

MATHEMATICAL MODELING FOR GENOMIC SELECTION IN SERBIAN DAIRY CATTLE

Radmila BESKOROVAJNI¹, Rade JOVANOVIĆ¹, Lato PEZO², Nikola POPOVIĆ¹, Nataša TOLIMIR¹, Ljubiša MIHAJLOVIĆ³, Gordana ŠURLAN-MOMIROVIĆ⁴

¹ Institute for Science Application in Agriculture, Belgrade, Serbia

² Institute for General and Physical Chemistry, University of Belgrade, Belgrade, Serbia

³ Ministry of Agriculture, Forestry and Water Management, Belgrade, Serbia

⁴ University of Belgrade, Faculty of Agriculture, Belgrade, Serbia

Beskorovajni R., R. Jovanović, L. Pezo, N. Popović, N. Tolimir, Lj. Mihajlović, G. Šurlan-Momirović (2021). *Mathematical modeling for genomic selection in serbian dairy cattle*. - Genetika, Vol 53, No.3, 1105-1115.

This manuscript has come as a result of an efficient breeding program in Serbian cattle populations for some economically important traits. Genomic selection in the last two decades has been the main challenge in animal breeding programs and genetics. Many SNP markers are used in statistical analysis in predicting the accuracy of breeding values for young animals without their performance. The new breeding tendency in the selection of young animals allows their genetic progress with reducing cost. In this study, 92 Holstein cows from various regions in Serbia were analyzed based on SNP molecular markers. Within this investigation, an empirical model was developed for the prediction of Yield Traits and Fertility Traits variables, according to Key traits data for dairy cattle. The developed model gave a reasonable fit to the data and successfully predicted Yield Traits (such as Fat and Protein Percent, Cheese Merit, Fluid Merit, and Cow Livability) and Fertility Traits variables (such as Sire Calving Ease, Heifer Conception Rate, Cow Conception Rate, Daughter Stillbirth, Sire Stillbirth, and Gestation Length). A total of 92 dairy cattle data were used to build a prediction model for the prediction of Yield Traits and Fertility Traits variables. The artificial neural network model, based on the Broyden-Fletcher-Goldfarb-Shanno iterative algorithm, showed good prediction capabilities (the r^2 values during the training cycle for the before mentioned output variables were in the range between 0.444 and 0.989).

Corresponding author: Gordana Šurlan-Momirović, Faculty of Agriculture, University of Belgrade, Belgrade, Serbia, Tel: 064 31 70 511. E-mail: surlang@gmail.com

Keywords: Mathematical modeling, genetic evaluation, single-nucleotide polymorphism, SNP, haplotype, genotyping.

INTRODUCTION

Genomic selection is the latest method of selection in cattle breeding. It involves using the cattle genome (DNA) for an earlier and better description of its breeding value which refers to the value of an animal in a breeding program for a particular trait (VEERKAMP and CALUS, 2009). Selection based on estimated breeding values (BV), calculated on the basis of phenotypic performance and pedigree data, was very successful. With the development of genomic tools, such as single nucleotide polymorphism (SNP) chips, they have led to a new method of selection - genomic selection (GODDARD *et al.*, 2010).

Due to genomic selection, the selection of breeding candidates is increasingly based on genomic breeding value (GBV), rather than on estimated BV obtained from progeny testing. To determine the most accurate genomic value of a young animal, it is necessary to compare it with the reference population, which is a group of individuals with safe breeding values, based on progeny test data (conventional breeding value) and the examined DNA profile. When testing bulls for offspring, a long generation interval and preferential treatment of bull mothers limit genetic progress. Genomic selection eliminates these limitations because the breeding value of individuals of both sexes is determined at the earliest age by direct genome analysis. Data on the quality of male breeding heads obtained by progeny testing are available at the age of about five, while genomic selection significantly shortens this process (PRKA, 2017).

The accuracy of genomic predictions depends on characteristics of the reference populations, such as the number of animals, number of markers, the heritability of the recorded phenotype, and the extent of relationships between selection candidates and the reference population (CALUS, 2009; SCHEFER and WEIGEL, 2012).

The advantages of genomic selection are accurate identification of individuals and parental pairs, the possibility of making significantly improved insemination plans, and plans for genetic improvement of the herd. Inbreeding and the occurrence of recessive genes are significantly reduced or eliminated, genomic information of breeding values is available at a much earlier age of the throat compared to conventional selection (BOUQUET and JUGA, 2013; IBÁÑEZ-ESCRICHE *et al.*, 2011; MEUWISSEN *et al.*, 2016; WIGGANS *et al.*, 2017).

The use of genomic selection can achieve better results for traits with low heritability, which can hardly be improved by the use of conventional (phenotypic) selection, as well as for traits whose phenotype is difficult to measure (longevity, disease resistance), or measurement is not feasible in candidates for selection (JOVANOVAČ, 2013). A significant advantage of genomic selection is the potential to estimate GBV with high accuracy for several generations without re-phenotyping, resulting in lower costs and shorter generation intervals (KEGALJ, 2015).

The objective of this report was to study the possibility of predicting the 5 Yield Traits (Fat %, Protein %, Cheese Merit, Fluid Merit, Cow Livability – LIV) and 6 Fertility Traits prediction variables (Sire Calving Ease - SCE, Heifer Conception Rate - HCR, Cow Conception Rate - CCR, Daughter Stillbirth, Sire Stillbirth, and Gestation Length - GL), according to 9 Key Traits (Milk Yield, Fat Yield – Fat (lbs), Protein Yield – Protein (lbs), Somatic Cell Score - SCS, Productive Life - PL, Daughter Pregnancy Rate - DPR, Daughter Calving Ease - DCE, Final

Type - PTA Type, Genomic Future Inbreeding – GFI) to Serbian conditions and determine the potential benefits of Artificial neural network (ANN) for genomic selection in Serbian dairy cattle. Artificial neural network models were used for mathematical modeling and determining the potential benefits of genomic selection in Serbian dairy cattle.

MATERIAL AND METHOD

In our study, hair samples were taken from the tail of Holstein heifers and sent to a Laboratory in Scotland (Neogen Gene Seek www.neogen.com dairygenomics@neogen.com). In August and September of 2019, 92 analyzes of the genomic throat of the Holstein were done. Data were analyzed by the CDCB (Council of Dairy Breeding, USA). Genomic analyses were compared with the reference population of USA Holstein cattle. It enabled the identification of parental pairs (fathers) from the tested animals.

Key Traits of our data of reports for the Igenity Dairy Heifer Programme contains the Key Traits that are most often used for evaluation of animals such as Milk Yield (Milk), Fat Yield (Fat lbs), Protein Yield (Protein lbs), Somatic Cell Score (SCS), Productive Life (PL), Daughter Pregnancy Rate (DPR), Daughter Calvin Ease (DCE), Final Type (PTA Type) and Genomic Future Inbreeding (GFI).

In this work, we analyzed various Yield Traits, such as Milk Yield (number of pounds of milk in a standard 305-day lactation); Fat Yield (number of pounds of fat in a standard 305-day lactation) - Fat (%); Protein Yield (number of pounds of protein in a standard 305-day lactation) - Protein (%), Cheese Merit, Fluid Merit, Cow Livability (LIV).

Fertility Traits are intended to bring together several measures of reproductive success and include Sire Calving Ease (SCE), Heifer Conception Rate (HCR), Cow Conception Rate (CCR), Daughter Stillbirth, Sire Stillbirth, and Gestation Length (GL).

All investigated genotypes have haplotype status. A haplotype is a set of DNA variations, or polymorphisms, that tend to be inherited together. A haplotype can refer to a combination of alleles or to a set of single nucleotide polymorphisms (SNPs) found on the same chromosome. All animals in this report have "T" haplotypes which haplotypes associated with fertility. T= Tested free; C= Carrier; A= Homozygote affected with recessive genes.

Genetic analysis

The Igenity Dairy Heifer Programme family of products contains comprehensive, powerful, and easy-to-use tools for genetic evaluation, at any time in an animal's lifetime. All animals receive a Genomic Predicted Transmitting Ability (PTA) based on DNA tests that use from 5,000 to nearly 150,000 markers from the bovine genome. DNA and SNP molecular markers were isolated from animal tail hair.

A genomic PTA gives an accurate measure of animal's true genetic potential. The PTA is an estimate of the relative genetic superiority that a particular animal will pass to its offspring for a given trait. The genomic PTA contains information on the animal's parents, its relatives, any progeny records that might be available, as well as an estimate of the animal's genetic merit based on the direct examination of the genetic markers in its DNA. In North America, genomic information has been used in national genetic evaluations for routine calculation of PTA for production, conformation, and fitness of dairy cattle since January 2009 (WIGGANS *et al.*, 2009).

In the case of genomic data, the specific results from DNA markers directly predict genetic merit. What is particularly powerful in dairy heifers is that the information from the DNA markers is equivalent to many progeny records, when predicting the true merit of an animal (Neogen Corporation).

ANN modeling

A multi-layer perceptron model (MLP), which consisted of three layers (input, hidden, and output) were used for modeling an artificial neural network model (ANN) for prediction of Yield Traits (Fat%, Protein%, Cheese Merit, Fluid Merit, LIV) and Fertility Traits prediction variables (SCE, HCR, CCR, Daughter Stillbirth, Sire Stillbirth, and GL). In the known literature, the ANN model was proven as quite capable of approximating nonlinear functions (GÖRGÜLÜ, 2011; SHAHINFAR *et al.*, 2012; YUN *et al.*, 2013; EHRET *et al.*, 2015; KLEIJNEN, 2018; GHOTBALDINI *et al.*, 2019; NAYERI *et al.*, 2019). Before the calculation, both input and output data were normalized to improve the behavior of the ANN. During this iterative process, input data were repeatedly presented to the network (KOLLO and VON ROSEN, 2005). Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm was used as an iterative method for solving unconstrained nonlinear optimization during the ANN modeling.

The experimental database for ANN was randomly divided into training, cross-validation, and testing data (with 60%, 20%, and 20% of experimental data, respectively). The training data set was used for the learning cycle of ANN and the evaluation of the optimal number of neurons in the hidden layer and also the weight coefficient of each neuron in the network. A series of different topologies were used, in which the number of hidden neurons varied from 15 to 20, and the training process of the network was run 100,000 times with random initial values of weights and biases. The optimization process was performed based on validation error minimization. It was assumed that successful training was achieved when learning and cross-validation curves approached zero.

Coefficients associated with the hidden layer (weights and biases) were grouped in matrices W_1 and B_1 . Similarly, coefficients associated with the output layer were grouped in matrices W_2 and B_2 . It is possible to represent the neural network by using matrix notation (Y is the matrix of the output variables, f_1 and f_2 are transfer functions in the hidden and output layers, respectively, and X is the matrix of input variables; KOLLO and VON ROSEN, 2005; GHOTBALDINI *et al.*, 2019):

$$Y = f_1(W_2 \cdot f_2(W_1 \cdot X + B_1) + B_2) \quad (1)$$

Weight coefficients (elements of matrices W_1 and W_2) were determined during the ANN learning cycle, which updated them using optimization procedures to minimize the error between the network and experimental outputs (KOLLO and VON ROSEN, 2005;), according to the sum of squares (SOS) and BFGS algorithm, used to speed up and stabilize convergence (TAYLOR, 2006). The coefficients of determination were used as parameters to check the performance of the obtained ANN model.

Global sensitivity analysis

Yoon's interpretation method was used to determine the relative influence of Key Traits and Yield Traits on Fertility Traits prediction variables (YOON *et al.*, 2017). This method was applied based on the weight coefficients previously calculated using the developed ANN model.

The accuracy of the model

The numerical verification of the developed model was tested using the coefficient of determination (r^2), reduced chi-square (χ^2), mean bias error (MBE), root mean square error (RMSE) and mean percentage error (MPE). These commonly used parameters can be calculated as follows (AĆIMOVIĆ *et al.*, 2020):

$$\chi^2 = \frac{\sum_{i=1}^N (x_{\text{exp},i} - x_{\text{pre},i})^2}{N - n}, \quad RMSE = \left[\frac{1}{N} \cdot \sum_{i=1}^N (x_{\text{pre},i} - x_{\text{exp},i})^2 \right]^{1/2},$$

$$MBE = \frac{1}{N} \cdot \sum_{i=1}^N (x_{\text{pre},i} - x_{\text{exp},i}), \quad MPE = \frac{100}{N} \cdot \sum_{i=1}^N \left(\frac{|x_{\text{pre},i} - x_{\text{exp},i}|}{x_{\text{exp},i}} \right) \quad (2)$$

where $x_{\text{exp},i}$ stands for the experimental values and $x_{\text{pre},i}$ are the predicted values calculated by the model, N and n are the number of observations and constants, respectively.

RESULTS AND DISCUSSION

ANN model

The acquired optimal neural network model showed a good generalization capability for the experimental data and could be used to accurately predict Yield Traits and Fertility Traits prediction based on the Key Traits data from a broad range of input parameters. According to 11 developed ANN performances, the optimal numbers of neurons in the hidden layer for Fat%, Protein%, Cheese Merit, Fluid Merit, LIV, SCE, HCR, CCR, Daughter Stillbirth, Sire Stillbirth and GL calculation were: 11 (network MLP 9-11-11) to obtain the highest values of r^2 (during the training cycle r^2 for output variables were: 0.951; 0.947; 0.989; 0.985; 0.902; 0.887; 0.676; 0.953; 0.590; 0.647 and 0.444, respectively), Table 1.

Table 1. Artificial neural network model summary (performance and errors), for training, testing, and validation cycles

Network name	Performance			Error			Training algorithm	Error function	Hidden activation	Output activation
	Train.	Test.	Valid.	Train.	Test.	Valid.				
MLP 9-11-11	0.830	0.725	0.812	1.7E+08	3.0E+08	2.1E+08	BFGS 57	SOS	Exponential	Identity

Performance term represents the coefficients of determination, while error terms indicate a lack of data for the ANN model

The obtained ANN model for prediction of output variables (Fat%, Protein%, Cheese Merit, Fluid Merit, LIV, SCE, HCR, CCR, Daughter Stillbirth, Sire Stillbirth, and GL) was complex (276 weights-biases) because of the high nonlinearity of the observed system (MONTGOMERY, 1984; ADAMCZYK *et al.*, 2021).

The goodness of fit between experimental measurements and model-calculated outputs, represented as ANN performance (sum of r^2 between measured and calculated Fat%, Protein%, Cheese Merit, Fluid Merit, LIV, SCE, HCR, CCR, Daughter Stillbirth, Sire Stillbirth, and GL), during training, testing and validation steps, are shown in Table 2.

Table 2. The "goodness of fit" tests for the developed ANN model

Output variable	χ^2	RMSE	MBE	MPE
Fat	0.000	0.003	0.000	0.134
Protein	0.000	0.003	0.000	1.923
Cheese Merit	0.190	0.434	0.048	0.000
Fluid Merit	0.166	0.405	0.078	0.436
LIV	0.656	0.806	0.099	9.440
SCE	1.9×10^7	4.3×10^3	-2.5×10^2	9.5×10^3
HCR	0.723	0.846	-0.023	-11.565
CCR	0.313	0.556	0.019	22.942
Daughter Stillbirth	1.5×10^8	1.2×10^4	3.1×10^2	3.5×10^4
Sire Stillbirth	2.7×10^7	5.2×10^3	-2.9×10^2	9.8×10^3
GL	0.786	0.882	-0.037	4.031

The ANN model predicted experimental variables (Fat, Protein, Cheese Merit, Fluid Merit, GM, LIV, SCE, HCR, CCR, Daughter Stillbirth, Sire Stillbirth, and GL) reasonably well for a broad range of the process variables (as seen in Figure 1, where the experimentally measured and ANN model predicted values of output variables are presented).

The accuracy of the ANN model could be visually assessed by the dispersion of points from the diagonal line in the graphics presented in Figure 1. For the ANN model, the predicted values were very close to the measured values in most cases, in terms of r^2 values (shown in Figure 1). *SOS* values obtained with the ANN model were of the same order of magnitude as experimental errors for output variables reported in the literature (KOLLO and VON ROSEN, 2005).

The ANN model had an insignificant lack of fit tests, which means the model satisfactorily predicted Yield Traits and Fertility Traits prediction variables. A high r^2 is indicative that the variation was accounted for and that the data fitted the proposed model satisfactorily (EHRET *et al.*, 2015; GHOTBALDINI *et al.*, 2019; NAYERI *et al.*, 2019).

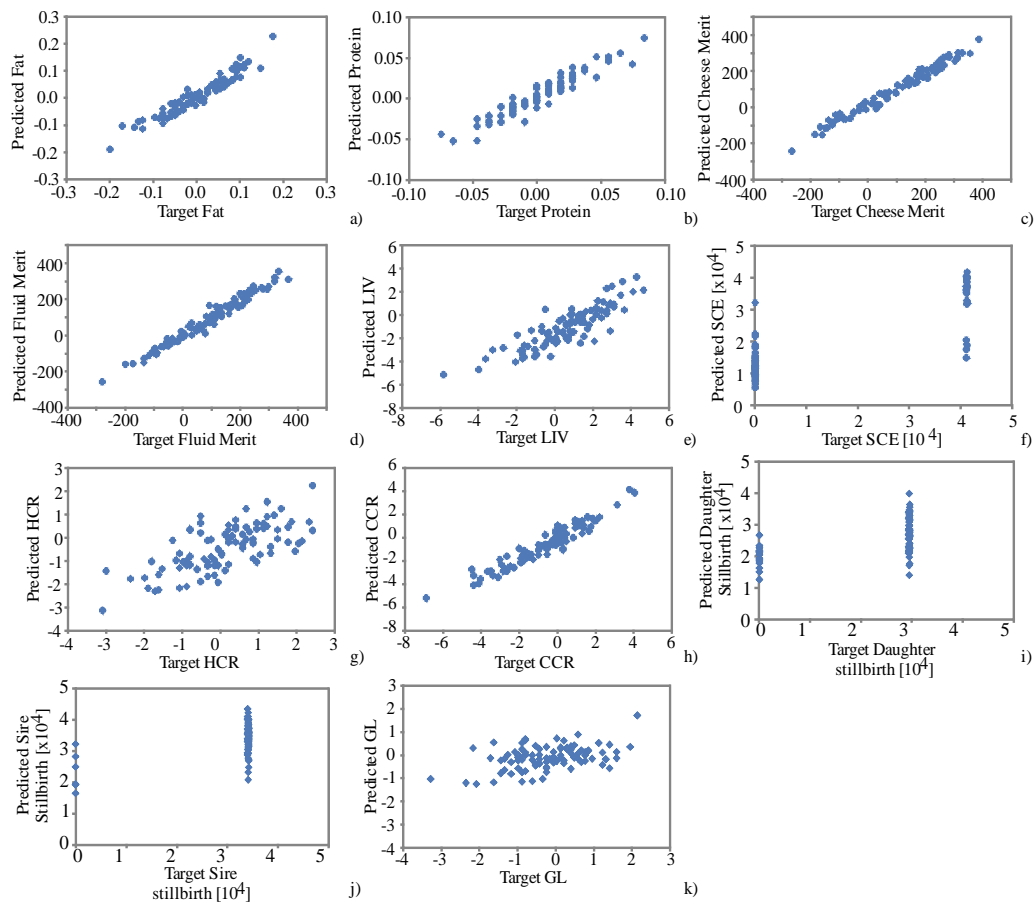


Figure 1. Comparison of experimentally obtained and ANN predicted values of Fat, Protein, Cheese Merit, Fluid Merit, LIV, SCE, HCE, CCR, Daughter Stillbirth, Sire Stillbirth and GL

Global sensitivity analysis- Yoon's interpretation method

In this section, the influence of the most important input variables on Fat%, Protein%, Cheese Merit, Fluid Merit, LIV, SCE, HCR, CCR, Daughter Stillbirth, Sire Stillbirth, and GL was studied. According to Figure 2, DCE was the most important variable for calculation, with 17.29% relative importance. Protein and Milk yield was also important, with relative importance 12.47% and 10.70%, respectively, while SCS, Fat, PL, DPR, GFI, and PTA Type were almost equally important for calculation, with relative importance between 9.81% and 10.04.

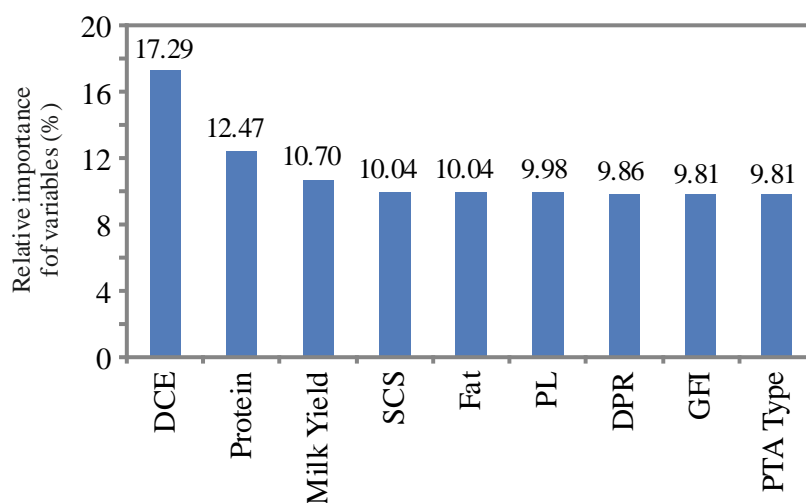


Figure 2. The relative importance of Key Traits and Yield Traits according to Milk Yield, Fat, Protein, SCS, PL, DPR, DCE, PTA Type and GFI, determined using Yoon interpretation method

CONCLUSION

The results of this study indicate that Key Traits and Yield Traits data could be used for the prediction of Fertility Traits prediction variables. The artificial neural network model showed to be adequate for the prediction of Fat Percentage, Protein Percentage, Cheese Merit, Fluid Merit, Cow Livability, Sire Calving Ease, Heifer Conception Rate, Cow Conception Rate, Daughter Stillbirth, Sire Stillbirth, and Gestation Length (the r^2 values during training cycle for these variables were: 0.951; 0.947; 0.989; 0.985; 0.902; 0.887; 0.676; 0.953; 0.590; 0.647 and 0.444, respectively). The ANN model (network MLP 9-11-11) exhibited a good fit for experimental results. The sensitivity analysis, calculated based on the ANN results, showed the influence of Key Traits data on Yield Traits and Fertility Traits prediction variables.

DCE was the most important variable for calculation, Protein and Milk yield were also important, while the influence of SCS, Fat, PL, DPR, GFI, and PTA Type was minor, with almost equal importance for calculation.

ACKNOWLEDGMENTS

This paper is the result of the research within the contract on the implementation and financing of scientific research in 2020, between the Institute for Science Application in Agriculture, Belgrade, and the Ministry of Education, Science and Technological Development of the Republic of Serbia, contract number: 451-03-68/2020-14/200045.

Received, December 28th, 2020

Accepted May 18th, 2021

REFERENCES

- AĆIMOVIĆ, M., L., PEZO, V., TEŠEVIĆ, I., ČABARKAPA, M., TODOSIJEVIĆ (2020): QSRR Model for predicting retention indices of Satureja kitaibelii Wierzb. ex Heuff. essential oil composition, *Industrial Crops and Products*, *154*:112752.
- ADAMCZYK, K., W., GRZESIAK, D., ZABORSKI (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. *Animal*, *11*: 721.
- BOUQUET, A. and J., JUGA (2013): Integrating genomic selection into dairy cattle breeding programmes: a review. *Animal*, *7*(5): 705-730.
- CALUS, M.P.L. (2009): Genomic breeding value prediction: methods and procedures. *Animal*, *4*(2): 157–164.
- EHRET, A., D., HOCHSTUHL, D., GIANOLA, G., THALLER (2015): Application of neural networks with back-propagation to genome-enabled prediction of complex traits in Holstein-Friesian and German Fleckvieh cattle, *Genetic Selection Evolution*, *47*(1): 22.
- GHOTBALDINI, H., M., MOHAMMADABADI, H., NEZAMABADI-POUR, O., IVANIVNA BABENKO, M., VITALIIVNA BUSHTRUK, S., VASYLIOVYCH TKACHENKO (2019): Predicting breeding value of body weight at 6-month age using Artificial Neural Networks in Kermani sheep breed, *Acta Scientiarum Animal Sciences*.
- GODDARD, M.E, B.J, HAYES, T.H.E., MEUWISSIEN (2010): Genomic selection in livestock populations. *Genetics research*, *92* (56): 413 – 421.
- GÖRGÜLÜ, Ö. (2011): Prediction of 305-day milk yield in Brown Swiss cattle using artificial neural networks. *South African Journal of Animal Science*, *42*(3): 280-287.
- IBAÑEZ-ESCRICHE, N., O., GONZALES-RECLÓ (2011): Review. Promises, pitfalls and challenges of genomic selection in breeding programs. *Spanish Journal of Agricultural Research*, *9* (2): 404-413.
- JOVANOVAČ, S. (2013): Principi uzgoja životinja, Poljoprivredni fakultet, Sveučilište J.J. Strossmayera u Osijeku, 356.
- KEGALI, A., M., KONJAČIĆ, M., VRDOLJAK, M., KRVAVICA (2015): Genomska selekcija u govedarstvu. *Stočarstvo*, *69* (3-4): 65-77.
- KLEIJNEN, J.P.C. (2015): Design and Analysis of Simulation Experiments. In: Pilz J, Rasch D., Melas V., Moder K. (eds) *Statistics and Simulation. IWS 2015. Springer Proceedings in Mathematics & Statistics 2018*, vol. *231*. Springer, Cham.
- KOLLO, T., D., VON ROSEN (2005): *Advanced Multivariate Statistics with Matrices* (Springer, Dordrecht).
- MEUWISSEN, T.H.E., B.J., HAYES, M.E., GODDARD (2016): Genomic selection: A paradigm shift in animal breeding, *Animal Frontiers*, *6* (1): 5-14.
- MONTGOMERY, D.C. (1984): *Design and Analysis of Experiments*, 2nd edn. 1984 (John Wiley and Sons, New York).
- NAYERI S., M., SARGOLZAEI, D., TULPAN (2019): A review of traditional and machine learning methods applied to animal breeding, *Animal Health Research Reviews*, *20*(1): 31 - 46.
- PRKA, I. (2017): Prednosti i mane genomskog ocenjivanja priplodnih bikova. *Veterinarski žurnal Republike Srpske*, *XVII* (2): 203–214.
- SHAHINFAR, S., H., MEHRABANI-YEGANEH, C., LUCAS, A., KALHOR, M., KAZEMIAN, K.A., WEIGEL (2012): Prediction of Breeding Values for Dairy Cattle Using Artificial Neural Networks and Neuro-Fuzzy Systems, *Computational and Mathematical Methods in Medicine*.
- SCHEFERS, J.M. and K.A., WEIGEL (2012): Genomic selection in dairy cattle: Integration of DNA testing into breeding programs, *Animal Frontiers*, *2* (1): 4–9.
- TAYLOR, B.J. (2006): *Methods and Procedures for the Verification and Validation of Artificial Neural Networks* (Springer Science & Business Media, New York).

- VEERKAMP, R. and M.P.L., CALUS (2009): Genomics revolution. *Veepro Magazine*, 71, 4–6.
- WIGGANS, G.R., P.M., VAN RADEN, L.R., BACHELLER, F.A., ROSS, T.S., SONSTEGARD, G., TE MEERMAN, C.P., VAN TASSELL (2009): Transition of genomic evaluation from a research project to a production system. *J. Anim. Sci.*, 87(E-Suppl. 2) / *J. Dairy Sci.* 92(E-Suppl. 1): 313–314(abstr. 278).
- WIGGANS, G.R., J.B., COLE, S.M., HUBBARD, T.S., SONSTEGARD (2017): Genomic selection in dairy cattle: The USDA experience. *Ann. Rev. Animal Biosci.*, 5: 309-327.
- YUN, T.S., Y.J., JEONG, T.S., HAN, K.S., YOUM (2013): Evaluation of thermal conductivity for thermally insulated concretes, *Energy and Buildings*, 61: 125-132.
- YOON, Y., G., SWALES, T.M., MARGAVIO (2017): A Comparison of Discriminant Analysis versus Artificial Neural Networks. *Journal of the Operational Research Society*, 44 (1): 51-60.

MATEMATIČKO MODELIRANJE ZA GENOMSKU SELEKCIJU MLEČNIH GOVEDA U SRBIJI

Radmila BESKOROVAJNI¹, Rade JOVANOVIĆ¹, Lato PEZO², Nikola POPOVIĆ¹,
Nataša TOLIMIR¹, Ljubiša MIHAJLOVIĆ³, Gordana ŠURLAN-MOMIROVIĆ⁴

¹Institut za primenu nauke u poljoprivredi, Beograd, Srbija

²Institut za opštu i fizičku hemiju, Univerzitet u Beogradu, Beograd, Srbija

³Ministarstvo poljoprivrede, šumarstva i vodoprivrede, Beograd, Srbija

⁴Univerzitet u Beogradu, Poljoprivredni fakultet, Beograd, Srbija

Izvod

Ovaj rad je nastao kao rezultat efikasnog uzgojnog programa u populacijama goveda u Srbiji, za neke ekonomski važne osobine. Genomska selekcija je u poslednje dve decenije bila glavni izazov u programima genetike i oplemenjivanja životinja. U statističkoj analizi za predviđanje stvarnih priplodnih vrednosti mladih životinja koristio se veliki broj SNP markera koji ne uključuju performanse životinja. Nova tendencija uzgoja u selekciji mladih životinja omogućava njihov genetički napredak uz smanjenje troškova. U ovom radu analizirane su 92 krave rase holštajn iz različitih regiona Srbije, na osnovu SNP molekularnih markera. U okviru ovog istraživanja razvijen je empirijski model za predviđanje varijabli osobina prinosa i plodnosti, prema podacima o ključnim osobinama za mlečna goveda. Razvijeni model je dao razumno prilagođavanje podacima i uspešno predvideo osobine prinosa (kao što su udeo masti i proteina, indeks za proizvodnju sira, indeks za proizvodnju mleka i životna sposobnost krava) i varijable osobina plodnosti (kao što su lakoća teljenja po očevima, stopa začeća junica, stopa začeća krava, mrtvorodenost po kćerima, mrtvorodenost po očevima i dužina bremenitosti). Ukupno 92 podatka o mlečnim govedima su korišćena za izgradnju modela za predviđanje varijabli osobina prinosa i plodnosti. Model veštačke neuronske mreže, zasnovan na Broiden-Fletcher- Goldfarb-Shanno iterativnom algoritmu, pokazao je dobre mogućnosti predviđanja (vrednosti r^2 tokom ciklusa obuke za prethodno pomenute izlazne varijable bile su u opsegu između 0,444 i 0,989).

Primljeno 28.12., 2020

Odobreno 18.5. 2021