International Journal of Innovative Research and Scientific Studies, 8(3) 2025, pages: 3220-3233

IJIRSS
International Journal of Innovative Research & Scientific Studies

ISSN: 2617-6548

URL: www.ijirss.com

Intelligent System for Predicting Freeze-Drying Parameters for Camel Milk Powder Production Using Sensors and Machine Learning

Ayaulym Rakhmatulina[1], Nurdaulet Tasmurzayev[1,2,3*], Bibars Amangeldy[1,2,3], Zhanar Omirbekova[1,5*], Adilet Kaharov[2,4]

[1]*Joldasbekov Institute of Mechanics and Engineering, Almaty 050000, Kazakhstan.*
[2]*Faculty of Information Technologies, Al-Farabi Kazakh National University, Almaty 050040, Kazakhstan.*
[3]*LLP «DigitAlem», Almaty 050040, Kazakhstan.*
[4]*LLP «Kazakhstan R&D Solutions», Almaty 050056, Kazakhstan.*
[5]*Satbayev University, Almaty 050013, Republic of Kazakhstan.*

Corresponding author: Nurdaulet Tasmurzayev (*Email: tasmurzayev.n@gmail.com*)

Abstract

Producing high-quality camel milk powder requires precise control of vacuum sublimation drying parameters due to the product's sensitivity to thermal stress and oxidation. This study addresses the optimization of the drying process by integrating advanced machine learning (ML) techniques to predict optimal drying conditions. In this research, four ML algorithms were developed and evaluated: Logistic Regression, Random Forest, Decision Tree, and Extreme Gradient Boosting (XGBoost). Each algorithm was trained on detailed sensor data, including chamber temperatures, vacuum pressure, and milk thickness, with drying success defined by maintaining temperature deviations within ±5°C and completing drying cycles within 24 hours. The XGBoost model exhibited the best performance, achieving an accuracy of 99.3%, precision and recall of 96.4%, and the highest F1-score. Temperature parameters, particularly in specific chamber locations, emerged as critical predictors of successful drying outcomes. By enabling accurate forecasting and real-time parameter adjustments, this ML-driven approach significantly enhances drying efficiency, product quality, and sustainability, offering substantial economic and logistical benefits, particularly for remote regions in Kazakhstan where camel milk production is prominent.

Keywords: Camel Milk, Drying optimization, Food processing, Machine learning, Predictive modeling, Sustainable drying processes, Vacuum sublimation drying.

DOI: 10.53894/ijirss.v8i3.7221
Funding: This research was funded by the Committee of Science of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No. BR21881957).
History: Received: 28 March 2025 / **Revised:** 5 May 2025 / **Accepted:** 7 May 2025 / **Published:** 20 May 2025

Competing Interests: The authors declare that they have no competing interests.
Authors' Contributions: All authors contributed equally to the conception and design of the study. All authors have read and agreed to the published version of the manuscript.
Transparency: The authors confirm that the manuscript is an honest, accurate, and transparent account of the study; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained. This study followed all ethical practices during writing.
Publisher: Innovative Research Publishing

1. Introduction

The growing demand for high-quality dairy products has led to the exploration of advanced drying techniques to preserve the nutritional integrity of milk. Among these, vacuum sublimation drying, also known as lyophilization, stands out for its ability to protect heat-sensitive nutrients, such as proteins, vitamins, and bioactive compounds. Camel milk, recognized for its rich nutritional profile and therapeutic properties, has gained significant attention for its potential in both the food industry and nutraceutical applications. However, the challenge remains in optimizing the drying process to maintain the bioactive properties of camel milk while ensuring efficiency and cost-effectiveness.

This study addresses this challenge by applying machine learning (ML) techniques to optimize camel milk powder production's vacuum sublimation drying process. Traditional methods of drying optimization often rely on trial-and-error experimentation, which can be time-consuming and resource-intensive. In contrast, ML offers a more efficient approach by analyzing large datasets and predicting optimal drying parameters, thus reducing the need for extensive physical trials.

In this research, we evaluate several ML algorithms, including Logistic Regression, Random Forest, Decision Tree, and Extreme Gradient Boosting (XGBoost), to identify the best approach for predicting drying success based on various process parameters. By examining key factors such as temperature, vacuum pressure, and milk thickness, the study provides valuable insights into how ML can enhance drying efficiency, product quality, and sustainability in camel milk powder production.

In the following sections, we will review the relevant literature on vacuum sublimation drying and machine learning in food processing, present the experimental setup and methodology, and discuss the results and their implications for future improvements in drying technologies.

2. Literature Review

2.1. Overview of Vacuum Sublimation Drying (Lyophilization)

Vacuum sublimation drying, or lyophilization, is a process widely used to remove moisture from products under low pressure, where ice transitions directly from solid to vapor, bypassing the liquid phase [1, 2]. Initially developed for pharmaceutical applications, this technique is highly valued for preserving heat-sensitive components such as proteins and vitamins, as well as maintaining the original taste, color, and aroma of products [3]. Over time, the use of lyophilization has expanded into the food industry, particularly for producing fruit, vegetable, and dairy powders [4, 5].

2.2. Importance of Parameter Control in Lyophilization

Successful vacuum sublimation is heavily dependent on optimizing various parameters such as shelf temperature, vacuum pressure, temperature increase rates, and freezing modes. Slight deviations from the optimal conditions can lead to structural collapse and detrimental effects on the biological properties of the product [6, 7]. Research has shown that cryoprotectants, such as trehalose and mannitol, play a critical role in stabilizing proteins during the lyophilization process [8, 9], ensuring that their functional properties are retained. Additionally, while lyophilization is an energy- and time-intensive process, it provides superior quality powders compared to thermal drying methods, making it particularly beneficial for high-value products like biologically active supplements [10, 11].

2.3. Camel Milk and Its Unique Nutritional Value

Camel milk has gained significant attention due to its rich content of bioactive compounds such as lactoferrin, immunoglobulins, and insulin-like growth factors [12]. These proteins have demonstrated antidiabetic, anti-inflammatory, and immunomodulatory effects, but they are highly susceptible to degradation under traditional drying methods. Lyophilization has proven effective in preserving the structure of key proteins like caseins and whey proteins in camel milk, thus maintaining its nutritional and functional value [13].

2.4. Comparison of Drying Technologies

Although lyophilization offers numerous advantages, other drying techniques, such as tray and spray drying, are also widely used in the food and pharmaceutical industries. Tray drying is simpler and more cost-effective but can lead to the oxidation of polyphenols, protein denaturation, and vitamin loss due to high thermal exposure [14, 15]. Spray drying, while highly productive, involves high air temperatures that can damage heat-sensitive components such as whey proteins [16, 17]. In the case of camel milk, these losses are undesirable, as they compromise the nutritional value of the final product [18].

2.5. The Role of Machine Learning in Drying Process Optimization

Recent advancements in machine learning (ML) have introduced powerful tools for optimizing drying processes. ML algorithms analyze parameters such as temperature, pressure, and product composition to predict optimal drying conditions. By dynamically adjusting process variables based on real-time data, ML models help minimize the risk of thermal damage, improve energy efficiency, and enhance product quality [19, 20]. These models have shown great promise in predicting the residual moisture content and the preservation of bioactive compounds, making them a valuable asset in optimizing drying technologies, including lyophilization.

2.6. Spray Freeze Drying and Hybrid Techniques

In some cases, spray freeze-drying (SFD), a combination of spraying products into a cryogenic medium followed by lyophilization, is employed. This method, often used in the pharmaceutical industry for producing highly soluble powders, suffers from high energy consumption and equipment complexity, limiting its large-scale application in food production [4]. Similarly, methods such as ultrasound or microwave-assisted drying, when used in combination with infrared heating, aim to speed up the moisture removal process but can lead to uneven drying or local overheating [17, 21].

2.7. Integration of Machine Learning with Lyophilization

The key advantage of integrating ML with lyophilization technology is its ability to optimize drying across various methods, from tray to vacuum sublimation. ML systems can combine data on raw material structure, drying dynamics, and energy consumption, enabling more precise control of the process [22, 23]. For high-value products like camel milk, this integration helps ensure the preservation of bioactive substances while reducing processing time and resource use.

2.8. Our Contribution

Despite the established benefits of lyophilization, there remains a need for further optimization, particularly in the drying of high-value products like camel milk. Our research introduces an innovative approach that leverages machine learning to predict optimal drying conditions for camel milk, based on detailed experimental data from six trials. By training ML models on this data, we can identify the ideal parameters for shelf temperature, vacuum pressure, and freezing rate, reducing the need for labor-intensive trial-and-error adjustments. The result is a more efficient, resource-conserving method for producing high-quality dry camel milk, preserving its bioactive components while minimizing energy consumption.

3. Methodology

3.1. System Architecture

This section describes the architecture and description of the system. Figure 1 shows the overall process flowchart of vacuum freeze-drying, which consists of three main steps: PLC-based automatic control, process data collection and storage, and intelligent analysis using machine learning.

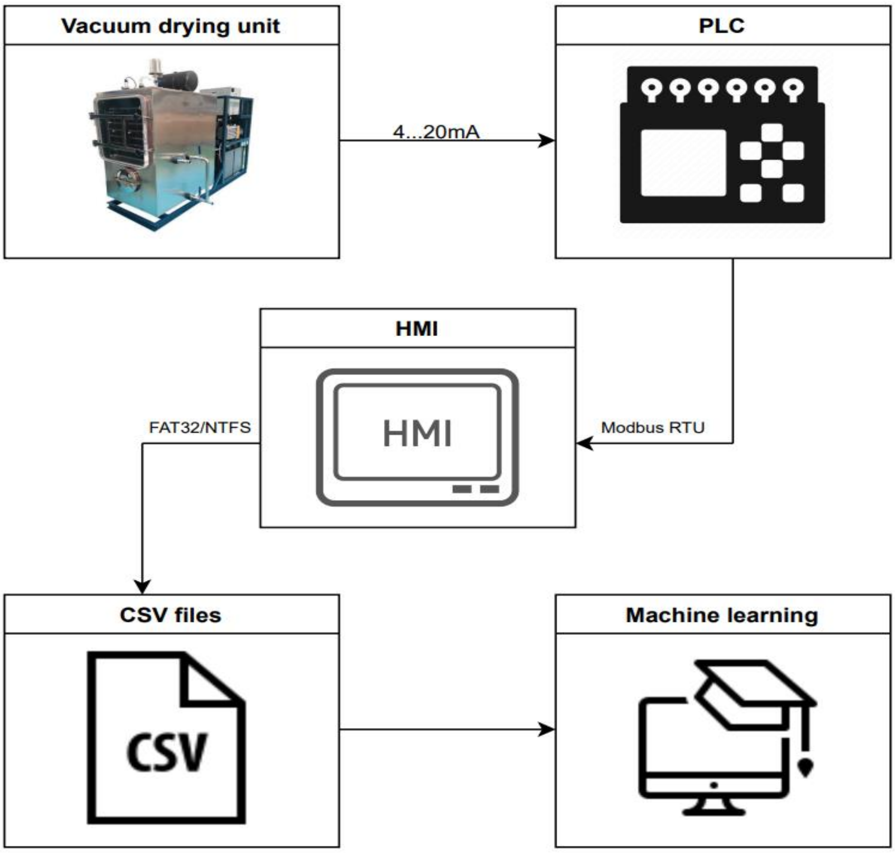

Figure 1.
Description of the architecture of the vacuum freeze-drying system technological process.

In the first stage, the drying process is controlled by a programmable logic controller (PLC) that collects and processes input data from temperature and pressure sensors, which come in the form of analog signals ranging from 4 to 20 milliamps. Based on this data, the PLC automatically adjusts the unit's operating parameters—temperature, vacuum degree, and drying time—to ensure optimal conditions for moisture removal from the product.

The operator interacts with the system in the second stage through the Human Machine Interface (HMI) panel. The HMI allows the operator to monitor current process parameters, receive fault notifications and warnings, and adjust setpoint settings. In parallel, data from sensors and equipment are continuously recorded into CSV files, which are stored on external media with FAT32 or NTFS file systems.

The accumulated data are intelligently analyzed at the third stage using machine learning algorithms. The main purpose of machine learning is to predict possible events and deviations in the drying process. The implementation of machine learning allows for the identification of potential problems in advance and corrects the process, minimizing the probability of negative consequences, as well as optimizing technological parameters, contributing to the improvement of quality and preservation of biologically active components in the finished product.

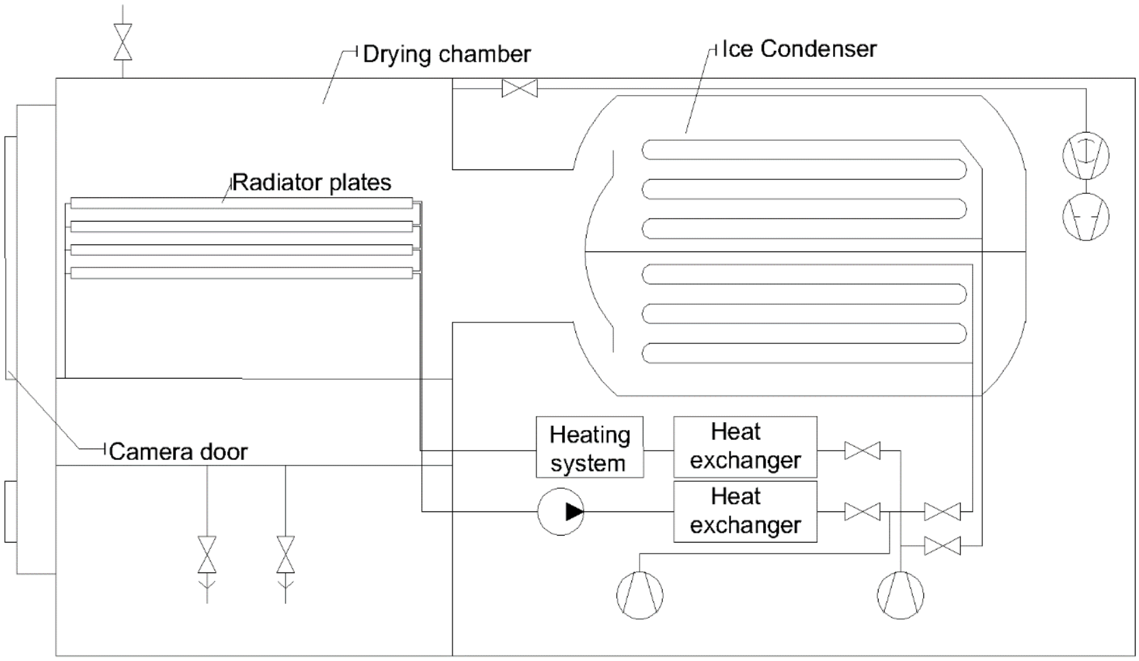

Figure 2.
Structural diagram of vacuum freeze-drying system.

Figure 2 presents the scheme of the vacuum freeze dryer. It includes the main functional blocks: a drying chamber with trays for the material, a cooling system for pre-freezing, a vacuum pump to create low pressure, and a heating system to control the temperature inside the chamber. An important system element is the condensing unit, which condenses water vapor, preventing it from entering the vacuum pump and maintaining process efficiency. Temperature and pressure sensors transmit data to the PLC, which regulates the operation of all plant components.

Figure 3.
General view of the vacuum freeze dryer ZLGJ-300.

Figure 3 presents the ZLGJ-300 freeze dryer, a vacuum freeze-drying equipment consisting of several components, each of which performs a specific function in the process of removing moisture from the material. The drying chamber is the main working volume of the unit where the samples are placed for drying. Inside the chamber are trays on which the material is distributed and thermal panels that provide heat for controlled sublimation. The condensation unit is designed for the condensation of water vapor released during sublimation. It prevents moisture from entering the vacuum pump, ensuring efficient operation of the entire system. The vacuum pump creates and maintains the required vacuum level in the drying chamber, facilitating the conversion of ice into vapor without intermediate transformation into liquid. The compressor is responsible for cooling the condensing unit and maintaining the low temperature required for efficient condensation of water vapor. The heating system regulates the temperature of the shelves inside the drying chamber, providing gradual and controlled heating of the material for even moisture removal. The control valves are used to regulate air supply, system vacuuming, and condensate drainage, ensuring the correct functioning of all process steps. All elements of the system work together to ensure effective vacuum freeze-drying, which removes moisture from the material without destroying its structure and loss of useful properties.

A set point is a target value that the system strives to reach during the process control. In the sublimation drying of camel milk, set points determine the temperature stages, ensuring a controlled heating and moisture removal process. Correctly setting the temperature regimes stabilizes the process and improves the quality of the final product by optimally removing moisture while preserving the structure of the milk.

3.2. Experimental Description

This section presents an analysis of several drying experiments, distinguishing between successful and unsuccessful cases. Successful drying processes are characterized by a consistent and smooth correlation between the predefined set points and actual temperature sensor readings.

Figure 4 provides a detailed comparison between the defined set points and the measured temperature values from sensors installed within the drying chamber. At the initiation of the process, the temperature sharply decreases to approximately -50°C, signifying the freezing stage. Starting at 03:00 on April 14, a gradual temperature increase is observed, corresponding to the sublimation phase, where ice transitions directly to vapor without passing through the liquid phase. This experiment demonstrates a strong correlation between the set temperature values and their respective durations and the actual sensor data. In the concluding phase, the temperature successfully reaches and maintains a stable level of 50°C for a defined period, indicating effective and complete moisture removal from the product.

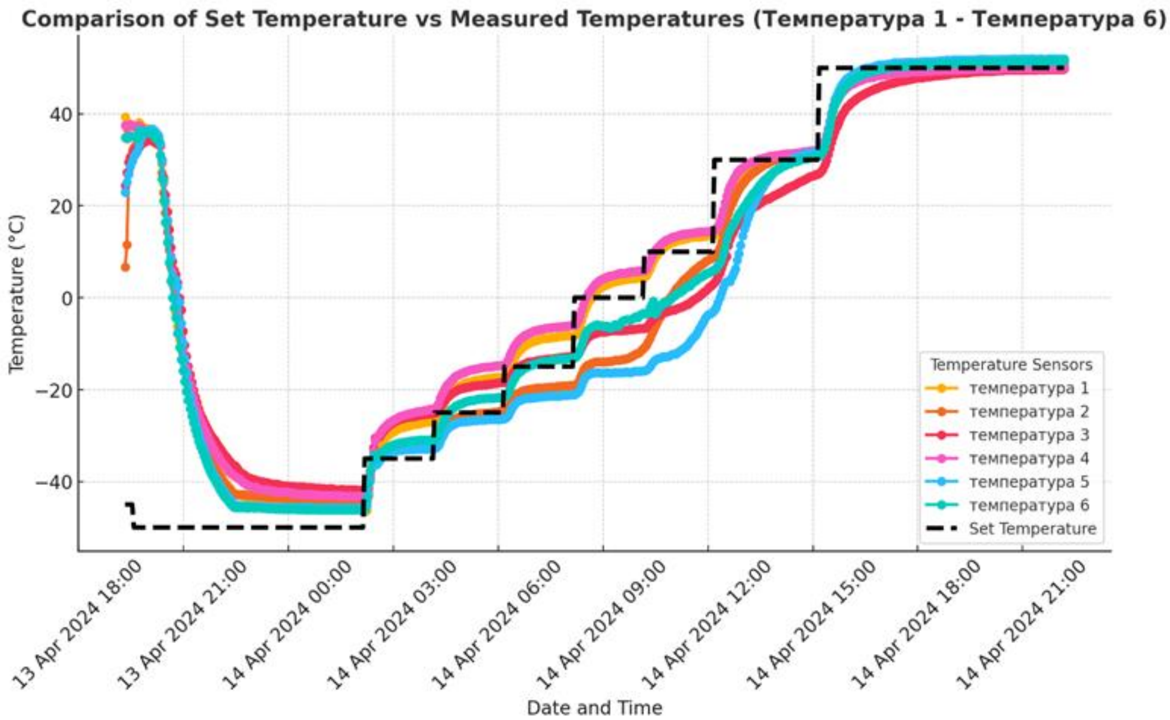

Figure 4.
Comparison of defined setpoints and actual temperature sensor readings during a successful drying process.

Figure 5 illustrates the temperature regime of an unsuccessful drying process, demonstrating a similar dynamic temperature progression. During the initial stage, uniform cooling was observed; however, during the subsequent stepwise heating phase and transition between temperature setpoints, the actual sensor readings lagged behind the predefined values. These discrepancies may be attributed to heating inertia, duration of temperature steps, uneven heat distribution within the drying chamber, and the influence of vacuum pressure conditions. In the final stage of the drying process, although temperature readings stabilized, they remained consistently below the intended set points, indicating possible inadequacies in heat transfer efficiency and incomplete moisture removal.

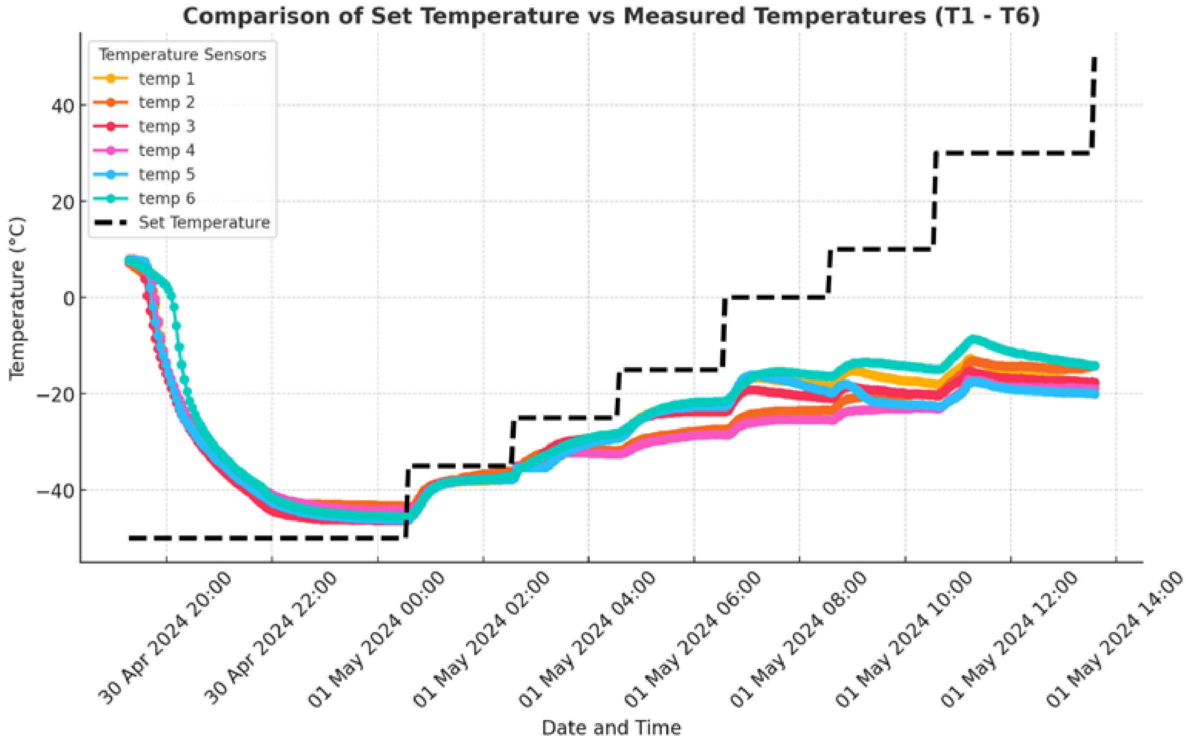

Figure 5.
Temperature profile demonstrating deviations between setpoints and sensor readings in an unsuccessful drying process.

3.3. Dataset Preprocessing

The dataset used in this study was collected from ten experimental trials of vacuum sublimation drying of camel milk. However, after filtering out the low-quality experiments, six trials were retained: three successful and three unsuccessful. Each experiment involved data collection from multiple sensors measuring temperatures in different parts of the chamber (temp_1 - temp_6), vacuum pressure levels, and cold trap indicators. Additionally, the thickness of the camel milk was taken into account. To assess the relationships between key process parameters, a correlation matrix was calculated. This analysis helps determine which variables have the most significant linear correlation with drying success.

The success of the drying process was evaluated based on two main criteria. First, the temperature deviation threshold in the chamber (temp_1 - temp_6) was set to 5°C. If the deviation of any cell's temperature from the set temperature exceeded 5°C at the end of each set, the process was considered unsuccessful. Second, the drying process was required to be completed within 24 hours. If the process exceeded this duration, it was also classified as unsuccessful.

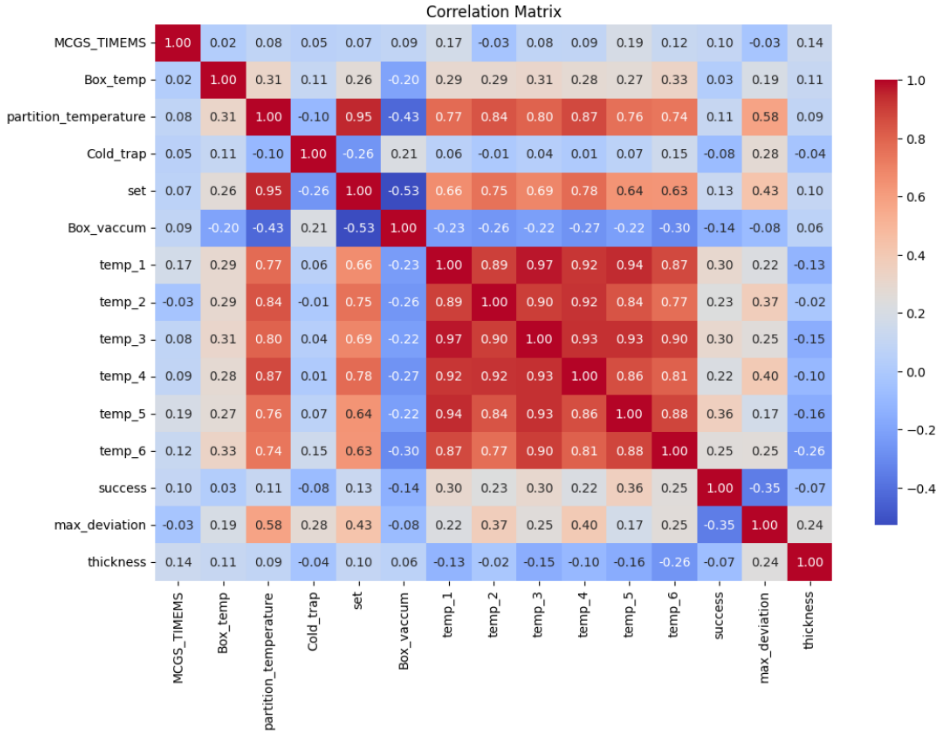

Figure 6.
Correlation Matrix of Key Parameters for Vacuum Sublimation Drying Process.

Significant relationships are observed between the temperature parameters inside the chamber (temp_1 – temp_6), indicating the importance of maintaining a uniform temperature regime within the chamber. To ensure the reliability of the analysis, the dataset was preprocessed by removing missing values and duplicates. The data was then split into training (80%) and testing (20%) subsets for effective model evaluation. The temperature at different levels of the chamber (temp_1 – temp_6) exhibits varying degrees of correlation with the success of the process, with the most significant values observed for temp_3 (0.25) and temp_6 (0.26), confirming the influence of temperature distribution on the final product quality. These observations highlight the importance of controlling the temperature regime and maintaining stable pressure within the chamber to enhance the success of vacuum sublimation drying.

3.4. Machine Learning Modeling

Our approach to modeling the prediction of vacuum sublimation drying success involves testing a range of machine learning algorithms, employing rigorous evaluation metrics, and optimizing model hyperparameters using cross-validation and grid search.

Four classical machine learning algorithms were chosen for this task: Extreme Gradient Boosting, Random Forest, Logistic Regression, and Decision Tree.

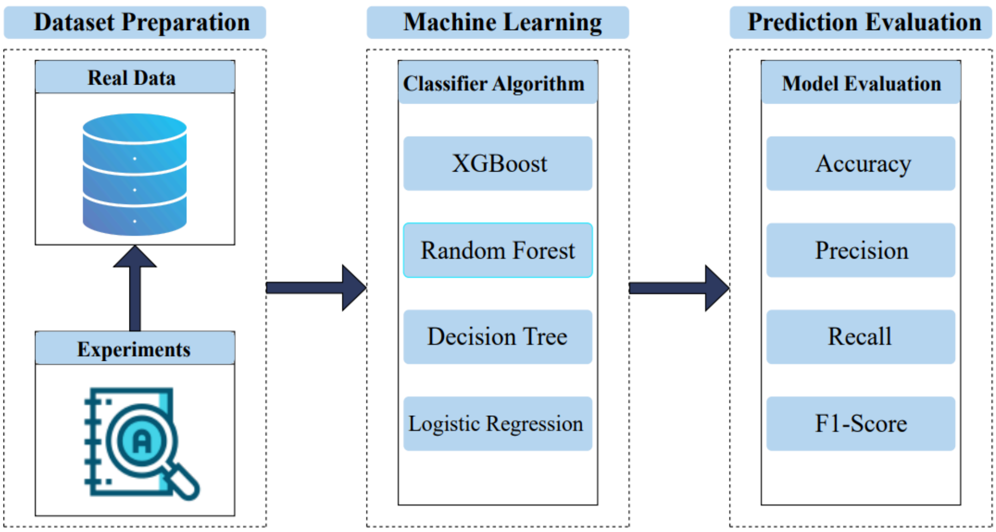

Figure 7.
Intelligent system for metallurgical process optimization.

3.4.1. Logistic Regression

Logistic Regression – a simple yet interpretable statistical model used for binary classification. It assumes a linear relationship between input features and the probability of the positive class. The model predicts the probability that an observation belongs to a particular class (usually the positive class). The logistic function (sigmoid) is used to map predicted values to probabilities. The formula for logistic regression is:

$$p(X) = \frac{1}{1+e^{-(\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_n x_n)}}, \quad (1)$$

where P(x) represents the probability of the positive class given the input features x_1, x_2,...,x_3, β_0 is the intercept term, and β_1,β_2,...,β_n are the coefficients for the corresponding input features. The coefficients β_1, β_2,...,β_n determine the contribution of each feature to the final predicted probability, while β_0 representing the bias or offset of the decision boundary. These coefficients provide insights into the model's decision-making process and the importance of each feature in predicting the target variable [24].

3.4.2. Random Forest

Random Forest – an ensemble learning method that constructs multiple decision trees and aggregates their predictions. Each tree in the forest is trained on a random subset of the data, and its predictions are combined (usually by averaging or majority vote) to make the final prediction. Random Forest is robust against overfitting and captures complex relationships between features. The formula for Random Forest is:

$$\hat{y} = \frac{1}{T}\sum_{t=1}^{T} f_t(x), \quad (2)$$

where $\hat{y}$ is the final prediction of the model, T is the number of trees in the ensemble, and $f_t(x)$ refers to the prediction of the t-th for the input data x [25].

3.4.3. Decision Tree

Decision Tree – A tree-based algorithm that splits the dataset into hierarchical decision rules. It is highly interpretable and effective for discovering key thresholds in drying conditions that determine process success. However, standalone decision trees are prone to overfitting, making them less generalizable compared to ensemble methods [26].

Trees are constructed by minimizing impurities at each node of the tree. At each step of the algorithm, the best feature and threshold are selected that minimize impurities in additional nodes. The formula for calculating Gini impurities:

$$Gini(D) = 1 - \sum_{i=1}^{C} {p_i}^2, \quad (3)$$

where D is the set of data in the node, C is several classes p_i , and is the proportion of objects of class i in the node D.

The goal of a decision tree is to minimize the Gini impurity at each node when splitting data. The lower the Gini impurity, the "purer" the nodes are.

The formula for calculating information (for classification using Entropy):

$$Entropy(D) = -\sum_{i=1}^{C} p_i log_2 p_i, \quad (4)$$

where D is the set of data in the node, p_i the probability that an element from the node D belongs to class i. The algorithm selects features and thresholds that minimize entropy and thus make the nodes as homogeneous as possible.

3.4.4. Extreme Gradient Boosting

Extreme Gradient Boosting (XGBoost) – A powerful gradient boosting algorithm optimized for structured data and high performance. It sequentially builds decision trees, correcting errors made by previous models, making it ideal for capturing complex, nonlinear dependencies. XGBoost also incorporates regularization techniques to prevent overfitting and improve model generalization [27].

The model update formula is:

$$\hat{y}_t = \hat{y}_{t-1} + \eta * f_t(x) \quad (5)$$

$\hat{y}_t$ is the prediction as step t, $\hat{y}_{t-1}$ is the prediction from the previous step, η is the learning rate, and $f_t(x)$ is the prediction from the t-th tree.

The loss function with regularization is:

$$L(\Theta) = \sum_{i=1}^{N} L(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k) \quad (6)$$

Each model was trained to classify the drying process outcome based on experimental sensor data. The PT-100 temperature sensor records the temperature conditions inside the drying chamber. Its platinum sensing element changes resistance in response to temperature, providing measurements from -200°C to +400°C. In the drying system, it monitors the heating of shelves and samples, transmitting data to a programmable logic controller that regulates the sublimation process.

Temperature fluctuations allow the drying conditions to be determined and possible deviations in the process to be recorded [28].

The ZJ-52T vacuum sensor measures the pressure inside the drying chamber in the range of 1×10^{-1} to 1×10^{5} Pa. It detects changes in the vacuum level, which affect the sublimation rate, and transmits the data to the control system. The vacuum pump adjusts the pressure according to the sensor readings, supporting the moisture removal process. The recording of pressure changes helps to detect leaks or insufficient sealing of the chamber [29].

To improve performance, hyperparameter tuning was conducted using grid search, while five-fold cross-validation ensured robust model evaluation. Additionally, feature scaling and preprocessing techniques were applied where necessary to optimize convergence and generalization. This systematic approach allows for an objective assessment of the most effective machine learning algorithm for predicting the success of vacuum sublimation drying.

4.2. Evaluation Metrics

To evaluate the performance of the models, the following metrics were chosen, namely: accuracy, precision, recall, and F1-score.

4.2.1. Accuracy

Accuracy – the proportion of correctly classified cases out of the total number of predictions. This metric provides a general measure of the model’s correctness but may be misleading in cases of class imbalance, where one class is significantly more frequent than another. In such scenarios, accuracy alone does not provide a full assessment of the model's performance. The formula for accuracy is as follows:

$$Accuracy = \frac{TP\ + TN}{TP + TN + FP + FN}, \tag{7}$$

4.5.2. Precision

Precision – the proportion of correctly predicted successful cases among all instances classified as successful. This metric is important for evaluating the model's ability to avoid false positives. A high precision value indicates that the model has a low false positive rate. The formula for precision is as follows:

$$Precision = \frac{TP}{TP + FP}, \tag{8}$$

3.5.3. Recall

Recall – the proportion of actual positive cases (successful instances) that are correctly identified by the model. This metric measures the model's ability to capture all relevant positive cases. A high recall indicates that the model successfully identifies most of the positive cases, reducing the likelihood of missing important instances. The formula for recall is as follows:

$$Recall = \frac{TP}{TP + FN}, \tag{9}$$

3.5.4. F1-Score

F1-score – the harmonic mean of precision and recall, providing a balanced evaluation of classification performance. This metric is especially valuable when there is a trade-off between precision and recall. It helps assess the model's overall reliability in distinguishing between successful and unsuccessful cases, ensuring that both false positives and false negatives are minimized. The formula for the F1-score is as follows:

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}, \tag{10}$$

3.5.5. Area Under the Curve

The area under the curve (AUC) is a commonly used metric to evaluate the performance of classification models. It is calculated by plotting the true positive rate against the false positive rate at various threshold levels, forming the Receiver Operating Characteristic (ROC) curve. The AUC is then obtained by numerically integrating this curve, typically using the trapezoidal rule. This single scalar value reflects the model’s ability to distinguish between classes, with higher values indicating better discriminative performance. As such, AUC provides a robust and interpretable summary of classification effectiveness across all possible threshold settings.

3.5.6. Confusion Matrix

The confusion matrix is a useful tool for evaluating the performance of classification models. It provides a tabular summary of prediction results by comparing actual class labels with those predicted by the model. The matrix includes four key components: true positives, true negatives, false positives, and false negatives. These values offer detailed insight into how well the model distinguishes between classes and allow for the calculation of various performance metrics such as accuracy, precision, recall, and F1-score. One can identify specific patterns of misclassification and better understand the strengths and weaknesses of the model by analyzing the confusion matrix.

4. Results and Discussion

4.1. Model Evaluation Results

The analysis of machine learning models for predicting the success of the vacuum drying process shows varying levels of accuracy and efficiency among the tested algorithms. While XGBoost demonstrated the highest accuracy, other methods also performed well, except for logistic regression, which significantly lagged in all key metrics.

The Random Forest model achieved 99.1% accuracy, along with high precision (96.3%), recall (94.5%), and F1-score (95.4%). These results indicate that the algorithm effectively classifies successful and unsuccessful drying processes with minimal errors. Due to its ensemble nature, the model is resistant to overfitting and can efficiently capture complex dependencies between process parameters.

The Decision Tree model performed slightly worse than the Random Forest, achieving 98.9% accuracy with similarly high precision and recall scores of 94.5%. Despite its strong performance, decision trees are prone to overfitting, especially when dealing with small datasets. While this model provides interpretability and helps identify the most influential factors in the drying process, its predictive ability is slightly lower than that of ensemble methods.

On the other hand, logistic regression was the least effective among all the tested algorithms. Its accuracy was 91.8%, with significantly lower precision (60.0%) and recall (38.2%) compared to the other models. This indicates that logistic regression struggles to accurately predict successful drying processes, leading to a high number of misclassifications. The primary limitation lies in the assumption of linear relationships between features, while the drying process likely follows complex, nonlinear dependencies that this model cannot capture effectively.

Among all tested algorithms, XGBoost demonstrated the best performance, achieving 99.3% accuracy and the highest precision, recall, and F1-score, as seen in Table 1. This indicates that XGBoost is the most reliable model for classifying drying success, minimizing errors, and providing highly accurate predictions. Ensemble methods, such as Random Forest and XGBoost, proved to be particularly effective in this task due to their ability to capture intricate relationships between process parameters. In contrast, the decision tree lacks robustness, and logistic regression exhibited the lowest accuracy, making it the least suitable approach for this classification problem. All model performance results can be seen in in Table 1 and Figure 8.

Table 1.
Performance of each model in predicting vacuum freeze-drying.

Model	Accuracy	Precision	Recall	F1-score
Logistic Regression	91.8%	60.0%	38.2%	46.7%
Random Forest	99.1%	96.3%	94.5%	95.4%
Decision Tree	98.9%	94.5%	94.5%	94.5%
XGBoost	99.3%	96.4%	96.4%	96.4%

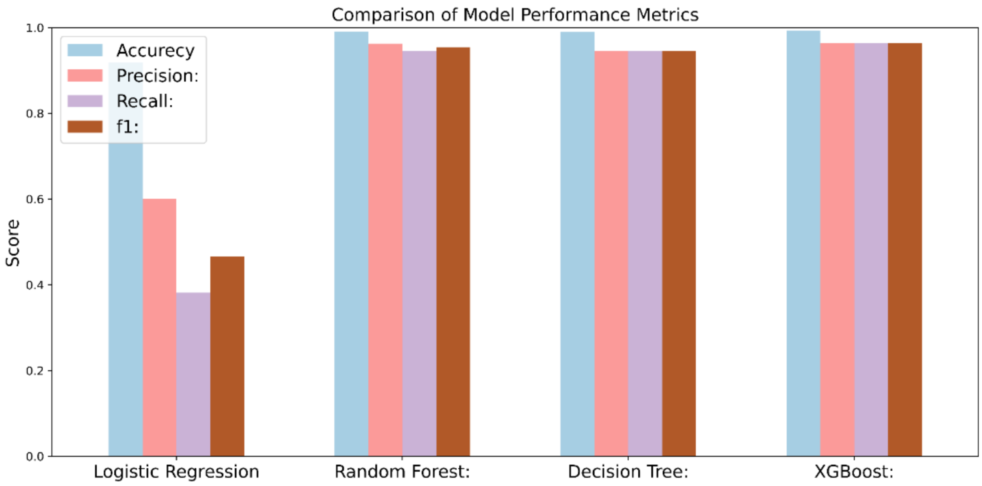

Figure 8.
Performance of each model in predicting vacuum freeze drying (Bar Chart).

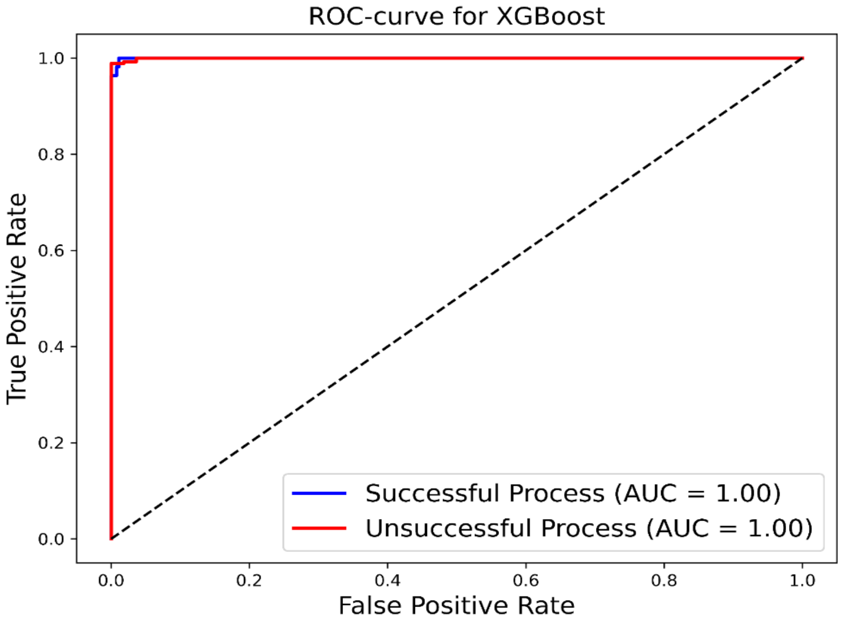

Figure 9.
ROC Curve for XGBoost Model: Performance in Classifying Successful and Unsuccessful Drying Processes.

In Figure 9, we can see the ROC curve for the XGBoost model, which demonstrates its performance at different classification thresholds. The graph shows the True Positive Rate (TPR) on the y-axis and the False Positive Rate (FPR) on the x-axis. The blue and red lines represent the classifications of successful and unsuccessful drying processes, respectively.

Both lines have an AUC (Area Under the Curve) = 1.00, indicating that the model perfectly classifies both classes without any errors. This is reflected in the sharp rise of the curve to the top left corner, demonstrating the model's strong ability to distinguish between successful and unsuccessful processes.

The black dashed line represents a random classifier with an AUC of 0.5, serving as a baseline for comparison. The position of the XGBoost curve well above this line confirms the model's superior performance in classification.

Thus, the ROC curve confirms that the XGBoost model excels in distinguishing between successful and unsuccessful drying processes, achieving perfect classification for both classes in this case.

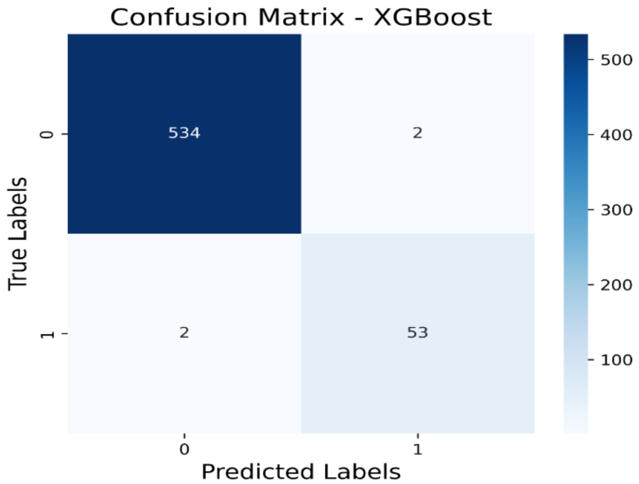

Figure 10.
Confusion Matrix for XGBoost Model: Performance in Classifying Successful and Unsuccessful Drying Processes.

In Figure 10, We can see the confusion matrix for the XGBoost model. The confusion matrix shows the number of correct and incorrect predictions made by the model for the two classes: successful (1) and unsuccessful (0) drying processes.

The top-left corner (True Negative) shows the number of correctly classified unsuccessful drying processes — 534. The top-right corner (False Positive) represents the number of instances where the model incorrectly classified an unsuccessful process as successful — 2. The bottom-left corner (False Negative) shows the number of instances where the model incorrectly classified a successful process as unsuccessful — 2. The bottom-right corner (True Positive) shows the number of correctly classified successful drying processes — 53.

The confusion matrix demonstrates that the XGBoost model makes very few errors: 2 false positives and 2 false negatives, which confirms its high accuracy and its ability to effectively classify both successful and unsuccessful drying processes.

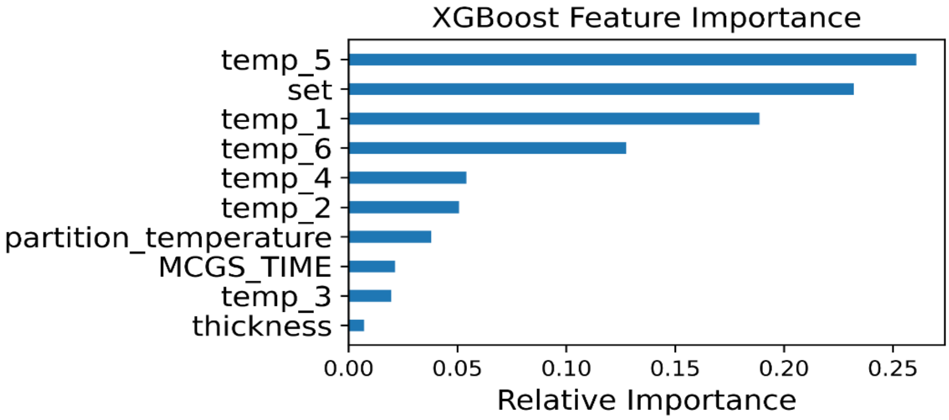

Figure 11.
XGBoost Feature Importance: Impact of Parameters on Drying Process Success Prediction.

In Figure 11, We see the feature importance for the XGBoost model. The chart displays the relative importance of each feature used for predicting the success of the drying process.

The most important feature of the model is temp_5, followed by set and temp_1, indicating their significant impact on the predicted outcomes. Other important features include temp_6, temp_4, and temp_2, which play a crucial role in the model when determining drying success.

Less important features include partition_temperature, MCGS_TIME, temp_3, and thickness, which have lower relative importance values. These results confirm that the temperature in different parts of the chamber, as well as the setup mode and some time parameters, has the most significant influence on the model's predictions.

The chart shows that XGBoost effectively utilizes key features related to temperature conditions for accurately classifying successful and unsuccessful drying processes.

4.2. Interpretation of Results

This study used machine learning (ML) models to optimize the vacuum sublimation drying process of camel milk, a product valued for its bioactive proteins and nutritional properties. The results obtained from the model evaluations provided meaningful insights into the dynamics of the drying process and the role of various parameters in determining the quality of the final product.

The study's primary focus was to identify the most crucial factors affecting the success of the drying process, particularly the chamber temperature, vacuum pressure, and milk thickness. Among the four ML algorithms tested, the XGBoost model performed the best, achieving an accuracy of 99.3% in predicting the successful outcomes of the drying process. The algorithm demonstrated high precision (96.4%), recall (96.4%), and F1-score (96.4%), indicating its effectiveness in classifying successful versus unsuccessful drying processes with minimal error.

A critical observation from the results was the importance of maintaining a uniform temperature distribution inside the drying chamber, especially the temperature readings at specific chamber locations (temp_5 and temp_1). The significant correlation between temperature distribution and drying success reaffirms the sensitivity of camel milk to temperature variations, which can influence the preservation of heat-sensitive bioactive compounds, such as proteins and vitamins.

The dataset analysis also revealed the complexity of the drying process, where slight deviations from the optimal conditions, such as temperature fluctuations exceeding ±5°C from the set points, led to failures in the drying process. This highlights the need for precise control and real-time adjustments, a capability enabled by machine learning models. ML algorithms, by dynamically adjusting drying parameters based on sensor data, minimize the risk of quality deterioration and allow for efficient moisture removal, reducing energy consumption and processing time.

Furthermore, the integration of ML also shows promising potential in enhancing the sustainability and economic feasibility of camel milk production. The intelligent systems not only predict optimal drying parameters but also help identify potential risks, such as protein denaturation or oxidation, which could compromise the nutritional integrity of the product. By optimizing the drying conditions, these systems contribute to the production of high-quality camel milk powder, meeting market demand for a nutritious and long-storable product.

In conclusion, the study's findings emphasize the transformative potential of machine learning in optimizing complex drying processes. The XGBoost model proved to be the most effective tool for predicting drying outcomes, providing valuable insights into the parameters that influence the success of vacuum sublimation drying. This approach can significantly reduce experimental effort, optimize energy consumption, and improve product quality, positioning ML-driven drying systems as a key technology for the food industry, especially in regions like Kazakhstan, where camel milk is a valuable agricultural product.

5. Conclusion

This research demonstrates the effectiveness of integrating advanced ML algorithms into vacuum sublimation drying processes for camel milk, significantly enhancing both product quality and process efficiency. Among the tested algorithms, Extreme Gradient Boosting (XGBoost) delivered the highest accuracy (99.3%), effectively capturing complex interactions between critical drying parameters. These findings underscore the importance of precise temperature and vacuum control in preserving valuable bioactive components and ensuring the production of a porous, easily soluble, and nutritious powder. The study highlights not only the potential to optimize drying parameters proactively but also the broader economic implications, particularly for Kazakhstan's arid regions, where efficient storage and transportation of camel milk are critical challenges. By reducing drying times and energy consumption through precise, ML-driven real-time adjustments, the proposed approach enhances both the sustainability and economic viability of camel milk processing. Future research should expand the experimental dataset to further validate and refine the ML models and incorporate standardized milk pre-processing techniques to better manage variations in milk composition. Additionally, integrating advanced sensor technologies, including moisture measurement systems, would provide greater accuracy in monitoring and controlling the drying process. Such developments will further consolidate the role of intelligent drying systems in transforming food processing practices, aligning economic goals with environmental sustainability and resource optimization.

Abbreviations:

DOAJ	Directory of open-access journals
TLA	Three letter acronym
LD	Linear dichroism
XGBoost	Extreme Gradient Boosting
TP	True Positive
TN	True Negative
FP	False Positive
FN	False Negative
HMI	Human Machine Interface
PLC	Programmable logic controller
FAT32	File Allocation Table
CSV	Comma-Separated Values
NTFS	New technology file system
ZJ-52T	Vacuum freeze dryer
PT-100	Temperature sensor
ROC	Receiver operating characteristic

References

[1] T. Noufeu *et al.*, "Overview of glycometabolism of lactic acid bacteria during freeze-drying: Changes, influencing factors, and application strategies," *Foods,* vol. 14, no. 5, p. 743, 2025. https://doi.org/10.3390/foods14050743

[2] A. Rakhmatulina, A. Altybay, N. Imanbayeva, S. Bagitova, and A. Baikonys, "Numerical Simulation and Experimental Analysis of Mare's Milk Sublimation Drying," *Processes,* vol. 13, no. 1, p. 206, 2025. https://doi.org/10.3390/pr13010206

[3] M. Ioannou Sartzi, D. Drettas, M. Stramarkou, and M. Krokida, "A comprehensive review of the latest trends in spray freeze drying and comparative insights with conventional technologies," *Pharmaceutics,* vol. 16, no. 12, p. 1533, 2024. https://doi.org/10.3390/pharmaceutics16121533

[4] V. Karunnanithy *et al.*, "Effectiveness of lyoprotectants in protein stabilization during lyophilization," *Pharmaceutics,* vol. 16, no. 10, p. 1346, 2024. https://doi.org/10.3390/pharmaceutics16101346

[5] A. Boruah *et al.*, "Impact of tray and freeze drying on physico-chemical and functional properties of underutilized Garcinia lanceifolia (Rupohi thekera)," *Foods,* vol. 14, no. 4, p. 705, 2025. https://doi.org/10.3390/foods14040705

[6] E. Jakubczyk, E. Tryzno-Gendek, A. Kot, A. Kamińska-Dwórznicka, and D. Nowak, "Pre-treatment impact on freeze-drying process and properties of dried blueberries," *Processes,* vol. 13, no. 2, p. 537, 2025. https://doi.org/10.3390/pr130205

[7] G. A. Jeong and C. J. Lee, "Quality characteristics of gluten-free noodles according to the drying method of pregelatinized rice flour," *Processes,* vol. 13, no. 2, p. 462, 2025. https://doi.org/10.3390/pr13020462

[8] J. Mazár, K. Albert, Z. Kovács, A. Koris, A. Nath, and S. Bánvölgyi, "Advances in spray-drying and freeze-drying technologies for the microencapsulation of instant tea and herbal powders: The role of wall materials," *Foods,* vol. 14, no. 3, p. 486, 2025. https://doi.org/10.3390/foods14030486

[9] C. T. Lai *et al.*, "GenV: Preservation of Human Milk for Biological Discovery," in *Proceedings,* 2025, vol. 112, no. 1: MDPI, p. 10.

[10] S. S. Hasson *et al.*, "In vitro apoptosis triggering in the BT-474 human breast cancer cell line by lyophilised camel's milk," *Asian Pacific Journal of Cancer Prevention,* vol. 16, no. 15, pp. 6651-6661, 2015. https://doi.org/10.7314/APJCP.2015.16.15.6651

[11] A. Al-Thaibani *et al.*, "Camel milk whey powder formulated using thermal (spray-drying process) and non-thermal (ultrasonication) processing methods: Effect on physicochemical, technological, and functional properties," *Ultrasonics Sonochemistry,* vol. 111, p. 107097, 2024. https://doi.org/10.1016/j.ultsonch.2024.107097

[12] B. Bouras *et al.*, "Effect of freeze-dried camel rennet extract on coagulation of camel–goat milk mixture and characterization of the cheese obtained," *Foods,* vol. 14, no. 3, p. 334, 2025. https://doi.org/10.3390/foods14030334

[13] A. B. Dogan *et al.*, "Stabilizing milk-derived extracellular vesicles (mEVs) through lyophilization: a novel trehalose and tryptophan formulation for maintaining structure and Bioactivity during long-term storage," *Journal of Biological Engineering,* vol. 19, no. 1, p. 4, 2025. https://doi.org/10.1186/s13036-024-00470-z

[14] N. Harizi, J. Madureira, A. Zouari, M. A. Ayadi, S. Cabo Verde, and N. Boudhrioua, "Effects of spray drying, freeze drying and gamma irradiation on the antioxidant activities of camel and cow milk fractions," *Processes*, vol. 11, no. 3, p. 897, 2023. https://doi.org/10.3390/pr11030897
[15] H. Li, J. Zhao, P. Jia, H. Ou, and W. Zhao, "Optimization of Bulk Cargo terminal unloading and outbound operations based on a deep reinforcement learning framework," *Journal of Marine Science and Engineering*, vol. 13, no. 1, p. 105, 2025. https://doi.org/10.3390/jmse13010105
[16] Z. Zou *et al.*, "Physicochemical properties and whey proteomes of camel milk powders produced by different concentration and dehydration processes," *Foods*, vol. 11, no. 5, p. 727, 2022. https://doi.org/10.3390/foods11050727
[17] F. Adoui *et al.*, "Unravelling the behaviour of camel milk caseins during hydrolysis and separation: Insights into proteomic profiles of caseins using high performance liquid chromatography and polyacrylamide gel electrophoresis," *Food and Humanity*, vol. 4, p. 100491, 2025. https://doi.org/10.1016/j.foohum.2024.100491
[18] D. Wang, L. Li, L. Bai, and Y. Chen, "Effects of Lacticaseibacillus casei Zhang addition on physicochemical properties and metabolomics of fermented camel milk during storage," *Food Chemistry: X*, vol. 26, p. 102318, 2025. https://doi.org/10.1016/j.fochx.2025.102318
[19] D. Donno *et al.*, "Freeze-drying for the reduction of fruit and vegetable chain losses: A sustainable solution to produce potential health-promoting food applications," *Plants*, vol. 14, no. 2, p. 168, 2025. https://doi.org/10.3390/plants14020168
[20] W. Kriangkrai, S. Puttipipatkhachorn, P. Sriamornsak, and S. Sungthongjeen, "Design and evaluation of new gel-based floating matrix tablets utilizing the sublimation technique for gastroretentive drug delivery," *Gels*, vol. 10, no. 9, p. 581, 2024. https://doi.org/10.3390/gels10090581
[21] Y. Liu, Z. Zhang, and L. Hu, "High efficient freeze-drying technology in food industry," *Critical Reviews in Food Science and Nutrition*, vol. 62, no. 12, pp. 3370-3388, 2022. https://doi.org/10.1080/10408398.2020.1865261
[22] P. Mudgil, M. Alalawi, M. Alghaithi, A. T. Mustapha, and S. Maqsood, "Modification of camel milk proteins through complexation with green solvent extracted date seed polyphenols: Effects on physiochemical, technofunctional properties and bioactive properties," *International Dairy Journal*, vol. 164, p. 106188, 2025. https://doi.org/10.1016/j.idairyj.2025.106188
[23] J. Wang *et al.*, "New insights into the acid gelation of camel milk casein," *Journal of Agriculture and Food Research*, p. 101746, 2025. https://doi.org/10.1016/j.jafr.2025.101746
[24] D. W. Hosmer Jr, S. Lemeshow, and R. X. Sturdivant, *Applied logistic regression*. John Wiley & Sons. https://doi.org/10.1002/9781118548387, 2013.
[25] L. Breiman, "Random forests," *Machine Learning*, vol. 45, pp. 5–32, 2001. http://dx.doi.org/10.1023/A:1010950718922
[26] R. J. Lewis, "An introduction to classification and regression tree (CART) analysis," in *Annual Meeting of the Society for Academic Emergency Medicine in San Francisco, California*, 2000, vol. 14: Citeseer.
[27] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd acm Sigkdd International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 785-794.
[28] Grundfos, "PT-100 temperature sensor," Retrieved: https://www.grundfos.com/uk/learn/research-and-insights/pt-100-sensor. [Accessed 10 March 2025], 2025.
[29] Aliexpress, "ZJ-52T Vacuum sensor," Retrieved: https://aliexpress.ru/item/1005008512745634.html. [Accessed 10 March 2025], 2025.

Appendix A.

Appendix A.1

	Mean	Std.	Min.	25%	50%	75%	Max.
MCGS_TIMEMS	278.94	29.03	219.0	271.00	284.00	305.00	314.0
Box_temp	-0.67	8.28	-41.6	-5.00	-0.10	5.10	29.0
partition_temperature	-3.91	35.87	-50.7	-39.00	-1.55	30.10	50.8
utection	925.70	0.00	925.7	925.70	925.70	925.70	925.7
Cold_trap	-57.19	29.79	-92.1	-72.93	-66.70	-56.20	38.6
set	-6.08	37.82	-50.0	-50.00	-15.00	30.00	50.0
Box_vaccum	17372.72	37959.72	0.1	0.35	2.30	15.00	110000.0
temp_1	-16.14	26.66	-49.5	-35.70	-22.60	-5.98	50.4
temp_2	-12.80	29.61	-51.4	-36.40	-22.40	6.50	51.2
temp_3	-16.42	26.27	-51.6	-35.90	-21.80	-7.68	51.0
temp_4	-12.22	30.87	-51.4	-36.70	-22.80	10.75	51.3
temp_5	-19.23	25.10	-50.8	-37.02	-22.20	-16.18	52.0
temp_6	-12.67	27.45	-46.8	-33.50	-18.95	0.80	51.6
success	0.09	0.29	0.0	0.00	0.00	0.00	1.0
thickness	27.27	22.09	1.5	8.30	19.55	41.72	89.5

Table A1. This is a table caption.