DOI: 10.29050/harranziraat.1464601
Harran Tarım ve Gıda Bilimleri Derg. 2024, 28(3): 374-384

Research Article/Araştırma Makalesi

Determination of the effect of climate change on small cattle milk yield in Iğdır province via machine learning

Iğdır ilinde iklim değişikliğinin küçükbaş süt verimi üzerine etkisinin makine öğrenmesi ile belirlenmesi

Ayça Nur ŞAHİN DEMİREL[1*], Taner ERİK[2]

[1]Iğdır University, Faculty of Agriculture, Department of Agricultural Economics, Iğdır, Türkiye
[2]Iğdır University, Institute of Graduate Studies, Department of Agricultural Economics, Iğdır, Türkiye

[1]https://orcid.org/0000-0003-2988-8448; [2]https://orcid.org/0009-0009-9443-2552

To cite this article:
Şahin Demirel, A. N. & Erik, T. (2024). Determination of the effect of climate change on small cattle milk yield in ığdır province via machine learning. Harran Tarım ve Gıda Bilimleri Dergisi, 28(3): 374-384

DOI: 10.29050/harranziraat.1464601

Address for Correspondence:
Ayca Nur SAHIN DEMIREL
e-mail:
aycanur.sahin@igdir.edu.tr

Received Date:
03.04.2024
Accepted Date:
09.07.2024

ABSTRACT

This study examines the potential impact of climate change on small cattle livestock and milk productivity in Iğdır province. The study takes into account various factors, including the effects of climate change on animal stress levels, nutrient quality in grazing areas, and the spread of parasites or diseases, which may indirectly affect milk productivity. To evaluate this impact, the study utilizes eXtreme Gradient Boosting (XGBoost) machine learning models with five different climate variables, analyzing the small cattle data from Iğdır province between 2004 and 2023. Two machine learning models were created to investigate the effect of climate variables on milk yield in small cattle in Iğdır province, using a dataset of 10820 rows and 16 columns. The machine learning models revealed that five different climate variables had no significant effect on milk yield. This finding is important for the economic welfare of the region, as cattle farming plays a crucial role in the economy of Iğdır province. The neutral effect of climate change is therefore evaluated positively for Iğdır province. The study suggests that there has been no significant change in milk productivity over the last 20 years due to the constant percentage of sheep that produce milk. It is recommended that farmers in Iğdır province consider increasing the number of lactating sheep to enhance overall cattle milk production.

Key Words: Diary Milk, Machine learning, Milk Yield, Sheep, Goat

ÖZ

Bu çalışma, iklim değişikliğinin Iğdır ilindeki küçükbaş hayvancılık ve süt verimliliği üzerindeki potansiyel etkisini incelemektedir. Çalışmada, iklim değişikliğinin hayvanların stres seviyeleri üzerindeki etkileri, otlatma alanlarındaki besin kalitesi ve süt verimliliğini dolaylı olarak etkileyebilecek parazitlerin veya hastalıkların yayılması gibi çeşitli faktörler dikkate alınmaktadır. Bu etkiyi değerlendirmek için çalışmada beş farklı iklim değişkeni ile eXtreme Gradient Boosting (XGBoost) makine öğrenimi modelleri kullanılmış ve Iğdır ilinin 2004-2023 yılları arasındaki küçükbaş hayvan verileri analiz edilmiştir. Iğdır ilindeki küçükbaş hayvanlarda iklim değişkenlerinin süt verimi üzerindeki etkisini araştırmak için 10820 satır ve 16 sütundan oluşan bir veri seti kullanılarak iki makine öğrenmesi modeli oluşturulmuştur. Makine öğrenimi modelleri, beş farklı iklim değişkeninin süt verimi üzerinde önemli bir etkisi olmadığını ortaya koymuştur. Büyükbaş hayvancılık Iğdır ilinin ekonomisinde önemli bir rol oynadığından, bu bulgu bölgenin ekonomik refahı açısından önemlidir. Dolayısıyla iklim değişikliğinin nötr etkisi Iğdır ili için olumlu olarak değerlendirilmektedir. Çalışma, süt üreten koyun oranının sabit kalması nedeniyle son 20 yılda süt verimliliğinde önemli bir değişiklik olmadığını göstermektedir. Iğdır ilindeki çiftçilere, toplam sığır sütü üretimini artırmak için süt veren koyun sayısını artırmayı düşünmeleri önerilmektedir.

Anahtar Kelimeler: Günlük Süt, Makine öğrenimi, Süt Verimi, Koyun, Keçi

Introduction

Iğdır is a province located in the eastern region of Turkey with a relatively low population and surface area (Koc et al., 2019; Öztürk et al., 2023). The summers are generally hot and dry, while the winters are cold and snowy (Türkeş and Tatli, 2011). These climatic conditions have the potential to affect livestock activities, which in turn can impact milk yield. In Iğdır province, small cattle milk production is primarily obtained from sheep and goats (Yilmaz, 2022). These species are usually domestic and hybrid breeds that can adapt well to the geographical and climatic conditions of the region.

To achieve high milk yield in small cattle, it is crucial to provide them with proper nutrition (Garg et al., 2013). Ensuring that the animals have access to adequate amounts of water and a balanced feeding program not only promotes their health but also increases milk yield. Additionally, it is especially important to feed high-quality feed to promote the growth of calves and milk production in mothers. However, it is important to consider that climate change may lead to warmer and drier conditions, which can result in reduced or inefficient grazing areas. This may have a negative impact on the nutrition of animals, potentially leading to a reduction in milk yield. Additionally, drought, particularly in the summer, can further reduce grazing areas and limit animals' access to natural feed resources, which can further impact the quality and quantity of nutrition, ultimately leading to a reduction in milk yield.

Climate change may have an impact on the spread and distribution of diseases and pests, which may in turn affect animal health and milk yields (Baumgard et al., 2012; Grace et al., 2015). In certain regions, warmer weather conditions may accelerate the spread of parasites and disease-carrying insects, leading to heat stress in animals and reduced milk production. It is important to take into account the potential consequences of climate change on animal agriculture. Changes to temperature and humidity levels have been known to promote the spread of bacteria and fungi, which can have a negative impact on animal health and milk yield.

Additionally, climate variables such as wind, temperature, humidity, and precipitation have been observed to affect milk yield in small cattle (Hill and Wall, 2015; Marumo et al., 2022). It is important to note that strong winds can cause stress in animals, which may lead to reduced milk yield. Temperature conditions can have a notable effect on animal welfare and milk production (Polsky and von Keyserlingk, 2017). Cold winds, for instance, may lower body temperature and cause dehydration, leading to a decrease in milk yield. Therefore, it is crucial to monitor and manage temperature conditions to ensure the best possible milk production.

High humidity can potentially increase the risk of heat stress in animals (Thornton et al., 2021). Sweating efficiency may be reduced in humid environments, which can make it challenging for animals to cool down. This may have a negative impact on animal health and potentially reduce milk yield. Furthermore, high humidity may increase the risk of microbial contamination of milk, which could lead to a decrease in milk quality. In addition, muddy or waterlogged grazing areas during wet weather may impede animal movement and further reduce milk yield. During extended periods of wet weather, it may be necessary for animals to spend more time outside, which could potentially increase the risk of heat stress.

This study examines the correlation between climate change and milk production yield of small cattle in Iğdır province. Meteorological factors, including temperature, precipitation, wind speed, and humidity, were analyzed as climate variables data over the last 20 years to determine the impact of climate change on milk yield in Iğdır province using a machine learning algorithm. There are many different algorithms in machine learning studies. The machine learning algorithms that are most commonly used today are support vector machines, decision trees, linear regression, logistic regression, gradient boosting, and XGBoost (Castro and Ferreira, 2023; Kurani et al., 2023;

Noorunnahar et al., 2023; Soori et al., 2023). Among these algorithms, XGBoost, also known as "extreme gradient boosting," is a highly effective and efficient supervised learning algorithm used in machine learning. XGBoost achieves its results by combining multiple decision trees through an "ensemble" method and correcting the errors of previous trees with each new tree. In addition, XGBoost can provide faster results and more accurate predictions when working with complex data in small data sets compared to other machine learning models. This algorithm was chosen due to its ability to model non-linear relationships, which is particularly important given the non-linear relationships between climate variables and milk yield. By utilizing a tree-based algorithm, the study was able to achieve a more flexible structure and more accurate analysis of the impacts of climate change on milk yield.

Material and Method

Material

To model machine learning, data on small cattle numbers and milk production in Iğdır province was obtained from the Iğdır Provincial Directorate of Agriculture (TOB, 2024). The dataset covers small cattle milk production from 2004 to 2023, including the total number of sheep and goats, as well as the number of sheep and goats that produce milk. The study assessed the milk production efficiency of sheep and goats by calculating the percentages of animals giving milk.

Furthermore, an analysis of climate change was conducted using data on wind speed U and V components, 2-meter temperature values, precipitation amounts, and humidity change. The data set comprised information from 2004-2023 annually for a total of five different climate factors. Data on climate change factors were obtained from the European Union Copernicus Climate Change database (Copernicus, 2024). Upon the integration of all data prior to analysis, a data set comprising 16 columns and 10,820 rows was generated.

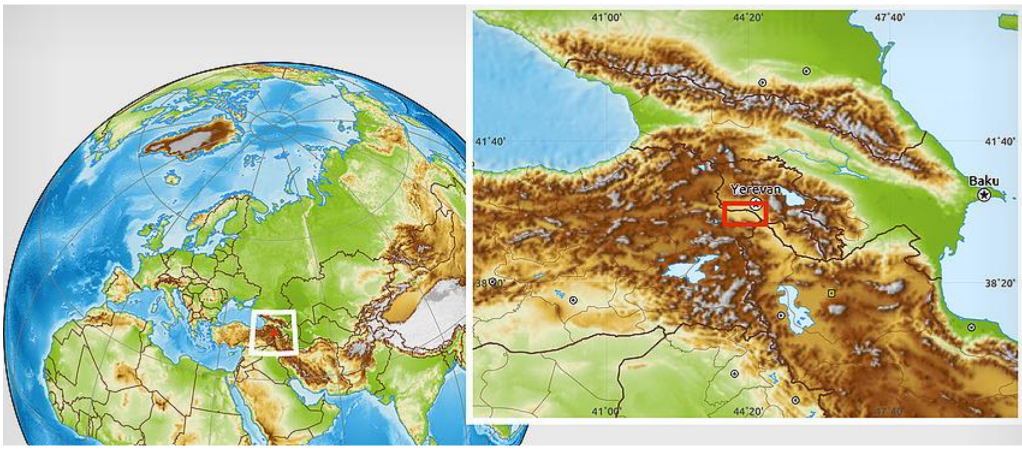

Figure 1. Iğdır province geographical location (Maphill, 2024).

Geographical coordinates of Iğdır province were used to obtain climate change data. Iğdır province coordinates are defined to the system as North 40°, West 43°, South 39°, East 45°. If we define the regions covered by these coordinates in detail; It covers the mountainous region in the north of Tuzluca district in the north of Iğdır province, the Aras River in the west of Aralık district in the west, the Iğdır Plain in the south of Iğdır Central district in the south and the Karasu River in the east of Karakoyunlu district in the east.

Method

The free and open-source RStudio software was used for machine learning modeling. XGBoost algorithm was preferred for the machine learning algorithm. XGBoost (eXtreme Gradient Boosting) is a high-performance version of the Gradient Boosting algorithm that has been optimized with various adjustments(Sahin Demirel, 2024). XGBoost was introduced by Tianqi Chen and Carlos Guestrin in the article "XGBoost: A Scalable Tree Boosting System" published in 2016. With its many advantages, XGBoost has become one of the most preferred algorithms in machine learning today(Noorunnahar et al., 2023). XGBoost provides high accuracy in data prediction (Anne and Gueye, 2024). Minimizes overlearning through careful editing. Effectively processes missing data. It is one of the best decision tree based algorithms (Anne and Gueye, 2024; Chen and Guestrin, 2016).

The working logic of XGBoost is similar to Gradient Boosting (Bui et al., 2021; Liang et al., 2020; Natekin and Knoll, 2013). With Base Score, an initial estimate of the modeling is made. This estimate is used in subsequent steps to get closer to the correct result. By default, this estimate is 0.5 (Mohamed et al., 2020; Huang et al., 2015). The errors (residual) of the first prediction are analyzed. Errors are the difference between the observed value and the predicted value (Bonavita and Laloyaux, 2020; Shrestha and Solomatine, 2006). As in Gradient Boosting, a decision tree is created that estimates the errors. A similarity score is calculated for each tree branch. This shows how well the data is grouped in the branches (Xie et al. 2019). To determine which tree is better, gain is calculated. As a result, the machine learning model ranks the importance of the variables. In the final stage, it gives the success values of the overall model. The mathematical formula of the XGBoost algorithm forms the basis of this algorithm, which belongs to the Gradient Boosting family (Chen and Guestrin, 2016). Objective Function:

$$\{obj\}(t) = \sum_{i=1}^{n} \ (y\_i - (y\_i^{(t-1)} + f\_t(x\_i)))^2 + \sum_{i=1}^{t} \ \omega\,(f\_i) \qquad (1)$$

Where *y_i* is the true target value; $y\_i^{(t-1)}$ is the prediction at step (t-1); *f_t(x_i)* is the new learner (decision tree) prediction at step t; *ω(f_i)* is the regularization term (e.g. tree depth or number of leaf nodes) and {obj}(t) is the objective function at iteration t.

In the XGBoost algorithm, it is also very important to define 7 different parameters in order to make a prediction close to the real values. In this stage, called hyperparameter tunning, value ranges are defined for seven different parameters and the ideal parameter values for the machine learning model are determined in the tunning process(Sahin Demirel, 2024).

The efficacy of XGBoost and other algorithms in making successful predictions has led to the widespread use of machine learning modelling in the literature on milk production and factors affecting milk production. Becker et al. 2021 employed logistic regression, random forest, and Gaussian naïve Bayes algorithms to analyse the levels of animals affected by environmental variables in milk production. In another study by Ji et al. 2022 employed the XGBoost algorithm to investigate the factors influencing milk production in cows. In a similar vein, Ebrahimie et al., 2018 utilised a decision tree algorithm to identify effective milking methods. Furthermore, Kamphuis et al. 2010 employed a decision tree algorithm to investigate the impact of automatic milking and clinical factors on milk production, with promising results. The aforementioned studies demonstrate the efficacy of machine learning algorithms in numerous aspects pertaining to milk production. It is anticipated that the XGBoost algorithm will yield successful outcomes in our study.

Results and Discussion

Small cattle milk production analysis in Iğdır province

Cattle farming on a small scale is a prevalent practice in the region of Iğdır, and milk production plays a crucial role in these activities. In Iğdır, small-scale cattle milk production holds considerable economic significance and plays a vital role in the local economy. The region's milk and dairy products are not only consumed locally but also marketed to neighboring provinces, contributing to the overall economic growth of the region. Figure 2 displays the graphs of small cattle livestock and milk productivity in Iğdır province. Specifically, Figure 2-a illustrates the relationship between the number of sheep, milk yield, and percentage of dairy sheep from 2004 to 2024. The figure highlights changes in the number of sheep, average milk yield of dairy sheep, and percentage of total sheep herd consisting of dairy sheep over time.

The data shows that the number of sheep was approximately 200,000 in 2004. It increased steadily until 2020, reaching around 600,000. However, from 2021 to 2023, there was a downward trend, and the number decreased to an average of 300,000 in 2023. When evaluating milk yield, it was observed that although there was a variable yield graph, the values were very close to each other. Over a period of 20 years, sheep in Iğdır province produced an average of 11.98 tons of milk annually.

It is important to note that milk productivity remained consistent despite fluctuations in the number of sheep between 2004 and 2023. To gain a better understanding of this complex situation, it is crucial to examine the change in milking sheep. Examining the percentage values of dairy sheep over a 20-year period can help us understand their consistent milk yield. The percentage change of lactating sheep fluctuated only slightly, between 88.9% and 89.0%, during this time. This suggests that the percentage of lactating sheep remained almost constant each year, which may contribute to the consistent milk productivity of sheep.

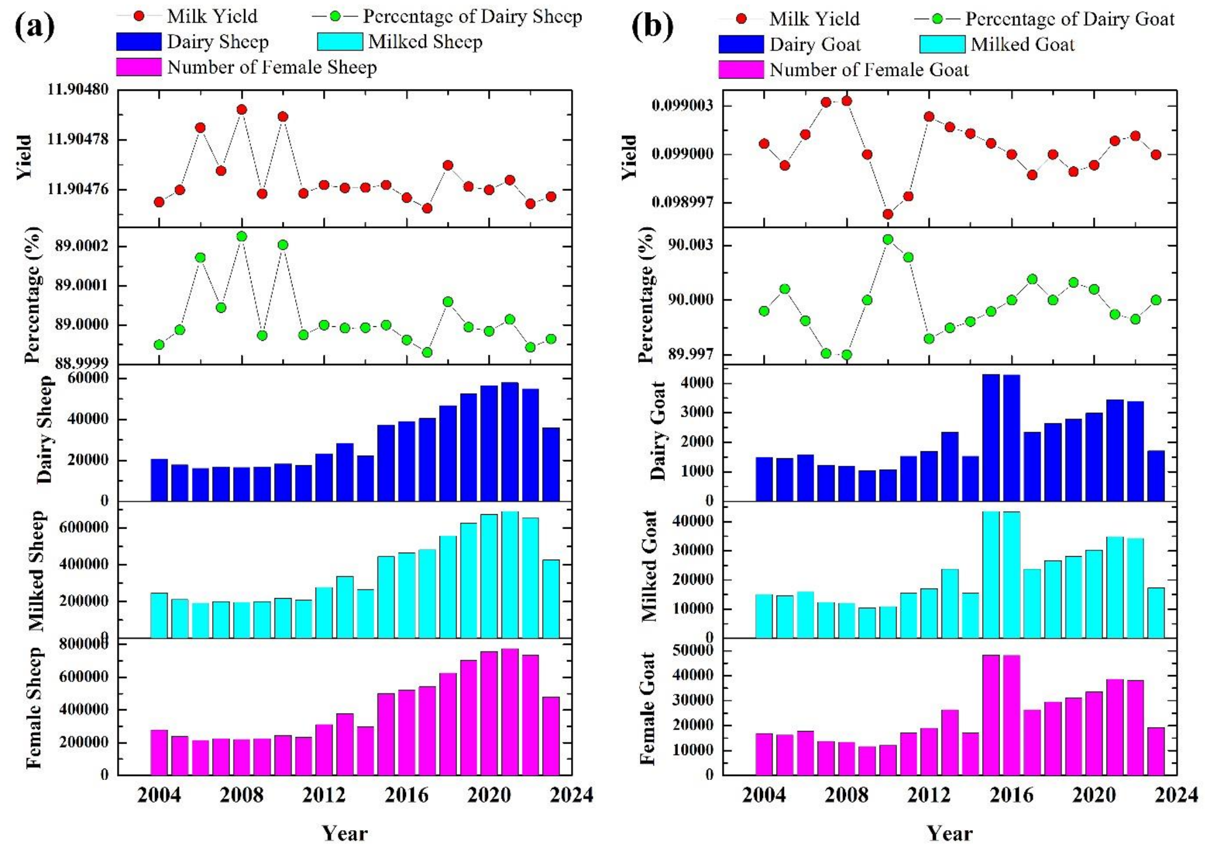

Figure 2. Small cattle milk production performance data for Iğdır province. a) sheep, b) goat

Figure 2-b displays milk production performance data for goats. A similar trend is observed when analyzing the change in the total number of goats, except for 2015 and 2016. It can be interpreted that the tendency of small cattle farmers to feed goats increased in 2015 and 2016. However, it is important to note that enterprises mostly choose sheep for cattle breeding and production. When analyzing the milk production efficiency and percentage values of goats and sheep, a similar trend is observed. However, it is worth noting that goats have a lower annual milk productivity of 0.098 tons compared to sheep's 11 tons. This may explain why some enterprises prefer sheep over goats.

Climate change analysis and machine learning models

Figure 2 shows a consistent trend in milk yield values. However, it would be beneficial to further investigate the reasons behind the lack of increase or decrease in this trend. Specifically, it is crucial to thoroughly examine the various factors that may affect milk productivity in small cattle in Iğdır province.

It is worth noting that there are differences in milk productivity between sheep and goat breeds when considering the factors that affect milk productivity in small cattle. Milk yield can vary among breeds due to differences in genetic potential (Haenlein, 2007).

It is important to ensure a balanced diet for optimal milk production, as it provides the necessary energy, protein, vitamins, and minerals (Pereira, 2014). In addition, proper grazing areas and feeding programs are essential for maintaining animal health, which directly impacts milk productivity (Hennessy et al., 2020). Parasites, diseases, and infections can have a negative impact on milk production. Therefore, it is recommended to schedule regular veterinary check-ups, vaccinations, and implement appropriate health measures to ensure the well-being of the animal (Perri et al. 2011). Milk yield reaches its highest level during the lactation period of the animal, and it is important to provide proper nutrition and management throughout this period (Gramu, 2019). While milk yield typically increases in the first few months after birth, it may decrease over time. It is important to use correct breastfeeding and milking methods to optimize milk yield. Proper suckling and regular milking are important for optimal milk production in calves. Additionally, environmental factors, such as temperature, humidity, and lighting, can have an impact on milk productivity (Amin Sheikh et al., 2017). Therefore, it is recommended to maintain a comfortable environment for the animals and take appropriate acclimatization measures.

Although there are several factors that can affect milk productivity, the environmental climate factor is considered to be the most significant. This is because the climate has a fundamental impact on various aspects, including possible animal stress, changes in food quality in grazing areas, and the spread of parasites or diseases. Therefore, it is important to take into account the climate change factor when assessing milk productivity in cattle (Becker et al., 2021).

To investigate the impact of climate change on milk production through machine learning, utilized various meteorological parameters such as 10-meter u-wind, 10-meter v-wind, 2-meter temperature, instantaneous moisture flux, and total precipitation values. The performance values and parameters of the machine learning models developed are presented in Table 1.

Table 1. Machine learning Hyperparameters and model performance values.

	Yield for Sheep	Yield for Goat
Nrounds	500	1000
Max_depth	3	3
Eta	0.05	0.1
Gamma	0	0
Colsample_bytree	0.7	0.7
Min_child_weight	1	1
Subsample	0.7	0.8
RMSE	0.322	0.0113
MAE	0.322	0.0113
R^2	Not Available	Not Available

Table 1 presents the hyperparameter and the model performance parameter values. The parameter 'Nrounds' refers to the number of rounds the model has been trained on. Increasing the number of rounds can improve the model's generalization ability. In this case, using 1000 rounds resulted in a higher goat yield. The parameter 'Max_depth' was determined to be appropriate for the ideal model in both cases, with a maximum depth of 3 for the trees. Deeper trees may allow the model to learn more complex relationships, but they may also increase the risk of overfitting. The learning rate, or 'Eta', can be adjusted to control how the model updates its weights in each round. A smaller learning rate can result in a slower learning process but a more robust model. In this case, it seems that a smaller learning rate produces better results for sheep yield. Gamma is a parameter that controls the growth decision of the trees. A value of zero indicates that this parameter is not used. 'Colsample_bytree' specifies the proportion of randomly sampled columns for each tree. Using the same values indicates that there is no difference in this parameter. Using the same values indicates that there is no difference in this parameter. 'Min_child_weight' specifies the minimum amount of weight required for a node to be split in the tree. 'Subsample' specifies the random sampling rate of a subset of the training data. It has been observed that utilizing a larger sample size can enhance the model's capacity to generalize. The present case suggests that a larger sample size is associated with improved results for goat yield (Dalal et al., 2022; Putatunda and Rama, 2018).

When evaluating model performance, it is important to consider the RMSE (Root Mean Square Error) which measures the distance between the model's predictions and the actual values (Sahin Demirel, 2024). It is worth noting that the RMSE value for goat yield is significantly smaller than that for sheep yield, indicating better performance for goat yield. MAE (Mean Absolute Error) measures the average distance between the model's predictions and the true values (Sahin Demirel, 2024). Both models performed well, with low MAE values indicating that the predictions are close to the actual values. The model for goat yields outperformed the model for sheep yields, but both were successful.

Finally, analyzing R^2 (R-Square) values reveals a complex situation. It is important to note that R^2 represents the percentage of the variance of the dependent variable explained by the independent variables. An R^2 value approaching 1 indicates a better fit of the model to the data (Sahin Demirel, 2024). However, in this case, R^2 values were not obtained for both models. This may indicate that there is no relationship between the independent variables and the dependent variables. It may be advisable to consider conducting sensitivity analysis for each variable in the machine learning model to help ensure accuracy and reliability.

Table 2. Feature importance and sensitivity analyses for both ML models.

Yield for Sheep			Yield for Goat		
No	**Variable**	**Mean Dropout Loss**	**No**	**Variable**	**Mean Dropout Loss**
1	***Full model***	0.3224164	**1**	***Full model***	0.01133641
2	10 meter u-wind	0.3224164	**2**	10 meter u-wind	0.01133641
3	10 meter v-wind	0.3224164	**3**	10 meter v-wind	0.01133641
4	2-meter Temperature	0.3224164	**4**	2-meter Temperature	0.01133641
5	Humidity	0.3224164	**5**	Humidity	0.01133641
6	Precipitation	0.3224164	**6**	Precipitation	0.01133641
7	***Baseline***	0.3224164	**7**	***Baseline***	0.01133641

Table 2 presents the feature importance and sensitivity analysis values. It is worth noting that all variables have a constant Mean Dropout Loss (MDL) value. Ideally, the MDL values of the variables in this table should differ from each other, resulting in different rankings (Dong et al., 2022). Nevertheless, as there was no correlation between the dependent and independent variables, the MDL values were calculated as shown in Table 2.

The complete model displays the model's performance using all features. Dropout loss measures the effect of removing a feature from the model on its performance. A low dropout loss indicates that a feature has little impact on the model's performance. It appears that a high dropout loss indicates that this feature may be significant to the model's performance. The baseline forecast represents the model's performance without utilizing any features. In this context, the fact that all variables have the same value in both the baseline and the full model means that these variables are ineffective in the model.

When considering Table 1 and Table 2 collectively, it is possible to gather insights on the impact of independent climate variables on milk yield. While the lack of R^2 value calculation under normal conditions may indicate a potential error, the computation of feature importance and sensitivity values provided clarification on why the R^2 value could not be determined. Based on the MDL values, if each variable in the models has the same value, it may suggest that the variables do not significantly affect the model.

However, it is important to note that the low RMSE and MAE values obtained demonstrate the success of the machine learning models. Based on the results of machine learning models, it has been suggested that climatic changes in Iğdır province may not have a significant impact on milk productivity. Furthermore, a correlation analysis was conducted.

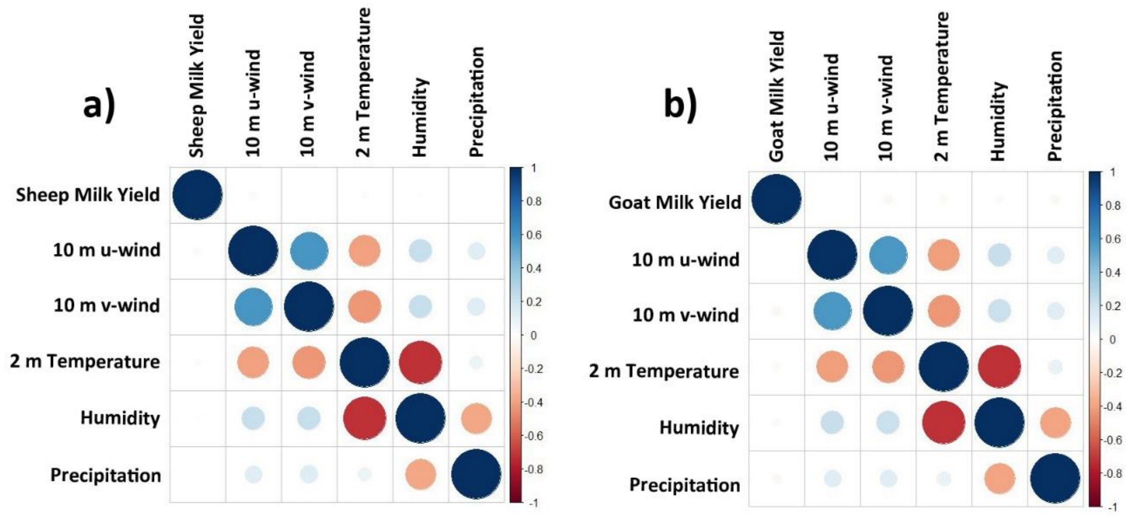

Figure 3. Milk Yield and climate variables correlations. a) for sheep, b) for goat.

Figure 3 displays the correlation matrice graphs, and, the correlation values in Figure 3a and b suggest that there is no correlation between goat and sheep milk data and climate variables, which is supported by machine learning.

This study examines the potential impact of climate change on small cattle livestock and milk productivity in Iğdır province. The results suggest that, at present, there is no significant impact on sheep and goat milk productivity, as indicated by the machine learning models used in the study.

However, it is important to note that these findings do not necessarily imply that climate change has no impact on milk productivity. However, it is important to note that climate change can impact the stress levels of animals, the nutrient quality of grazing areas, and the spread of parasites or diseases, which may indirectly affect milk productivity. It is possible that the models used in the current study did not fully capture these effects.

The study found that milk productivity remained constant over the years because the percentage of sheep giving milk remained almost the same. In conclusion, it is suggested that further research is necessary to develop more comprehensive and precise models to determine the impact of climate change on milk productivity. Additionally, to fully understand the potential impacts of climate change on milk productivity, it is recommended that more comprehensive

studies with direct measurements of these impacts be conducted. This study aims to enhance our comprehension of the consequences of climate change on small-scale cattle farming and milk production. By doing so, we can develop effective strategies to manage these impacts.

Conclusion

This study examines the potential relationship between climate change and small cattle livestock and milk productivity in Iğdır province. The study reveals that climate variables have no significant effect on milk productivity in sheep and goats. While in similar studies conducted for different geographical regions in the literature, it was determined that climatic factors such as temperature, precipitation and humidity affect milk production, this result regarding the neutral effect of five climatic factors for Iğdır province was very remarkable. However, the fact that such a conclusion was reached as a result of the evaluation of only five climatic factors in the study revealed that more research is needed to fully understand the potential effects of climate change on livestock and milk production in the region. The study provides valuable information for farmers and policymakers to better understand the factors affecting milk productivity and develop more effective strategies. It is important to note that milk productivity plays a critical role in the economic welfare of the region. According to this study, the stability of milk productivity can be attributed to the consistent percentage of sheep that produce milk over the years. To improve their flock's overall productivity, farmers may want to consider increasing the number of lactating sheep. This study highlights the benefits of utilizing advanced techniques, including machine learning and sensitivity analysis, to comprehend the intricate interplay between agricultural economics and livestock management. Future studies could focus on conducting long-term research to better understand the long-term impacts of climate change on milk productivity. In order to provide more comprehensive information, several recommendations can be made. It is possible to utilise different machine learning algorithms in order to capture non-linear relationships within a dataset more effectively than linear regression. This can be achieved through the use of algorithms such as Random Forest. Furthermore, polynomial features can be generated in order to capture potential non-linear relationships between variables. In order to enhance the reliability of the findings, the consistency of the findings can be evaluated on distinct data subsets through the application of the cross-validation method. Alternative subsets of data or alternative models can be employed to assess the robustness of the results. Additionally, it would be beneficial to extend the geographical coverage to other regions and integrate improved prediction models with additional variables such as soil quality (for feed quality), diseases, pests and water availability. It is also recommended to investigate the impact on other livestock species, analyse socio-economic factors and develop climate adaptation strategies for livestock management.

Declarations

Conflict of Interest: The authors declare that there is no conflict of interest between them.

Author Contribution: TE obtained the study data and organized the data set, designed the ANSD study, analyzed the data set, and wrote the manuscript.

References

Amin Sheikh, A., Tajamul Islam, S., Rashid Dar, R., Ahmad Sheikh, S., Mohammad Wani, J., Dogra, P., Amir Amin Sheikh, C., & Bhagat, R. (2017). Effect of climate change on reproduction and milk production performance of livestock: A review. *Journal of Pharmacognosy and Phytochemistry*, *6*(6), 2062–2064.

Anne, S., & Gueye, A. D. (2024). XGBoost Algorithm to Predict a Patient's Risk of Stroke. Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, LNICST, 541, 151–160. https://doi.org/10.1007/978-3-031-51849-2_10

Baumgard, L. H., Rhoads, R. P., Rhoads, M. L., Gabler, N. K., Ross, J. W., Keating, A. F., Boddicker, R. L., Lenka, S., & Sejian, V. (2012). Impact of Climate Change on

Livestock Production. Environmental Stress and Amelioration in Livestock Production, 9783642292057, 413–468. https://doi.org/10.1007/978-3-642-29205-7_15

Becker, C. A., Aghalari, A., Marufuzzaman, M., & Stone, A. E. (2021). Predicting dairy cattle heat stress using machine learning techniques. Journal of Dairy Science, 104(1), 501–524. https://doi.org/10.3168/JDS.2020-18653

Bonavita, M., & Laloyaux, P. (2020). Machine Learning for Model Error Inference and Correction. Journal of Advances in Modeling Earth Systems, 12(12), e2020MS002232.

Bui, Q. T., Chou, T. Y., Hoang, T. Van, Fang, Y. M., Mu, C. Y., Huang, P. H., Pham, V. D., Nguyen, Q. H., Anh, D. T. N., Pham, V. M., & Meadows, M. E. (2021). Gradient Boosting Machine and Object-Based CNN for Land Cover Classification. Remote Sensing, 13(14), 2709. https://doi.org/10.3390/RS13142709

Castro, H. M., & Ferreira, J. C. (2023). Linear and logistic regression models: when to use and how to interpret them? Jornal Brasileiro de Pneumologia, 48(6), e20220439. https://doi.org/10.36416/1806-3756/E20220439

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794. https://doi.org/10.1145/2939672.2939785

Copernicus, (2024). Climate Change. Retrieved from: https://climate.copernicus.eu/

Dalal, S., Onyema, E. M., & Malik, A. (2022). Hybrid XGBoost model with hyperparameter tuning for prediction of liver disease with better accuracy. World Journal of Gastroenterology, 28(46), 6551. https://doi.org/10.3748/WJG.V28.I46.6551

Dong, J., Chen, Y., Yao, B., Zhang, X., & Zeng, N. (2022). A neural network boosting regression model based on XGBoost. Applied Soft Computing, 125, 109067. https://doi.org/10.1016/J.ASOC.2022.109067

Ebrahimie, E., Ebrahimi, F., Ebrahimi, M., Tomlinson, S., & Petrovski, K. R. (2018). Hierarchical pattern recognition in milking parameters predicts mastitis prevalence. Computers and Electronics in Agriculture, 147, 6–11. https://doi.org/10.1016/J.COMPAG.2018.02.003

Garamu, K. (2019). Significance of Feed Supplementation on Milk Yield and Milk Composition of Dairy Cow. Journal of Dairy & Veterinary Sciences, 13(2), 1-9.

Garg, M. R., Sherasia, P. L., Bhanderi, B. M., Phondba, B. T., Shelke, S. K., & Makkar, H. P. S. (2013). Effects of feeding nutritionally balanced rations on animal productivity, feed conversion efficiency, feed nitrogen use efficiency, rumen microbial protein supply, parasitic load, immunity and enteric methane emissions of milking animals under field conditions. Animal Feed Science and Technology, 179(1–4), 24–35. https://doi.org/10.1016/J.ANIFEEDSCI.2012.11.005

Grace, D., Bett, B. K., Lindahl, J. F., & Robinson, T. P. (2015). Climate and livestock disease: assessing the vulnerability of agricultural systems to livestock pests under climate change scenarios. https://hdl.handle.net/10568/66595

Haenlein, G. F. W. (2007). About the evolution of goat and sheep milk production. Small Ruminant Research, 68(1–2), 3–6. https://doi.org/10.1016/J.SMALLRUMRES.2006.09.021

Hennessy, D., Luc, D., Agnes van den P-v. D., & Laurence S.. (2020). Increasing Grazing in Dairy Cow Milk Production Systems in Europe. Sustainability, 12(6), 2443. https://doi.org/10.3390/su12062443

Hill, D. L., & Wall, E. (2015). Dairy cattle in a temperate climate: the effects of weather on milk yield and composition depend on management. Animal, 9(1), 138–149. https://doi.org/10.1017/S1751731114002456

Huang, J., Li, Y-F. & Xie, M. (2015). An empirical analysis of data preprocessing for machine learning-based software cost estimation. Information and Software Technology, 67, 108-127.

Ji, B., Banhazi, T., Phillips, C. J. C., Wang, C., & Li, B. (2022). A machine learning framework to predict the next month's daily milk yield, milk composition and milking frequency for cows in a robotic dairy farm. Biosystems Engineering, 216, 186–197. https://doi.org/10.1016/J.BIOSYSTEMSENG.2022.02.013

Kamphuis, C., Mollenhorst, H., Feelders, A., Pietersma, D., & Hogeveen, H. (2010). Decision-tree induction to detect clinical mastitis with automatic milking. Computers and Electronics in Agriculture, 70(1), 60–68. https://doi.org/10.1016/J.COMPAG.2009.08.012

Koc, A., Turk, S., & Şahin, G. (2019). Multi-criteria of wind-solar site selection problem using a GIS-AHP-based approach with an application in Igdir Province/Turkey. Environmental Science and Pollution Research, 26(31), 32298–32310. https://doi.org/10.1007/S11356-019-06260-1/TABLES/10

Kurani, A., Doshi, P., Vakharia, A., & Shah, M. (2023). A Comprehensive Comparative Study of Artificial Neural Network (ANN) and Support Vector Machines (SVM) on Stock Forecasting. Annals of Data Science, 10(1), 183–208. https://doi.org/10.1007/S40745-021-00344-X/TABLES/2

Liang, W., Luo, S., Zhao, G., & Wu, H. (2020). Predicting Hard Rock Pillar Stability Using GBDT, XGBoost, and LightGBM Algorithms. Mathematics, 8(5), 765. https://doi.org/10.3390/MATH8050765

Maphill, (2024). Physical 3D Map of Iğdır. Retrieved from: http://www.maphill.com/turkey/kars/igdir/3d-maps/physical-map/

Marumo, J. L., Lusseau, D., Speakman, J. R., Mackie, M., & Hambly, C. (2022). Influence of environmental factors and parity on milk yield dynamics in barn-housed dairy cattle. Journal of Dairy Science, 105(2), 1225–1241. https://doi.org/10.3168/JDS.2021-20698

Mohamed, S., Rosca, M., Figurnov, M. & Mnih A. (2020). Monte Carlo Gradient Estimation in Machine Learning. Journal of Machine Learning Research, 21(132), 1−62.

Natekin, A., & Knoll, A. (2013). Gradient boosting machines, a tutorial. Frontiers in Neurorobotics, 7, 63623. https://doi.org/10.3389/FNBOT.2013.00021/BIBTEX

Noorunnahar, M., Chowdhury, A. H., & Arefeen, F. (2023). A tree based eXtreme Gradient Boosting (XGBoost) machine learning model to forecast the annual rice production in Bangladesh. PLOS ONE, 18(3), e0283452. https://doi.org/10.1371/JOURNAL.PONE.0283452

Öztürk, Y., Yulu, A., & Turgay, O. (2023). Remote sensing supported analysis of the effect of wind erosion on local air pollution in arid regions: a case study from Iğdır province in eastern Türkiye. Environmental Systems Research, 12(1), 1–23. https://doi.org/10.1186/S40068-023-00294-8

Pereira, P. C. (2014). Milk nutritional composition and its role in human health. Nutrition, 30(6), 619–627. https://doi.org/10.1016/J.NUT.2013.10.011

Perri, A.F., Mejía, M.E., Licoff, N., Lazaro, L., Miglierina, M., Ornstein, A., Becu-Villalobos, D., Lacau-Mengido, I.M. (2011). Gastrointestinal parasites presence during the peripartum decreases total milk production in grazing dairy Holstein cows. Veterinary Parasitology, 178(3–4), 311-318. https://doi.org/10.1016/j.vetpar.2010.12.045.

Polsky, L., & von Keyserlingk, M. A. G. (2017). Invited review: Effects of heat stress on dairy cattle welfare. Journal of Dairy Science, 100(11), 8645–8657. https://doi.org/10.3168/JDS.2017-12651

Putatunda, S., & Rama, K. (2018). A comparative analysis of hyperopt as against other approaches for hyper-parameter optimization of XGBoost. ACM International Conference Proceeding Series, 6–10. https://doi.org/10.1145/3297067.3297080

Sahin Demirel, A.N. (2024). Investigating the effect of climate factors on fig production efficiency with machine learning approach. Journal of the Science of Food and Agriculture. https://doi.org/10.1002/JSFA.13619

Shrestha, D.L. & Solomatine, D.P. (2006). Machine learning approaches for estimation of prediction interval for the model output. Neural Networks, 19(2), 225-235.

Soori, M., Arezoo, B., & Dastres, R. (2023). Artificial intelligence, machine learning and deep learning in advanced robotics, a review. Cognitive Robotics, 3, 54–70. https://doi.org/10.1016/J.COGR.2023.04.001

Thornton, P., Nelson, G., Mayberry, D., & Herrero, M. (2021). Increases in extreme heat stress in domesticated livestock species during the twenty-first century. Global Change Biology, 27(22), 5762–5772. https://doi.org/10.1111/GCB.15825

TOB, (2024). Iğdır İl Tarım ve Orman Müdürlüğü. Retrieved from: https://igdir.tarimorman.gov.tr/

Türkeş, M., & Tatli, H. (2011). Zirai Meteorolojik Açıdan Iğdır İklim Etüdü. Journal of the Institute of Science and Technology, 1(1), 97–104. https://doi.org/10.1002/JOC.1862

Xie, Y., Jiang, B., Gong, E., Li, Y., Zhu, G., Michel, P., Wintermark, M., & Zaharchuk, G. (2019).Use of Gradient Boosting Machine Learning to Predict Patient Outcome in Acute Ischemic Stroke on the Basis of Imaging, Demographic, and Clinical Information. American Journal of Roentgenology, 212(1), 1-232.

Yılmaz, A.C.I. (2022). Milk Yield, Fertility, Udder Characteristics, and Raw Milk Somatic Cell Count of the Damascus Goats Reared in Iğdır Conditions. International Journal of Agricultural and Wildlife Sciences, 8(2), 358–367. https://doi.org/10.24180/IJAWS.1090613