remained stable, while $\sigma^{2}$š ranged from 170.446 to 170.785. Heritability estimates varied from 0.223 in model VIII to 0.289 in model V, demonstrating the models' capacity to capture genetic contributions effectively.

For ICL, the inclusion of ó$^{2}$˜ alongside additive $\sigma^{2}$ in models III and IV significantly affected the magnitude of additive effects and heritability compared to the simpler animal models (Models I and II). The highest log-likelihood values and increased heritability estimates were observed in models V and VI, which incorporated $\sigma^{2}$ ˜ . The $\sigma^{2}$› in models VII and VIII was modest (0.270 and 0.259, respectively), while $\sigma^{2}$‘ and $\sigma^{2}$š remained consistent. Heritability ranged from 0.023 in model VIII to 0.001 in model V, highlighting the models' ability to represent genetic influences on ICL effectively.

For DD, $\sigma^{2}$ varied significantly across models, ranging from 1.064 in model VII to 4.548 in model II, with the latter indicating a stronger genetic influence. The $\sigma^{2}$˜, present in models III through VIII, ranged from 4.553 to 5.126. models V and VI showed negative ó$^{2}$˜ values. The $\sigma^{2}$š and $\sigma^{2}$‘ were stable across models, and $\sigma^{2}$š remained consistent. Heritability estimates were low, reflecting limited genetic contributions to DD.

For GP, the inclusion of $\sigma^{2}$˜ alongside $\sigma^{2}$ in Models III and IV resulted in significant changes in additive genetic effects and heritability compared to simpler models. Incorporating $\sigma^{2}$˜ in models V and VI showed the highest log-likelihood values and increased heritability estimates. The $\sigma^{2}$› in models VII and VIII was notable, with values of 0.476 and 0.491, respectively, while additive genetic and direct heritability estimates were minimal in models with sire effects. Residual and phenotypic variances remained consistent, with heritability ranging from 0.027 in model VIII to 0.001 in model V.

For SP, model VI provided the best fit, with a log-likelihood of -17,436.43, indicating its superior ability to capture the genetic variability. Heritability estimates across all models

Table 4.9: (Co)Variance structure and genetic parameters for different models applied on reproduction traits in Vrindavani while employing various animal models with inclusion or exclusion of different fixed and random effects

| Traits | MODEL | $\sigma^2_a$ | $\sigma^2_m$ | $\sigma^2_{am}$ | $\sigma^2_s$ | $\sigma^2_{pe}$ | $\sigma^2_e$ | $\sigma^2_p$ | $h^2$ | $m^2$ | $s^2$ | $pe^2$ | $r_{am}$ | Log L |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **AFC** | Model I | 41.96 | - | - | - | - | 124.99 | 167.03 | 0.25 | - | - | - | - | -8821.87 |
| **AFC** | Model II | 42.88 | - | - | - | - | 124.98 | 167.85 | 0.26 | - | - | - | - | -8820.67 |
| **AFC** | Model III | 44.34 | 0.10 | - | - | - | 125.03 | 168.94 | 0.26 | 0.00 | - | - | - | -8630.67 |
| **AFC** | Model IV | 44.49 | 0.10 | - | - | - | 125.57 | 169.35 | 0.26 | 0.00 | - | - | - | -8630.96 |
| **AFC** | Model V | 49.03 | 5.25 | -4.16 | - | - | 126.25 | 170.62 | 0.29 | 0.03 | - | - | -0.86 | -8624.95 |
| **AFC** | Model VI | 49.47 | 5.26 | -4.01 | - | - | 126.56 | 170.79 | 0.29 | 0.03 | - | - | -0.87 | -8624.45 |
| **AFC** | Model VII | 38.04 | 0.10 | - | 7.01 | - | 123.65 | 170.45 | 0.22 | 0.00 | 0.04 | - | - | -8639.67 |
| **AFC** | Model VIII | 38.47 | 0.10 | - | 7.09 | - | 123.45 | 170.65 | 0.23 | 0.00 | 0.04 | - | - | -8637.31 |
| **ICL** | Model I | 0.15 | - | - | - | 2.86 | 27.75 | 30.76 | 0.01 | - | - | 0.09 | - | -13839.03 |
| **ICL** | Model II | 0.15 | - | - | - | 2.84 | 27.76 | 30.85 | 0.01 | - | - | 0.09 | - | -13836.92 |
| **ICL** | Model III | 0.05 | 0.28 | - | - | 2.68 | 27.75 | 30.76 | 0.01 | 0.01 | - | 0.09 | - | -13838.10 |
| **ICL** | Model IV | 0.06 | 0.27 | - | - | 2.66 | 27.76 | 30.75 | 0.01 | 0.01 | - | 0.09 | - | -13836.03 |
| **ICL** | Model V | 0.72 | 1.15 | -0.91 | - | 2.04 | 27.77 | 30.77 | 0.02 | 0.04 | - | 0.07 | -0.93 | -13833.92 |
| **ICL** | Model VI | 0.69 | 1.11 | -0.87 | - | 2.06 | 27.77 | 30.76 | 0.02 | 0.04 | - | 0.07 | -0.94 | -13832.00 |
| **ICL** | Model VII | 0.01 | 0.27 | - | 0.17 | 2.62 | 27.74 | 30.80 | 0.01 | 0.01 | 0.01 | 0.09 | - | -13836.81 |
| **ICL** | Model VIII | 0.01 | 0.26 | - | 0.14 | 2.61 | 27.75 | 30.76 | 0.01 | 0.01 | 0.01 | 0.09 | - | -13834.71 |
| **DD** | Model I | 0.15 | - | - | - | 2.86 | 27.75 | 30.76 | 0.01 | - | - | 0.09 | - | -13839.03 |
| **DD** | Model II | 0.15 | - | - | - | 2.84 | 27.76 | 30.85 | 0.01 | - | - | 0.09 | - | -13836.92 |
| **DD** | Model III | 0.05 | 0.28 | - | - | 2.68 | 27.75 | 30.76 | 0.01 | 0.01 | - | 0.09 | - | -13838.10 |
| **DD** | Model IV | 0.06 | 0.27 | - | - | 2.66 | 27.76 | 30.75 | 0.01 | 0.01 | - | 0.09 | - | -13836.03 |
| **DD** | Model V | 0.72 | 1.15 | -0.91 | - | 2.04 | 27.77 | 30.77 | 0.02 | 0.04 | - | 0.07 | -0.97 | -13833.92 |
| **DD** | Model VI | 0.69 | 1.11 | -0.87 | - | 2.06 | 27.77 | 30.76 | 0.02 | 0.04 | - | 0.07 | -0.95 | -13832.00 |
| **DD** | Model VII | 0.01 | 0.27 | - | 0.17 | 2.62 | 27.74 | 30.80 | 0.01 | 0.01 | 0.01 | 0.09 | - | -13836.81 |
| **DD** | Model VIII | 0.01 | 0.26 | - | 0.14 | 2.61 | 27.75 | 30.76 | 0.01 | 0.01 | 0.01 | 0.09 | - | -13834.71 |
| **GP** | Model I | 4.46 | - | - | - | 6.89 | 134.01 | 145.36 | 0.03 | - | - | 0.05 | - | -16100.78 |

**Table 4.9: Contd...**

| Traits | MODEL | $\sigma^2_a$ | $\sigma^2_m$ | $\sigma^2_{am}$ | $\sigma^2_s$ | $\sigma^2_{pe}$ | $\sigma^2_e$ | $\sigma^2_p$ | $h^2$ | $m^2$ | $s^2$ | $pe^2$ | $r_{am}$ | Log L |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **GP** | Model II | 4.55 | - | - | - | 6.64 | 134.11 | 145.30 | 0.03 | - | - | 0.05 | - | -16100.58 |
| **GP** | Model III | 1.58 | 4.75 | - | - | 4.33 | 134.52 | 145.18 | 0.01 | 0.03 | - | 0.03 | - | -16096.98 |
| **GP** | Model IV | 1.78 | 4.55 | - | - | 4.15 | 134.64 | 145.12 | 0.01 | 0.03 | - | 0.03 | - | -16097.01 |
| **GP** | Model V | 1.71 | 5.13 | -0.39 | - | 4.22 | 134.52 | 145.18 | 0.01 | 0.04 | - | 0.03 | -0.13 | -16096.96 |
| **GP** | Model VI | 1.91 | 4.90 | -0.36 | - | 4.04 | 134.63 | 145.12 | 0.01 | 0.03 | - | 0.03 | -0.12 | -16096.99 |
| **GP** | Model VII | 1.06 | 4.89 | - | 0.16 | 4.53 | 134.52 | 145.16 | 0.01 | 0.03 | 0.01 | 0.03 | - | -16096.97 |
| **GP** | Model VIII | 1.32 | 4.68 | - | 0.14 | 4.33 | 134.63 | 145.11 | 0.01 | 0.03 | 0.01 | 0.03 | - | -16096.99 |
| **SP** | Model I | 0.04 | - | - | - | 0.11 | 1.49 | 1.64 | 0.03 | - | - | 0.06 | - | -4402.32 |
| **SP** | Model II | 0.05 | - | - | - | 0.10 | 1.49 | 1.64 | 0.03 | - | - | 0.06 | - | -4397.55 |
| **SP** | Model III | 0.03 | 0.02 | - | - | 0.10 | 1.49 | 1.64 | 0.02 | 0.01 | - | 0.06 | - | -4401.41 |
| **SP** | Model IV | 0.03 | 0.02 | - | - | 0.10 | 1.49 | 1.64 | 0.02 | 0.01 | - | 0.06 | - | -4396.69 |
| **SP** | Model V | 0.02 | 0.01 | 0.01 | - | 0.10 | 1.49 | 1.64 | 0.01 | 0.01 | - | 0.06 | 0.93 | -4396.15 |
| **SP** | Model VI | 0.02 | 0.01 | 0.01 | - | 0.10 | 1.49 | 1.64 | 0.01 | 0.01 | - | 0.06 | 0.78 | -4401.01 |
| **SP** | Model VII | 0.04 | 0.01 | - | 0.01 | 0.10 | 1.48 | 1.64 | 0.02 | 0.01 | 0.01 | 0.03 | - | -4525.74 |
| **SP** | Model VIII | 0.35 | 0.01 | - | 0.12 | 0.10 | 1.48 | 1.63 | 0.02 | 0.01 | 0.01 | 0.03 | - | -4527.54 |

were very low, suggesting minimal contributions of additive genetic effects to SP. The $\sigma^{2\sim}$ was included in models III through VIII, while $\sigma^{2}$š ‘ remained consistent across models, highlighting stable environmental influences.

Overall, the results from Table 4.9 demonstrated the complex relationship of genetic and environmental factors affecting AFC, ICL, DD, GP, and SP. The variation in (co)variance components and genetic parameters across models underscores the importance of model selection in accurately estimating genetic contributions to these traits.

### 4.9.1 *Effect of inbreeding on reproduction traits across different models*

The effect of inbreeding on various models and traits of crossbred cattle showed notable differences when Fx is included or excluded. For all reproductive traits, models including Fx (models I, III, V, VII) generally showed lower additive genetic variance compared to their counterparts without Fx (models II, IV, IV, VIII), indicating that inbreeding reduces genetic variance. Overall, inbreeding effects lead to a reduction in genetic variance and heritability estimates across all reproduction traits, highlighting the detrimental impact of inbreeding on the genetic potential and performance of crossbred cattle. Specifically, the percentage decrease in additive genetic variance due to inbreeding ranged from 1.32 to 3.25% for AFC, 2.12 to 5.61% for ICP, 2.29 to 9.30% for DD, 1.34 to 3.54% for GP, and 5.46 to 28% for SP.

## 4.10 Comparison of Models for Growth, Production, and Reproduction Traits

The comparison of models based on log-likelihood values across various growth, production, and reproduction traits in Vrindavani cattle revealed significant differences. For growth traits like BW0, and BW6, model VII (include sire and dam effect as random effect with inbreeding as fixed covariable) consistently showed the highest log-likelihood values, indicating a better fit compared to the other models. Similarly, for BW3, model VI (with individual and dam additive effect as correlated without inbreeding as fixed covariable) exhibited higher log-likelihood values. For BW12, BW18, and BW24, model V (with individual and dam additive effect as correlated with inbreeding as fixed covariable) exhibited higher log-likelihood values. In production traits such as TLMY, and PY, model VI consistently proved

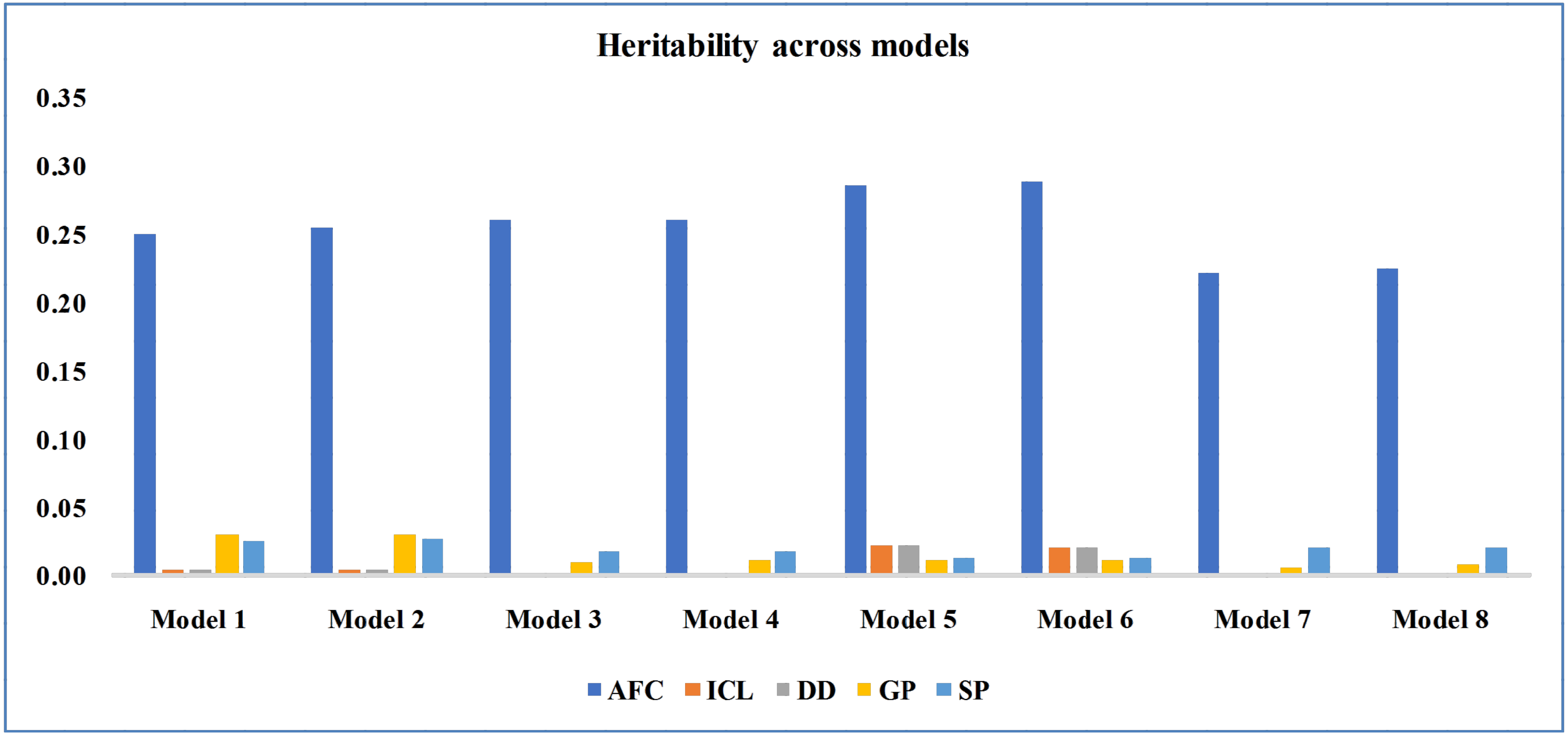


Fig. 4.6: Heritability estimates for reproduction traits across different models

## Table 4. 10: Comparison of models for growth, production, and reproduction traits based on log likelihood values

| Traits | Log Likelihood value | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Model I | Model II | Model III | Model IV | Model V | Model VI | Model VII | Model VIII |
| **BW0** | -16388.08 | -16383.76 | -16385.36 | -16381.06 | -16347.72 | -16350.07 | **-16346.86** | -16350.86 |
| **BW3** | -17663.19 | -17663.09 | -17659.49 | -17659.47 | -17578.43 | **-17560.08** | -17560.78 | -17561.25 |
| **BW6** | -16496.54 | -16496.71 | -16495.93 | -16496.08 | -16452.72 | -16452.71 | **-16448.17** | -16448.19 |
| **BW12** | -12234.25 | -12235.36 | -12231.97 | -12233.06 | **-12190.32** | -12191.36 | -12193.97 | -12195.01 |
| **BW18** | -9924.15 | -9928.44 | -9918.98 | -9922.84 | **-9872.81** | -9875.01 | -9883.08 | -9885.43 |
| **BW24** | -8782.17 | -8788.31 | -8776.40 | -8781.94 | **-8733.37** | -8736.90 | -8739.27 | -8742.87 |
| **TLMY** | -20327.37 | -20326.36 | -20327.37 | -20326.36 | -20323.87 | **-20322.61** | -20324.44 | -20323.45 |
| **305-DMY** | -8250.57 | -8248.54 | -8250.57 | -8248.55 | -8251.33 | -8249.33 | -8250.55 | **-8248.53** |
| **LL** | -17171.77 | -17168.91 | -17171.76 | -17168.90 | -17170.31 | -17167.39 | -17170.28 | **-17167.36** |
| **PY** | -12672.04 | -12670.18 | -12672.04 | -12670.19 | -12668.75 | **-12666.88** | -12670.52 | -12668.68 |
| **AFC** | -8821.87 | -8820.67 | -8630.67 | -8630.96 | -8624.95 | **-8624.45** | -8639.67 | -8637.31 |
| **ICP** | -13839.03 | -13836.92 | -13838.10 | -13836.03 | -13833.92 | **-13832.00** | -13836.81 | -13834.71 |
| **DD** | -16100.78 | -16100.58 | -16096.98 | -16097.01 | **-16096.96** | -16096.99 | -16097.96 | -16096.99 |
| **GP** | -4402.32 | -4397.55 | -4401.41 | -4396.69 | **-4396.15** | -4401.01 | -4525.74 | -4527.54 |
| **SP** | -17440.55 | -17441.16 | -17439.89 | -17440.50 | **-17435.87** | -17436.43 | -17439.06 | -17439.62 |

*Models with bold log-likelihood estimates were taken as best fit for the respective traits as the best fit, as evident by its higher log-likelihood values. Notably, for LL, and 305-DMY, model VIII (with sire and dam effect as random effect without inbreeding as fixed covariable) emerges as the best-fit model. Among reproduction traits, model VI exhibits the highest log-likelihood value for AFC, and ICP, while model V best fit for GP, DD and SP. Overall, the comparison underscores the importance of selecting appropriate models for specific traits. Models V, VI, VII, and VIII performed better across traits, reflecting their ability to effectively capture the genetic and environmental variability underlying the growth, production, and reproduction traits in Vrindavani cattle.

## 4.11 Artificial neural network for performance traits in Vrindavani cattle

Artificial neural network (ANN) architectures were employed to predict three performance traits: TLMY, 305-DMY, and PY. The analyses were conducted using the neural network toolbox in MATLAB (version R2024b). For TLMY, the model was developed using a dataset comprising 2,975 first-lactation records with eight input variables (SOB, POB, BW0, BW3, BW6, M1, M2, and M3). Similarly, the models for 305-DMY and LL utilized datasets containing 929 and 2,401 first-lactation records, respectively, with eight input variables (SOB, POB, BW0, BW3, BW6, M1, M2, and M3) for 305-DMY and six (SOB, POB, BW0, BW3, BW6, and M1) for LL. In all cases, the target output was the respective performance trait. The networks were trained using three algorithms: Levenberg-Marquardt (trainlm), scaled conjugate gradient (trainscg), Bayesian regularization (trainbr). The architectures were evaluated with 1 to 30 neurons in the hidden layer over 1000 epochs, and the results for training, validation, and testing datasets have been presented in Table 4.11a, b, and c and result for various ANN algorithm are described in fig 4.11a–4.11i and fig 4.11.1a – 4.11.1c.

The best performance was for TLMY achieved using 15 neurons in the hidden layer. Among the algorithms, the Levenberg-Marquardt method attained an R value of 0.8426 and a MSE of 2.89×10u in just 12 epochs. The scaled conjugate gradient algorithm performed slightly lower, achieving an R value of 0.8212 with an MSE of 3.22×10u after 64 epochs. However, Bayesian regularization method demonstrated the best performance, achieving an R value of 0.8683 and the lowest MSE of 2.48×10u after 505 epochs, reflecting superior predictive accuracy despite requiring significantly more computational iterations.

The best performance for 305-DMY was observed with 20 neurons in the hidden layer. The Levenberg-Marquardt algorithm achieved an R value of 0.9206 and an MSE of $8.94\times10^4$ in just 8 epochs, showcasing its computational efficiency. The scaled conjugate gradient algorithm attained an R value of 0.9164 with an MSE of $9.74\times10^4$ after 88 epochs, indicating moderate predictive performance. Bayesian Regularization method outperformed both, achieving the lowest MSE of $1.02\times10^5$ and an R value of 0.906 after 258 epochs, demonstrating improved predictive accuracy despite a slightly lower correlation coefficient compared to the Levenberg-Marquardt algorithm.

The best performance for PY was obtained using 12 neurons in the hidden layer. The Levenberg-Marquardt method achieved an R value of 0.870 and an MSE of 2.04 in only 11 epochs, highlighting its computational efficiency. The scaled conjugate gradient algorithm yielded an R value of 0.85 with an MSE of 2.27 after 95 epochs, reflecting moderate performance. Bayesian Regularization method emerged as the best-performing algorithm, achieving the lowest MSE of 1.94 and the highest R value of 0.876 after 247 epochs, indicating superior predictive accuracy despite requiring significantly more iterations.

Across all traits, Bayesian regularization consistently provided the best predictive accuracy, albeit at the cost of higher computational demands, while the Levenberg-Marquardt algorithm offered high efficiency with slightly lower predictive performance. The scaled conjugate gradient algorithm demonstrated balanced performance, with estimates between the other two methods in terms of accuracy and computational efficiency.

Table 4.11a: The Pearson's correlation coefficient (R), coefficient of determination ($R^2$) means absolute error (MAE), and root mean absolute error (RMAE), for training, validation and training dataset for TLMY

| Architecture | Epochs | Algorithm | Split Ratio | R | $R^2$ | MAE | RMAE |
|---|---|---|---|---|---|---|---|
| Single hidden layer with 15 neurons | 12 | Levenberg-Marquardt | Train (70%) | 0.84 | 0.71 | 2.89×10u | 537.59 |
| | | | Validation (15%) | 0.78 | 0.61 | 3.92×10u | 626.01 |
| | | | Test (15%) | 0.77 | 0.59 | 4.11×10u | 641.09 |
| | | | Overall | 0.82 | 0.67 | 3.92×10u | 626.01 |
| | 64 | Scaled conjugate gradient | Train (70%) | 0.82 | 0.67 | 3.22×10u | 567.45 |
| | | | Validation (15%) | 0.80 | 0.64 | 3.75×10u | 612.37 |
| | | | Test (15%) | 0.77 | 0.59 | 4.43×10u | 665.58 |
| | | | Overall | 0.81 | 0.66 | 3.75×10u | 612.37 |
| | 505 | Bayesian regularization | Train (70%) | 0.87 | 0.76 | 2.40×10u | 489.9 |
| | | | Validation | NA | NA | NA | NA |
| | | | Test (30%) | 0.77 | 0.59 | 4.06×10u | 637.18 |
| | | | Overall | 0.85 | 0.72 | 2.48×10u | 497.00 |

**Table 4.11b: The Pearson's correlation coefficient (R), coefficient of determination ($R^2$) means absolute error (MAE), and root mean absolute error (RMAE), for training, validation and training dataset for 305-DMY**

| Architecture | Epochs | Algorithm | Split Ratio | R | $R^2$ | MAE | RMAE |
|---|---|---|---|---|---|---|---|
| Single hidden layer with 20 neurons | 8 | Levenberg-Marquardt | Train (70%) | 0.92 | 0.85 | $8.94\times10^{4}$ | 298.00 |
| | | | Validation (15%) | 0.90 | 0.81 | 1.07×10u | 327.11 |
| | | | Test (15%) | 0.87 | 0.76 | 1.31×10u | 361.94 |
| | | | Overall | 0.91 | 0.83 | 1.07×10u | 327.10 |
| | 88 | Scaled conjugate gradient | Train (70%) | 0.92 | 0.85 | $9.74\times10^{4}$ | 312.09 |
| | | | Validation (15%) | 0.87 | 0.76 | 1.20×10u | 397.49 |
| | | | Test (15%) | 0.84 | 0.71 | 1.49×10u | 386.00 |
| | | | Overall | 0.91 | 0.83 | 1.20×10u | 346.41 |
| | 258 | Bayesian regularization | Train (70%) | 0.91 | 0.83 | 1.02×10u | 319.37 |
| | | | Validation | NA | NA | NA | NA |
| | | | Test (30%) | 0.90 | 0.81 | 1.07×10u | 327.10 |
| | | | Overall | 0.91 | 0.83 | $1.03\times10^{5}$ | 320.93 |

**Table 4.11c: The Pearson's correlation coefficient (R), coefficient of determination ($R^2$) means absolute error (MAE), and root mean absolute error (RMAE), for training, validation and training dataset for PY**

| Architecture | Epochs | Algorithm | Split Ratio | R | $R^2$ | MAE | RMAE |
|---|---|---|---|---|---|---|---|
| Single hidden layer with 12 neurons | 11 | Levenberg-Marquardt | Train (70%) | 0.87 | 0.74 | 2.04 | 1.43 |
| | | | Validation (15%) | 0.82 | 0.67 | 2.70 | 1.64 |
| | | | Test (15%) | 0.80 | 0.64 | 2.93 | 1.71 |
| | | | Overall | 0.85 | 0.72 | 2.70 | 1.64 |
| | 95 | Scaled conjugate gradient | Train (70%) | 0.85 | 0.72 | 2.41 | 1.55 |
| | | | Validation (15%) | 0.78 | 0.61 | 3.14 | 1.85 |
| | | | Test (15%) | 0.84 | 0.71 | 2.57 | 1.60 |
| | | | Overall | 0.84 | 0.71 | 3.14 | 1.85 |
| | 247 | Bayesian regularization | Train (70%) | 0.88 | 0.77 | 1.94 | 1.39 |
| | | | Validation | NA | NA | NA | NA |
| | | | Test (30%) | 0.77 | 0.59 | 3.27 | 1.81 |
| | | | Overall | 0.86 | 0.74 | 1.94 | 1.39 |

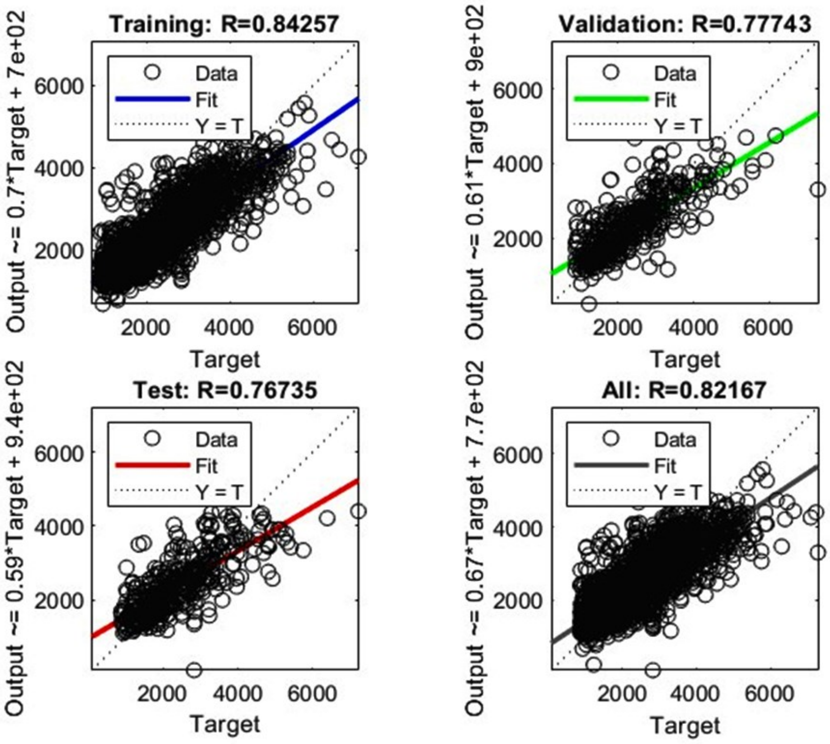


**Fig. 4.7a:** Correlation coefficients for training, validation, and testing datasets using the Levenberg-Marquardt algorithm for TLMY

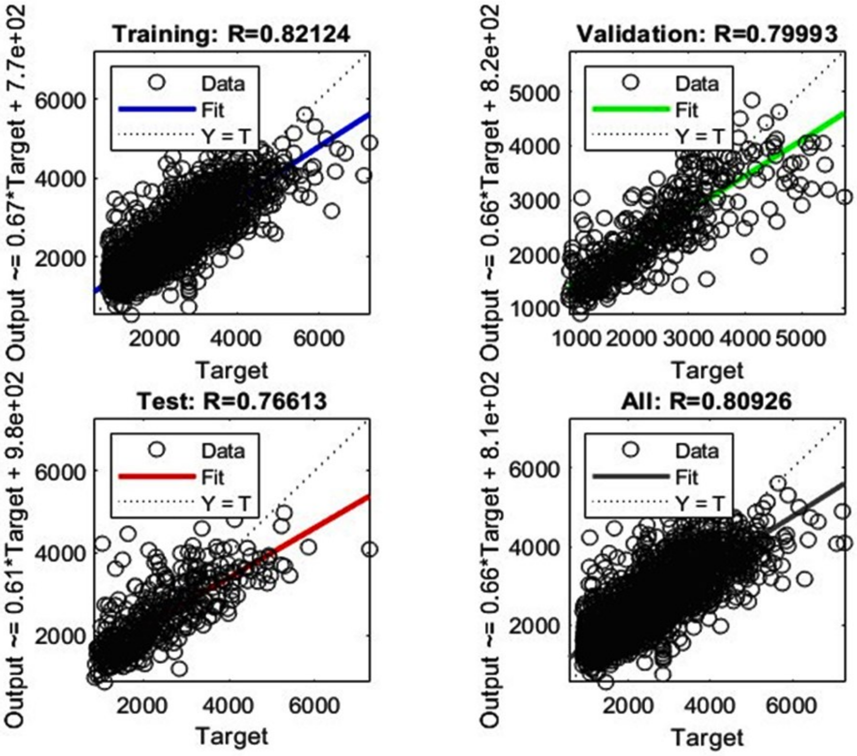


**Fig. 4.7b:** Correlation coefficients for training, validation, and testing datasets using the scaled conjugate gradient algorithm for TLMY

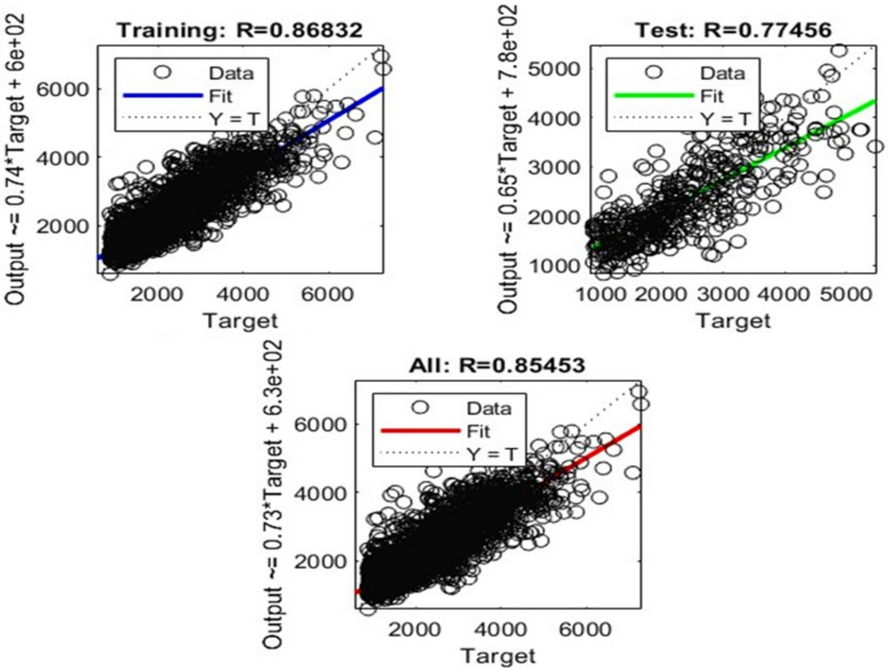


**Fig. 4.7c:** **Correlation coefficients for training, validation, and testing datasets using the Bayesian regularization algorithm for TLMY**

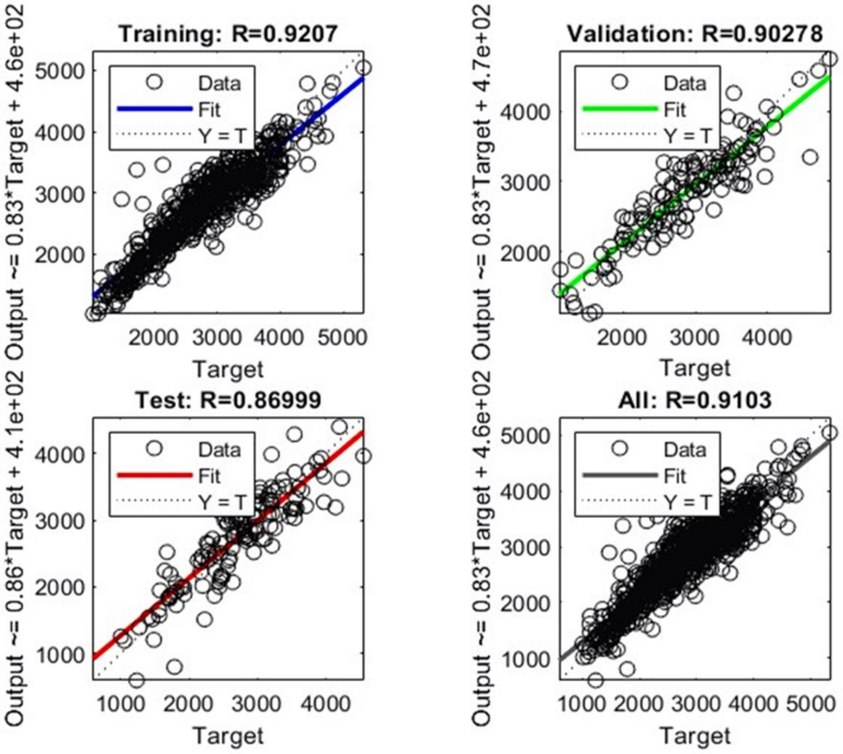


**Fig. 4.7d:** **Correlation coefficients for training, validation, and testing datasets using the Levenberg-Marquardt algorithm for 305-DMY**

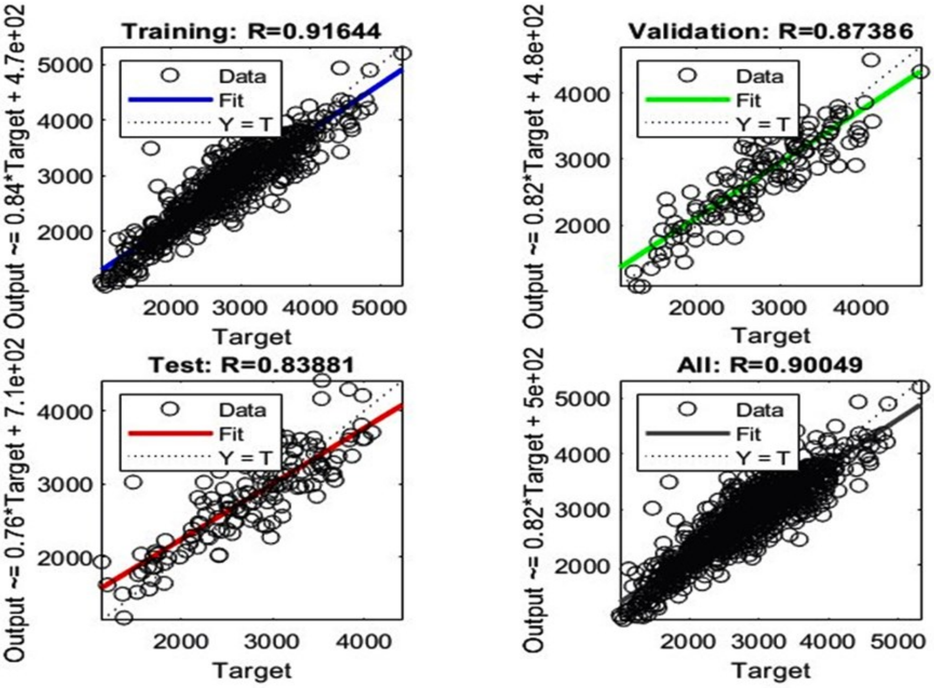


**Fig. 4.7e:** **Correlation coefficients for training, validation, and testing datasets using the scaled conjugate gradient algorithm for 305-DMY**

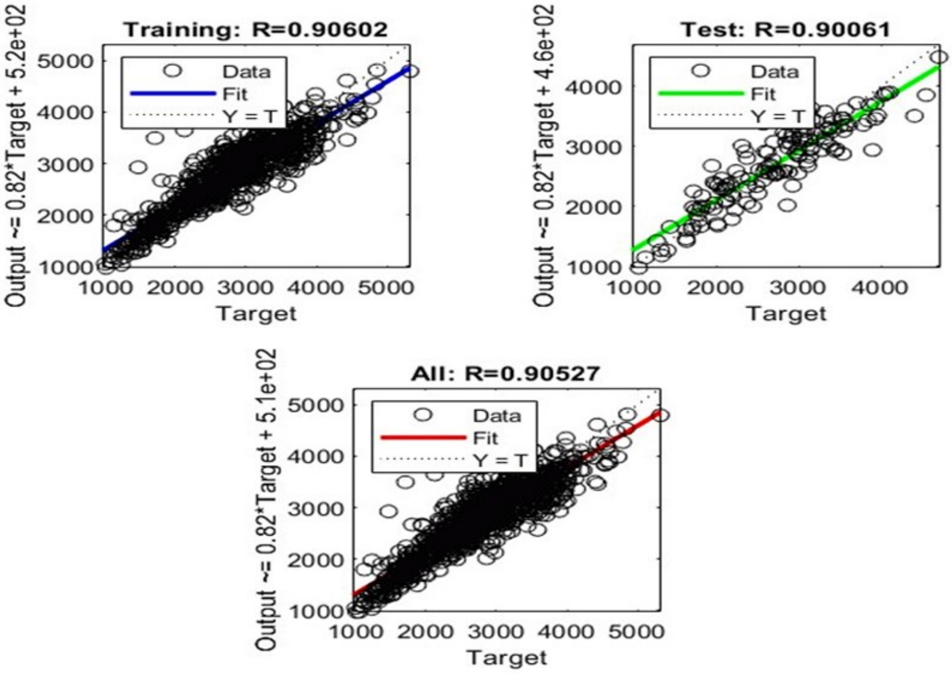


**Fig. 4.7f:** **Correlation coefficients for training, validation, and testing datasets using the Bayesian regularization algorithm for 305-DM**

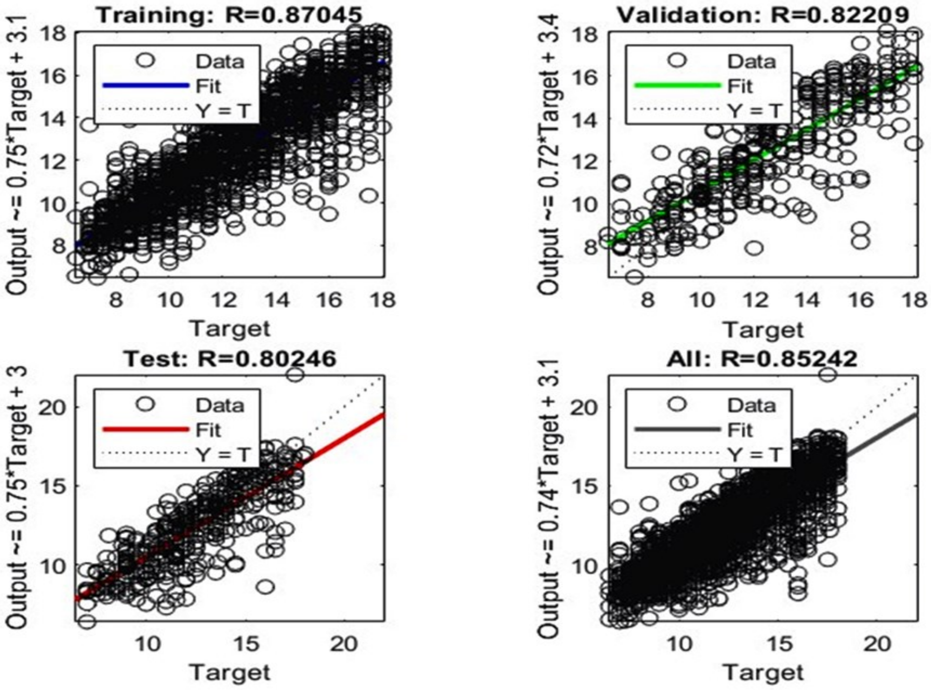


**Fig. 4.7g:** Correlation coefficients for training, validation, and testing datasets using the Levenberg-Marquardt algorithm for PY

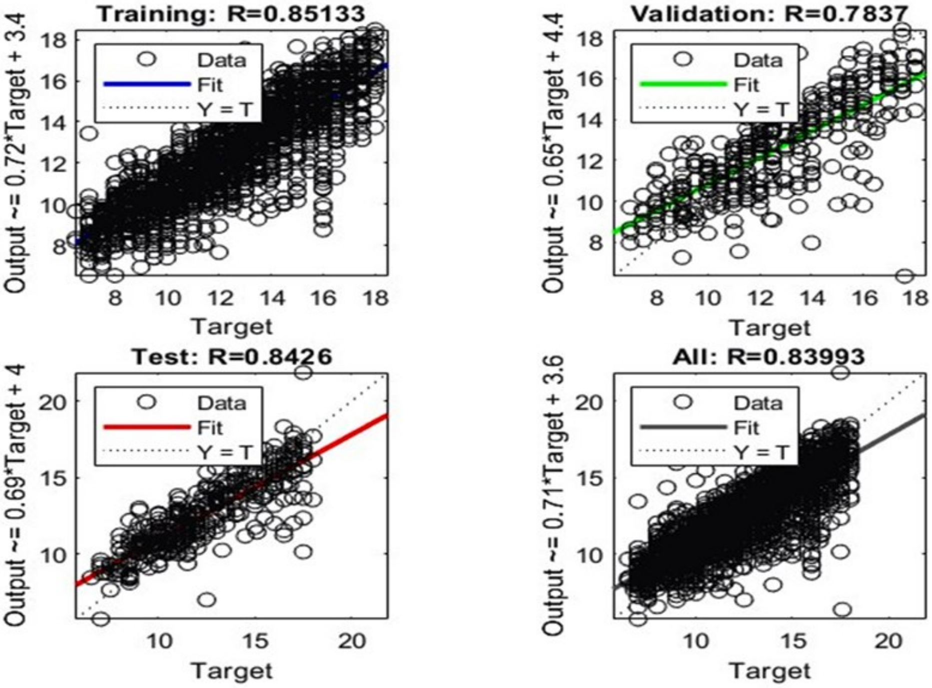


**Fig. 4.7h:** Correlation coefficients for training, validation, and testing datasets using the scaled conjugate gradient algorithm for PY

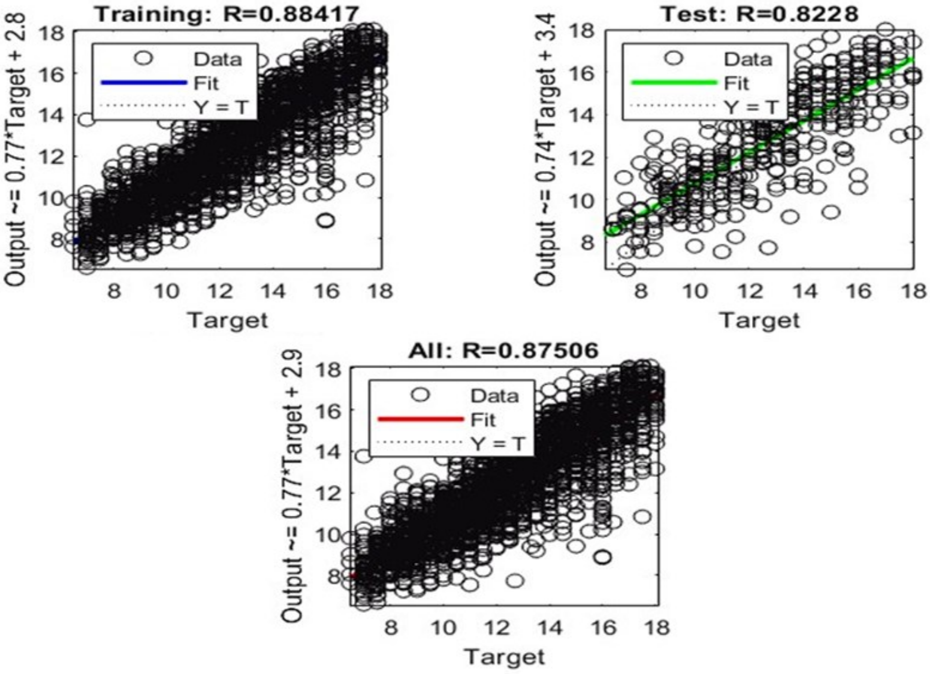


**Fig. 4.7i:** **Correlation coefficients for training, validation, and testing datasets using the Bayesian regularization algorithm for PY**

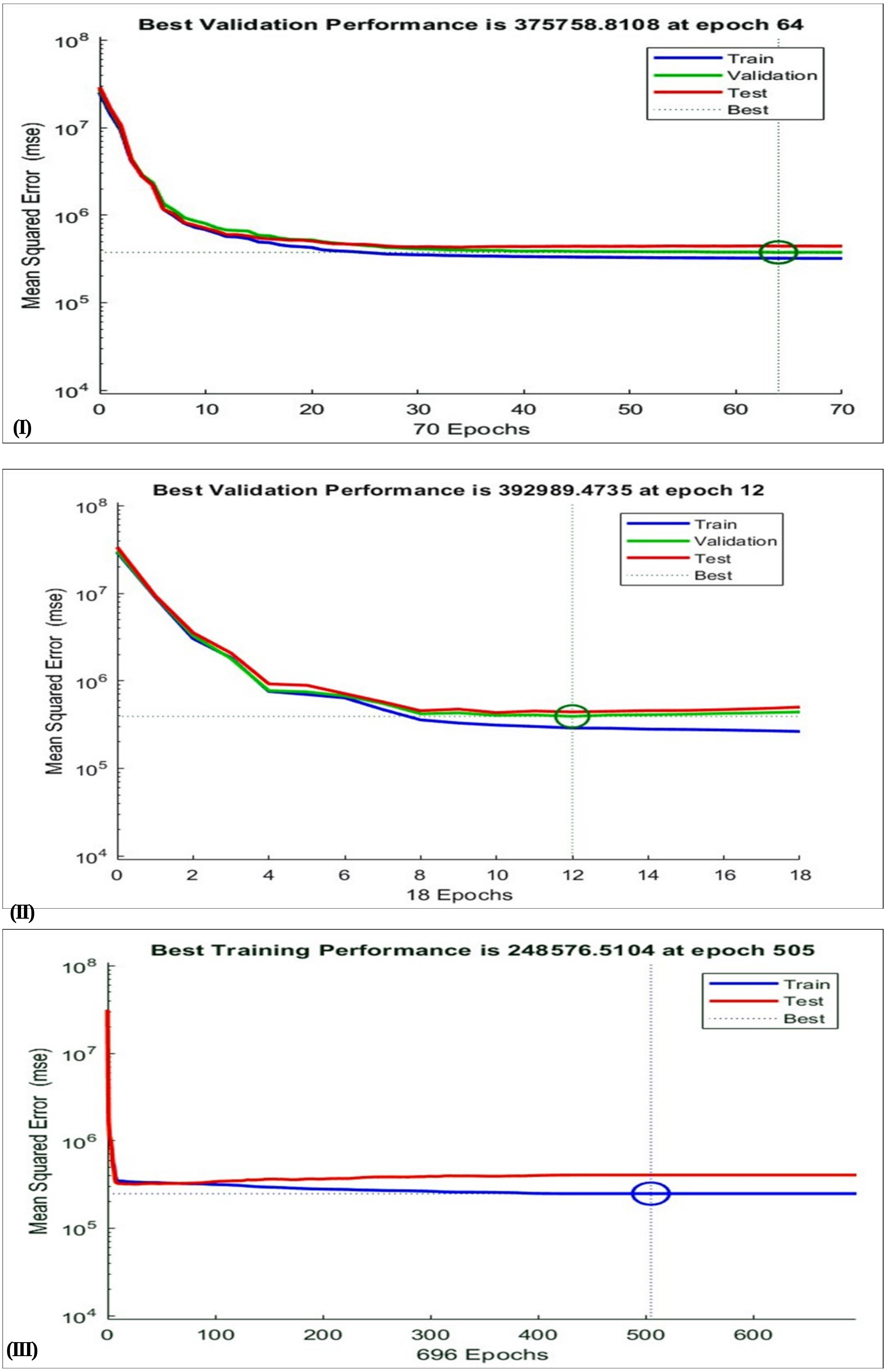


**Fig 4.8a:** (I, II, III): Validation performance curve for TLMY using Levenberg-Marquardt, scaled conjugate gradient, and Bayesian regularization algorithms in a neural network

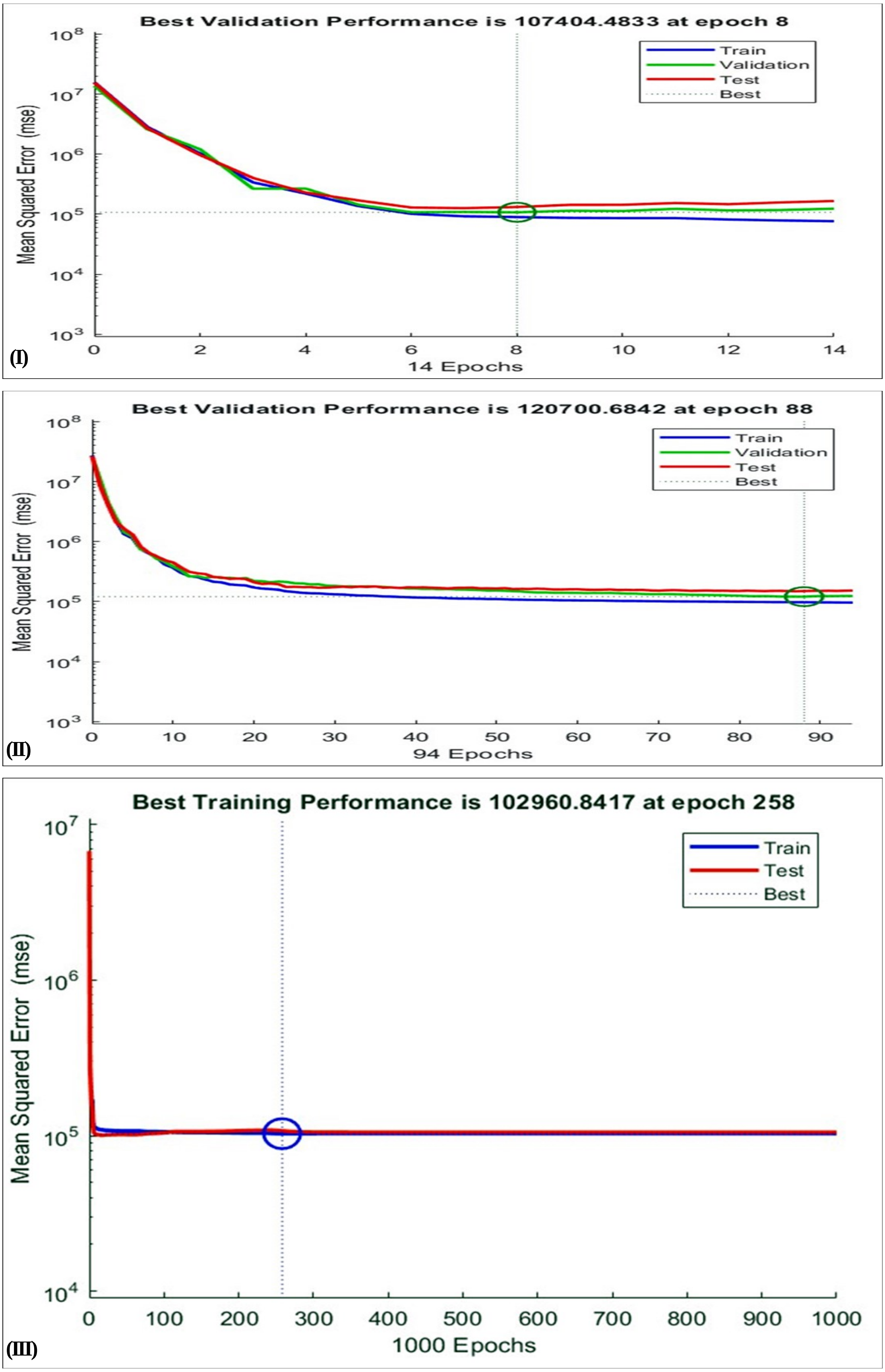


**Fig 4.8b:** (I, II, III): Validation performance curve for 305-DMY using Levenberg-Marquardt, scaled conjugate gradient, and Bayesian regularization algorithms in a neural network

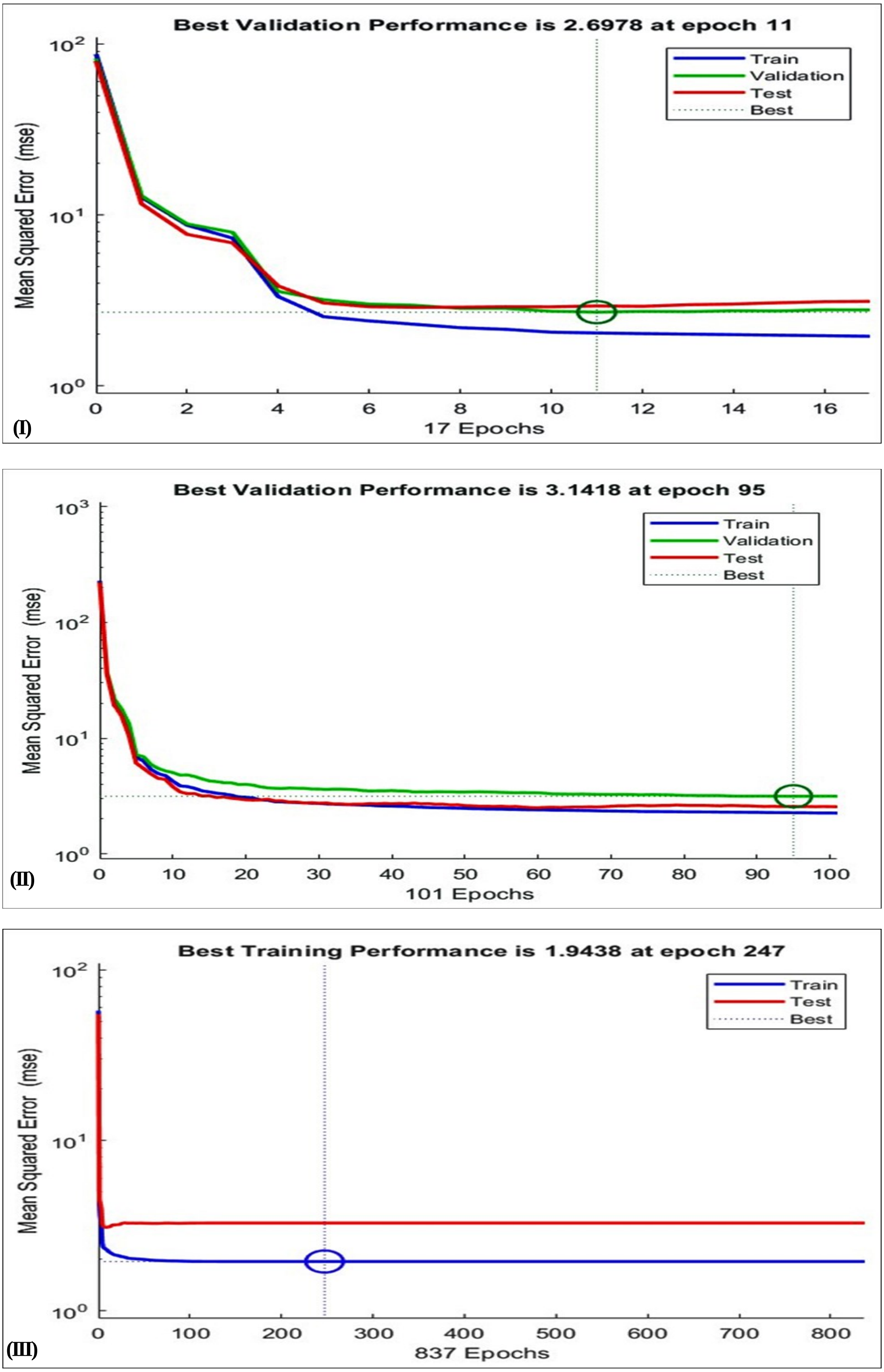


**Fig 4.8c:** **(I, II, III): Validation performance curve for PY using Levenberg-Marquardt, scaled conjugate gradient, and Bayesian regularization algorithms in a neural network**

## 4.12 Bayesian neural network for performance traits in Vrindavani cattle

Bayesian neural network **(**BNN) architectures were employed to predict three performance traits: TLMY, 305-DMY, and PY. The analyses were conducted using the Python programme (version 4.0). The model for TLMY was developed using a dataset comprising 2,975 first-lactation records with nine input variables. Similarly, the models for 305-DMY and PY utilized datasets containing 929 and 2,401 first-lactation records, respectively, with nine input variables for 305-DMY and seven for LL. In all cases, the target output was the respective performance trait. To better capture non-linear interactions among the features, polynomial features were generated up to the second degree, expanding the input space for the model.

The dataset was split into training and test sets, with 85% allocated for training and 15% for testing. During model training, the training data was further divided into training and validation subsets in an 85:15 ratio to monitor performance and mitigate overfitting. All features were standardized using a standard scaler to normalize the data and ensure consistent scaling, which is crucial for improving the model performance. The input layer processed 66 features, including the polynomial-expanded features. The first hidden layer consisted of 28 neurons with ReLU activation, followed by a dropout layer with a 30% dropout rate to prevent overfitting. The second hidden layer comprised 56 neurons with ReLU activation, also followed by a dropout layer with a 30% dropout rate. Finally, the output layer consisted of a single neuron for regression, enabling the model to predict performance traits as a continuous variable. The model was compiled using the Adam optimizer and a MSE loss function. Training was conducted with a batch size of 52 over a maximum of 1000 epochs, and early stopping was employed with a patience of 40 epochs to halt training if the validation loss ceased improving.

The model achieved an R-squared value of 0.67 for TLMY, explaining 67.21% of the variance in the target variable. The MAE was 420.68, reflecting the average deviation between predicted and actual values, while the MSE and RMSE were calculated as 3.3×10u and 577.4, respectively. Training for TLMY was completed at epoch 159, with early stopping ensuring optimal performance without overfitting.

Similarly, the BNN model for 305 DMY exhibited even stronger predictive capability, achieving an R-squared value of 0.81, indicating that 81.36% of the variance in 305-DMY was accounted for by the input features. The MAE was 253.71, with an MSE of 1.26×10u and an RMSE of 356.08. The training process stopped at epoch 272, demonstrating the model's ability to generalize well while avoiding overfitting.

The model for PY also performed exceptionally, achieving an R-squared value of 0.78 and explaining 77.82% of the variance in the target variable. The MAE for PY was 1.90, while the MSE and RMSE were 7.02 and 2.65, respectively. Training for PY was concluded at epoch 59 due to early stopping, optimizing the balance between model complexity and performance.

Overall, these results of BNN model point towards reliability and effectiveness in predicting multiple key performance traits with high accuracy, while leveraging early stopping to maintain a balance between underfitting and overfitting.

## 4.13 Comparison of ANN and BNN models for performance traits in Vrindavani cattle

Table 4.13 presents a comparative analysis of the performance metrics of ANN and BNN models in predicting three traits: TLMY, 305-DMY, and PY. The ANN models were trained using three optimization algorithms: trainlm, trainscg, and trainbr, while the BNN represents a separate model approach using common input parameters.

The ANN models for TLMY showed $R^2$ values of 0.71 (trainlm), 0.67 (trainscg), and 0.76 (trainbr), indicating that approximately 67–76% of the variance in TLMY was explained by the input features. The RMAE for these models ranged from 497.00 to 626.01, with trainbr performing the best in terms of RMAE (497.00). The BNN model had a slightly lower $R^2$ of 0.67, explaining 67% of the variance, but it showed an RMAE of 577.39, demonstrating competitive predictive performance.

The ANN models for 305-DMY achieved $R^2$ values of 0.85 (trainlm), 0.85 (trainscg), and 0.83 (trainbr), indicating that the ANN models explained a significant portion (83–85%)

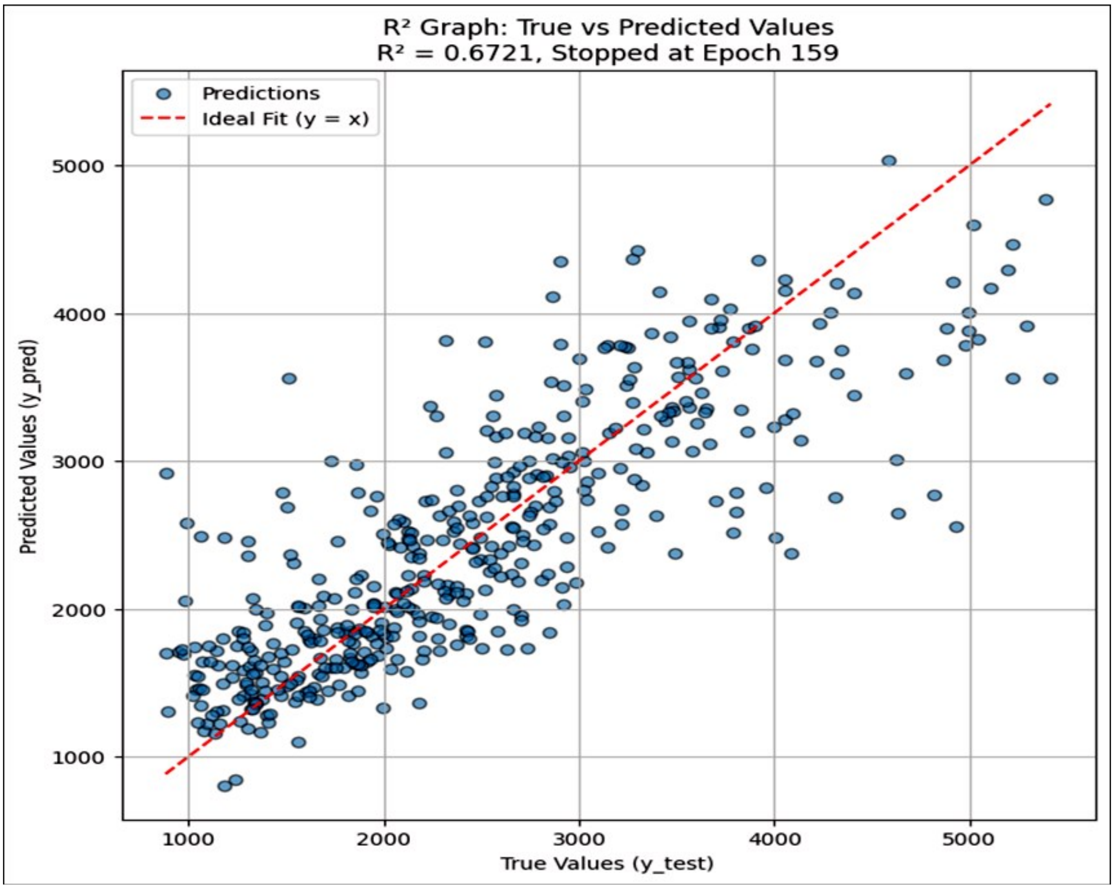


Fig. 4.9a: BNN performance: true vs predicted for TLMY

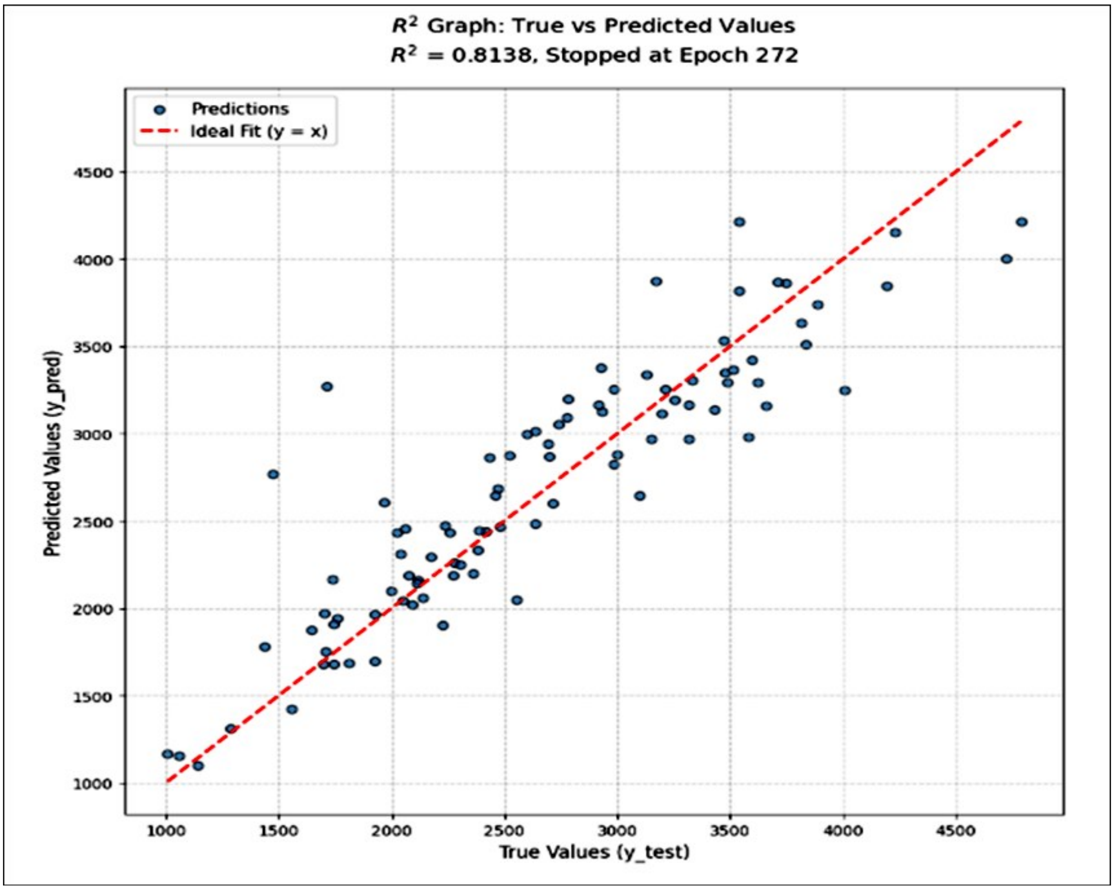


Fig. 4.9b: BNN performance: true vs predicted for 305-DMY

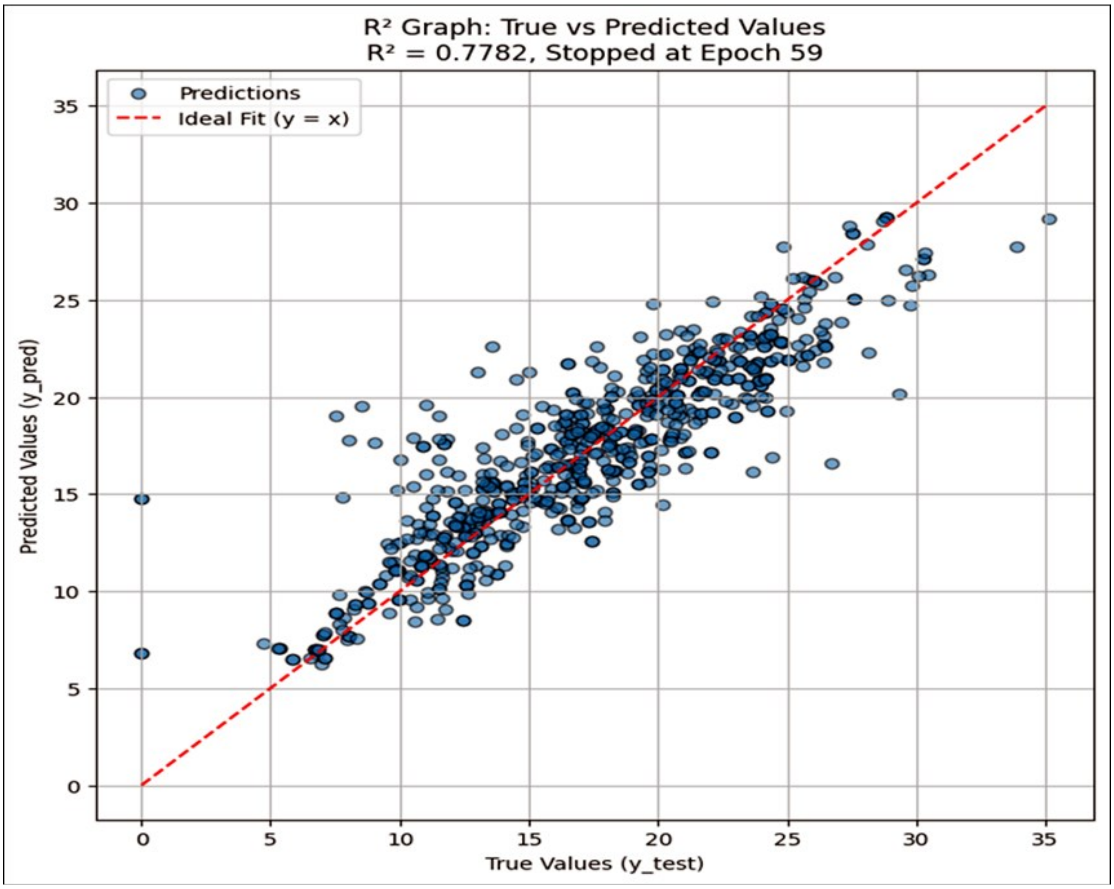


Fig. 4.9c: BNN performance: true vs predicted for PY

### Table 4.12: The Pearson's correlation coefficient (R), coefficient of determination ($R^2$) means absolute error (MAE), mean squared error (MSE), and root mean absolute error (RMAE), for performance traits in BNN model

| Traits | Architecture | Epochs | R | $R^2$ | MAE | MSE | RMAE |
|---|---|---|---|---|---|---|---|
| TL | Two hidden layers, with 28 neurons in the first and 56 in the second | 159 | 0.86 | 0.67 | 420.68 | 3.3×10u | 577.39 |
| MY | | 272 | 0.86 | 0.81 | 253.71 | 1.26×10u | 356.08 |
| PY | | 59 | 0.83 | 0.78 | 1.90 | 7.03 | 2.65 |

### Table 4.13. Comparison of ANN and BNN for performance traits on the basis of $R^2$ and RMAE

| Traits | ANN | | | | | | BNN | |
|---|---|---|---|---|---|---|---|---|
| | trainlm | | trainscg | | trainbr | | | |
| | $R^2$ | RMAE | $R^2$ | RMAE | $R^2$ | RMAE | $R^2$ | RMAE |
| **TLMY** | 0.71 | 626.01 | 0.67 | 612.37 | 0.76 | 497.00 | 0.67 | 577.39 |
| **305-DMY** | 0.85 | 327.10 | 0.85 | 346.41 | 0.83 | 320.93 | 0.81 | 357.24 |
| **PY** | 0.74 | 1.64 | 0.72 | 1.85 | 0.77 | 1.39 | 0.78 | 2.65 |

of the variance in 305-DMY. The RMAE ranged from 320.93 (trainbr) to 327.10 (trainscg), with trainbr again showing the best performance. The BNN model exhibited an $R^2$ of 0.81, explaining 81% of the variance, and an RMAE of 357.24, which falls within a reasonable range compared to the ANN models.

For PY, the ANN models showed $R^2$ values of 0.74 and 0.72 (trainlm and trainscg) and 0.77 (trainbr), indicating that the trainbr algorithm explained the most variance in PY. The RMAE for the ANN models ranged from 1.39 (trainbr) to 1.85 (trainscg), with trainbr achieving the best performance. The BNN model achieved an $R^2$ of 0.78, slightly higher than the trainbr algorithm, and an RMAE of 2.65, which was higher compared to the ANN models.

In summary, the ANN models, particularly those trained with Bayesian regularization, generally outperformed BNN in terms of RMAE while showing comparable or slightly higher $R^2$ values. However, the BNN model demonstrated robust predictive performance and served as a competitive alternative to ANN, particularly in capturing complex relationships within the data.

* * *

# Discussion

# 5

# DISCUSSION

The objectives of present thesis were threefold. First, it aimed to assess the population structure, demographic patterns, and genetic diversity of the Vrindavani cattle herd to better understand its genetic foundation and evolutionary dynamics. Second, it sought to estimate the (co)variance components and evaluate the efficiency of different statistical models for genetic evaluation of growth, production, and reproduction traits, thereby providing insights into their heritability and other genetic parameters. Finally, the study evaluated the efficiency of advanced computational techniques, including ANN and BNN, in predicting performance genetic traits, offering a modern, data-driven approach to optimize genetic improvement strategies for cattle breeding.

Vrindavani is a crossbred population, developed by crossing Hariana cows with sires of exotic inheritance (Holstein-Friesian, Jersey, and Brown Swiss) in various combinations, with subsequent *interse* mating during later generations. Constant assessment of population structure, genetic architecture, and other genealogical parameters, along with their dynamics is essential for the planning and implementation of an efficient breeding programme. To the best of our knowledge, this is the first study to thoroughly assess and report the population dynamics and genetic diversity parameters of Vrindavani cattle using pedigree records. The number of founders with unknown pedigrees (both parents unknown) was 2620 out of the total 12718 animals present in the data, thus accounting for ~20.60% of the total population. The information pertaining to the founder population is significant for the management of inbreeding in livestock populations (Gowane *et al.*, 2013a). The number of founders for the Vrindavani population was much higher than that reported for the Sahiwal nucleus herd (188) maintained under Indian conditions by Mumtaz *et al.* (2021).

Completeness of ancestral generations decreased with an increase in pedigree depth. Pedigree records are a fundamental component of breeding plans as they provide a historical and genetic framework for making informed decisions about mating pairs, selecting desirable traits, and maintaining genetic diversity within a population. These records are essential for the success and sustainability of breeding programmes across the various livestock species (Farquharson *et al.*, 2017). The accuracy of estimates and predictions made about a population is directly dependent on the completeness of its pedigree information. Estimates that are based on genealogical data such as inbreeding coefficient, effective population size, generation interval, and estimates of gene flow are directly related to pedigree completeness. A large amount of missing pedigree information can lead to underestimation of certain parameters, like the inbreeding coefficient, and overestimation of others, including the effective population size (Cervantes *et al.*, 2011).

The percentage of animals with a complete known pedigree in the immediately preceding generation was 80.87%, revealing a near-optimal degree of pedigree completeness, which could be used for analysis of the population structure and genetic architecture of this population. The pedigree completeness level was 52.56, 32.87, and 15.97% in the second, third, and fourth generations of the whole pedigree. The pedigree of recent generations was more detailed as compared to older generations. A similar trend has been reported earlier in other cattle populations of indicine and taurine inheritance (Santana *et al.*, 2014; Muasya *et al.*, 2011; Mumtaz *et al.*, 2021). Furthermore, the completeness of pedigree information was lower for male animals as compared to their female counterparts. This was as expected, with a lesser number of males being used in farm animals, including ruminants. The information pertaining to the maternal line of inheritance was increasingly complete for the latest generations. Similar to the present study, several investigations have reported decreasing completeness of pedigree information in livestock populations with increasing depth (Cervantes *et al.*, 2011). Pedigree recording has gained increased significance over the last few decades. With increased stress on planned breeding and efficient animal improvement programmes.

The number of founders in the whole population was 2620, while the effective number of founders was 118, representing around 4.48% of the founders. Similarly, the total number of founders in the reference population was 235, other hand the effective number of founders was 16, representing 6.80% of the founder animals in the cohort. The number of ancestors describing 50, 75, and 100 percent of the gene pool in the whole population was 58, 149, and 2194, respectively. The effective number of founders was lower than the total number of animals in the population. This is attributable to the lesser number of sires within the herd and the effects of the selection process within the crossbred population. A lower number of effective founders has been reported to result in reduced genetic variation within the population, lowered estimates of effective population size ($N_e$), increased homozygosity and other similar effects through genetic drift.

The study revealed a relatively smaller number of founders (36) explaining 50% of the genetic variation within the Vrindavani population. On the other hand, the effective number of ancestors ($f_a$) in the reference population was 15. The extent of the genetic bottleneck was assessed based on $f_a$ and $f_e$ estimates and expressed as a ratio of $f_e$ to $f_a$ ($f_e/f_a$). Ideally, this ratio should be equal to one (Pedrosa *et al.*, 2010). The ratio in the Vrindavani population was equal to 1.19, thus revealing evidence of a fewer bottlenecks in the genetic history of the crossbred population (Sakthivel *et al.*, 2019). In other words, the balance between ancestors and founders was optimal in the Vrindavani population. The decline in genetic variability due to genetic drift, predicted by $1/2\ f_a$, was 0.47 %. These losses in population were inevitable for genetic improvement because selected sires were utilized in mating.

The founder genome equivalents ($f_g$) are directly linked to genetic variability loss caused by genetic drift in subsequent generations, whereas the non-founder genome equivalents ($f_{ng}$) help in assessing the amount of genetic drift that may have occurred across the population since its inception. The estimated $f_g$ and $f_{ng}$ values in the reference population were 9.44 and 0.17, respectively. The lower $f_g$ estimate indicates that a small portion of founders' genes were present in the population. Loss in genetic variability is more due to the accumulation of genetic drift in the non-founder generation when the value of $f_e$ is greater than $f_{ng}$ (Baneh and Ahmadpanah 2020). In studies conducted in Brazil, a reduction in genetic diversity over the years has been detected in several breeds, such as polled Nellore, Indubrasil, Gir, and Guzera, by pedigree analysis using estimates based on the theory of the probability of gene origin

(Vozzi *et al.*, 2006; Faria *et al.*, 2009; Brito *et al.*, 2013). A similar response regarding genetic drift, i.e., genetic bottleneck, but at a lower intensity, has also been reported by Faria *et al.* (2009) and Santana *et al.* (2016) for a Brahman population in Brazil.

The mean GI of the whole Vrindavani population was 6.90 years. In concordance with our results, Muasya *et al.* (2011) reported an average generation interval of 6.9 years in Sahiwal cattle maintained in Kenya. Higher estimates of average generation interval have been reported for other indicine breeds, including Brazilian Gir, Guzerat, and Nelore cattle (Faria *et al.*, 2009); dairy Guzerat cattle (Peixoto *et al.*, 2010); dairy Gir cattle (Reis Filho *et al.*, 2010); and Sahiwal cattle (Mumtaz *et al.*, 2021), with estimates ranging between 7.48 and 8.25 years. Similarly, Santana *et al.* (2014) reported a mean generation interval of 9 years in Gir cattle maintained over a 73-year period between 1938 and 2010. However, a lower estimate of the average generation interval (6.3 years) has been reported in Sahiwal cattle of by Kamiti *et al.* (2016). The GI estimate was the highest for sire to daughter (8.49 years), while it was the lowest for dam to daughter (5.29 years) path. Lower average estimates of generation interval for different paths in Vrindavani cattle are attributable to the implementation of artificial insemination program for propagation of generations, using the semen of elite bulls.

Various earlier studies have reported increased GI estimates for sire-related pathways (sire to daughter and sire to sire) in different cattle populations (Mc Parland *et al.*, 2007; Fabbri *et al.*, 2019; Mumtaz *et al.*, 2021). This was partly attributable to the use of a smaller number of sires for breeding purposes and the increased duration of time taken for their selection in progeny testing programs. On the other hand, a smaller generation interval involving females was attributable to the availability of a larger number of females for selection, their selection at early stages (young age), and their early replacement, once they turn old (Mumtaz *et al.*, 2021).

GI is an important parameter that is significant for the sustainable success of any selection programme, as it is directly related to the annual genetic gain that is realized in the population. GI represents the time that is required for transmission of genes from parental to offspring generation; the larger the GI, the lower the genetic gain per unit time, which may result in genetic lag within the population (Mueller and Van Eenennaam, 2022). However, the relation between GI and genetic gain is not straightforward; GI is related to an increase in annual inbreeding rates with an indirect effect on effective population size, population structure, and genetic diversity maintained within the population (Santana *et al.*, 2016). Lower GI in recent years has enhanced the annual genetic gain for the targeted traits, which increases the economic return of the breeding programme because GI measures the time required for genes to be passed from parents to their offspring. The estimates of GI vary from herd to herd within a breed, which might be due to variation in agro-climatic conditions, the genetic structure of the population, and management practices, besides other factors.

The mean inbreeding coefficient in the whole population dataset was 1.1%, with an average increase of 3.44% over generations. The total inbred population was 25.23%, with 48.57% males and the rest females. Males had a slightly higher inbreeding coefficient (1.19%) than females (1.05%). This shows an intense selection of males over females. The herd was under constant selection pressure, and accordingly, males were stringently selected based on the dams' yield. Due to the closed flock structure with no inheritance from outside gene pool, the average inbreeding coefficient and percentage of inbred animals increased with increasing generation number.

The mean average relatedness (AR) for the Vrindavani population was 1.16%, with an estimate of 5.49%. The inbreeding parameters should consider flock and pedigree depth. AR was greater than the mean inbreeding coefficient, with increases of 0.43, 0.89, and 1.09 percent. The mean inbreeding coefficient, percentage of inbred animals, and average relatedness among animals increased across the completed generations. A higher average relatedness combined with a lower inbreeding coefficient indicates a high degree of relatedness among all individuals in the pedigree dataset. This could lead to difficulties while trying to avoid matings between unrelated or distantly related individuals. About 5.52 % of total matings were highly inbred. Among these, 1.15% (n=146) were half-sib and 2.14% (n=272) were parent-offspring mating. The mean inbreeding coefficient increased irregularly from the second generation (0.29) onwards during subsequent generations. The estimate was the highest in the 13th generation (6.36%). The increased coefficient of inbreeding with a higher number of generations indicates the effects of closed herd mating with minimal to no introduction of foreign inheritance. The results also revealed a regular, increasing trend in the coefficient of AR among individuals as the generation number increased. Assessment of AR gains significance to limit the higher frequency of crosses between highly related animals and conserve the genetic variability in a population (Gowane *et al.*, 2013b). AR estimates can also be used to select male animals for breeding and estimate long-term inbreeding in the population under consideration. Chhotaray *et al.* (2021) analyzed ROH segments in crossbred Vrindavani cattle using genomic data, and observed a 3% inbreeding coefficient in the >20 Mb region, indicating significant inbreeding within ROH segments. They also found that the highest genomic coverage was on chromosome 25, with 50.86%, while chromosome 1 had the least coverage, at 12.42%. The mean inbreeding value were in similar with the findings reported by Rege and Wakhungu (1992) in Sahiwal cattle of Kenya. However, previous studies conducted by Muasya *et al.* (2011) and Kamiti *et al.* (2016) have reported lower values as compared to the current study.

The average Ne estimates from maximum, equivalent, and complete generations were 115.56, 56.42, and 46.02, respectively. The effective population size based on regression and log regression on birth date was 77.40 and 71.24, respectively. The effective population size of the founders was 174.92. Similar to our results, Chhotaray *et al.* (2021) and Singh *et al.* (2021) have reported Ne estimates of 40 and 46, respectively in Vrindavani crossbred cattle genome wide SNP. The effective population size was within the recommended critical value of Food and Agriculture Organization (FAO 1998). Normally, the value should lie between 50 and 100 and should not be below 50 to sustain the genetic diversity for selection programmes and conservation (FAO 1998; Bijma, 2000). Maintenance of genetic variability should be given proper emphasis in the population preferably via national breeding policy. The reported estimates of effective population size were not constant and did change with time according to the level of inbreeding in the flock. It is because the value of effective population size illustrates the number of breeding animals required to obtain the average change in inbreeding rate and does not evaluate the total reduction in allelic diversity or changes in breeding structure from year to year. Gutiérrez *et al.* (2003) reported positive relationship between the depth and quality of a pedigree with both effective population size and levels of inbreeding.

Individuals would typically get equal contributions from all the founding descendants of a population, leading to better values for animal conservation. An animal's conservation value increases with its GCI score. The mean values of GCI for all individuals in the population improved continuously over time and reached a maximum value of 15.03, which was estimated for one individual. The proportion of animals with a GCI greater than 10 was 3.22%. If genetic diversity is to be preserved, superior animals that are used for breeding must get an equal genetic contribution from founders in a population. Accordingly, the greater GCI value leads to greater founder's genetic contribution to an individual (Gutiérrez and Goyache 2005). The genetic contribution of each individual from the founder generation grew as the number of the maximum generation increased, which was in concordance with earlier reports (Gutiérrez and Goyache. 2005).

The estimated value of GD in the reference population relative to the base population or Nei expected heterozygosity was 0.947, indicating that 5.3 % of the genetic diversity in the base population was lost during the study period. However, the genetic diversity estimated for the base population ($GD^*$) was 0.969. In the founder generation, losses in heterozygosity due to genetic drift and bottleneck effects (1-GD) and uneven contributions of founder alleles (1-GD*) were calculated as 0.053 and 0.031, respectively. The present findings were similar to the earlier reports of Baneh *et al.* (2020) in Adani goats and Mandal *et al.* (2000) in Muzaffarnagri sheep, who also observed the loss of 3 % and 3.2 % genetic diversity in the base population. However, the loss was dependent on multiple factors.

The values of Ne, *fe,* and $f_a$ are always greater than the $f_g$. The effective number of founder genome equivalents ($f_g$) compensates for asymmetrical founder contributions as well as the fraction of founder genomes lost to genetic drift during bottlenecks. Unequal contributions of the founder genes were confirmed because the current reference population showed low $f_g$ estimates. Furthermore, the discrepancy between $f_g$ and $f_a$ also revealed that alleles from founder animals were lost at random, and the reference population had a small percentage of founder genes. According to the $f_g/f_a$ ratio, the reference population still retained 62.93 % of its ancestral genetic diversity. The realized Ne was higher than the calculated values of $f_e$, $f_a$ and $f_g$ indicating that the population was increasing in effective size.

The variation in different estimates pertaining to population dynamics, structure, and genetic diversity in Vrindavani cattle when compared to other populations was attributable to differences in genealogical parameters like the completeness of the pedigree, breeding and management decisions taken over time, and the pattern of use of breeding animals (males and females).

The current study revealed significant variation, with coefficients of variation ranging from 4.17% to 33.35 % for the traits examined. These values suggested a wide variation among the animals in the current herd. Zambrano and Echeverri (2014) conducted a similar study on a Holstein-Friesian and Jersey dairy herds, revealing a coefficient of variation of 18 to 35% for the growth, production and reproduction traits.

Co (variance) components and heritability estimates using eight animal models for BW0, BW3, BW6, BW12, BW18 and BW24 were elucidated. Based on the log-likelihood ratio tests (Meyer, 1991), model V and VI were emerged as best fit. The selected models were used to estimate the heritability and additive genetic variance in all growth traits. The direct heritability of BW0 obtained in this study was 0.13 to 0.21 across different models. The results align with other findings in different population (Tesfa *et al.,* 2014). However, the value was lower than that reported by Schoeman and Jordaan (1999) for multi-breed cattle in South Africa, and larger than the value previously reported (Zeleke *et al.*,2016) and Horro cattle by Demeke *et al.* (2003). Similar heritability value for BW3, as that of present study has been reported for Holstein Friesian crosses (0.24 ± 0.07) by Zeleke *et al.* (2016). Our estimated $h^2$ value for BW12 was lower than that reported by Laureano *et al.* (2011). We observed similar results in BW18 and BW24; which was lower to those of Albuquerque and Meyer (2001) and Laureano *et al.* (2011) (0.26 and 0.24, respectively). The heritability estimates for growth traits revealed that the phenotypic value for these characters might be useful for genetic gains. Increased in weight gain in these periods can be achieved by using these traits as selection criteria.

The additive genetic variance was the highest in the model, that included covariance structure of additive and maternal genetic effects for all the production traits, due to its ability to capture more genetic influences from both individuals and dam, resulting in better fit of underlying data.

This finding aligned with the results reported by ABO Elfadl and Radwan (2016) in the first and third lactations of Holstein cows. El-Awady *et al.* (2011) published a conflicting opinion, using single trait analysis to estimate variance and ignoring maternal genetic effects, leading to the largest estimates for direct additive variance. Edriss *et al.* (2006) conducted a study using six animal models to assess the impact of direct and maternal genetic factors on milk yield traits in Iranian Holstein cows. They obtained the highest additive genetic variances for milk and fat yield traits from the animal model with cov (a,m). The additive genetic variance in this study were higher than those reported by several authors, including Schutz *et al.* (1992) in the USA, Olori *et al.* (2002) in Ireland, Khattab *et al.* (2005) on Friesian cows in Egypt, Edriss *et al.* (2006) on Holstein Friesian cows in Iran, and Mostafa *et al.* (2013) on Egyptian native cows. Radwan *et al.* (2017) found higher additive and maternal variances in models that included maternal effects for lactation length, total lactation milk yield, and 305 days milk yield equivalent traits, which were similar to the current study. High additive genetic variance indicated that a significant portion of observed phenotypic variation was attributed to genetic factors, rather than environmental factors. This suggested that selective breeding has substantial potential to improve traits in animal populations.

We used eight univariate repeated animal models in this study with heritability estimates ranging from 0.131± 0.023 to 0.247±0.029 for TLMY, 0.145±0.019 to 0.195±0.025 for 305DMY, 0.085±0.009 to 0.11±0.009 for LL, and 0.111±0.008 to 0.230±0.018 for PY. These findings were similar to Singh and Barwal (2012) and Sharma *et al.* (2024) in Murrah buffaloes. Sahin *et al.* (2014) reported direct heritability values of 0.10 for lactation lengh and total lactation milk yield for production variables in Brown Swiss cattle. These results were similar to our findings. In contrast to the findings of this study, Rehman *et al.* (2008) reported a lower heritability estimate of 0.11 for TLMY in Sahiwal cows. The results were more or less similar to the findings of Eid *et al.* (2012), who reported a heritability estimate of 0.04 for Holstein Friesian cows for LL. The results also align with Rehman *et al.* (2008) heritability estimate of 0.09 ± 0.03 for the LL trait in Sahiwal cows in Pakistan. Deb *et al.* (2008) found higher direct heritability for TLMY (0.404) and LL (0.333) than the current study. The results from each model were in alignment with 0.11 from dairy cattle raised in Egypt Holstein Friesian

(Dematawewa *et al.,*1998). On the other hand, various studies from different production environments reported higher estimates (Kadarmideen *et al.,*2000; Ojango *et al.,*2001).

This study found that maternal effects, except for lactation length (LL), had very low effect on production traits. This could be due to genetic or environmental factors overshadowing the influence of maternal effects. Despite the relatively small maternal heritabilities in this study, the inclusion of direct-maternal genetic covariance (from models V to VI) led to an increase in heritabilities for all production records. Elfadl and Radwan (2016) conducted a study that compared three multi-trait animal models and found that the inclusion of maternal additive genetic effects increased maternal heritability and additive maternal variance in first and third lactation in dairy cows. The study also revealed that the inclusion of the maternal component inflated heritability for certain traits. This result is in disagreement with the findings of Zulkadirs *et al.* (2009) and A.S *et al.* (2015). However, the model V estimates showed that the direct-maternal genetic covariances and correlations were consistently negative, indicating a strong inverse relationship between additive and maternal effects. These findings suggested that this relationship may influence the observed improvement in maternal heritabilities from model V to model VI. Various studies have documented a negative correlation between additive and maternal additive genetic effects: -0.45 by Lee *et al.* (2003) and Khattab *et al.* (2005), and -0.58 by Mostafa *et al.* (2013). Conversely, research by Edriss *et al.* (2006) found positive correlations and covariances between these effects.

The present study suggested that continuous selection of high-producing Vrindavani animals would lead to higher milk yield and longer lactation length in future generations. The sire random effect in models VII and VIII greatly reduced the heritability estimaties. Hammoud and Salem (2013) reported conflicting results on lactation traits in Egyptian Holstein cows. They calculated heritability estimates for total milk yield, 305-day milk yield, and lactation length using the sire-dam mixed model, which were lower than the current study.

Heritability estimates for AFC were range from 0.223±0.029 to 0.289±0.031 for different models in Vrindavani cattle. The high heritability estimates for AFC in this study can be attributed to the limiting selection pressure applied to this trait, in contrast to the strong selection pressure for milk production. Additionally, the presence of genetic variability in the herd can be attributed to the use of different sires from various sources during the study period. The current findings were similar with the estimates reported by Ayalew *et al.* (2017) for AFC traits in Holstein Friesian cattle. The present estimate for AFC was higher than the findings of Das *et al.* (2013), who reported values of 0.26 and 0.23 for two and three breed crosses, and Belay (2014), who reported 0.149 for Fogera x Holstein Friesian (HF) crosses. Additionally, lower heritability estimates for this trait were reported by Gutierrez *et al.* (2002) and Oyama *et al.* (2002), with values of 0.235 and 0.215, respectively. Conversely, higher heritability values were reported in Ethiopia: 0.44 by Demeke *et al.* (2004), 0.7 by Haile *et al.* (2009), and 0.4 by Effa *et al.* (2011). The moderate heritability value of AFC found in this study suggests potential for improvement through direct selection or other correlated traits (indirect selection).

Heritability estimates for ICP in Vrindavani cattle were very low, ranging from 0.0001±0.00 to 0.023±0.009 for different models. These estimates for ICP remained consistent with the literature values for the traits. The present findings were similar with these reported by Solemani-Baghshah *et al.* (2014) and Faraji-Arough *et al.* (2011) who found heritability estimates of 0.019 ± 0.007 for Isfahan cows and 0.19 ± 0.005 for Iranian Holstein Friesian cows. Rehman *et al.* (2008) also reported similar estimates (0.02 ± 0.019) for Sahiwal cows in Pakistan. The current estimates are lower than that reported for ICP trait for Iraqi Friesian cows respectably (Hermiz *et al.*,2005). However, it is higher than the estimates reported in other tropical countries (Ojango *et al.*,2001; Zambrano *et al.*,2014). Gaikwad and Narayankhedkar (2000) also reported higher estimates (0.39 ± 0.28) for crossbreed cows (Gir × Holstein Friesian and Gir × Jersey) in India. This relatively lower heritability estimates for ICP could be explained by large environmental variance. ICP is crucial for reproductive efficiency. Longer ICP can also result from poor management and extended dry days. Efficient herd management can shorten the calving interval by reducing dry days. Therefore, improvements in nutrition and reproductive management should lead to a considerable decrease in length of ICP than making purely genetic selection alone.

Heritability of dry period was low in Vrindavani cattle and ranged from 0.007±0.001 to 0.031±0.008 across different models. Heritability estimates of 0.03 for Jersey cows reported by Suhail *et al.* (2010) is similar to present study results. Rehman *et al.* (2008) studied Sahiwal cows and reported low estimate of heritability (0.05 ± 0.02) for dry period, which was slightly higher than the present findings. Proper nutrition and breeding can help shorten the post-partum service length, ultimately reducing dry period.

The heritability estimate for the GP was low and ranged between 0.014±0.001 and 0.028±0.002 across various models in Vrindavani cattle. Ulhoa *et al.* (2016) reported a heritability estimate averaging 0.10 for GL in Nellore cows, which was similar with the present findings. In contrast, Lopez *et al.* (2019) found higher heritability estimates of 0.13 for Korean Cattle. The heritability estimates for the SP ranged between 0.001 and 0.027 across various models. Foxworthy *et al.* (2019) found a similar heritability estimate of 0.06 ± 0.02 for SP trait in Angus cattle. Estimates from other studies, such as Bormann *et al.* (2006), Ghiasi *et al.* (2011), Peters *et al.* (2013), and Rahbar *et al.* (2016), varied widely based on factors like breed, age, and sample size of the cattle evaluated. For instance, Peters *et al.* (2013) reported a higher heritability of 0.18 compared to the current study for SP in Brangus heifers, whereas Rahbar *et al.* (2016) reported lower heritability for first service success and other fertility traits in Holstein dairy cows.

The effect of inbreeding on various models and traits of crossbred Vrindavani cattle showed notable differences when Fx was included or excluded. For TLMY, models including Fx (models I, III, V, VII) generally showed lower additive genetic variance compared to their counterparts without Fx (models II, IV, VI, VIII), indicating that inbreeding reduced genetic variance. Similar trends were observed in the 305-DMY and LL traits, where models accounting for Fx also showed slightly reduced genetic variances and slightly lower heritability estimates. Overall, inbreeding effects led to reduction in genetic variance and heritability estimates across most traits, highlighting the detrimental impact of inbreeding on the genetic potential and performance of crossbred cattle. Specifically, the percentage decrease in additive genetic variance due to inbreeding in different animal models ranges from 1.32% to 5.25% for TLMY, 0.23% to 0.61% for 305-DMY, and 0.29% to 1.30% for LL.

Inbreeding significantly impacted additive genetic variance and heritability estimates across most of growth production and reproductive traits in Vrindavani cattle. For growth traits models including Fx showed a decrease in genetic variance (ranging from 3.75% to

12.3%). The models including Fx show a decrease in genetic variance (ranging from 0.23% to 5.25%) for traits TLMY, 305-DMY, LL, and PY. Similarly, reproductive traits such as AFC, ICP, DD, GP, and SP exhibit a reduction in genetic variance due to inbreeding, with decreases ranging from 1.32% to 28%. These results are supported by Makanjuola *et al.* (2020), who reported that pedigree inbreeding effects to decrease the phenotypic means of TLMY and PY by 0.49%, and 0.47%, respectively. Similarly, Doekes *et al.* (2019) reported decreases of 0.47%, and 0.45% in these traits (TLMY and PY), respectively. Similarly, Faria *et al.* (2009) reported inbreeding levels ranging from 1.75% to 2.28% in Brazilian Zebu cattle, and Brito *et al.* (2013) found an inbreeding level of 1.73% in the Nellore breed. Continuous increase in inbreeding is linked to the rise of deleterious homozygous recessive alleles. This rise in inbreeding level has been reported to negatively impacts additive genetic variance and leads to a reduction in overall production and reproductive traits in various dairy breeds (Martikainen *et al.*,2018). This emphasizes the need to manage inbreeding levels to preserve genetic variance and improve performance in dairy and crossbred cattle.

ANN are machine learning models that model non-nlinear systems using information learned through experience and can handle noisy data. They are considered an alternative to conventional predictive models due to their ability to capture non-linear relationships between predictors and responses. The performance of ANN is determined by network architecture, including hidden layers and neurons, training type, linear or non-linear transformation process, and input data type. To improve ANN performance, factors such as network dimensionality must be considered, as network overfitting or underfitting can lead to reduced performance predictions (Gianola *et al.*,2011; Ehret *et al.*,2015). The study found that a single hidden layer model with 10 to 20 neurons was the best non-linear model for performance traits in cattle, with high $R^2$ value (0.74 to 0.85).

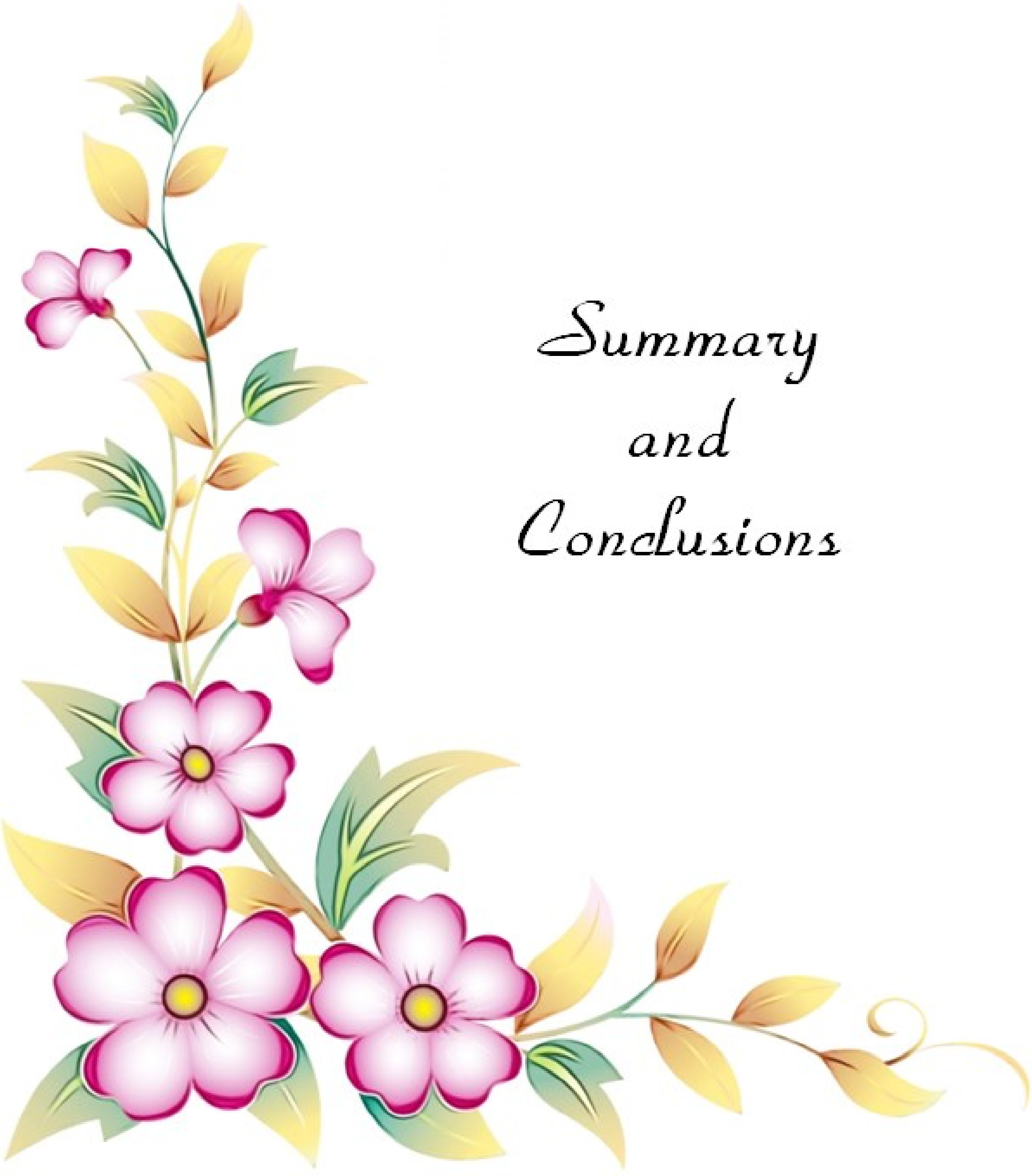
Study performed by Görgülü *et al.* (2012) deduced that ANN could be used as a substitute to the multiple regression model to forecast milk production corrected at 305 days lactation period in Brown Swiss dairy cattle and cumulative milk yield in crossbred cattle. Similarly, many have authors compared ANN algorithm to classical statistical methods such as fuzzy logic (Atil and Akilli,2014), K-means for clustering milk-producing cattle (Atil and Akilli, 2016), and multiple linear regressions for the anticipation of body weight in hair-bearing goats (Akkol *et al.,*2017). They found that ANN approach has a better performance in prediction, especially when the association between variables is complicated. In addition, Behzadi and Aslaminej (2010) reported that ANN had been used for both classification and prediction data in several knowledge fields. Chaturvedi *et al.* (2013) stated that ANN is a strong predictor of future milk production, relying on early expressed merits. Another study conducted by Ali *et al.* (2015) to compare neural networks to traditional statistical approaches, reported that the ANN made the best of testing the most complicated associations between input and output variables.

BNN models offer a probabilistic approach to modeling, providing both predictions and uncertainty estimates, which are particularly useful in scenarios with limited or noisy data. Unlike traditional ANN models, BNN algorithm incorporate prior information into the learning process, enhancing their ability to generalize, especially for traits with sparse or imbalanced datasets. In this study, BNN models showed competitive performance compared to ANN (76%, 85%, and 77%) with $R^2$ values of 67%, 81%, and 78% for TLMY, 305-DMY, and peak yield, respectively, suggesting their robustness in certain scenarios. Previous studies, such as González-Camacho *et al.* (2018) and Perez-Rodriguez *et al.* (2020), have demonstrated the advantages of BNN models in genomic selection and complex traits with limited data, where uncertainty quantification was critical. BNN models were particularly superior for peak yield prediction in this study, indicating their strength in handling traits with lesser variability and complexity. However, for traits like 305-DMY, where the data is less noisy and more linear, ANN performed better, achieving an $R^2$ of 85% compared to 81% for BNN. This aligns with findings by Montesinos-López *et al.* (2021), who highlighted that ANN models often outperform BNN algorithm for traits with strong and consistent non-linear patterns, while BNN excel in uncertain or less deterministic traits due to their probabilistic framework. Together, these findings emphasize the complementary strengths of ANN and BNN, with BNN being advantageous in handling uncertainty and sparse data while ANN excelling in accurately predicting performance traits with well-defined relationships.

This study provides a detailed investigation of the performance traits and estimation of genetic and performance parameters of Vrindavani cattle, beginning with an assessment of population dynamics, genetic diversity, and demographic patterns. Pedigree analysis revealed that while the population maintained a moderate level of genetic diversity, effective management of inbreeding and founder contributions is essential for sustainable breeding programs. The genetic evaluation of growth, production, and reproduction traits highlighted significant variation within the herd, with heritability estimates indicating potential for genetic improvement through selection. Advanced machine learning tools, including ANN and BNN, were employed to predict key performance traits. ANN models demonstrated superior performance in predicting traits with well-defined nonlinear patterns, achieving high $R^2$ values, while BNN model excelled in handling uncertainty and variability in traits with noisy or sparse data. These models complement traditional genetic evaluation methods by offering more precise predictions and insights. The findings underscore the importance of integrating computational techniques into breeding programs to optimize genetic improvement strategies. By balancing selection for higher productivity with measures to preserve genetic diversity, this study provides a framework for enhancing the genetic potential and long-term sustainability of Vrindavani cattle populations.

* * *

# Summary and Conclusions



# 6 SUMMARY AND CONCLUSIONS

The present research attempted a comprehensive analysis of the genetic evaluation, population structure, and performance modelling of traits in Vrindavani cattle using advanced statistical and computational tools. Vrindavani cattle, a synthetic breed developed in India by crossbreeding indigenous Hariana cattle with Holstein-Friesian, Jersey, and Brown Swiss, has shown the capacity to produce over 3,000 kg of milk per lactation. The study utilized extensive pedigree and performance data spanning 53 years (1970–2023) from a population of 12,718 Vrindavani cattle to examine the population dynamics using ENDOG software. The analysis revealed a high level of genetic diversity, with key parameters like the pedigree completeness index and effective population size playing crucial roles in understanding the genetic value of the population. The study also revealed that the effective population size was nearing critical thresholds, which indicated a need for targeted breeding strategies to minimize genetic drift.

The Vrindavani population had a pedigree completeness of 80.87% in the preceding generation, but declined in older generations. The population had 2,620 founders, with an effective number of 118, representing around 4.48% of the founders. The ratio of the effective number of founders to effective number of ancestors was 1.19, indicating few bottlenecks in the genetic history. The mean generation interval for the Vrindavani population was 6.90 years, with the sire to daughter path having the longest interval (8.49 years) and the dam to daughter path the shortest (5.29 years). The mean inbreeding coefficient for the entire population was 1.1%, showing a 3.44% increase over generations. The mean average relatedness was 1.16% for the whole population and 3.49% for the reference cohort. 5.52% of total matings were highly inbred, including half-sib and parent-offspring types. The genetic conservation index

(GCI) improved over time, reaching a maximum value of 15.03. The genetic diversity (GD) in the reference population was 0.947, reflecting a 5.3% loss of genetic diversity compared to the base population.

The research also explores the (co)variance components and genetic parameters for growth, production, and reproduction traits using different animal models and statistical methods. The results emphasize the complex interaction between additive genetic effects, maternal influences, and environmental factors in determining the phenotypic variance in crossbred cattle. For growth traits, heritability estimates varied, with models incorporating inbreeding showing lower additive genetic variance leading to genetic erosion. The heritability for production traits ranged from 0.131 to 0.247 for TLMY, with maternal effects having a limited impact on production traits, except for lactation length. The study revealed that models including sire effects heritability estimates, and the permanent environmental effects significantly impacted the total variance, especially for lactation length and peak yield. Heritability estimates for reproduction traits were generally low, suggesting minimal contributions of additive genetic effects and prolong impact of environment.

Furthermore, the study evaluated the effectiveness of advanced computing techniques, specifically ANN and BNN, for predicting performance traits. ANN models were trained using the Levenberg-Marquardt algorithm and achieved high prediction accuracy for TLMY, 305-DMY, and PY, demonstrating their ability to capture non-linear relationships within the data. BNN models, which incorporated Bayesian inference to quantify uncertainty in predictions, also showed competitive performance. The study revealed that BNNs were advantageous in handling uncertainty and sparse data, while ANNs excel in accurately predicting performance traits with well-defined relationships. A comparison of the models showed that, for example, BNNs achieved an R-squared value of 0.81 for 305-DMY, while for the same trait, ANN had an R-squared of 0.85.

In conclusion, this study demonstrated the potential of combining traditional and advanced methodologies to address the challenges of dairy cattle breeding in India. The research emphasized the importance of detailed pedigree and performance data, along with cutting-edge computational tools, to enhance the productivity, profitability, and genetic health of dairy cattle populations. By implementing targeted breeding strategies, managing inbreeding, and utilizing advanced machine learning models, the study contributed to the broader goal of improving dairy cattle breeding programs in India and beyond.

* * *

# Mini Abstract

# 7

# MINI ABSTRACT

The present research evaluated the genetic structure, population dynamics, and performance optimization of Vrindavani cattle, a synthetic breed developed in India by crossbreeding indigenous Hariana cattle with Holstein-Friesian, Jersey, and Brown Swiss breeds. Using a 53-year dataset of pedigree and performance records from 12,718 individuals, the research integrated traditional statistical methods and advanced computational tools, including ANN and BNN, to assess key aspects of dairy cattle breeding. Population analysis revealed high genetic diversity, with a near-optimal pedigree completeness index of 80.87% in the most recent generation. Key findings included an effective number of 118 founders out of 2,620, a mean generation interval of 6.90 years, and an inbreeding coefficient of 1.1% for the whole population, which increased by 3.44% in the reference population. Effective population size ranged from 46.02 to 115.56, indicating the need for targeted breeding strategies. Genetic diversity in the reference population was 0.947, with a 5.3% loss compared to the base population. A higher genetic conservation index of 15.03 highlighted the presence of animals with high conservation value. Genetic parameters for growth, production, and reproduction traits were also analyzed. Inbreeding reduced genetic variance for growth traits, while heritability estimates for production traits like TLMY ranged from 0.131 to 0.247. Sire effects reduced heritability, and permanent environmental effects significantly impacted variance in lactation length and peak yield. Reproduction traits showed low heritability estimates, suggesting minimal additive genetic influence. The effectiveness of ANN and BNN models in predicting performance traits was also evaluated. ANN models achieved a R-squared value of 0.85 for 305DMY, while BNN models, incorporating Bayesian inference, achieved an R-squared value of 0.81. BNN models showed better performance in handling uncertainty, while ANN performed well for traits with well-defined relationships. In conclusion, this study emphasized the importance of combining traditional genetic evaluation methods with machine learning tools like ANN and BNN to optimize dairy cattle breeding programs. It provided valuable insights into genetic health, inbreeding management, and performance prediction, offering implications for enhancing productivity, profitability, and genetic sustainability in dairy cattle populations.

# लघु सारांश

# 8

# लघु सारांश

वर्तमान शोध में वृंदावनी गायों की आनुवंशिक संरचना, जनसंख्या गतिकी, और प्रदर्शन अनुकूलन का मूल्यांकन किया गया है। वृंदावनी एक संश्लेषित नस्ल है, जिसे भारत में देशी हरियाणा गायों को होल्स्टीन–फ्रीजियन, जर्सी, और ब्राउन स्विस नस्लों के साथ क्रॉसब्रीडिंग करके विकसित किया गया है। इस शोध में 53 वर्षों के दौरान 12.718 पशुओं की वंशावली और प्रदर्शन रिकॉर्ड का उपयोग किया गया, जिसमें पारंपरिक सांख्यिकीय विधियों और उन्नत कम्प्यूटेशनल टूल्स जैसे आर्टिफिशियल न्यूरल नेटवर्क (ANN) और बायोज़ियन न्यूरल नेटवर्क (BNN) को शामिल किया गया है, ताकि डेयरी मवेशी प्रजनन के प्रमुख पहलुओं का आकलन किया जा सके। जनसंख्या विश्लेषण में उच्च आनुवंशिक विविधता का पता चला, जिसमें नवीनतम पीढ़ी में वंशावली पूर्णता सूचकांक 80.87% था। प्रमुख निष्कर्षों में 2,620 संस्थापकों में से 118 की प्रभावी संख्या, औसत पीढ़ी अंतराल 6.90 वर्ष, और पूरी आबादी के लिए 1.1% का इनब्रीडिंग गुणांक शामिल हैं, जो संदर्भ आबादी में 3.44% तक बढ़ गया। प्रभावी जनसंख्या आकार के अनुमान 46.02 से 115.56 के बीच थे, जो खित प्रजनन रणनीतियों की आवश्यकता को दर्शाते हैं। संदर्भ आबादी में आनुवंशिक विविधता 0.947 थी, जो आधार आबादी की तुलना में 5.3% की हानि दर्शाती है। 15.03 के उच्च आनुवंशिक संरक्षण सूचकांक ने उच्च संरक्षण मूल्य वाले जानवरों की उपस्थिति को उजागर किया। वृद्धि, उत्पादन, और प्रजनन लक्षणें के लिए आनुवंशिक मानदंडों का भी विश्लेषण किया गया। इनब्रीडिंग ने वृद्धि लक्षणों के लिए आनुवंशिक परिवर्तनशीलता को कम कर दिया, जबकि उत्पादन लक्षणों जैसे कुल लैक्टेशन दूध उत्पादन (TLMY) के लिए हेरिटेबिलिटी अनुमानों की सीमा 0.131 से 0.247 तक थी। सायर प्रभावों ने हेरिटेबिलिटी को कम किया, और स्थायी पार्यावरणीय प्रभावों ने लैक्टेशन अवधि और पीक यील्ड में परिवर्तनशीलता पर महत्वपूर्ण प्रभाव डाला। प्रजनन लक्षणों ने निम्न हेरिटेबिलिटी अनुमान दिखाए, जो न्यूनतम एडिटिव आनुवंशिक प्रभाव का संकेत देते हैं। प्रदर्शन लक्षणों की भविष्यवाणी में ANN और BNN मॉडलों की प्रभावशीलता का भी मूल्यांकन किया गया। ANN मॉडलों ने 305 दिनों के दूध उत्पादन (305DMY) के लिए 0.85 का आर–स्क्वेयर मान प्राप्त किया, जबकि बायेजियन अनुमानों को शामिल करने वाले BNN मॉडलों ने 0.81 का आर–स्क्वेयर मान प्राप्त किया। BNN मॉडल अनिश्चितता को संभालने में बेहतर प्रदर्शन करते हैं, जबकि ANN मॉडलों ने उन लक्षणों के लिए अच्छा प्रदर्शन किया जिनमें स्पष्ट संबंध थे। निष्कर्षत, यह अध्ययन पारंपरिक आनुवंशिक मूल्यांकन विधियों को ANN और BNN जैसे मशीन लर्निंग टूल्स के साथ संयोजित करने के महत्व को रेखांकित करता है। ताकि डेयरी मवेशी प्रजनन कार्यक्रमों का अनुकूलन किया जा सके। यह अध्ययन आनुवंशिक स्वास्थ्य, इनब्रीडिंग प्रबंधन, और प्रदर्शन भविष्यवाणी के क्षेत्र में महत्वपूर्ण अंतदृष्टि प्रदान करता है, जो डेयरी मवेशी आबादी में उत्पादकता, लाभप्रदता, और आनुवंशिक स्थिरता बढ़ाने के लिए प्रभावशाली है।

# References

# 9

# REFERENCES

Abbas, A., Ullah, M. A., & Waheed, A. (2021). Body weight prediction using different data mining algorithms in Thalli sheep: A comparative study. *Veterinary World*, 2332–2338. https://doi.org/10.14202/vetworld.2021.2332-2338

Adamczyk, K., Grzesiak, W., & Zaborski, D. (2021). The Use of Artificial Neural Networks and a General Discriminant Analysis for Predicting Culling Reasons in Holstein-Friesian Cows Based on First-Lactation Performance Records. *Animals*, *11*(3), 721. https://doi.org/10.3390/ani11030721

Adrion, J. R., Galloway, J. G., & Kern, A. D. (2020). Predicting the Landscape of Recombination Using Deep Learning. *Molecular Biology and Evolution*, *37*(6), 1790–1808. https://doi.org/10.1093/molbev/msaa038

Ahin, A., Uluta°, Z., Yýlmaz Adkinson, A., & W. Adkinson, R. (2014). Genetic parameters of first lactation milk yield and fertility traits in brown Swiss cattle. *Annals of Animal Science*, *14*(3). https://doi.org/10.2478/aoas-2014-0038

Akýllý, A., & Atýl, H. (2014). Artificial intelligence technology in dairy cattle: Fuzzy logic and artificial neural networks. *Animal Production*, 55(1), 39–45.

Akkol, S., Akilli, A., & Cemal, Ý. (2017). Kýl Keçilerinin Canlý Aðýrlýk Tahmininde Yapay Sinir Aðlarý ve Çoklu Doðrusal Regresyon Yöntemlerinin Karþýlaþtýrýlmasý. *Yüzüncü Yýl Üniversitesi Tarým Bilimleri Dergisi*, *27*(1), 21–29. https://doi.org/10.29133/yyutbd.263968

Akýllý, A., & Atýl, H. (2014). Süt Sýðýrcýlýðýnda Yapay Zeka Teknolojisi: Bulanýk Mantýk ve Yapay Sinir Aðlarý. *Hayvansal Üretim*, 39–45. https://doi.org/10.29185/hayuretim.363911

Albuquerque, L. G., & Meyer, K. (2001). Estimates of covariance functions for growth from birth to 630 days of age in Nelore cattle. *Journal of Animal Science*, *79*(11). https://doi.org/10.2527/2001.79112776x

Alderson, G. L. H. (1992). A system to maximize the maintenance of genetic variability in small populations. In *Genetic conservation of domestic livestock.*

Ali, M., Eyduran, E., Tariq, M. M., Tirink, C., Abbas, F., Bajwa, M. A., Baloch, M. H., Nizamani, A. H., Waheed, A., Awan, M. A., Shah, S. H., Ahmad, Z., & Jan, S. (2015). Comparison of artificial neural network and decision tree algorithms used for predicting live weight at post weaning period from some biometrical characteristics in harnai sheep. *Pakistan Journal of Zoology*, *47*(6).

Alves, K., Schenkel, F. S., Brito, L. F., & Robinson, A. (2018). Estimation of direct and maternal genetic parameters for individual birth weight, weaning weight, and probe weight in Yorkshire and Landrace pigs. *Journal of Animal Science*, *96*(7). https://doi.org/10.1093/jas/sky172

Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M., & Farhan, L. (2021). Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, *8*(1). https://doi.org/10.1186/s40537-021-00444-8

Angeles-Hernandez, J. C., Castro-Espinoza, F. A., Peláez-Acero, A., Salinas-Martinez, J. A., Chay-Canul, A. J., & Vargas-Bello-Pérez, E. (2022). Estimation of milk yield based on udder measures of Pelibuey sheep using artificial neural networks. *Scientific Reports*, *12*(1), 9009. https://doi.org/10.1038/s41598-022-12868-0

Antwi, P., Li, J., Boadi, P. O., Meng, J., Shi, E., Deng, K., & Bondinuba, F. K. (2017). Estimation of biogas and methane yields in an UASB treating potato starch processing wastewater with backpropagation artificial neural network. *Bioresource Technology*, *228*. https://doi.org/10.1016/j.biortech.2016.12.045

Atil, H., & Akilli, A. (2015). Investigation of dairy cattle traits by using artificial neural networks and cluster analysis. *CEUR Workshop Proceedings*, *1498*.

Atýl, H., & Akýllý, A. (2016). Comparison of artificial neural network and K-means for clustering dairy cattle. *International Journal of Sustainable Agricultural Management and Informatics*, *2*(1), 40-52.

Ayalew, W., Aliy, M., & Negussie, E. (2017). Estimation of genetic parameters of the productive and reproductive traits in Ethiopian Holstein using multi-trait models. *Asian-Australasian Journal of Animal Sciences*, *30*(11). https://doi.org/10.5713/ajas.17.0198

Ayalew, W., Chu, M., Liang, C., Wu, X., & Ping, Y. (2021). Adaptation mechanisms of yak (Bos grunniens) to high-altitude environmental stress. In *Animals* (Vol. 11, Issue 8). https://doi.org/10.3390/ani11082344

BAHS. (2023). *Basic Animal Husbandry Statistics-2023*. Https://Dahd.Nic.in/Sites/Default/Filess/BAHS2023.Pdf.

Ballou, J. D., Lees, C., Faust, L. J., Long, S., Lynch, C., B, L. L., & Foose, T. J. (2010). Demographic and Genetic Management of Captive Populations. *Wild Mammals in Captivity, Principles and Techniques*.

Baltenweck, I., Enahoro, D., Frija, A., & Tarawali, S. (2020). Why Is Production of Animal Source Foods Important for Economic Development in Africa and Asia? *Animal Frontiers/ : The Review Magazine of Animal Agriculture*, *10*(4), 22–29. https://doi.org/10.1093/af/vfaa036

Baneh, H., & Ahmadpanah, J. (2020). Mgenetic evaluation of composite reproductive traits in ghezel sheep. *Journal of Animal and Plant Sciences*, *30*(1).

Baneh, H., Javanrouh, A., Sadeghi, S. A. T., Yazdanshenas, M. S., Mandal, A., Ahmadpanah, J., & Mohammadi, Y. (2020). Characterization of population structure and genetic diversity of adani goats. *Journal of Livestock Science and Technologies*, *8*(1). https://doi.org/10.22103/jlst.2020.15775.1315

Bautista-Díaz, E., Mezo-Solis, J. A., Herrera-Camacho, J., Cruz-Hernández, A., Gomez-Vazquez, A., Tedeschi, L. O., Lee-Rangel, H. A., Vargas-Bello-Pérez, E., & Chay-Canul, A. J. (2020). Prediction of Carcass Traits of Hair Sheep Lambs Using Body Measurements. *Animals*, *10*(8), 1276. https://doi.org/10.3390/ani10081276

Behzadi, M. R. B., & Aslamineja, A. A. (2010). A Comparison of Neural Network and Nonlinear Regression Predictions of Sheep Growth. *Journal of Animal and Veterinary Advances*, *9*(16), 2128–2131. https://doi.org/10.3923/javaa.2010.2128.2131

Belay Zeleke. (2014). Estimation of genetic parameters for growth and reproductive traits of Fogera x Holstein Friesian crossbred cattle at Metekel Ranch, Amhara Region, Ethiopia. MSc Thesis, Haramaya University, Haramaya,Ethiopia

Bijma, P. (2000). Long-term Genetic Contributions - Prediction of Rates of Inbreeding and Genetic Gain in Selected Populations. In *Animal Breeding and Genetic Group, Department of Animal Sciences.*

Billah, M., Wang, X., Yu, J., & Jiang, Y. (2022). Real-time goat face recognition using convolutional neural network. *Computers and Electronics in Agriculture*, *194*, 106730. https://doi.org/10.1016/j.compag.2022.106730

Blum, M. G. B., & François, O. (2010). Non-linear regression models for Approximate Bayesian Computation. *Statistics and Computing*, *20*(1), 63–73. https://doi.org/10.1007/s11222-009-9116-0

Blundell, C., Cornebise, J., Kavukcuoglu, K., Com, W., & Deepmind, G. (2025). Weight Uncertainty in Neural Networks Daan Wierstra. *ICML'15: Proceedings of the 32nd International Conference on International Conference on Machine Learning - Volume 37July*.

Boichard, D., Maignel, L. & Verrier, É (1997). The value of using probabilities of gene origin to measure genetic variability in a population. *Genetics Selection Evolution* 29, 5. https://doi.org/10.1186/1297-9686-29-1-5

Bormann, J. M., Totir, L. R., Kachman, S. D., Fernando, R. L., & Wilson, D. E. (2006). Pregnancy rate and first-service conception rate in Angus heifers. *Journal of Animal Science*, *84*(8). https://doi.org/10.2527/jas.2005-615

Brito, F. V., Sargolzaei, M., Braccini Neto, J., Cobuci, J. A., Pimentel, C. M., Barcellos, J., & Schenkel, F. S. (2013). In-depth pedigree analysis in a large Brazilian Nellore herd. *Genetics and Molecular Research*, *12*(4). https://doi.org/10.4238/2013.November.22.2

Bureš, D., & Bartoò, L. (2012). Growth performance, carcass traits and meat quality of bulls and heifers slaughtered at different ages. *Czech Journal of Animal Science*, *57*(1), 34–43. https://doi.org/10.17221/5482-CJAS

Caballero*, A., & Toro, M. A. (2002). Interrelations between effective population size and other pedigree tools for the management of conserved populations. *Conservation Genetics*, *3*(3), 289–299. https://doi.org/10.1023/A:1019956205473

Calvin, O. L., & McDowell, J. J. (2016). Extending unified-theory-of-reinforcement neural networks to steady-state operant behavior. *Behavioural Processes*, *127*, 52–61. https://doi.org/10.1016/j.beproc.2016.03.016

Cervantes, I., Goyache, F., Molina, A., Valera, M., & Gutiérrez, J. P. (2011). Estimation of effective population size from the rate of coancestry in pedigreed populations. *Journal of Animal Breeding and Genetics*, *128*(1), 56–63. https://doi.org/10.1111/j.1439-0388.2010.00881.x

Cervantes, I., Gutiérrez, J. P., & Meuwissen, T. H. E. (2016). Response to selection while maximizing genetic variance in small populations. *Genetics Selection Evolution*, *48*(1), 69. https://doi.org/10.1186/s12711-016-0248-3

Chafai, N., Hayah, I., Houaga, I., & Badaoui, B. (2023). A review of machine learning models applied to genomic prediction in animal breeding. In *Frontiers in Genetics* (Vol. 14). https://doi.org/10.3389/fgene.2023.1150596

Chakraborty, S., & Ghosh, M. (2012). Applications of Bayesian Neural Networks in Prostate Cancer Study. In *Handbook of Statistics* (Vol. 28). https://doi.org/10.1016/B978-0-44-451875-0.00009-9

Chaturvedi, S., Gupta, A. K., Yadav, R. L., & Sharma, A. K. (2013). Life Time Milk Amount Prediction in Dairy Cows using Artificial Neural Networks. In *International Journal of Recent Research and Review: Vol. V.*

Chavlis, S., & Poirazi, P. (2021). *Drawing Inspiration from Biological Dendrites to Empower Artificial Neural Networks*. https://doi.org/10.1016/j.conb.2021.04.007

Chhotaray, S., Panigrahi, M., Pal, D., Ahmad, S. F., Bhanuprakash, V., Kumar, H., Parida, S., Bhushan, B., Gaur, G. K., Mishra, B. P., & Singh, R. K. (2021). Genome-wide estimation of inbreeding coefficient, effective population size and haplotype blocks in Vrindavani crossbred cattle strain of India. *Biological Rhythm Research*, *52*(5), 666–679. https://doi.org/10.1080/09291016.2019.1600266.

Chimonyo, M., Dzama, K., & Bhebhe, E. (2008). Genetic determination of mothering ability and piglet growth in indigenous Mukota sows of Zimbabwe. *Livestock Science*, *113*(1). https://doi.org/10.1016/j.livsci.2007.02.014.

Chitra, P. T., Ponnuvel, P., Pan Perunthalaivar, S., Krishi, K., & Kendra, V. (2016). Bovine population dynamics of West Bengal-A case study. *International Journal of*

*Advanced Research in Biological Sciences,* 3(2). http://s-o-i.org/1.15/ijarbs-2016-3-2-11

Çilek, S., & Kaygisiz, A. (2008). Breeding value estimation of dairy cattle using test day milk yields for Brown Swiss cows reared at Ulaº State Farm. *Journal of Animal and Veterinary Advances*, *7*(6).

Collier, R. J., Dahl, G. E., & Vanbaale, M. J. (2006). Major advances associated with environmental effects on dairy cattle. *Journal of Dairy Science*, *89*(4). https://doi.org/10.3168/jds.S0022-0302(06)72193-2

Das, A., Miah, G., Das Gupta, M., & Islam Khan, M. K. (2013). Genetic parameters of holstein crossbred on commercial dairy farms in Chittagong, Bangladesh. *Indian Journal of Animal Research*, *47*(4).

de los Campos, G., Hickey, J. M., Pong-Wong, R., Daetwyler, H. D., & Calus, M. P. L. (2013). Whole-Genome Regression and Prediction Methods Applied to Plant and Animal Breeding. *Genetics*, *193*(2), 327–345. https://doi.org/10.1534/genetics.112.143313

Deb, G., Mufti, M., Mostari, M., & Huque, K. (2008). Genetic Evaluation Of Bangladesh Livestock Research Institute Cattle Breed-1/ : Heritability And Genetic Correlation. *Bangladesh Journal of Animal Science*, *37*(2). https://doi.org/10.3329/bjas.v37i2.9878

Dematawewa, C. M. B., & Berger, P. J. (1998). Genetic and Phenotypic Parameters for 305-Day Yield, Fertility, and Survival in Holsteins. *Journal of Dairy Science*, *81*(10). https://doi.org/10.3168/jds.S0022-0302(98)75827-8

Demeke, S., Neser, F. W. C., & Schoeman, S. J. (2003). Variance components and genetic parameters for early growth traits in a mixed population of purebred Bos indicus and crossbred cattle. *Livestock Production Science*, *84*(1). https://doi.org/10.1016/S0301-6226(03)00074-5

Demeke, S., Neser, F. W. C., & Schoeman, S. J. (2004). Estimates of genetic parameters for Boran, Friesian, and crosses of Friesian and Jersey with the Boran cattle in the tropical highlands of Ethiopia: milk production traits and cow weight. *Journal of Animal Breeding and Genetics*, *121*(3), 163–175. https://doi.org/10.1111/j.1439-0388.2004.00446.x

Diop, M., Dodenhoff, J., & Van Vleck, L. D. (1999). Estimates of direct, maternal and grandmaternal genetic effects for growth traits in Gobra cattle. *Genetics and Molecular Biology*, *22*(3). https://doi.org/10.1590/S1415-47571999000300013

Doekes, H. P., Veerkamp, R. F., Bijma, P., De Jong, G., Hiemstra, S. J., & Windig, J. J. (2019). Inbreeding depression due to recent and ancient inbreeding in Dutch Holstein-Friesian dairy cattle. *Genetics Selection Evolution*, *51*(1). https://doi.org/10.1186/s12711-019-0497-z

Dongre, V. B., & Gandhi, R. S. (2016). Applications of artificial neural networks for enhanced livestock productivity: A review. *The Indian Journal of Animal Sciences*, *86*(11). https://doi.org/10.56093/ijans.v86i11.62970

Dunner, S., Checa, M. L., Gutierrez, J. P., Martin, J. P., & Cañon, J. (1998). Genetic analysis and management in small populations: the Asturcon pony as an example. *Genetics Selection Evolution*, *30*(4), 397. https://doi.org/10.1186/1297-9686-30-4-397

Dutt, T., & Kumar, N. S. (1999). Genetic divergence among Hariana crosses with exotic dairy breeds. *Indian Journal of Animal Sciences*, *69*(10).

Èandek-Potokar, M., Prevolnik, M., Škrlep, M., Font-i-Furnols, M., & Noviè, M. (2015). An Attempt to Predict Conformation and Fatness in Bulls by Means of Artificial Neural Networks Using Weight, Age and Breed Composition Information. *Italian Journal of Animal Science*, *14*(1), 3198. https://doi.org/10.4081/ijas.2015.3198

Edriss, M. A., Nilforooshan, M. A., & Sadeghi, J. M. (2006). Estimation of direct genetic and maternal effects for production traits of Iranian Holstein cows using different animal models. *Pakistan Journal of Biological Sciences*, *9*(4). https://doi.org/10.3923/pjbs.2006.636.640

Effa, K., Wondatir, Z., Dessie, T., & Haile, A. (2011). Genetic and environmental trends in the long-term dairy cattle genetic improvement programmes in the central tropical highlands of Ethiopia. *Journal of Cell and Animal Biology*, *5*(6).

Ehret, A., Hochstuhl, D., Gianola, D., & Thaller, G. (2015). Application of neural networks with back-propagation to genome-enabled prediction of complex traits in Holstein-Friesian and German Fleckvieh cattle. *Genetics Selection Evolution*, *47*(1). https://doi.org/10.1186/s12711-015-0097-5

Eid, I. I., Elsheikh, M. O., & Yousif, I. A. (2012). Estimation of Genetic and Non-Genetic Parameters of Friesian Cattle under Hot Climate. *Journal of Agricultural Science*, *4*(4). https://doi.org/10.5539/jas.v4n4p95

El-Awady, H.G., E.Z.M., O., N.A., S., M.N., E. A., & H.R., M. (2011). Genetic improvement study on pre-weaning body weight of Egyptian Rahmani lambs under a pure breeding production system. *Options Méditerranéennes. Séries A. Mediterranean Seminars*, *100*(100).

Elfadl, E. A. A., & Radwan, H. A. (2016). Model Comparisons and Genetic Parameters Estimates of Productive Traits in Holstein Cows. *Journal of Applied Sciences*, *16*(8). https://doi.org/10.3923/jas.2016.380.387

Ewing, G. B., & Jensen, J. D. (2016). The consequences of not accounting for background selection in demographic inference. *Molecular Ecology*, *25*(1), 135–141. https://doi.org/10.1111/mec.13390

Fabbri, M. C., Gonçalves de Rezende, M. P., Dadousis, C., Biffani, S., Negrini, R., Souza Carneiro, P. L., & Bozzi, R. (2019). Population Structure and Genetic Diversity of Italian Beef Breeds as a Tool for Planning Conservation and Selection Strategies. *Animals/ : An Open Access Journal from MDPI*, *9*(11). https://doi.org/10.3390/ani9110880

Falconer, D. S., & Mackay, T. F. C. (1996). Introduction to Quantitative Genetics (Fourth Edition). In *Trends in Genetics* (Vol. 12).

FAO (2007) Criteria for minimum population sizes. Food and Agricultural Organization

FAO. (1998). Secondary guidelines for development of national farm animal genetic resources management plans: management of small populations at risk. *Food and Agricultural Organization*.

Faraji Arough, H., Aslaminejad, A. A., & Farhangfar, H. (2011). Estimation of genetic parameters and trends for age at first calving and calving interval in Iranian holstein cows. *Journal of Research in Agricultural Science, 7(1)*, 79–87

Faria, F. J. C., Filho, A. E. V., Madalena, F. E., & Josahkian, L. A. (2009). Pedigree analysis in the Brazilian Zebu breeds. *Journal of Animal Breeding and Genetics*, *126*(2). https://doi.org/10.1111/j.1439-0388.2008.00767.x

Farquharson, K. A., Hogg, C. J., & Grueber, C. E. (2017). Pedigree analysis reveals a generational decline in reproductive success of captive Tasmanian devil (Sarcophilus harrisii): implications for captive management of threatened species. *Journal of Heredity*, *108*(5), 488–495. https://doi.org/10.1093/jhered/esx030

Fekadu, A., Kassa, T., & Belehu, K. (2011). Study on reproductive performance of Holstein-Friesian dairy cows at Alage Dairy Farm, Rift Valley of Ethiopia. *Tropical Animal Health and Production*, *43*(3). https://doi.org/10.1007/s11250-010-9734-8

Fernández, C., Soria, E., Sánchez-Seiquer, P., Gómez-Chova, L., Magdalena, R., Martín-Guerrero, J. D., Navarro, M. J., & Serrano, A. J. (2007). Weekly milk prediction on dairy goats using neural networks. *Neural Computing and Applications*, *16*(4–5), 373–381. https://doi.org/10.1007/s00521-006-0061-y

Ferraz, P. C., Malhado, C. H. M., Ramos, A. A., Carneiro, P. L. S., Carrillo, J. A., & Malhado, A. C. M. (2015). Population structure and genetic variability of a closed Jaffarabadi buffalo herd from Brazil. *Buffalo Bulletin*, *34*(2).

Flagel, L., Brandvain, Y., & Schrider, D. R. (2019). The unreasonable effectiveness of convolutional neural networks in population genetic inference. *Molecular Biology and Evolution*, *36*(2). https://doi.org/10.1093/molbev/msy224

Foxworthy, H. M., Enns, R. M., Thomas, M. G., & Speidel, S. E. (2019). The estimation of heritability and repeatability of first service conception and first cycle calving in Angus cattle. *Translational Animal Science*, *3*. https://doi.org/10.1093/tas/txz047

Frankham, R. (1995). Conservation Genetics. *Annual Review of Genetics*, *29*(1), 305–327. https://doi.org/10.1146/annurev.ge.29.120195.001513

Gaikwad-Inamdar, S., & Narayankhedkar, S. G. (2000). Estimation of heritability of reproduction traits in crossbred cows. *Indian Veterinary Journal*, *77*(9).

Gandhi, R. S., Raja, T. V., Ruhil, A. P., & Kumar, A. (2009). Prediction of lifetime milk production using artificial neural network in Sahiwal cattle. *Indian Journal of Animal Sciences*, *79*(10).

Gangwar, M., Kumar, S., Ahmad, S. F., Singh, A., Agrawal, S., Anitta, P. L., & Kumar, A. (2023). Identification of genetic variants affecting reproduction traits in Vrindavani cattle. *Mammalian Genome*. https://doi.org/10.1007/s00335-023-10023-2

Ghiasi, H., Pakdel, A., Nejati-Javaremi, A., Mehrabani-Yeganeh, H., Honarvar, M., González-Recio, O., Carabaño, M. J., & Alenda, R. (2011). Genetic variance components for female fertility in Iranian Holstein cows. *Livestock Science*, *139*(3). https://doi.org/10.1016/j.livsci.2011.01.020

Ghotbaldini, H., Mohammadabadi, M., Nezamabadi-Pour, H., Babenko, O. I., Bushtruk, M. V., & Tkachenko, S. V. (2019). Predicting breeding value of body weight at 6-month age using artificial neural networks in Kermani sheep breed. *Acta Scientiarum - Animal Sciences*, *41*(1). https://doi.org/10.4025/actascianimsci.v41i1.45282

Gianola, D., Okut, H., Weigel, K. A., & Rosa, G. J. (2011). Predicting complex quantitative traits with Bayesian neural networks: a case study with Jersey cows and wheat. *BMC Genetics*, *12*(1), 87. https://doi.org/10.1186/1471-2156-12-87

Gilmour, a R., Gogel, B. J., Cullis, B. R., & Thompson, R. (2009). ASReml user guide release 3.0. *VSN International Ltd.*

Gilmour, A. R., Thompson, R. and Cullis, B. R. (1995) Biometrics 51:1440

Gonçalves, D. N., Weber, V. A. de M., Pistori, J. G. B., Gomes, R. da C., de Araujo, A. V., Pereira, M. F., Gonçalves, W. N., & Pistori, H. (2021). Carcass image segmentation using CNN-based methods. *Information Processing in Agriculture*, *8*(4). https://doi.org/10.1016/j.inpa.2020.11.004

González-Camacho J. M., Ornella L., Pérez-Rodríguez P., Gianola D., Dreisigacker S., Crossa J. (2018). Applications of machine learning methods to genomic selection in breeding wheat for rust resistance. *Plant Genome 11*:170104. 10.3835/plantgenome2017.11.0104

González-Recio, O., López De Maturana, E., & Gutiérrez, J. P. (2007). Inbreeding depression on female fertility and calving ease in spanish dairy cattle. *Journal of Dairy Science*, *90*(12). https://doi.org/10.3168/jds.2007-0203

Gorgulu, O. (2012). Prediction of 305-day milk yield in Brown Swiss cattle using artificial neural networks. *South African Journal of Animal Science*, *42*(3). https://doi.org/10.4314/sajas.v42i3.10

Gowane, G. R., Prakash, V., Chopra, A., & Prince, L. L. L. (2013). Population structure and effect of inbreeding on lamb growth in Bharat Merino sheep. *Small Ruminant Research*, *114*(1), 72–79. https://doi.org/10.1016/j.smallrumres.2013.06.002

Goyache, F., del Coz, J. J., Quevedo, J. R., López, S., Alonso, J., Ranilla, J., Luaces, O., Alvarez, I., & Bahamonde, A. (2001). Using artificial intelligence to design and implement a morphological assessment system in beef cattle. *Animal Science*, *73*(1), 49–60. https://doi.org/10.1017/S1357729800058045

Grzesiak, W., B³aszczyk, P., & Lacroix, R. (2006). Methods of predicting milk yield in dairy cows—Predictive capabilities of Wood's lactation curve and artificial neural networks (ANNs). *Computers and Electronics in Agriculture*, *54*(2), 69–83. https://doi.org/10.1016/j.compag.2006.08.004

Grzesiak, W., Zaborski, D., Sablik, P., ¯ukiewicz, A., Dybus, A., & Szatkowska, I. (2010). Detection of cows with insemination problems using selected classification models. *Computers and Electronics in Agriculture*, *74*(2), 265–273. https://doi.org/10.1016/j.compag.2010.09.001

Gutiérrez, J. P., & Goyache, F. (2005). A note on ENDOG: a computer program for analysing pedigree information. *Journal of Animal Breeding and Genetics*, *122*(3), 172–176. https://doi.org/10.1111/j.1439-0388.2005.00512.x

Gutiérrez, J. P., Altarriba, J., Díaz, C., Quintanilla, R., Cañón, J., & Piedrafita, J. (2003). Pedigree analysis of eight Spanish beef cattle breeds. *Genetics Selection Evolution*, *35*(1). https://doi.org/10.1051/gse:2002035

Gutiérrez, J. P., Alvarez, I., Fernández, I., Royo, L. J., Díez, J., & Goyache, F. (2002). Genetic relationships between calving date, calving interval, age at first calving and type traits in beef cattle. *Livestock Production Science*, *78*(3). https://doi.org/10.1016/S0301-6226(02)00100-8

Gutiérrez, J. P., Cervantes, I., & Goyache, F. (2009). Improving the estimation of realized effective population sizes in farm animals. *Journal of Animal Breeding and Genetics*, *126*(4), 327–332. https://doi.org/10.1111/j.1439-0388.2009.00810.x

Gutiérrez, J. P., Cervantes, I., Molina, A., Valera, M., & Goyache, F. (2008). Individual increase in inbreeding allows estimating effective sizes from pedigrees. *Genetics Selection Evolution*, *40*(4), 359–378. https://doi.org/10.1051/gse:2008008

Haile, A., Joshi, B. K., Ayalew, W., Tegegne, A., & Singh, A. (2009). Genetic evaluation of Ethiopian Boran cattle and their crosses with Holstein Friesian in central Ethiopia: reproductive traits. *The Journal of Agricultural Science*, *147*(1), 81–89. https://doi.org/10.1017/S0021859608008095

Hamadani, A., Ganai, N. A., & Bashir, J. (2023). Artificial neural networks for data mining in animal sciences. *Bulletin of the National Research Centre*, *47*(1). https://doi.org/10.1186/s42269-023-01042-9

Hammoud, M. H., & Salem, M. M. I. (2013). The Genetic Evaluation of Some First Lactation Traits of Holstein Cows in Egypt. *J. Agric. Res*, *58*(1).

Hayes, B. J., Visscher, P. M., & Goddard, M. E. (2009). Increased accuracy of artificial selection by using the realized relationship matrix. *Genetics Research*, *91*(1), 47–60. https://doi.org/10.1017/S0016672308009981

Hermiz, H. N., Juma, K. H., Khalaf, S. S., & Aldoori, T. S. (2005). Genetic parameters of production, reproduction and growth traits of Holstein cows. *Dirasat. Agricultural Sciences*, *32*(2).

Huma, Z. E., and Iqbal, F. (2019). Predicting the body weight of Balochi sheep using a machine learning approach. *Turkish Journal of Veterinary and Animal Sciences*, *43*(4), 500–506. https://doi.org/10.3906/vet-1812-23

Ibrahim, A. K., Moumouni, I., & Mogueza, C. (2024). Genetic profile of milk production traits and analysis of correlations with reproductive performance in the Azawak Zebu in Niger. *Veterinary and Animal Science*, *25*, 100365. https://doi.org/10.1016/J.VAS.2024.100365

*ICAR-NBAGR registered 8 new livestock and poultry breeds | ICAR*. (2024). Retrieved December 31, 2024, from https://icar.org.in/icar-nbagr-registered-8-new-livestock-and-poultry-breeds

Iqbal, F., Waheed, A., Zil-E-Huma, & Faraz, A. (2022). Comparing the Predictive Ability of Machine Learning Methods in Predicting the Live Body Weight of Beetal Goats of Pakistan. *Pakistan Journal of Zoology*, *54*(1). https://doi.org/10.17582/JOURNAL.PJZ/20191003081007

Johnson, D. L. and Thompson, R. (1995) J. Dairy Sci., 78:449

Jospin, L. V., Laga, H., Boussaid, F., Buntine, W., & Bennamoun, M. (2022). Hands-On Bayesian Neural Networks - A Tutorial for Deep Learning Users. *IEEE Computational Intelligence Magazine*, *17*(2). https://doi.org/10.1109/MCI.2022.3155327

Kadarmideen, H. N., Thompson, R., & Simm, G. (2000). Linear and threshold model genetic parameters for disease, fertility and milk production in dairy cattle. *Animal Science*, *71*(3). https://doi.org/10.1017/S1357729800055338

Kamiti, D., Ilatsia, E., Bett, R., & Kahi, A. (2016). Population structure and demographic trends of the registered Sahiwal cattle in Kenya. *Tropical Animal Health and Production*, *48*(5). https://doi.org/10.1007/s11250-016-1055-0

Kaya, I., Akbaº, Y., & Uzmay, C. (2003). Estimation of breeding values for dairy cattle using test-day milk yields. *Turkish Journal of Veterinary and Animal Sciences*, *27*(2).

Khan, M. S., Rehman, Z. U. R., Khan, M. A., & Sohail, A. (2011). Genetic resources and diversity in Pakistani cattle. *Pakistan Veterinary Journal*, *28*(2).

Khattab, A. S., Atil, H., & Badawy, L. (2005). Variances of direct and maternal genetic effects for milk yield and age at first calving in a herd of Friesian cattle in Egypt. *Archives Animal Breeding*, *48*(1). https://doi.org/10.5194/aab-48-24-2005

Kliœ, P., Piwczyñski, D., Sawa, A., & Sitkowska, B. (2021). Prediction of Lactational Milk Yield of Cows Based on Data Recorded by AMS during the Periparturient Period. *Animals*, *11*(2), 383. https://doi.org/10.3390/ani11020383

Kufel, J., Bargie³-£¹czek, K., Kocot, S., KoŸlik, M., Bartnikowska, W., Janik, M., Czogalik, £., Dudek, P., Magiera, M., Lis, A., Paszkiewicz, I., Nawrat, Z., Cebula, M., & Gruszczyñska, K. (2023). What Is Machine Learning, Artificial Neural Networks and Deep Learning?—Examples of Practical Applications in Medicine. In *Diagnostics* (Vol. 13, Issue 15). https://doi.org/10.3390/diagnostics13152582

Kumar, A., & Mandal, A. (2021). Evaluation of animal models to explore the influence of maternal genetic and maternal permanent environment effect on reproductive performance of Jersey crossbred cattle. *Reproduction in Domestic Animals*, *56*(3). https://doi.org/10.1111/rda.13889

Kwon, K.-D., Lee, A., Lim, J., Cho, S., Lee, W., Cho, B.-K., & Seo, Y. (2020). Quality grading of Hanwoo (Korean native cattle breed) sub-images using convolutional neural network. *Korean Journal of Agricultural Science*, *47*(4).

Lacy, R. C. (1989). Analysis of founder representation in pedigrees: Founder equivalents and founder genome equivalents. *Zoo Biology*, *8*(2), 111–123. https://doi.org/10.1002/zoo.1430080203

Lacy, R. C. (1995). Clarification of genetic terms and their use in the management of captive populations. *Zoo Biology*, *14*(6), 565–577. https://doi.org/10.1002/zoo.1430140609

Laureano, M. M. M., Boligon, A. A., Costa, R. B., Forni, S., Severo, J. L. P., & Albuquerque, L. G. (2011). Estimativas de herdabilidade e tendênciasgenéticas para características de crescimento e reprodutivasembovinos da raçaNelore: Estimates of heritability and genetic trends for growth and reproduction traits in Nelore cattle. *Arquivo Brasileiro de Medicina Veterinária e Zootecnia*, *63*(1). https://doi.org/10.1590/s0102-09352011000100022

Lee, D.H., K.J. Han and B.H. Park (2003). Genetic relationship between milk yields, calving ease and days open at first lactation of Holstein cows in Korea. *Interbull Bull. 31*: 122–129.

Leem, S., Jeong, H. H., Lee, J., Wee, K., & Sohn, K. A. (2014). Fast detection of high-order epistatic interactions in genome-wide association studies using information theoretic measure. *Computational Biology and Chemistry*, *50*. https://doi.org/10.1016/j.compbiolchem.2014.01.005

Leven, S. (1996). The roots of backpropagation: From ordered derivatives to neural networks and political forecasting. *Neural Networks*, *9*(3). https://doi.org/10.1016/0893-6080(96)90015-5

Li, G., Huang, Y., Chen, Z., Chesser, G. D., Purswell, J. L., Linhoss, J., & Zhao, Y. (2021). Practices and applications of convolutional neural network-based computer vision systems in animal farming: A review. In *Sensors* (Vol. 21, Issue 4). https://doi.org/10.3390/s21041492

Li, Z., Lei, X., & Liu, S. (2022). A lightweight deep learning model for cattle face recognition. *Computers and Electronics in Agriculture*, *195*, 106848. https://doi.org/10.1016/j.compag.2022.106848

Lopez, B. I., Son, J. H., Seo, K., & Lim, D. (2019). Estimation of genetic parameters for reproductive traits in Hanwoo (Korean Cattle). *Animals*, *9*(10). https://doi.org/10.3390/ani9100715

Maignel, L., Boichard, D., & Verrier, E. (1996). Genetic Variability of French Dairy Breeds Estimated From Pedigree Information. *Interbull Bulletin*, *14*.

Makanjuola, B. O., Maltecca, C., Miglior, F., Schenkel, F. S., & Baes, C. F. (2020). Effect of recent and ancient inbreeding on production and fertility traits in Canadian Holsteins. *BMC Genomics*, *21*(1). https://doi.org/10.1186/s12864-020-07031

Malhado, C. H. M., Malhado, A. C. M., Carneiro, P. L. S., Ramos, A. A., Ambrosini, D. P., & Pala, A. (2012). Population structure and genetic variability in the Murrah dairy breed of water buffalo in Brazil accessed via pedigree analysis. *Tropical Animal Health and Production*, *44*(8). https://doi.org/10.1007/s11250-012-0153-x

Mandal, A., Singh, L. B., & Rout, P. K. (2000). The Muzaffarnagari sheep, a mutton breed in India. *Animal Genetic Resources Information*, *28*, 19–25. https://doi.org/10.1017/S1014233900001322

Martikainen, K., Sironen, A., & Uimari, P. (2018). Estimation of intrachromosomal inbreeding depression on female fertility using runs of homozygosity in Finnish Ayrshire cattle. *Journal of Dairy Science*, *101*(12). https://doi.org/10.3168/jds.2018-14805

Matthews, D., Pabiou, T., Evans, R. D., Beder, C., & Daly, A. (2022). Predicting carcass cut yields in cattle from digital images using artificial intelligence. *Meat Science*, *184*, 108671. https://doi.org/10.1016/j.meatsci.2021.108671

Mc Parland, S., Kearney, J. F., Rath, M., & Berry, D. P. (2007). Inbreeding trends and pedigree analysis of Irish dairy and beef cattle populations. *Journal of Animal Science*, *85*(2). https://doi.org/10.2527/jas.2006-367

Mcdowell, R. M. (2016). Genomic Selection with Deep Neural Networks. *Graduate Theses and Dissertations*.

McPhee, M. J., Walmsley, B. J., Skinner, B., Littler, B., Siddell, J. P., Cafe, L. M., Wilkins, J. F., Oddy, V. H., & Alempijevic, A. (2017). Live animal assessments of rump fat and muscle score in Angus cows and steers using 3-dimensional imaging1. *Journal of Animal Science*, *95*(4), 1847–1857. https://doi.org/10.2527/jas.2016.1292

Meuwissen, T. H. E., Hayes, B. J., & Goddard, M. E. (2001). Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps. *Genetics*, *157*(4), 1819–1829. https://doi.org/10.1093/genetics/157.4.1819

Meuwissen, T., & Luo, Z. (1992). Computing inbreeding coefficients in large populations. *Genetics Selection Evolution*, *24*(4), 305. https://doi.org/10.1186/1297-9686-24-4-305

Meyer, K. (2007). WOMBAT: a tool for mixed model analyses in quantitative genetics by restricted maximum likelihood (REML). *Journal of Zhejiang University. Science. B.*, *8*(11). https://doi.org/10.1631/jzus.2007.B0815

Meyer, K., 1991. DFREML: programs to estimate variance components by restricted maximum likelihood using a derivative-free algorithm. User notes, Ver. 2.0

Mieth, B., Kloft, M., Rodríguez, J. A., Sonnenburg, S., Vobruba, R., Morcillo-Suárez, C., Farré, X., Marigorta, U. M., Fehr, E., Dickhaus, T., Blanchard, G., Schunk, D., Navarro, A., & Müller, K.-R. (2016). Combining Multiple Hypothesis Testing with Machine Learning Increases the Statistical Power of Genome-wide Association Studies. *Scientific Reports*, *6*(1), 36671. https://doi.org/10.1038/srep36671

Miraei-Ashtiani, S. R., Seyedalian, S. A. R., & Moradi Shahrbabak, M. (2007). Variance components and heritabilities for body weight traits in Sangsari sheep, using univariate and multivariate animal models. *Small Ruminant Research*, *73*(1–3). https://doi.org/10.1016/j.smallrumres.2006.12.002

Misztal, I., Lawlor, T. J., & Short, T. H. (1993). Implementation of Single- and Multiple-Trait Animal Models for Genetic Evaluation of Holstein Type Traits. *Journal of Dairy Science*, *76*(5). https://doi.org/10.3168/jds.S0022-0302(93)77473-1

Moawed, S., & El-Bayomi, K. (2015). A Comparison of Linear Models for Estimating Co-Variance Components and Genetic Parameters in Holstein Dairy Cattle. *Journal of Animal, Poultry & Fish Production*, *4*(1). https://doi.org/10.21608/japfp.2015.7424

Mohammadi, H., Shahrebabak, M. M., Shahrebabak, H. M., Bahrami, A., & Dorostkar, M. (2013). Model comparisons and genetic parameter estimates of growth and the Kleiber ratio in Shal sheep. *Archives Animal Breeding*, *56*(1). https://doi.org/10.7482/0003-9438-56-026

Mohammadi, K., Beigi Nassiri, M. T., Rahmatnejad, E., Sheikh, M., Fayazi, J., & Karimi Manesh, A. (2013). Phenotypic and genetic parameter estimates for reproductive traits in Zandi sheep. *Tropical Animal Health and Production*, *45*(2), 671–677. https://doi.org/10.1007/s11250-012-0276-0

Mondal, S. K., Kumar, A., Dubey, P. P., Sivamani, B., & Dutt, T. (2014). Estimation of variance and genetic parameters for pre-weaning weights of individual Landrace

X Desi synthetic piglets. *Journal of Applied Animal Research*, *42*(3). https://doi.org/10.1080/09712119.2013.875901

Montesinos-López, O.A., Montesinos-López, A., Pérez-Rodríguez, P. *et al.* (2021). A review of deep learning applications for genomic selection. *BMC Genomics* **22**, 19. https://doi.org/10.1186/s12864-020-07319-x

Mostafa, M., Bader, E., & khattab, A. (2013). Variances of Direct and Maternal Genetic Effects for Milk Yield and Its Composition in a Herd of Friesian Cows in Egypt. *Journal of Animal and Poultry Production*, *4*(8). https://doi.org/10.21608/jappmu.2013.71512

Muasya, T. K., Kariuki, J. N., & Muia, J. M. K. (2011). Population structure of the Sahiwal breed in Kenya. *Livestock Research for Rural Development*, *23*(9).

Mueller, M. L., & Van Eenennaam, A. L. (2022). Synergistic power of genomic selection, assisted reproductive technologies, and gene editing to drive genetic improvement of cattle. In *CABI Agriculture and Bioscience* (Vol. 3, Issue 1). BioMed Central Ltd. https://doi.org/10.1186/s43170-022-00080-z

Mumtaz, S., Mukherjee, A., Pathak, P., & Parveen, K. (2021). In-depth Pedigree Analysis and Inbreeding Effect on Reproduction Traits in Sahiwal Nucleus Herd . *Indian Journal of Animal Research*, *Of*. https://doi.org/10.18805/ijar.b-4431

Musingi, M., Muasya, T. K., Ilatsia, E. D., & Kahi, A. K. (2018). Effect of inbreeding on traits of economic importance in Kenyan Sahiwal cattle. *Livestock Research for Rural Development*, *30*(1).

Nelson, R. M., Pettersson, M. E., & Carlborg, Ö. (2013). A century after Fisher: time for a new paradigm in quantitative genetics. *Trends in Genetics*, *29*(12), 669–676. https://doi.org/10.1016/j.tig.2013.09.006

Neser, F. W. C., Erasmus, G. J., & Van Wyk, J. B. (2000). Genetic parameter estimates for pre-weaning growth traits in Dorper sheep. *South African Journal of Animal Science*, *30*(4). https://doi.org/10.4314/sajas.v30i4.3928

Neto, M. R. T., da Cruz, J. F., Ramos, A. A., Carneiro, P. L. S., Azevêdo, D. M. M. R., Bozzi, R., & Malhado, C. H. M. (2012). Genetic variability in Mediterranean buffalos evaluated by pedigree analysis. *Ciencia Rural*, *42*(11). https://doi.org/10.1590/S0103-84782012001100021

Nusrat, I., & Jang, S. B. (2018). A comparison of regularization techniques in deep neural networks. *Symmetry*, *10*(11). https://doi.org/10.3390/sym10110648

Ojango, J. M. K., & Pollott, G. E. (2001). Genetics of milk yield and fertility traits in Holstein-Friesian cattle on large-scale Kenyan farms. *Journal of Animal Science*, *79*(7). https://doi.org/10.2527/2001.7971742x

Olori, V. E., Meuwissen, T. H. E., & Veerkamp, R. F. (2002). Calving interval and survival breeding values as measure of cow fertility in a pasture-based production system with seasonal calving. *Journal of Dairy Science*, *85*(3). https://doi.org/10.3168/jds.S0022-0302(02)74125-8

Oyama, K., Katsuta, T., Anada, K., & Mukai, F. (2002). Heritability and repeatability estimates for reproductive traits of Japanese Black cows. *Asian-Australasian Journal of Animal Sciences*, *15*(12). https://doi.org/10.5713/ajas.2002.1680

Pablo Gutiérrez, J., Goyache, F., & Cervantes, I. (2010). *User's Guide A Computer Program for Monitoring Genetic Variability of Populations Using Pedigree Information*.

Panda, S., Gaur, G. K., Sahoo, N. R., Ahmad, S. F., Devamma Ramachandrappa, P., & Bharti, P. K. (2020). Evaluation of animal models for genetic analysis of growth performance in Landlly pigs. *Turkish Journal of Veterinary and Animal Sciences*, *44*(4). https://doi.org/10.3906/vet-1908-33

Paré, G., Mao, S., & Deng, W. Q. (2017). A machine-learning heuristic to improve gene score prediction of polygenic traits. *Scientific Reports*, *7*(1), 12665. https://doi.org/10.1038/s41598-017-13056-1

Park, C.-H., Koo, Y.-M., Kim, B.-W., Sun, D.-W., Kim, J.-I., Song, C.-E., Lee, K.-H., Lee, J.-Y., Jeoung, Y.-H., & Lee, J.-G. (2012). Repeated Records Animal Model to Estimate Genetic Parameters of Ultrasound Measurement Traits in Hanwoo Cows. *Journal of Animal Science and Technology*, *54*(2). https://doi.org/10.5187/jast.2012.54.2.71

Parveen, K. (2016). Performance trends of economic traits in Sahiwal cattle. Ph.D. Thesis. ICAR-National Dairy Research Institute, Deemed University, Karnal, India.

Pastur-Romay, L., B. Porto-Pazos, A., Cedron, F., & Pazos, A. (2016). Parallel Computing for Brain Simulation. *Current Topics in Medicinal Chemistry*, *17*(14). https://doi.org/10.2174/1568026617666161104105725

Pedrosa, V. B., Santana, M. L., Oliveira, P. S., Eler, J. P., & Ferraz, J. B. S. (2010). Population structure and inbreeding effects on growth traits of Santa Inês sheep in Brazil. *Small Ruminant Research*, *93*(2–3), 135–139. https://doi.org/10.1016/j.smallrumres.2010.05.012

Peixoto, M. G. C. D., Poggian, C. F., Verneque, R. S., Egito, A. A., Carvalho, M. R. S., Penna, V. M., Bergmann, J. A. G., Viccini, L. F., & Machado, M. A. (2010). Genetic basis and inbreeding in the Brazilian Guzerat (Bos indicus) subpopulation selected for milk production. *Livestock Science*, *131*(2–3). https://doi.org/10.1016/j.livsci.2010.03.015

Pérez-Rodríguez P, Flores-Galarza S, Vaquera-Huerta H, Montesinos-López OA, del Valle-Paniagua DH, Crossa J (2020). Genome-based prediction of Bayesian linear and non-linear regression models for ordinal data. *Plant Genome*.13(1):e20021.

Peters, S. O., Kizilkaya, K., Garrick, D. J., Fernando, R. L., Reecy, J. M., Weaber, R. L., Silver, G. A., & Thomas, M. G. (2013). Heritability and bayesian genome-wide association study of first service conception and pregnancy in Brangus heifers. *Journal of Animal Science*, *91*(2). https://doi.org/10.2527/jas.2012-5580

Pinna, D., Sara, G., Todde, G., Atzori, A. S., Artizzu, V., Spano, L. D., & Caria, M. (2023). Advancements in combining electronic animal identification and augmented reality technologies in digital livestock farming. *Scientific Reports*, *13*(1). https://doi.org/10.1038/s41598-023-45772-2

Poyato-Bonilla, J., Perdomo-González, D. I., Sánchez-Guerrero, M. J., Varona, L., Molina, A., Casellas, J., & Valera, M. (2020). Genetic inbreeding depression load for morphological traits and defects in the Pura Raza Española horse. *Genetics Selection Evolution*, *52*(1). https://doi.org/10.1186/s12711-020-00582-2

Prevolnik, M., Skorjanc, D., Candek-Potokar, M., & Novic, M. (2011). Application of Artificial Neural Networks in Meat Production and Technology. In *Artificial Neural Networks - Industrial and Control Engineering Applications*. InTech. https://doi.org/10.5772/15077

Radwan, H. A., A. ABO Elf, E., & El-Bayoumi, K. M. (2017). Different Single-trait Animal Models for Estimating Direct-maternal Covariance Components in Holstein Cows. *Asian Journal of Animal and Veterinary Advances*, *12*(2). https://doi.org/10.3923/ajava.2017.80.87

Rahbar, R., Aminafshar, M., Abdullahpour, R., & Chamani, M. (2016). <b>Genetic analysis of fertility traits of Holstein dairy cattle in warm and temperate climate. *Acta Scientiarum. Animal Sciences*, *38*(3). https://doi.org/10.4025/actascianimsci.v38i3.31377

Raj, M. R., & Sreeja, A. (2018). Analysis of Computational Gene Prioritization Approaches. *Procedia Computer Science*, *143*, 395–410. https://doi.org/10.1016/j.procs.2018.10.411

Rege, J. E. O., & Wakhungu, J. W. (1992). An evaluation of a long term breeding programme in a closed Sahiwal herd in Kenya. *Journal of Animal Breeding and Genetics*, *109*(1–6), 374–384. https://doi.org/10.1111/j.1439-0388.1992.tb00417.x

Rehman, Z. U., Sajjad Khan, M. S., Bhatti, S. A., Iqbal, J., & Iqbal, A. (2008). Factors affecting first lactation performance of Sahiwal cattle in Pakistan. *Archives Animal Breeding*, *51*(4). https://doi.org/10.5194/aab-51-305-2008

Reis Filho, J. C., Lopes, P. S., Verneque, R. da S., Torres, R. de A., Teodoro, R. L., & Carneiro, P. L. S. (2010). Population structure of Brazilian Gyr dairy cattle. *Revista Brasileira de Zootecnia*, *39*(12), 2640–2645. https://doi.org/10.1590/S1516-35982010001200012

Riesel, D. (2007). A brief Introduction on Neural Networks. *Springer-Verlag, Berlin*.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, *323*(6088). https://doi.org/10.1038/323533a0

Safari, E., Fogarty, N. M., Gilmour, A. R., Atkins, K. D., Mortimer, S. I., Swan, A. A., Brien, F. D., Greeff, J. C., & van der Werf, J. H. J. (2007). Across population genetic parameters for wool, growth, and reproduction traits in Australian Merino sheep. 2. Estimates of heritability and variance components. *Australian Journal of Agricultural Research*, *58*(2), 177. https://doi.org/10.1071/AR06162

ªahin, A., Ulutaº, Z., Yýlmaz Adkinson, A., & W. Adkinson, R. (2014). Genetic parameters of first lactation milk yield and fertility traits in brown Swiss cattle. *Annals of Animal Science*, *14*(3). https://doi.org/10.2478/aoas-2014-0038

Sakthivel, M., Tamilmani, G., Abdul Nazar, A. K., Jayakumar, R., Sankar, M., Rameshkumar, P., Anikuttan, K. K., Samal, A. K., Anbarasu, M., & Gopakumar, G. (2019). Genetic variability of a small captive population of the cobia (Rachycentron canadum) through pedigree analyses. *Aquaculture*, *498*, 435–443. https://doi.org/10.1016/j.aquaculture.2018.08.047

Santana, M. L., Aspilcueta-Borquis, R. R., Bignardi, A. B., Albuquerque, L. G., & Tonhati, H. (2011). Population structure and effects of inbreeding on milk yield and quality of Murrah buffaloes. *Journal of Dairy Science*, *94*(10). https://doi.org/10.3168/jds.2011-4377

Santana, M. L., Pereira, R. J., Bignardi, A. B., Ayres, D. R., Menezes, G. R. O., Silva, L. O. C., Leroy, G., Machado, C. H. C., Josahkian, L. A., & Albuquerque, L. G. (2016). Structure and genetic diversity of Brazilian Zebu cattle breeds assessed by pedigree analysis. *Livestock Science*, *187*. https://doi.org/10.1016/j.livsci.2016.02.002

Santana, M. L., Pereira, R. J., Bignardi, A. B., El Faro, L., Tonhati, H., & Albuquerque, L. G. (2014). History, structure, and genetic diversity of Brazilian Gir cattle. *Livestock Science*, *163*(1). https://doi.org/10.1016/j.livsci.2014.02.007

Sarker, I. H. (2021). Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions. In *SN Computer Science* (Vol. 2, Issue 6). Springer. https://doi.org/10.1007/s42979-021-00815-1

Schaeffer, L. R. (2019). Competition model for international comparisons of livestock. *Journal of Animal Breeding and Genetics*, *136*(6). https://doi.org/10.1111/jbg.12440

Schoeman, S. J., & Jordaan, G. F. (1999). Multitrait estimation of direct and maternal (co)variances for growth and efficiency traits in a multibreed beef cattle herd. *South African Journal of Animal Science*, *29*(3). https://doi.org/10.4314/sajas.v29i3.44205

Schrider, D. R., & Kern, A. D. (2018). Supervised Machine Learning for Population Genetics: A New Paradigm. *Trends in Genetics*, *34*(4), 301–312. https://doi.org/10.1016/j.tig.2017.12.005

Schrider, D. R., Mendes, F. K., Hahn, M. W., & Kern, A. D. (2015). Soft Shoulders Ahead: Spurious Signatures of Soft and Partial Selective Sweeps Result from Linked Hard Sweeps. *Genetics*, *200*(1), 267–284. https://doi.org/10.1534/genetics.115.174912

Schutz, M. M., Freeman, A. E., Beitz, D. C., & Mayfield, J. E. (1992). The Importance of Maternal Lineage on Milk Yield Traits of Dairy Cattle. *Journal of Dairy Science*, *75*(5). https://doi.org/10.3168/jds.S0022-0302(92)77884-9

Sebastian, A., Pendurthi, R., Kozhakhmetov, A., Trainor, N., Robinson, J. A., Redwing, J. M., & Das, S. (2022). Two-dimensional materials-based probabilistic synapses and reconfigurable neurons for measuring inference uncertainty using Bayesian neural networks. *Nature Communications*, *13*(1). https://doi.org/10.1038/s41467-022-33699-7

Seyedsharifi, R., Eskandari-Nasab, M. P., & Sobhani, A. (2008). Estimation of genetic parameters and breeding values for test-day and 305-days milk yields in some Iranian Holstein herd. *Journal of Animal and Veterinary Advances*, *7*(11).

Shalaby, N. A., Moawed, S. A., & El-Bayomi, K. M. (2015). A Comparison of Linear Models for Estimating Co-Variance Components and Genetic Parameters in Holstein Dairy Cattle. *Journal of Animal, Poultry & Fish Production; Suez Canal University, 4(1), 7-15. https://japfp.journals.ekb.eg/article_7424_bb79633bad39afc9ab027fc987d1c832.pdf*

Sharma, S., Dhaka, S. S., & Patil, C. S. (2024). Estimation of additive and maternal covariance of production traits in Murrah buffalo. *Journal of Animal Breeding and Genetics*, *141*(4). https://doi.org/10.1111/jbg.12852

Sheehan, S., & Song, Y. S. (2016). Deep Learning for Population Genetic Inference. *PLoS Computational Biology*, *12*(3). https://doi.org/10.1371/journal.pcbi.1004845

Shi, J., Zhu, Y., Khan, F., & Chen, G. (2019). Application of Bayesian Regularization Artificial Neural Network in explosion risk analysis of fixed offshore platform. *Journal of Loss Prevention in the Process Industries*, *57*. https://doi.org/10.1016/j.jlp.2018.10.009

Shine, J. M., Li, M., Koyejo, O., Fulcher, B., & Lizier, J. T. (2021). Nonlinear reconfiguration of network edges, topology and information content during an artificial learning task. *Brain Informatics*, *8*(1), 26. https://doi.org/10.1186/s40708-021-00147-z

Simonsen, K. L., Churchill, G. A., & Aquadro, C. F. (1995). Properties of statistical tests of neutrality for DNA polymorphism data. *Genetics*, *141*(1), 413–429. https://doi.org/10.1093/genetics/141.1.413

Singh, A., Kumar, A., Mehrotra, A., A, K., Pandey, A. K., Mishra, B. P., & Dutt, T. (2021). Estimation of linkage disequilibrium levels and allele frequency distribution in crossbred Vrindavani cattle using 50K SNP data. *PloS One*, *16*(11), e0259572. https://doi.org/10.1371/journal.pone.0259572

Singh, C. V., & Barwal, R. S. (2012). Use of Different Animal Models in Prediction of Genetic Parameters of First Lactation and Herd Life Traits of Murrah Buffaloes. *Indian Journal of Dairy Science*, *65*(5). https://epubs.icar.org.in/index.php/IJDS/article/view/25298

Singh, N. P., Dutt, T., Usman, S. M., Baqir, M., Tiwari, R., & Kumar, A. (2022). Prediction of first lactation 305 days milk yield using artificial neural network in Murrah buffalo. *The Indian Journal of Animal Sciences*, *92*(9). https://doi.org/10.56093/ijans.v92i9.117570

Singh, N., Usman, S., Maurya, V., Dutt, T., bhatt, N., & Kumar, A. (2020). Comparative analysis of artificial neural network algorithms for prediction of FL305DMY in Murrah buffalo. *International Journal of Livestock Research*, *0*, 1. https://doi.org/10.5455/ijlr.20200704062936

Singh, R. R., Dutt, T., Kumar, A., Tomar, A. K. S., & Singh, M. (2011). On-farm characterization of Vrindavani cattle in India. *Indian Journal of Animal Sciences*, *81*(3).

Solemani Baghshah, S., Ansari Mahyari, S., Edriss, M. A., & Nanaei, H. A. (2014). Estimation of genetic and phonotypic trends for age at first calving, calving interval, days open and number of insemina tion to conception for Isfahan holstein cows. *International Journal of Advanced Biological and Biomedical Research, 2(5),* 1307–1314.

Song, S., Liu, T., Wang, H., Hasi, B., Yuan, C., Gao, F., & Shi, H. (2022). Using Pruning-Based YOLOv3 Deep Learning Algorithm for Accurate Detection of Sheep Face. *Animals*, *12*(11), 1465. https://doi.org/10.3390/ani12111465

Špehar, M., Ramljak, J., & Kasap, A. (2022). Estimation of genetic parameters and the effect of inbreeding on dairy traits in Istrian sheep. *Italian Journal of Animal Science*, *21*(1), 331–342. https://doi.org/10.1080/1828051X.2022.2031320

Suhail, S. M., Ahmad, I., Hafeez, A., Ahmed, S., Jan, D., Khan, S., & Rehman, A. (2010). Genetic study of some reproductive of jersey cattle under subtropical conditions. *Sarhad Journal of Agriculture, 26(1),* 87–91.

Sun, C., Madsen, P., Nielsen, U. S., Zhang, Y., Lund, M. S., & Su, G. (2009). Comparison between a sire model and an animal model for genetic evaluation of fertility traits in Danish Holstein population. *Journal of Dairy Science*, *92*(8). https://doi.org/10.3168/jds.2008-1918

Szanda³a, T. (2021). Review and comparison of commonly used activation functions for deep neural networks. In *Studies in Computational Intelligence* (Vol. 903). https://doi.org/10.1007/978-981-15-5495-7_11

Taye, M. M. (2023). Understanding of Machine Learning with Deep Learning: Architectures, Workflow, Applications and Future Directions. *Computers*, *12*(5), 91. https://doi.org/10.3390/computers12050091

Tealab, A. (2018). Time series forecasting using artificial neural networks methodologies: A systematic review. *Future Computing and Informatics Journal*, *3*(2). https://doi.org/10.1016/j.fcij.2018.10.003

Tesfa, A., & Garikipati, D. K. (2014). Genetic and non-genetic parameter estimates of dairy cattle in Ethiopia: A review. *Online Journal of Animal and Feed Research*, *4*(3).

Tomiyama, M., Kanetani, T., Tatsukawa, Y., Mori, H., & Oikawa, T. (2010). Genetic parameters for preweaning and early growth traits in Berkshire pigs when creep feeding is used. *Journal of Animal Science*, *88*(3). https://doi.org/10.2527/jas.2009-2072

Torada, L., Lorenzon, L., Beddis, A., Isildak, U., Pattini, L., Mathieson, S., & Fumagalli, M. (2019). ImaGene: a convolutional neural network to quantify natural selection from genomic data. *BMC Bioinformatics*, *20*(S9), 337. https://doi.org/10.1186/s12859-019-2927-x

Tusell, L., Pérez-Rodríguez, P., Forni, S., Wu, X.-L., & Gianola, D. (2013). Genome-enabled methods for predicting litter size in pigs: a comparison. *Animal*, *7*(11), 1739–1749. https://doi.org/10.1017/S1751731113001389

UlhôaMagnabosco, C., Brito Lopes, F., de Magalhaes Rosa, G. J., & Sainz, R. D. (2016). Bayesian estimates of genetic parameters for reproductive traits in Nellore cows raised on pasture in tropical regions. *Revista Colombiana de CienciasPecuarias*, *29*(2). https://doi.org/10.17533/udea.rccp.v29n2a05

Ulutas, Z., & Sezer, M. (2009). Genetic Study of Milk Production and Reproduction Traits of Local Born Simmental Cattle in Turkey Yerli Simmental Sýðýrlarýnýn Süt ve Döl Verim Özelliklerine ait Genetik. *GOU. Ziraat Fakültesi Dergisi*, *26*(1).

van Bergen, G. H. H., Duenk, P., Albers, C. A., Bijma, P., Calus, M. P. L., Wientjes, Y. C. J., & Kappen, H. J. (2020). Bayesian neural networks with variable selection for prediction of genotypic values. *Genetics, Selection, Evolution/ : GSE*, *52*(1). https://doi.org/10.1186/s12711-020-00544-8

van der Werf, J. (2013). *Genomic Selection in Animal Breeding Programs* (pp. 543–561). https://doi.org/10.1007/978-1-62703-447-0_26

Vitsios, D., & Petrovski, S. (2019). Stochastic semi-supervised learning to prioritise genes from high-throughput genomic screens. *BioRxiv*, *May*.

Vozzi, P. A., Marcondes, C. R., Magnabosco, C. de U., Bezerra, L. A. F., & Lôbo, R. B. (2006). Structure and genetic variability in Nellore (Bos indicus) cattle by pedigree analysis. *Genetics and Molecular Biology*, *29*(3), 482–485. https://doi.org/10.1590/S1415-47572006000300014

Wakchaure, R., Ganguly, S., Para, P. A., Praveen, P. K., Kumar, A., & Sharma, S. (2008). Development of Crossbred Cattle in India: A Review. *International Journal of Emerging Technology and Advanced Engineering*, *5*(10).

Widyaningrum, R., Budisatria, I. G. S., Mukhtar, & Maharani, D. (2021). Estimation output of Aceh cattle in livestock breeding and forage center of Indrapuri. *IOP Conference Series: Earth and Environmental Science*, *788*(1). https://doi.org/10.1088/1755-1315/788/1/012003

Wright, S. (1922). Coefficients of Inbreeding and Relationship. *The American Naturalist*, *56*(645). https://doi.org/10.1086/279872

Zamani, P., & Mohammadi, H. (2008). Comparison of different models for estimation of genetic parameters of early growth traits in the Mehraban sheep. *Journal of Animal Breeding and Genetics*, *125*(1). https://doi.org/10.1111/j.1439-0388.2007.00676.x

Zambrano, J. C., & Echeverri, J. (2014). Genetic and environmental variance and covariance parameters for some reproductive traits of Holstein and Jersey cattle in Antioquia (Colombia). *Revista Brasileira de Zootecnia*, *43*(3). https://doi.org/10.1590/S1516-35982014000300005

Zamorano-Algandar, R., Medrano, J. F., Thomas, M. G., Enns, R. M., Speidel, S. E., Sánchez-Castro, M. A., Luna-Nevárez, G., Leyva-Corona, J. C., & Luna-Nevárez, P. (2023). Genetic Markers Associated with Milk Production and Thermotolerance in Holstein Dairy Cows Managed in a Heat-Stressed Environment. *Biology*, *12*(5). https://doi.org/10.3390/biology12050679

Zeleke, B., Kebede, K., & Banerjee, A. K. (2016). Estimation of Genetic Parameters for Growth Traits of Fogera and Holstein Friesian Crossbred Cattle at Metekel Ranch, Amhara Region, Ethiopia. *Global Veterinaria*, *17*(3).

Zhao, T., Fernando, R., & Cheng, H. (2021). Interpretable artificial neural networks incorporating Bayesian alphabet models for genome-wide prediction and association studies. *G3: Genes, Genomes, Genetics*, *11*(10). https://doi.org/10.1093/g3journal/jkab228

Zulkadir, U., Aytekin, I., & Pala, A. (2009). Genetic analyses for milk yield, lactation period and fat percentage in Brown Swiss cattle. *Journal of Animal and Veterinary Advances*, *8*(5). https://doi.org/10.3923/javaa.2009.857.862

* * *

# VITAE

| | | |
|---|---|---|
| **Name** | : | **Dr. Munish Gangwar** |
| Father's Name | : | Shri Dinesh Kumar Gangwar |
| Mother's Name | : | Smt. Sushila Devi |
| Date of Birth | : | 12.03.1995 |
| Permanent Address | : | Vill. & Post : Ruppur Kamalu, Philibhit |
| Mobile No. | : | 9457747743 |
| Email ID | : | gangwarmunish72@gmail.com |

## Educational Qualifications

| Degree | Board/university | Year of passing | OGPA |
|---|---|---|---|
| B.V.Sc. & A.H. | SDAU, Dantiwada, Gujarat | 2019 | 6.99 |
| M.V.Sc. | IVRI, Izatnagar, U.P. | 2021 | 8.25 |
| PhD | IVRI, Izatnagar, U.P. | 2025 | 7.79 |

## Awards and Honours

- Recipient of National Talent Scholarship **(NTS)** during B.V.Sc. & A.H. programme
- Recipient of Institute Fellowship **(JRF)** for MVSc degree
- Recipient of CSIR- Senior Research Fellowship **(SRF)** for PhD degree
- ICAR NET (2021 & 2023)
- CSIR NET (2023)

## Life Membership of Professional Societies

- Veterinary Council of India
- Uttar Pradesh Veterinary Council