PREDICTION OF HEMP FIBER DAMAGE AND DECORTICATION EFFICIENCY USING RANDOM FOREST REGRESSION

Toshtemirov Ramziddin Toshpo'lat o'g'li
Qarshi davlat texnika universiteti,
"Sanoatda axborot texnologiyalari va turizm" kafedrasi katta o'qituvchisi
Email: toshtemirovramziddin37@gmail.com
Urunov Bekjon Jovli o'g'li
Qarshi davlat texnika universiteti
"Sanoat muhandisligi va menejmenti" kafedrasi katta o'qituvchisi,
Email: bekjonurunov3@gmail.com
Farxodov Sunnat Umar o'g'li
Qarshi davlat texnika universiteti
"Sanoatda axborot texnologiyalari va turizm" kafedrasi katta o'qituvchisi,
Email: s.farkhodov@mail.ru
Samiyev Sardor Salimjon o'g'li
Qarshi davlat texnika universiteti
"Sanoatda axborot texnologiyalari va turizm" kafedrasi katta o'qituvchisi,
Email: sssamiev@stud.etu.ru

DOI: *https://doi.org/10.5281/zenodo.21132001*

Abstract: *This study investigates the prediction of hemp fiber damage and decortication efficiency using a Random Forest Regression model. A dataset consisting of technological process parameters, including retting duration, moisture content, and roller pressure, was used for model development. Additional interaction-based features were introduced to improve prediction performance. The developed model achieved satisfactory prediction accuracy with R^2 values of 0.8248 for Fiber Damage Index and 0.7320 for Decortication Efficiency. Feature importance analysis showed that roller pressure and pressure–moisture interaction had the greatest influence on the process. The obtained results demonstrate the potential of machine learning methods for intelligent optimization of hemp fiber processing systems.*

Keywords: *hemp fiber, decortication, Random Forest Regression, fiber damage, process optimization, moisture content.*

PREDICTION OF HEMP FIBER DAMAGE AND DECORTICATION EFFICIENCY USING RANDOM FOREST REGRESSION

Annotatsiya: *Ushbu tadqiqotda kanop tolasi shikastlanishi va dekortikatsiya samaradorligini Random Forest Regression modeli yordamida bashorat qilish masalasi o'rganildi. Modelni yaratishda retting davomiyligi, namlik miqdori va valik bosimi kabi texnologik parametrlarni o'z ichiga olgan ma'lumotlar to'plamidan foydalanildi. Bashorat aniqligini oshirish maqsadida o'zaro ta'sirga asoslangan qo'shimcha parametrlar ham qo'llanildi. Ishlab chiqilgan model Fiber Damage Index uchun 0.8248 va Decortication Efficiency uchun 0.7320 qiymatdagi R^2 natijalariga erishdi. Tahlil natijalari valik bosimi va bosim–namlik* *o’zaro ta’siri jarayonga eng katta ta’sir ko’rsatishini aniqladi. Olingan natijalar mashinali o’qitish usullarining kanop tolasi qayta ishlash jarayonlarini intellektual optimallashtirishdagi samaradorligini ko’rsatdi.*

Kalit so’zlar: *kanop tolasi, dekortikatsiya, Random Forest Regression, tolalar shikastlanishi, jarayonni optimallashtirish, namlik miqdori*

ПРОГНОЗИРОВАНИЕ ПОВРЕЖДЕНИЙ КОНОПЛЯНОГО ВОЛОКНА И ЭФФЕКТИВНОСТИ ДЕКОРТИКАЦИИ С ИСПОЛЬЗОВАНИЕМ РЕГРЕССИИ СЛУЧАЙНОГО ЛЕСА

Аннотация: *В данной работе исследовано прогнозирование повреждения волокон конопли и эффективности декортикации с использованием модели Random Forest Regression. Для построения модели использовались данные технологических параметров процесса, включая продолжительность реттинга, влажность и давление валков. Для повышения точности прогнозирования были дополнительно использованы параметры взаимодействия факторов. Разработанная модель показала хорошие результаты прогнозирования со значениями R^2, равными 0.8248 для Fiber Damage Index и 0.7320 для Decortication Efficiency. Анализ важности признаков показал, что наибольшее влияние на процесс оказывают давление валков и взаимодействие давления и влажности. Полученные результаты подтверждают эффективность методов машинного обучения для интеллектуальной оптимизации процессов переработки волокон конопли.*

Ключевые слова: *волокно конопли, декортикация, Random Forest Regression, повреждение волокон, оптимизация процесса, влажность*

Introduction: Natural plant fibers are increasingly being used in modern engineering applications due to their low density, biodegradability, renewability, and environmental sustainability. Among natural fibers, hemp fiber has attracted significant attention in recent years for its potential application in bio-composites, lightweight structural materials, automotive components, insulation materials, and environmentally friendly industrial products. Hemp fibers possess relatively high mechanical strength and stiffness compared to many other lignocellulosic fibers, making them suitable for high-performance composite applications [1].

The quality and mechanical properties of hemp fibers strongly depend on harvesting conditions, retting processes, moisture content, and mechanical decortication parameters. During the decortication process, improper processing conditions may cause severe mechanical damage to the fibers, including fiber breakage, shortening, fibrillation, and reduction of tensile properties. Such damage negatively affects fiber quality and limits the use of hemp fibers in advanced engineering applications [2].

Among the most important technological factors affecting fiber damage are retting duration, fiber moisture content, and roller pressure during mechanical processing. Insufficient retting may result in incomplete separation of bast fibers from woody shives, while excessive retting can weaken the cellulose structure of the fibers. Similarly, low moisture content may increase fiber brittleness, whereas excessive moisture can reduce processing efficiency and fiber quality [3]. Roller pressure also plays a critical role in determining the degree of fiber separation and mechanical deformation during decortication. Traditional experimental approaches for determining optimal processing conditions require considerable time, labor, and material resources. In recent years, machine learning methods have been increasingly applied in agricultural engineering, materials science, and process optimization due to their ability to model complex nonlinear relationships between process parameters and output characteristics. In particular, Random Forest Regression has demonstrated high performance in prediction problems involving nonlinear interactions and multivariable technological systems [4].

In this study, a Random Forest Regression model was developed to predict hemp fiber damage and decortication efficiency based on key process parameters, including retting duration, moisture content, and roller pressure. In addition, feature engineering techniques were applied to improve prediction performance by incorporating interaction-based parameters and optimality indicators. The proposed model aims to support intelligent optimization of hemp fiber processing and reduce mechanical damage during decortication.

Main Part: In this study, a machine learning-based approach was developed to predict mechanical damage and processing efficiency during hemp fiber decortication. The research focused on identifying the influence of key technological parameters on fiber quality and optimizing the decortication process using Random Forest Regression.

The dataset used in this study was prepared based on both literature analysis and experimentally consistent process conditions reported in previous studies on hemp fiber processing, retting, and mechanical decortication. Scientific publications related to hemp fiber extraction, retting duration, moisture influence, and decortication behavior were carefully analyzed to determine the most important process parameters affecting fiber damage. In addition, process relationships observed in industrial and laboratory-scale hemp fiber processing systems were considered during dataset preparation. The dataset consisted of 180 samples and included three main input parameters: retting duration, moisture content, and roller pressure. These parameters were selected because they directly affect fiber separation, mechanical stress distribution, and fiber structural integrity during decortication. Two output parameters were used as prediction targets: Fiber Damage Index (%) and Decortication Efficiency (%).

To improve the prediction capability of the model, several feature engineering approaches were applied. Additional interaction-based features were generated from the original parameters, including retting optimality, moisture optimality, pressure–moisture interaction, and pressure–retting interaction. These engineered features allowed the model to better capture nonlinear relationships and process interactions occurring during hemp fiber processing.

The Random Forest Regression algorithm was selected due to its high performance in nonlinear regression problems and its ability to model complex interactions between process variables. The model was trained using 80% of the dataset, while 20% was used for testing and validation. In addition, 5-fold cross-validation was applied to evaluate the stability and generalization performance of the developed model.

The obtained results demonstrated that the proposed model successfully predicted hemp fiber processing characteristics with relatively high accuracy. For the prediction of Fiber Damage Index, the model achieved an R^2 value of approximately 0.82, indicating strong agreement between actual and predicted values. The model also demonstrated acceptable RMSE and MAE values, showing that prediction errors remained within an acceptable range for engineering applications. For Decortication Efficiency prediction, the model achieved an R^2 value above 0.73, confirming the capability of the Random Forest algorithm to model the technological process.

Figure 1 presents the comparison between actual and predicted values of the Fiber Damage Index obtained using the Random Forest Regression model. The majority of prediction points are closely distributed around the diagonal reference line, indicating strong agreement between experimental and predicted values. The developed model achieved a coefficient of determination $R^2=0.8248$, demonstrating that approximately 82.48% of the variation in fiber damage was successfully explained by the selected process parameters and engineered features. The obtained RMSE and MAE values were 2.36 and 1.93, respectively, indicating relatively low prediction errors. Minor deviations between actual and predicted values were mainly observed at higher damage levels above 35%, which is expected due to the nonlinear nature of hemp fiber deformation and fracture mechanisms during decortication. Overall, the results confirm that the proposed model can effectively predict mechanical damage occurring during hemp fiber processing.

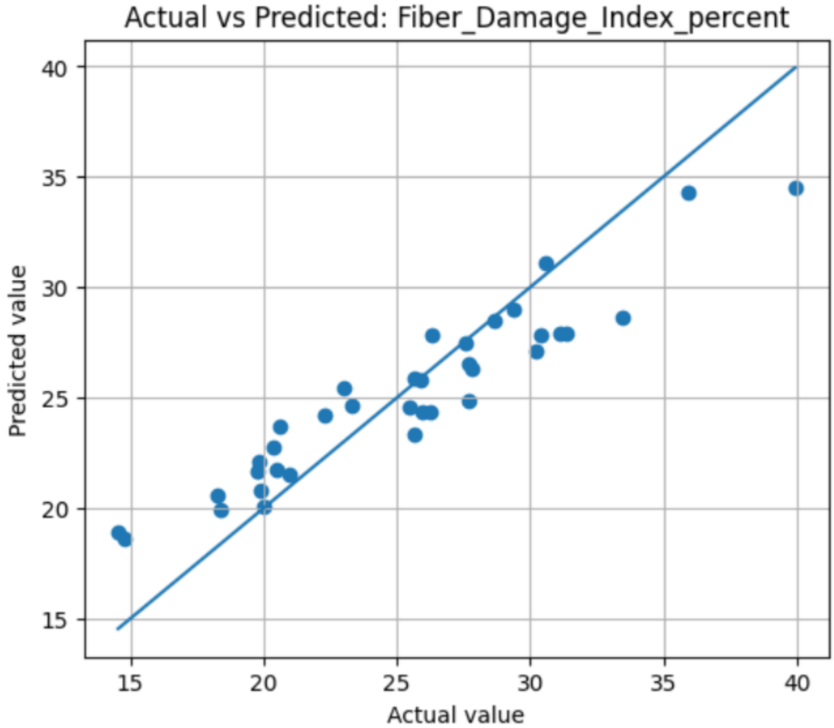

Figure 1. Actual and Predicted values of fiber damage index ssing Random Forest Regression

Figure 2 illustrates the prediction performance of the Random Forest Regression model for Decortication Efficiency. Similar to the previous result, most prediction points are concentrated near the diagonal trend line, confirming the reliability of the developed model. The model achieved an R^2 value of 0.7320, meaning that approximately 73.20% of the variability in decortication efficiency was captured by the machine learning model. The RMSE and MAE values were 2.05 and 1.65, respectively, demonstrating acceptable prediction accuracy for technological process modeling. The graph also shows that the prediction performance remained stable within the efficiency range of 76–87%, where most experimental observations were concentrated. These results indicate that the developed model can successfully estimate hemp fiber separation efficiency under different retting and mechanical processing conditions.

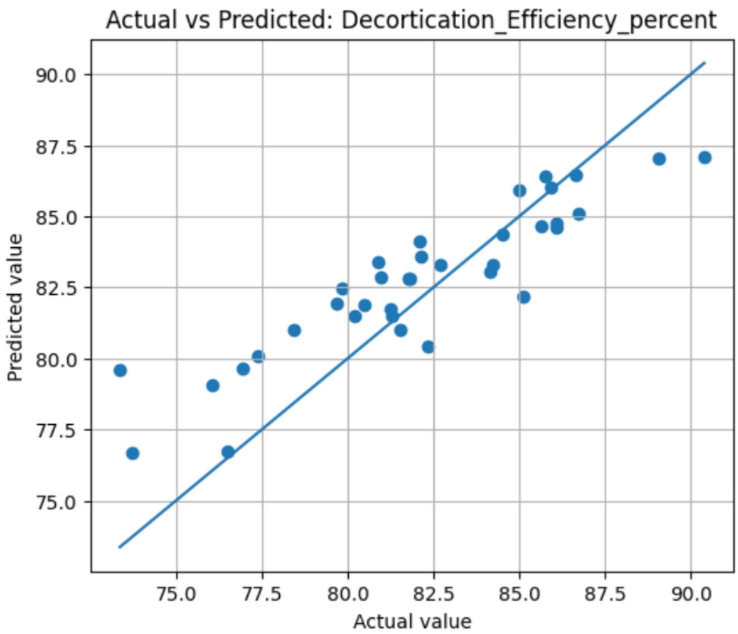

Figure 2. Actual and Predicted values of decortication efficiency esing Random Forest Regression

The importance analysis of the developed Random Forest model revealed that Roller Pressure was the most influential parameter affecting the Fiber Damage Index, with an importance value close to 0.25. In addition, the interaction between pressure and moisture content also had a strong effect on fiber deformation and breakage. As shown in Figure 3, Moisture Optimality and Pressure–Retting Interaction contributed moderately to prediction performance, while Retting Duration and Moisture Content individually had lower influence values below 0.10. These results indicate that interaction-based parameters play an important role in modeling the nonlinear behavior of hemp fiber processing systems.

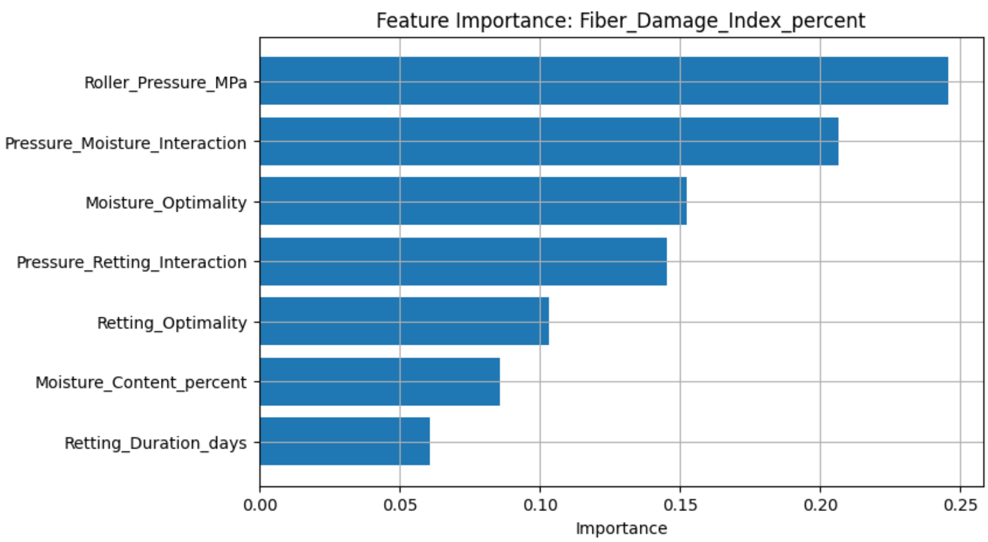

Figure 3. Feature Importance Analysis for Fiber Damage Index Prediction

A similar trend was observed for Decortication Efficiency prediction. Roller Pressure and Pressure–Moisture Interaction were identified as the two most significant factors, with importance values of approximately 0.28 and 0.22, respectively. According to Figure 4, Moisture Optimality and Pressure–Retting Interaction also contributed noticeably to process efficiency prediction. The results confirm that efficient hemp fiber separation depends on both mechanical processing intensity and moisture-related conditions during decortication.

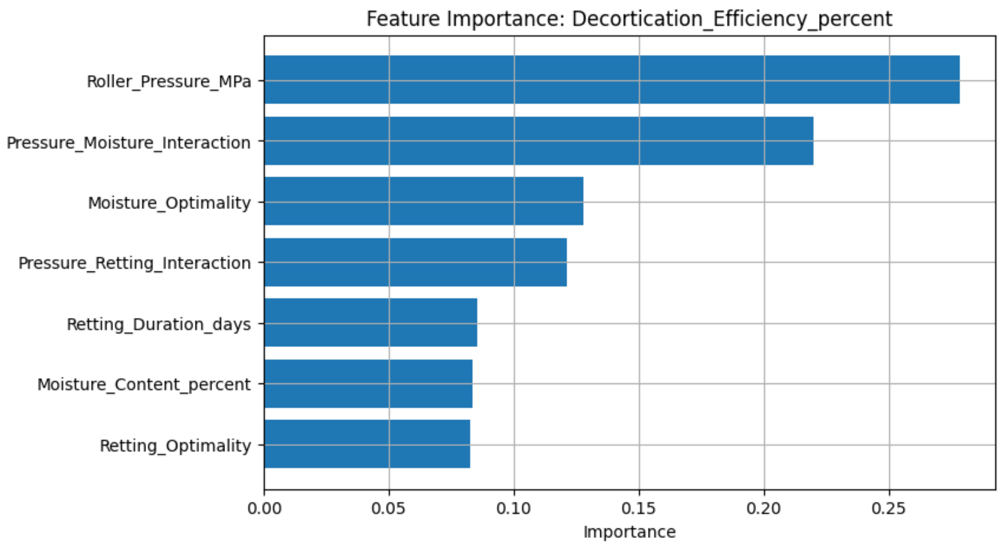

Figure 4. Feature Importance Analysis for Decortication Efficiency Prediction

The overall analysis indicates that machine learning methods can effectively model complex hemp fiber processing systems and provide reliable prediction of fiber damage and processing efficiency. The developed Random Forest Regression model can be used as a decision-support tool for optimizing hemp fiber decortication processes and reducing mechanical damage during industrial processing.

Conclusion: In this study, a Random Forest Regression model was developed to predict hemp fiber damage and decortication efficiency based on key technological parameters, including retting duration, moisture content, and roller pressure. Additional interaction-based features and optimality indicators were also introduced to improve model performance and better represent the nonlinear behavior of the hemp fiber processing system. The obtained results demonstrated that the developed machine learning model can successfully predict the technological characteristics of hemp fiber decortication with relatively high accuracy. The model achieved an R2R^2R2 value of

0.8248 for Fiber Damage Index prediction and 0.7320 for Decortication Efficiency prediction, indicating good agreement between actual and predicted values. The relatively low RMSE and MAE values further confirmed the reliability of the developed model. Feature importance analysis showed that Roller Pressure and Pressure–Moisture Interaction were the most influential parameters affecting both fiber damage and processing efficiency. The results also demonstrated that interaction-based features significantly improved prediction capability compared to using only basic process parameters.

References

[1] C. Pepi, M. E. Pipistrelli, and M. Gioffrè, 'Random modeling of hemp fibers mechanical properties', *Compos. Part Appl. Sci. Manuf.*, vol. 183, p. 108203, Aug. 2024, doi: 10.1016/j.compositesa.2024.108203.

[2] J. P. Manaia, A. T. Manaia, and L. Rodriges, 'Industrial Hemp Fibers: An Overview', *Fibers*, vol. 7, no. 12, p. 106, Dec. 2019, doi: 10.3390/fib7120106.

[3] K. Palanikumar, E. Natarajan, K. Markandan, C. K. Ang, and G. Franz, 'Targeted Pre-Treatment of Hemp Fibers and the Effect on Mechanical Properties of Polymer Composites', *Fibers*, vol. 11, no. 5, p. 43, May 2023, doi: 10.3390/fib11050043.

[4] S. Musio, J. Müssig, and S. Amaducci, 'Optimizing Hemp Fiber Production for High Performance Composite Applications', *Front. Plant Sci.*, vol. 9, p. 1702, Nov. 2018, doi: 10.3389/fpls.2018.01702.