

Article

Reliability Reserve: A Markov Chain-Based Metric for Real-Time Operator Decision Support in Ayrar Fermentation

Zhanagul Doumchariyeva ^{1,2}, Jamalbek Tussupov ¹, Madina Sambetbayeva ¹, Tamara Zhukabayeva ¹, Madina Yessenaliyeva ³, Begzhan Kalemshariv ⁴, Sagi Issayev ^{3,*} and Munaram Khassanova ^{3,*}

¹ Department of Information Systems, L.N. Gumilyov Eurasian National University, Astana 010000, Kazakhstan; doumchariyeva@gmail.com (Z.D.); tussupov@mail.ru (J.T.); sambetbayeva_ma_1@enu.kz (M.S.); tamara.kokenovna@gmail.com (T.Z.)

² Department of Information Technology, Turan–Astana University, Astana 010000, Kazakhstan

³ Department of Applied Informatics and Programming, Taraz University Named After M.Kh. Dulaty, Taraz 080000, Kazakhstan; esenalieva_madina@mail.ru

⁴ Department of Food Technology and Processing Products, Saken Seifullin Kazakh Agrotechnical Research University, Astana 010000, Kazakhstan; begjan.ae@gmail.com

* Correspondence: sagiissayev.1993kz@gmail.com (S.I.); munaram89@gmail.com (M.K.)

Abstract

This study presents a Markov chain-based metric called the Reliability Reserve (τ), designed to estimate the time available for operator intervention during the ayran fermentation process. This indicator can be integrated into a digital twin forecast management system. The fermentation process was obtained using a DTMC (discrete-time Markov chain) and divided into five states according to pH (S1–S5). Laboratory samples were prepared from premium-grade cow's milk sourced from the Zher-Ana farm and divided into three experimental groups: Control (without additives), Opt1 (3% additive), and Opt2 (4% additive). A sequence of states was created for the three studied groups (Control, Opt1, and Opt2), and the transition states of the matrix were calculated. The Reliability Reserve quantifies how much time is left before the system transitions from the target state to the acidification state. For the first group, P_{45} was 0.200, corresponding to $\tau = 26.9$ min. The incorporation of functional additives increased P_{45} to 0.250, reducing τ to 20.9 min and shortening the operator intervention window by approximately 6 min. Markov chains were constructed using 101 interpolated time points obtained from five experimental pH measurements for each group. The original pH values were recorded at 2, 4, 6, 8, and 10 h of fermentation and linearly interpolated with a step size of 0.1 h to improve temporal resolution. Model robustness was evaluated using sensitivity analysis (± 0.05 pH boundary shifts) and bootstrap resampling ($n = 1000$, 95% confidence intervals). The concept of Reliability Reserve is a practical decision-making tool in real time. It offers an alternative to traditional reliability indicators such as MTTF and allows integration into digital twin-based control systems.



Academic Editors: Sarawut Ninsawat and Jaturong Som-ard

Received: 30 April 2026

Revised: 19 June 2026

Accepted: 29 June 2026

Published: 3 July 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Keywords: markov chain-based metric; fermentation kinetics; reliability reserve; digital twin; pH monitoring; bioprocess reliability; automatic control

1. Introduction

Currently, the food industry is moving towards digitalization within the framework of Industry 4.0. One of the main directions of this change is the development of digital representations for production processes that allow efficient monitoring and management of technological systems [1–3]. These digital representations make it possible to create virtual

copies of physical processes and support repeated analysis and decision-making [4–6]. In addition, many existing digital twin models are still based on deterministic approaches. Such models often fail to reflect the stochastic nature of real systems, especially in bioprocesses where stochastic variability plays an important role [5–7]. In this context, integrating Markov chain models into digital twin applications opens up new opportunities, as it allows more efficient consideration of the probabilistic behavior of the system [3,8].

The concept of the digital twin has evolved significantly with the rapid development of Industry 4.0, the Internet of Things (IoT), big data, and artificial intelligence technologies, enabling widespread applications in industrial and bioprocess systems [3–5]. A digital twin is a virtual copy of a physical system [3]. Digital twin technologies enable the integration of physical and virtual systems through continuous data acquisition, simulation, and predictive analysis. Recent studies have demonstrated their application in manufacturing, biological systems, and bioprocess engineering for process monitoring, optimization, and decision support [3–6,9]. Grimstad et al. [10] and Krupitzer et al. [11] considered digital twins for maintenance. As for the bioprocess technology, De Giacomo et al. [1] linked Markov decision-making processes with digital twins, and Ray et al. [12] linked artificial intelligence with yogurt fermentation.

The dairy industry is one of the industries that is actively involved in digital transformation. Ayran is a dairy product containing probiotics. An important object for such studies is the traditional low-viscosity fermented milk product, which is currently widely used in Central Asia and the Middle East [13,14]. The fermentation stage is the key stage in the production of ayran and involves complex biochemical processes. The quality of the product during fermentation is determined, first of all, by the pH dynamics, which depends on the permeability of microbes, the composition of raw materials and the course of processing. Even minor changes in this process can lead to instability or product defects [15,16]. In recent years, several studies have focused on the microbiological and technological characteristics of Kazakhstani ayran [14]. Abilkhadirov et al. reported that traditional Kazakh ayran contains unique strains of *Lactobacillus* and *Bifidobacteria* with high probiotic potential [17]. Shalabi et al. examined the microstructure and rheological properties of ayran produced using commercial starter cultures and reported a correlation between starter culture composition and the viscosity of the final product [18]. These findings indicate the need for specialized and adapted ayran fermentation control systems that consider the microbiological and physicochemical characteristics of traditional ayran [14,17,18].

From an industrial perspective, the proposed metric holds direct relevance for the dairy sector of Kazakhstan and other Central Asian countries, where traditional fermented products such as ayran and related dairy beverages remain important components of food production. Maintaining consistent product quality during fermentation is challenging because pH dynamics are affected by microbial activity, raw material variability, and environmental conditions. In many small- and medium-scale production facilities, process supervision still relies largely on periodic measurements and operator experience. The Reliability Reserve metric introduced in this study provides an interpretable estimate of the remaining intervention time before the process reaches an undesirable over-acidification state. When integrated into a digital twin framework, this indicator can support proactive process management, reduce product losses caused by fermentation deviations, improve batch-to-batch consistency, and facilitate the future adoption of predictive control technologies in industrial dairy fermentation systems.

In Kazakhstan, fermented dairy products constitute a significant segment of the dairy market. According to the Bureau of National Statistics, the provision of domestic production for yogurt, kefir, and other fermented products reached 84% in 2023 [19]. The total volume

of fermented (cultured) products produced in Kazakhstan amounted to 231,693 tons in 2023 [19], while yogurt and fermented milk production grew by 4% in 2024 [20].

However, the dairy industry faces significant challenges. Only 20% of produced raw milk complies with the technical regulations of the Customs Union [21], and dairy processing enterprises are loaded with raw materials at only 77% of their capacity [22]. The milk processing rate is only 33% [23], indicating substantial inefficiencies in the production chain. Furthermore, the industry suffers from low profitability: low prices for dairy products and high production costs limit the ability of farmers and producers to generate adequate returns [24]. The annual economic loss in the meat and dairy sector is estimated at approximately \$800 million [23].

The proposed Reliability Reserve metric aims to reduce quality-related losses and improve process efficiency by providing operators with a predictive intervention tool.

Three main approaches are used to model the fermentation of dairy products. Kinetic models based on logistic functions and the Gompertz model describe the average course of fermentation but do not account for stochastic variability [25,26]. Machine learning methods (neural networks, gradient boosting) show high accuracy in predicting pH changes [27,28]; however, the disadvantage is the “black box” principle and the difficulty of interpreting the results [29]. Hybrid models combine the advantages of both approaches. Kennedy and O’Hagan [27] developed Bayesian calibration methods, while Rasmussen and Williams [30] proposed Gaussian processes. pH monitoring is crucial in the dairy industry. Aydogdu et al. [15] examined the basics of pH, Arango et al. [31] investigated the inline observation method using infrared sensors, while Muncan et al. studied real-time observation using infrared spectroscopy [32].

Continuous time in the network of chemical reactions is described by the Markov chains [8,33]. However, Anderson and Kurtz [8] applied them to chemical reactions, and Fearnhead et al. [33] developed an applied forecast. In technical systems, the Markov model has become a standard tool for assessing reliability and predicting failures [10]. For example, Grimstad et al. [10] developed a danger and risk-based maintenance method. The same approach has also found successful application in agriculture: Markov chains are used to predict dairy cow diseases with 85–92% accuracy [34], to model lactation curves [35], and to determine the expected calving time with an accuracy of 89% [36].

Despite the proven effectiveness of Markov chains in modeling biological systems such as dairy cow disease prediction [34], lactation curve analysis [35], and calving time estimation [36], their application to dairy product fermentation processes—particularly traditional fermented milk products such as ayran—remains largely unexplored. Most fermentation studies rely on deterministic kinetic models [25,26] or machine-learning approaches [28,30,37], whereas probabilistic state-transition frameworks have received comparatively little attention. Recent advances in smart fermentation technologies have demonstrated the value of real-time monitoring and predictive process management [38]. However, no previous study has proposed a Markov-derived indicator specifically designed to quantify the remaining operator intervention window during fermentation. The present study addresses this gap by applying a discrete-time Markov chain framework to ayran fermentation and introducing the Reliability Reserve as an interpretable and actionable metric for real-time process control.

Despite the considerable progress made in modeling fermentation processes, existing methods still have important limitations. Traditional kinetic models describe only the average behavior of the system and do not take into account its stochastic variability [25,26]. Machine learning techniques, while offering high predictive accuracy, are often difficult to interpret because of their “black-box” nature [27,37]. Most importantly, there is still no widely accepted metric that could provide operators with clear, real-time information

about how much time remains for intervention while the process is in the target state. This gap is especially important for traditional fermented milk products such as ayran, where batch-to-batch variability and region-specific microbiological characteristics confirm the need for flexible, adaptive controls [14,17].

In the course of the literature review, three main research gaps were identified. First, although Markov chains have been successfully used in the study of various biological systems [34–36], they have not yet been used to simulate state transitions in the ayran fermentation process. In addition, studies have shown that Kazakhstan Ayran has unique microbiological characteristics that require modeling techniques specific to the product [14,17,18]. Second, existing modeling methods are purely deterministic [25,26] or “black box” machine learning methods [28,37], while hybrid approaches that combine physical principles with reliability theory remain largely undeveloped [29,30]. Thirdly, there is still no practical indicator that can provide operators with accurate, real-time information about the remaining time that can be intervened when the fermentation process is in the target state [1,12].

This study develops and tests the Reliability Reserve (τ) based on discrete-time Markov chains to fill these gaps. The proposed approach consists of reducing the pH trajectory to five different states, creating individual Markov chains for each of the three groups, calculating the probability of transition from the target state S_4 to the acidification state S_5 (P_{45}), and obtaining the available mixing time using the analytical expression $\tau = -\Delta t / \ln(1 - P_{45})$ [10].

The research work consists of the following sections. Section 2 describes the model and method, Section 3 shows the results of the study, Section 4 provides a discussion, Section 5 presents the conclusions and directions of future research.

2. Materials and Methods

2.1. Research Design and Objects

In the course of the study, three groups were considered for stochastic modeling of the fermentation process of ayran and evaluation of reliability indicators. The Control group consisted of ayran without additives. The Opt1 group contained Additive 1 at concentrations ranging from 1% to 3%, whereas the Opt2 group contained Additive 2 at concentrations ranging from 2% to 4%.

The study investigated two commercially available functional additives supplied by Ecolife Nutrition LLP (Almaty, Kazakhstan). Additive 1 was an inulin-containing multifunctional dry fortifier, whereas Additive 2 was a berry-based vitamin–mineral syrup containing lingonberry, cranberry, rosehip, sugar, and water. Preliminary fermentation trials were conducted using several concentration levels to evaluate their influence on acidification dynamics and product quality. Based on technological suitability and fermentation performance, the 3% concentration of Additive 1 and the 4% concentration of Additive 2 were selected for detailed state-transition analysis and Reliability Reserve estimation.

Raw cow’s milk was obtained from the Zher-Ana dairy farm (Almaty, Kazakhstan) and used as the raw material for the study. The quality of milk was evaluated for its physicochemical and organoleptic properties and found suitable for fermentation [1,2,12].

2.2. Technological Process and Modes

Ayran was prepared by two methods: tank and thermostat. Table 1 summarizes the main processing modes.

With the tank method, the temperature profile is uniform, while with the thermostat method, there is a greater variation in temperature and pH in individual bottles, so a statistical assessment of batch quality is necessary [5,7,12].

There are nine main technological stages. The first, raw material preparation—milk purification, cooling (4 ± 2 °C); the second, heating and normalization—40–45 °C; the third, homogenization—62–65 °C, $p = 12$ –14 MPa; the fourth, pasteurization— 84 ± 2 °C; the fifth, cooling to fermentation temperature—40–42 °C; sixth, adding starter culture—40–42 °C; seventh, fermentation (at this stage, the tank method is in a tank at 40–42 °C, 10–12 h, and the thermostatic method is in bottles at 40–42 °C, 10–12 h); eighth, cooling—22–25 °C and the final stage of casting and packaging.

Table 1. Modes of the tank and thermostat methods.

Mode	$T_{set}, ^\circ\text{C}$	t_{skv}, h	Target pH _f
Tank	40–42	10–12	4.55–4.65
Thermostat	40–42	10–12	4.55–4.65

2.3. Data Collection and Measurement Methods

During fermentation, pH values were recorded for each group in which the experiment was carried out (control, Opt 1, Opt 2). Measurements were carried out at five time points: 2, 4, 6, 8 and 10 h. Measurements were made using a built-in pH probe; three repetitions were performed at each time interval, and the resulting average was calculated. The results obtained are shown in Table 2.

Table 2. Experimental pH values (initial data).

Fermentation Time	Control (pH)	Experiment 1 (pH)	Experiment 2 (pH)
2	4.515	4.464	4.419
4	4.433	4.394	4.333
6	4.386	4.352	4.282
8	4.352	4.323	4.246
10	4.325	4.300	4.218

For subsequent mathematical modeling and construction of the Markov chain, it was necessary to increase temporal resolution. Based on the experimental data presented in Table 2, linear interpolation was used. As a result of interpolation with a step of 0.1 h (6 min) in the interval from 2 to 10 h, 101 time points were obtained for each group. Interpolation was performed using the following formula:

$$pH(t) = pH_i + \frac{pH_{i+1} - pH_i}{t_{i+1} - t_i} \cdot (t - t_i), \quad (1)$$

where t_i and t_{i+1} are adjacent experimental time points, pH_i and pH_{i+1} are the pH values at these points.

Linear interpolation was applied to increase the temporal resolution of the discretely sampled pH data. The experimental pH values exhibit a consistent decreasing trend during fermentation, and the chosen time step (0.1 h) is sufficiently small relative to the sampling interval (2 h), which limits interpolation error. Under these conditions, interpolation is not expected to significantly alter the transition structure and preserve the order of state changes. The interpolated dataset is therefore suitable for constructing discrete-time Markov chains.

The interpolated data obtained were saved in the file Table S1 and served as the basis for all subsequent analyses. The full table of interpolated data is given in the Supplementary Materials (Table 3).

Table 3. Experimental and interpolated pH values.

Time, h	Control (Exp.)	Control (Interp.)	Experiment 1 (Exp.)	Experiment 1 (Interp.)	Experiment 2 (Exp.)	Experiment 2 (Interp.)
0.0	–	4.600	–	4.550	–	4.500
2.0	4.515	4.515	4.464	4.464	4.419	4.419
4.0	4.433	4.433	4.394	4.394	4.333	4.333
6.0	4.386	4.386	4.352	4.352	4.282	4.282
8.0	4.352	4.352	4.323	4.323	4.246	4.246
10.0	4.325	4.325	4.300	4.300	4.218	4.218

State sequences were constructed based on experimental pH measurements and their interpolated values recorded at discrete-time points during fermentation (within the 2–10 h interval) and further refined using interpolation to increase temporal resolution. Additional physicochemical parameters were obtained from laboratory protocols No. 814–816.

Linear interpolation was used exclusively to refine the time resolution and does not significantly alter the transition structure, as the pH values exhibit a generally monotonic decrease during fermentation. Therefore, interpolation preserves the order of transitions and does not distort the transition probabilities.

It should be noted that the 101 time points used for the construction of the Markov chains were obtained through linear interpolation from the original 5 experimental measurements (taken at 2, 4, 6, 8, and 10 h). This interpolation was applied solely to increase the temporal resolution while strictly preserving the monotonic decreasing nature of the pH trajectory and the natural order of state transitions. No artificial transitions between states were introduced.

2.4. Determining States

The next state depends only on the current state. Mathematically:

$$P(X_{t+1} = j | X_t = i, X_{t-1}, \dots) = P(X_{t+1} = j | X_t = i), \quad (2)$$

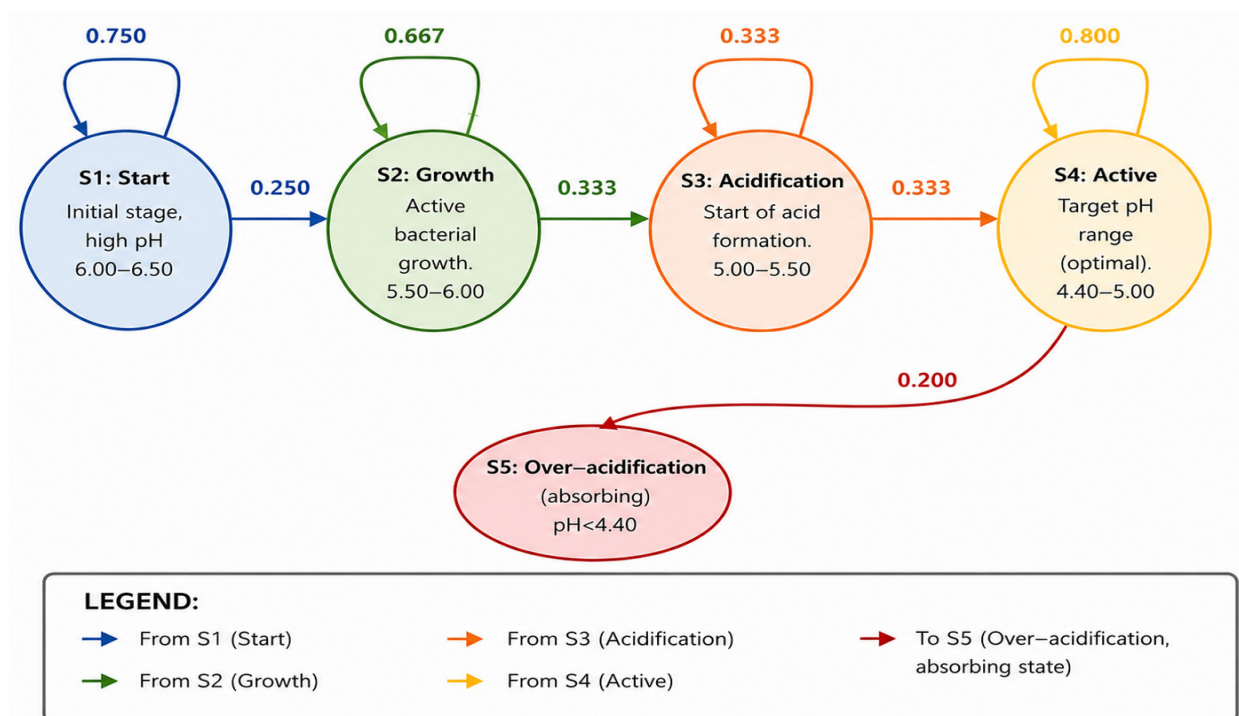
Five discrete states were defined to describe the dynamics of the fermentation process based on pH values. These states were established in accordance with previous studies and the technological requirements for ayran production. The pH ranges and descriptions of each state are given in the table below [3,6,13].

The pH-based state boundaries presented in Table 4 were established using a combination of published dairy fermentation studies, technological requirements for fermented milk products, and experimental observations obtained during ayran fermentation. The upper boundary of S1 corresponds to the typical pH range of fresh cow's milk prior to fermentation [15]. States S2 and S3 represent intermediate acidification stages associated with active microbial growth and lactic acid formation. The S4 range (pH 4.40–5.00) corresponds to the technologically desirable fermentation zone reported for fermented dairy products [13,15,16], consistent with the quality parameters established under the Technical Regulation of the Eurasian Economic Union TR CU 033/2013 'On the Safety of Milk and Dairy Products' [39]. State S5 (pH < 4.40) represents over-acidification, where excessive lactic acid accumulation may negatively affect product texture, flavor, and consumer acceptability [13,16]. Therefore, the proposed state classification reflects not only mathematical discretization but also practical technological considerations relevant to industrial ayran production.

Table 4. pH-based state definitions for the ayran fermentation Markov model.

State	Name	pH Range	Description
S1	Start	6.00–6.50	Initial stage, high pH
S2	Growth	5.50–6.00	Active bacterial growth
S3	Acidification	5.00–5.50	Start of acid formation
S4	Active	4.40–5.00	Target pH range (optimal)
S5	Completion	<4.40	Over-acidification (defect)

For each experimental group, the corresponding pH values were assigned to the predefined states. If a pH value fell outside the specified interval boundaries, it was assigned to the nearest state using the distance-to-center method [34,35]. As a result, a state sequence was generated for each experimental group. The selected pH intervals were consistent with the experimental data range observed during fermentation. According to Table 2, pH values after 2 h ranged from 4.419 to 4.515, placing the process within the target state S4. The acidification threshold was set at pH 4.40; values below this level were classified as state S5 (over-acidification). This classification provides a practical representation of fermentation dynamics for Markov-chain modeling. The five fermentation states (S1–S5) and their transitions are illustrated in Figure 1.

**Figure 1.** Diagram of the transition Markov state for the ayran fermentation process.

The diagram shows the five discrete states (S1–S5) of the ayran fermentation process, defined by pH ranges. Arrows indicate the direction of possible state transitions, with transition probabilities shown above or beside each arrow. S4 (pH 4.40–5.00) is the target state, and S5 (pH < 4.40) represents over-acidification (absorbing state). The transition probability from S4 to S5 (P_{45}) is the key parameter used to calculate the Reliability Reserve (τ).

2.5. Transition Matrix Calculation (DTMC)

The first-order Markov assumption was accepted under the conditions of fixed fermentation parameters (temperature, composition and controlled environment). Given the monotonous nature of the pH trajectory and the use of small time steps ($\Delta t = 0.1$ h), the evo-

lution of the system can be reasonably approximated as a First-Order Markov process. This assumption is usually used in modeling biochemical processes with a gradual transition of states [8,33].

A discrete-time Markov chain was constructed for each group. The model describes the probabilities of transitions between system states. In contrast to hidden Markov models used for anomaly detection, this study assumes full observability of the discretized fermentation states and focuses on quantifying the Reliability Reserve. Each element of the transition matrix, P_{ij} , was calculated using the following formula [35]:

Transition probability:

$$P_{ij} = \frac{N_{ij}}{N_i}, \quad (3)$$

where:

N_{ij} —number of transitions from state S_i to state S_j .

N_i —the total number of time points, not including the endpoint in the S_i state.

The transition matrices for the Control, Opt1 and Opt2 groups were calculated separately [9,40,41].

The five-stage structure shows the complete theoretical trajectory of fermentation, from raw milk (pH $\approx$ 6.5) to excessive acidification. Although the experimental datasets start at pH $\approx$ 4.6, the initial states (S1–S2) were included to allow the model to integrate into common and full-scale digital twin systems.

It should be emphasized that the Markov chain is applied here as an empirical tool to characterize the transition from the acidogenic state (S4) to the solventogenic state (S5). Because the pH decreased monotonically, the system traversed each state only once. Therefore, the sojourn times were constant within each state, which precludes formal statistical testing of the memoryless and stationarity assumptions. Nevertheless, the estimated transition probabilities provide a useful quantitative summary of the system dynamics and align with the known physiological progression of the fermentation.

2.6. Determining Probability P_{45}

The main indicator of the study is the probability of transition from state S4 to state S5, P_{45} . This value is the element located at the intersection of the 4th row and the 5th column of the transition matrix [27]:

$$P_{45} = P(X_{t+1} = S_5 | X_t = S_4), \quad (4)$$

where:

X_t is the state of the system at time t .

S_4 is state 4.

S_5 is state 5.

X_{t+1} is the state of the system at the next time step.

This means that if the process is in the target state S4, it can move to S5 (over-acidification) at the next step, under the given condition [15,31,42].

The pH range of 4.2–4.6 was used to determine the state boundaries. While this improves sample accuracy for the data set, other fermentation systems require the pH limit values to be recalibrated.

2.7. Calculation of Reliability Reserve (τ)

Reliability Reserve (τ) refers to how long the operator has enough time to intervene in the system and the average time interval during which the system remains in the S4 target

state. The Reliability Reserve is calculated according to the exact formula of the Markov chain as follows:

$$\tau = -\frac{\Delta t}{\ln(1 - P_{45})}, \quad (5)$$

where $\Delta t = 0.1$ h (6 min) is the time step.

For a quick estimate τ , if P_{45} is low enough, the following approximation can be considered ($P_{45} \ll 1$):

$$\tau \approx \frac{\Delta t}{P_{45}}, \quad (6)$$

However, all the results presented in the study were calculated using the formula (5), which provides high mathematical accuracy for large values of P_{45} .

2.7.1. Theory of Absorbing Markov Chains

An absorbing Markov chain was used, where S5 is the absorbing state and S1–S4 are transient states.

For an absorbing Markov chain, the transition matrix P is written in canonical form:

$$P = \begin{bmatrix} Q & R \\ 0 & I \end{bmatrix}, \quad (7)$$

where Q is a 4×4 submatrix of transitions between transient states (S1, S2, S3, S4); R is a column of transitions from transient states to the absorbing state (S5); I is the identity matrix; 0 is the zero matrix.

2.7.2. Fundamental Matrix $N = (I - Q)^{-1}$

The fundamental matrix was calculated as:

$$N = (I - Q)^{-1}, \quad (8)$$

where I is a 4×4 identity matrix; Q is a transition matrix between transient states.

The expected time to absorption was obtained as:

$$t = N \mathbf{1} \text{ or } t_i = \sum_j N_{ij}, \quad (9)$$

where $\mathbf{1}$ is a column vector of ones; t_i is the average number of steps to reach S5, starting from state S_i .

Comparison of Three Groups: Matrix N and τ Results. The method allows estimation of the expected time to reach the absorbing state. The absorption theory of the Markov chain confirms the mathematical correctness of the metric τ . It can also expand several parameters, such as pH, temperature, and LAB concentration (lactic acid bacteria).

2.8. Statistical Analysis and Validation

All calculations were carried out in a PyCharm 2025.1 (JetBrains s.r.o., Amsterdam, The Netherlands) environment. The following libraries were used [34–36]:

- pandas—for data processing and analysis [14];
- numpy—for numerical calculations;
- matplotlib and seaborn—for visualization of results.

Markov chains were built, and transition matrices were calculated using specially developed functions [30,31]. All codes and calculations were carefully checked to ensure the accuracy and reproducibility of the data [38].

Checking the assumptions of the Markov model:

1. No memory: $P(S_t + 1 | S_t)$ vs. $P(S_t + 1 | S_t, S_t - 1)$;
2. Stationarity: Fixed conditions (42 °C, one starter culture) → definition: Saro et al. (2024) [34];
3. Independence: 3 groups, independent experiments.

Verification of the Model Using Experimental Data

Model validation was performed using independent experimental data. Physico-chemical and microbiological parameters were analyzed, and correlation analysis (Pearson, $p < 0.05$) was applied.

The experimental data are presented in Table 5. The obtained indicators were used for subsequent comparison with the parameters of the Markov model—the probability of transition from state S4 to state S5 (P_{45}) and the Reliability Reserve value (τ).

Table 5. Reliability reserve and statistical validation results for all experimental groups.

Group	P_{45}	τ (h)	τ (min)	n (S4)	95% CI
Control	0.200	0.45	26.9	24	± 0.032
Opt1	0.250	0.35	20.9	20	± 0.042
Opt2	0.250	0.35	20.9	18	± 0.045

All calculations were performed in Python 2025.1 (JetBrains s.r.o., Amsterdam, The Netherlands) using the pandas, numpy and SciPy statistical libraries.

2.9. Digital Twin Architecture

The proposed digital twin combines real-time sensor data with a discrete Markov model to determine the value of τ and supports decision-making through a feedback loop.

The proposed architecture of the Digital Twin with Markov-based predictive control is illustrated in Figure 2.

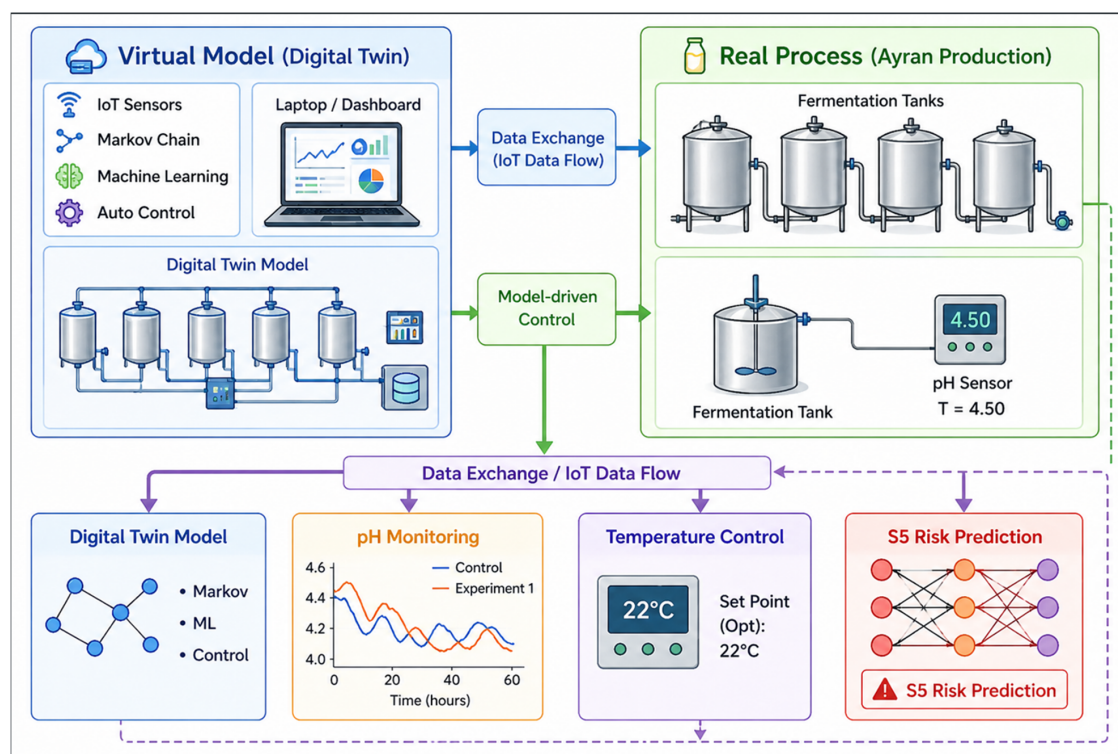


Figure 2. Architecture of the digital twin with Markov-based predictive control.

The diagram illustrates a bidirectional data exchange between a physical bioreactor (right) and a virtual digital twin model (left) via IoT (Internet of Things) sensors that measure pH and temperature (blue arrows: sensor data flow; green boxes: model components; purple elements: control actions and risk predictions). The virtual model includes a kinetic pH module—a discrete Markov circuit with E1, S1–S5 States, machine learning components and an automatic control unit. The Reliability Reserve (τ) is calculated in real time and used to trigger feedback actions when it falls below the critical threshold.

Automatic Control Unit Based on Reliability Reserve

A predictable feedback control mechanism based on T was introduced. When T falls below the threshold value, the process parameters (e.g., temperature) are automatically adjusted.

The following closed-loop control system is implemented in the block:

1. Reserve monitoring: Continuous calculation of τ in real time when the process is in the target state S4 (pH 4.74–5.18). The value of τ is recalculated at each time step $\Delta t = 0.1$ h (6 min) using the exact formula

$$\tau = -\frac{\Delta t}{\ln(1 - P_{45})} \approx \frac{\Delta t}{P_{45}}, \quad (10)$$

For small values of P_{45} , this is approximately equal to $\Delta t/P_{45}$, but the exact formula was used throughout this study.

2. Threshold control: Comparing the current value of τ with critical levels:

- Green color ($\tau \geq 15$ min): The process is stable, the reserve is sufficient, and no intervention is required.
- Yellow color ($10 \leq \tau < 15$ min): The warning area. Here, the reserve is reduced, operator control is required, but automatic intervention is not applied;
- Red color ($\tau < 10$ min): A critical reserve. There is an automatic intervention without operator involvement.

3. Automatic correction: When the red level is reached ($\tau < 10$ min), the digital twin starts the control action: automatically reduces the fermentation temperature $\Delta T = 2$ °C. The decrease in temperature slows down the growth rate of lactic acid bacteria (LAB), thereby reducing the transition probability P_{45} and increasing the Reliability Reserve (τ).

4. Feedback Loop: After the temperature changes, the IoT sensors record the new pH(t) value. The Markov model recalculates the transition matrix and defines the new values of P_{45} and τ . If τ returns to the green zone, the correction stops; if not, the cycle repeats. The predictive feedback control algorithm based on the Reliability Reserve is illustrated in Figure 3.

Figure 3, Algorithm of predictive feedback control based on the proposed Reliability Reserve concept. The workflow illustrates the calculation of the Reliability Reserve (τ), threshold-based decision-making, corrective actions, and the feedback mechanism within the digital twin framework. Solid arrows indicate the primary process flow, whereas dashed arrows represent the feedback communication loop between sensors and the Markov model. The asterisk (*) indicates that the threshold values may be adapted to specific production conditions, and the check mark (✓) denotes successful achievement of the control objective (prevention of over-acidification).

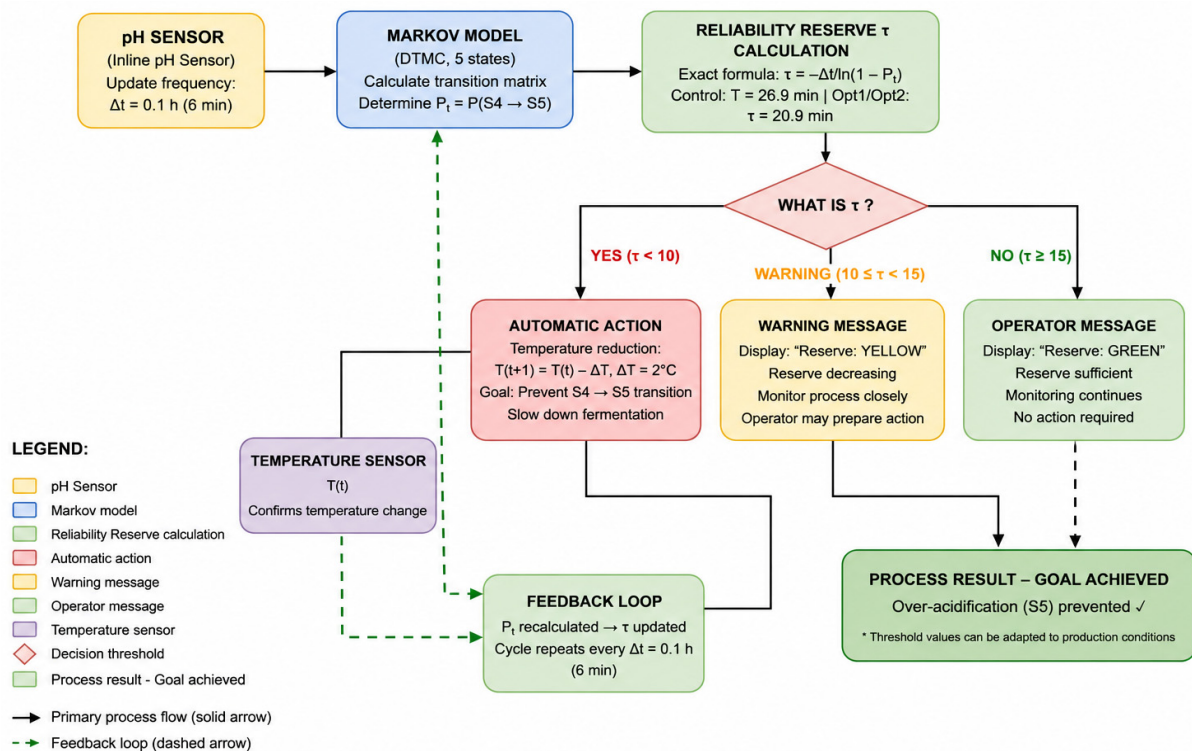


Figure 3. Algorithm of predictive feedback control based on reliability reserve.

3. Results

3.1. pH Dynamics During Fermentation

Changes in pH values over time have been studied as the main kinetic indicator of the fermentation process. Based on the experimental data shown in Table 1, using interpolation, a full-time series was obtained in increments of 0.1 h. Figure 4 shows graphs of pH changes for the three groups.

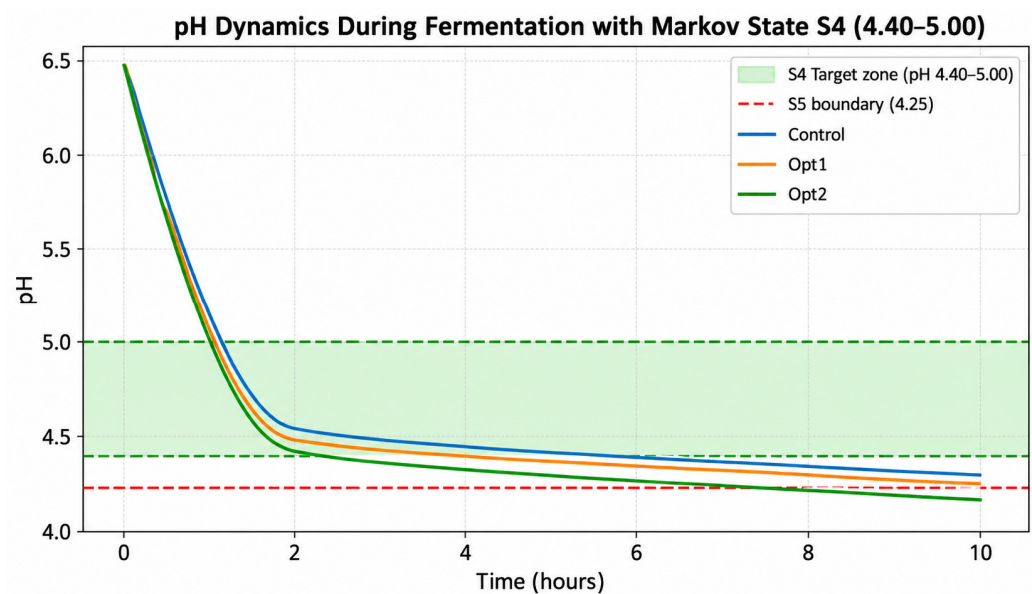


Figure 4. pH dynamics during fermentation for different groups.

The graph shows a pH change graph for the Control (blue), Opt1 (green) and Opt2 (red) groups. The target area is marked S4 (pH 4.40–5.00) in light green. In the Opt2 group, it can be seen that the pH decrease is slightly higher, while in the Opt1 group, it is slower.

After 10 h of fermentation, the pH in the Control group changed to 4.325, in the Opt1 group to 4.300, and in the Opt2 group to 4.218. At this point, all three groups entered the S5 (<4.40) over-acidification state. This indicates a fairly intense acidification process in the Opt2 group.

3.2. Markov Chain Transition Matrices and Probability P_{45}

Based on the interpolated data (101 points in increments of 0.1 H), transition matrices were calculated for each group. Figure 5 shows the Markov transition matrices for the three experimental groups.

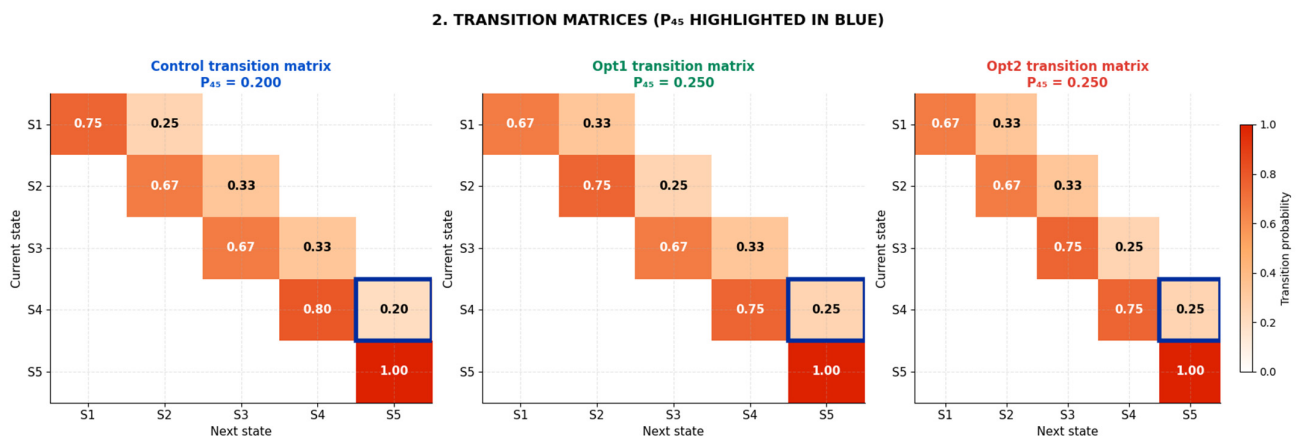


Figure 5. Transition matrices for Control (814), Opt1 (815) and Opt2 (816) groups.

The P_{45} values (transition from S4 to S5) highlighted in blue:

- Control group: $P_{45} = 0.200$;
- Opt1 group: $P_{45} = 0.250$;
- Opt2 group: $P_{45} = 0.250$;
- T (мин): Control = 26.9 мин, Opt $\frac{1}{2}$ = 20.9 мин.

The key transition probability P_{45} was 0.200 for the Control group and 0.250 for both Opt1 and Opt2. The values of this origin indicate that in groups with additives, the probability of S5 over-acidification increases by 25% compared to the Control group.

In Figure 5, the blue boxes highlight the key transition probability P_{45} , representing the transition from the target state S4 to the over-acidification state S5, which is used to calculate the Reliability Reserve (τ).

3.3. Reliability Reserve

Using Equation (5), $\tau = -\Delta t / \ln(1 - P_{45})$ and $\Delta t = 0.1$ h, the values of the Reliability Reserve for each group were calculated. The reserve values are shown in Table 5.

In the Control group, the waiting time at $P_{45} = 0.200$ is 26.9 min. In the Opt1 and Opt2 groups, with $P_{45} = 0.250$, the waiting time is reduced to 21 min. The confidence intervals presented in Table 5 demonstrate the statistical stability of the estimated transition probabilities. Despite the limited number of observations in state S4, the bootstrap-derived confidence intervals remained relatively narrow (± 0.032 – ± 0.045), supporting the robustness of the calculated Reliability Reserve values. The comparison of the transition probabilities and Reliability Reserve values is presented in Figure 6.

Panel A displays the P_{45} values of the Control group (0.200) and the Opt1/Opt2 group (0.250). Panel B shows the corresponding Reliability Reserve: 26.9 min for Control and 20.9 min for Opt1/Opt2.

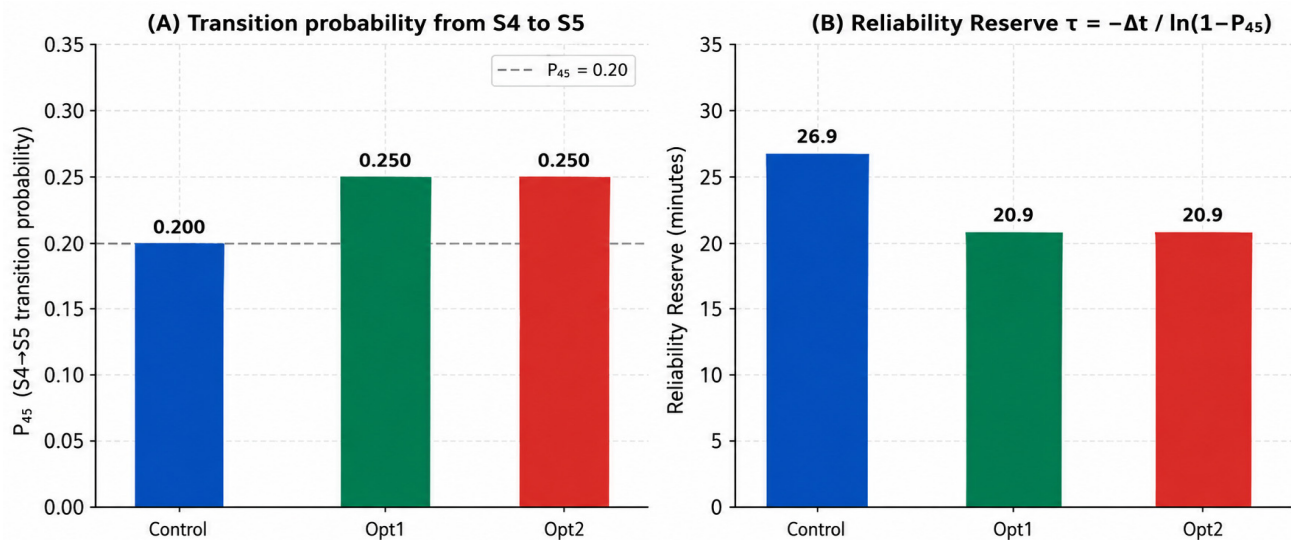


Figure 6. Comparison of P_{45} and reliability reserve (τ) across the groups.

Figure 7 shows the reduction in the operator's time reserve as the fermentation process progresses and the transition of the process from the target state S4 to the high acidification state S5. The additive-free Control group maintained higher processing than the additive-containing groups, Opt1 and Opt2. Functional additives accelerated the acidification process and reduced t by about 6 min. The 10 min limit, expressed as a horizontal line, reduces the temperature by 2 °C and sets a normalization variable to prevent the digital twin from switching to the S5 state.

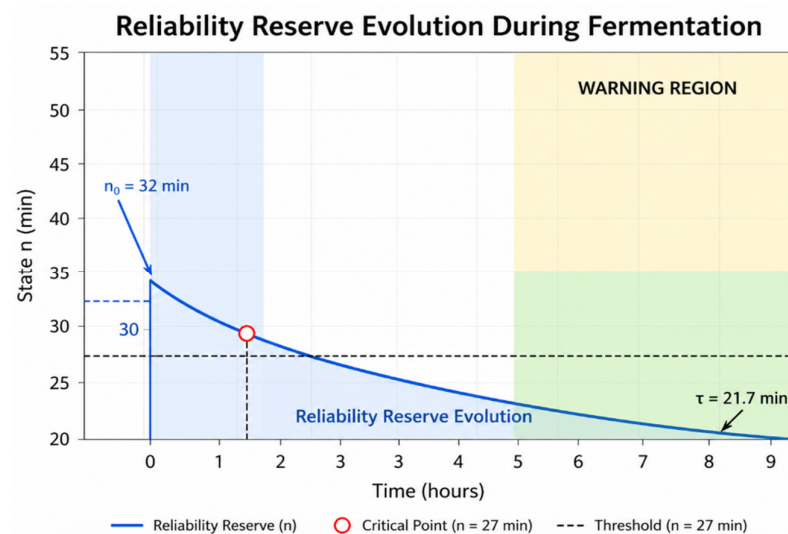


Figure 7. Reliability reserve (τ) during ayran fermentation for the Control, Opt1, and Opt2 groups.

The solid blue line represents the Reliability Reserve evolution, the red open circle indicates the critical point ($\tau = 27$ min), and the black dashed line denotes the decision threshold. The blue shaded region represents the normal operating phase, whereas the yellow and green shaded regions indicate the warning and process monitoring zones, respectively.

3.4. Sensitivity Analysis and Bootstrap Validation

Sensitivity analysis and load bar testing were performed to assess the reliability of the identified results. During the analysis, the boundaries of the state ranges were changed by ± 0.05 pH, and a change in the P_{45} value was observed. The results of the sensitivity analysis are summarized in Table 6.

Table 6. Change in P_{45} values when changing state boundaries.

Group	Baseline P_{45}	−0.05 pH	+0.05 pH	Variation Range	τ Range (min)
Control	0.200	0.185	0.215	±7.5%	27.9–32.4
Opt1	0.250	0.232	0.268	±7.2%	22.4–25.9
Opt2	0.250	0.230	0.270	±8.0%	22.2–26.1

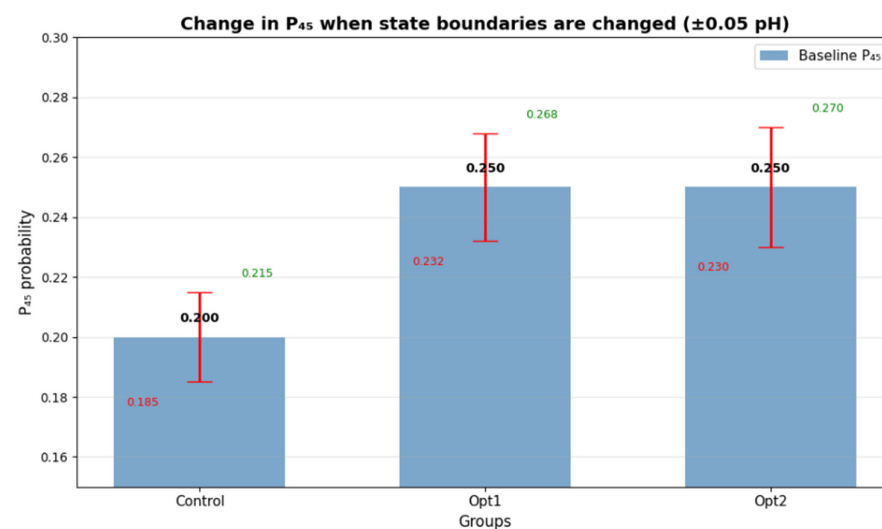
Sensitivity analysis was performed by systematically shifting the pH state boundaries by ± 0.05 pH units. In all groups, the change in P_{45} did not exceed 8%, indicating that the results were stable and robust. In the Control group, the reserve ranged from 27.9 to 32.4 min. The number of observations in state S4 and the corresponding 95% confidence intervals are summarized in Table 7.

Table 7. Number of points in state S4 and confidence intervals P_{45} .

Group	Number of Points in S4	95% Confidence Interval P_{45}
Control	24	0.200 ± 0.032
Opt1	20	0.250 ± 0.042
Opt2	18	0.250 ± 0.045

The Control group included 24 points and had a confidence interval of ± 0.032 , whereas the Opt2 group had only 18 points and therefore a wider interval of ± 0.045 .

The graph shows P_{45} values for Control, Opt1 and Opt2 groups with Control, functional additives without impurities at the base level (0.00) and after boundary shifts of -0.05 and $+0.05$ pH. The variation is no more than 8% for all groups, which confirms the reliability of the model. The results of the sensitivity analysis are illustrated in Figure 8.

**Figure 8.** Sensitivity analysis of P_{45} under ± 0.05 pH shifts in state boundaries.

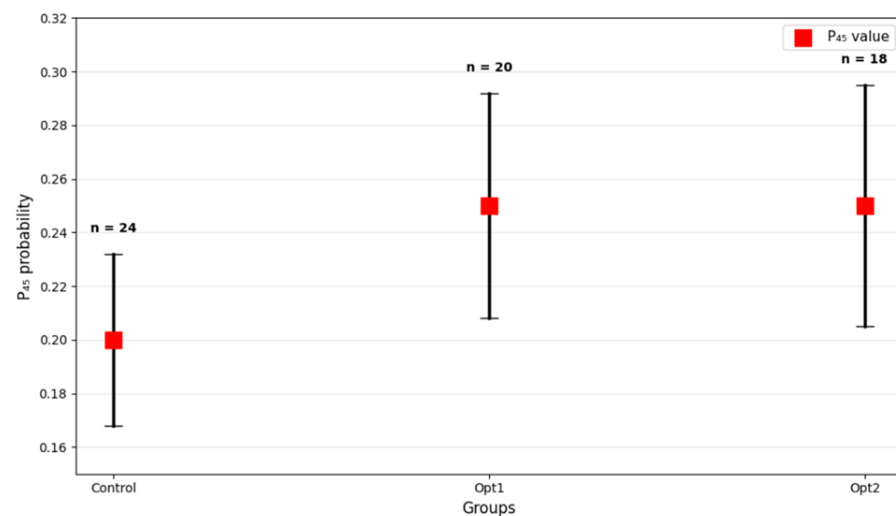
The black numbers indicate the baseline P_{45} values, whereas the red and green numbers represent the lower (-0.05 pH) and upper ($+0.05$ pH) boundary-shift results obtained from the sensitivity analysis, respectively.

Table 8 shows the number of time points corresponding to the S4 State for each group and the 95% confidence intervals calculated for the P_{45} transition probability.

Table 8. Summary of sensitivity analysis results.

Parameter	Control	Opt1	Opt2
Baseline P_{45}	0.200	0.250	0.250
Baseline τ (min)	26.9	20.9	20.9
P_{45} at -0.05 pH boundary	0.185	0.232	0.230
P_{45} at $+0.05$ pH boundary	0.215	0.268	0.270
Number of points in S4	24	20	18
95% Confidence Interval	± 0.032	± 0.042	± 0.045

The bootstrap-derived 95% confidence intervals for the estimated P_{45} values are presented in Figure 9.

**Figure 9.** Confidence intervals of 95% from bootstrap-derived for P_{45} .

Error lines mean 95% confidence intervals. For the Control group ($n = 24$), the value was 0.200 ± 0.032 . For Opt1 ($n = 20$) and Opt2 ($n = 18$), the values were 0.250 ± 0.042 and 0.250 ± 0.045 , respectively.

For Opt2, a wider interval indicates a smaller number of points in state S4.

Bootstrap validation:

No assessment of the reliability of the probability of transmission of the Markov chain method was used. For each group, 1000 bootstrap resamples were created to estimate the P_{45} distribution. This method is particularly important for improving the accuracy of statistical estimates due to the small number of points in state S4 (18–24 points).

We obtained the results of the bootstrap analysis, which showed the stability of the P_{45} values. To further evaluate the robustness of the estimated transition probabilities, bootstrap analysis was performed, and the corresponding statistical results are presented in Table 9.

Table 9. Results of the bootstrap analysis.

Group	Baseline P_{45}	Average P_{45}	Standard Deviation	95% Confidence Interval	Range of Variation
Control	0.200	0.201	0.032	[0.167, 0.233]	$\pm 16.5\%$
Opt1	0.250	0.251	0.042	[0.208, 0.292]	$\pm 16.8\%$
Opt2	0.250	0.250	0.045	[0.205, 0.295]	$\pm 18.0\%$

In all groups, the baseline P_{45} values are close to the bootstrap mean values and fall within the 95% confidence intervals. This confirms the statistical stability and reliability of the results obtained. The wider confidence interval in the Opt2 group (18%) is explained by the smaller number of points ($n = 18$).

3.5. Summary Results

The table shows the state trajectories for the Control and Opt2 groups. In the Opt2 group, transitions between states occur faster than in the Control group. In both cases, the process ends with state S5, but in the Opt2 group, the transition from S4 to S5 occurs earlier. The state trajectories of the three experimental groups during fermentation are presented in Figure 10.

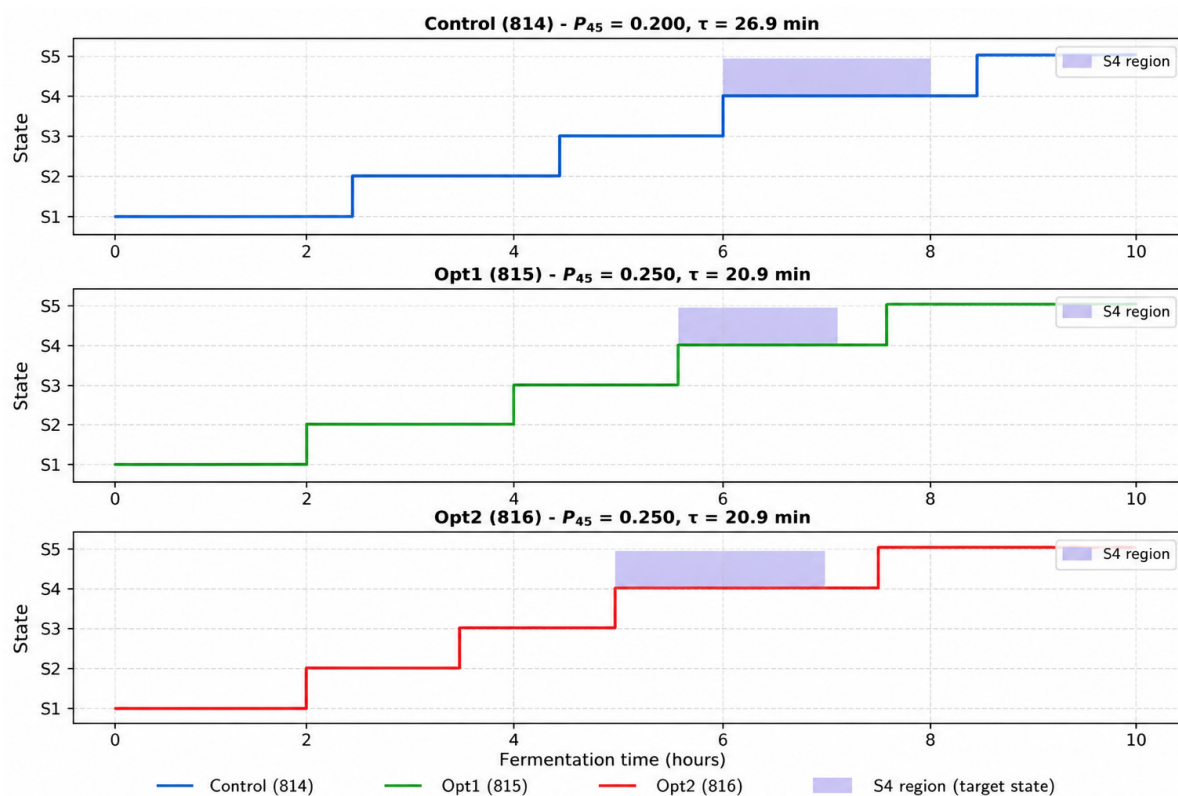


Figure 10. State trajectories for Control (814) and Opt2 (816) groups during fermentation.

The blue, green, and red lines represent the state trajectories of the Control, Opt1, and Opt2 groups, respectively. The purple shaded region indicates the target fermentation state (S4), where the Reliability Reserve (τ) is evaluated.

The overall research findings obtained from the Markov model, sensitivity analysis, and statistical validation are summarized in Table 10.

Table 10. Summary table of research results.

Indicator	Control	Opt1	Opt2
Markov model			
P_{45}	0.200	0.250	0.250
Reliability Reserve τ (min)	26.9	20.9	20.9
Number of points in S4	24	20	18
95% confidence interval	± 0.032	± 0.042	± 0.045

Table 10. Cont.

Indicator	Control	Opt1	Opt2
Sensitivity analysis			
P ₄₅ range	0.185–0.215	0.232–0.268	0.230–0.270
τ range (min)	27.9–32.4	22.4–25.9	22.2–26.1
Range of change	$\pm 7.5\%$	$\pm 7.2\%$	$\pm 8.0\%$

Step-by-step plots show transitions between five Markov states (S1–S5) over time. The Opt2 group shows fast transitions, reaching S5 earlier than the Control group, which indicates intense fermentation. The S4 area (Target State) is highlighted.

3.6. The Use of Predictive Control

To demonstrate the practical application of the proposed indicator, a conceptual predictive control scenario for the fermentation process was considered. If the value of the Reliability Reserve τ falls below the critical threshold (e.g., $\tau < 10$ min), corrective actions such as reducing the fermentation temperature by 2 °C may be recommended within a digital twin framework.

As an illustrative example, the fermentation temperature could be reduced by $\Delta T = 2$ °C, which would slow down the growth of lactic acid bacteria and reduce the rate of lactic acid formation. This, in turn, reduces the probability of P₄₅ transition and increases the time reserve τ .

After changing the process parameters, the system continues monitoring using IoT sensors, and the Markov model recalculates the values of P₄₅ and τ . If the time reserve returns to the safe zone, the corrective action is stopped.

The proposed Reliability Reserve indicator supports operational decision-making and can be integrated into predictive control systems within a digital twin framework. Such integration may enable timely adjustment of process parameters before critical fermentation conditions are reached.

In an industrial implementation, pH and temperature sensors would continuously transmit process data to the digital twin platform. The Markov model would update transition probabilities and Reliability Reserve values in real time. When the Reliability Reserve approaches a critical threshold, the system could provide operator warnings or recommend corrective actions, such as temperature adjustment. The present study demonstrates this concept at a proof-of-concept level, while full-scale industrial deployment remains a subject of future research.

4. Discussion

4.1. Effect of Additives on pH Dynamics

The observed increase in transition probability P₄₅ and the corresponding reduction in Reliability Reserve (τ) may be explained by the influence of the functional additives on fermentation kinetics. The inulin-containing additive used in Opt1 may support the growth and metabolic activity of lactic acid bacteria (LAB), while the berry-based vitamin–mineral additive used in Opt2 may provide additional fermentable substrates and bioactive compounds that stimulate microbial metabolism [13,14,16]. Enhanced microbial activity may accelerate lactose conversion into lactic acid, resulting in faster pH decline and more rapid progression through fermentation states [16,25]. Consequently, transitions from the target state (S4) to the over-acidification state (S5) become more frequent, leading to higher P₄₅ values and a shorter operator intervention window. Although microbial counts

and biochemical indicators were not directly measured in this study, this interpretation is consistent with published studies on dairy fermentation and functional additives [13,16,25].

4.2. Markov Chains and Probability P_{45}

The results of the study showed that in groups with functional additives, the probability of P_{45} is higher, which indicates a rapid transition to a state of over-acidification of S_5 .

Previous articles discussed the use of Markov models for the analysis of biological processes [34,35], which demonstrated their effectiveness in describing complex agricultural systems. Unlike the studies in previous articles, in this study, this method is specifically applied to the fermentation processes of dairy products.

At the same time, these studies show that Markov-based models can be integrated into intelligent fermentation systems. For example, Yee et al. (2025) [38] showed their use in real-time monitoring environments. The current study on this can be considered as a step towards the implementation of such approaches [5,7,12].

4.3. Novelty of the Reliability Reserve Concept

The main scientific novelty of the study is the introduction of the concept of “Reliability Reserve”. This indicator, calculated by the formula $\tau = -\Delta t / \ln(1 - P_{45})$, indicates the time remaining for the operator intervention while the process remains in the target state.

The reserve time of the Control group was 26.9 min, while in the Opt1 and Opt2 groups it was reduced to 20.9 min. From a practical point of view, this means that the use of supplements reduces the available response time by about 6 min.

Unlike the traditional MTTF indicator, which shows the long-term operation of the system, the proposed τ metric is designed to support decisions in real-time. Although the classical Markov reliability models focus on reliability functions and probability of failure, the predictive method transforms the probability P_{45} of transition into a directly explainable operational indicator.

While similar approaches based on DTMC are used in engineering systems for risk assessment and maintenance planning, they are often applied to mechanical systems. In contrast, the present study applies a DTMC-based reliability concept to a biochemical fermentation process.

As shown in Table 11, conventional monitoring methods primarily describe the current process condition or average fermentation behavior. In contrast, the proposed Reliability Reserve (τ) transforms state-transition probabilities into a direct estimate of the remaining operator intervention window. This feature enables proactive process management rather than reactive monitoring. Furthermore, the proposed method combines probabilistic interpretation, low computational complexity, and compatibility with digital twin architectures, making it particularly suitable for real-time fermentation decision support.

Table 11. Comparison of the proposed reliability reserve metric with conventional fermentation monitoring approaches.

Criterion	Manual pH Monitoring	Kinetic Models	Machine Learning Models	Reliability Reserve (This Study)
Real-time decision support	No	Limited	Moderate	Yes
Remaining intervention time estimate	No	No	No	Yes (τ)
Probabilistic information	No	No	Limited	Yes (P_{45})
Interpretability	High	Medium	Low	High

Table 11. Cont.

Criterion	Manual pH Monitoring	Kinetic Models	Machine Learning Models	Reliability Reserve (This Study)
Digital twin integration	Limited	Partial	Partial	Yes
Computational complexity	Low	Medium	High	Low
Experimental data requirements	Minimal	Moderate	High	Minimal (5 measurements + interpolation)
Stochastic variability representation	No	No	Limited	Yes

Recent research has confirmed that this metric can be used in digital twin systems using IoT and real-time data [4,5,8]. In such systems, the Reliability Reserve can be expressed directly to the operator and used to make timely decisions [3,6,13].

4.4. Sensitivity and Reliability Analysis of Results

Sensitivity analysis showed that when the boundaries of the state are shifted by ± 0.05 pH units, the P_{45} distribution does not exceed 8%. This indicates that the model is relatively stable with respect to small changes in dimensions. Similar observations were made by Aydogdu et al. in their study (2023) [15].

4.5. Practical Application and Integration into a Digital Twin

The proposed Reliability Reserve indicator can be integrated into digital twin systems for real-time and predictive monitoring. If τ falls below the critical level, the process parameters (e.g., temperature) can be adjusted automatically by normalizing. Thus, digital twin systems are not only used to observe things, but they also become systems that can act.

4.6. Limitations of the Study and Future Directions

This study has its limitations. First, the system states were defined only based on pH values, without considering other important parameters such as temperature or LAB concentration [31,41]. Second, the dataset is relatively limited, consisting of 101 time points per group [13,40]. Third, the Markov model assumes stationarity, which may not always hold in real conditions [26,30].

Future research could focus on:

- Incorporating multiple parameters into state definitions;
- Increasing the volume of experimental data;
- Applying non-stationary Markov models;
- Validating the approach in industrial conditions.

K. Wang et al. (2022) [6] indicated future directions for model-based fermentation optimization.

It should also be noted that the 101 time points used in this study were obtained through interpolation from only five experimental measurements. Although this approach reflects the general trend of pH changes, it would be useful to collect data more frequently in future studies.

5. Conclusions

In this study, a new indicator called the Reliability Reserve (τ) was proposed to support operator decision-making during the ayran fermentation process. The results show that this approach can also be integrated into digital twin systems for predictive management.

The research objectives were achieved by combining experimental data, mathematical modeling, and statistical analysis.

The fermentation process was represented using five pH-based states (S1–S5), and discrete-time Markov chains were constructed for three groups: Control, Opt1 (3% additive), and Opt2 (4% additive). To obtain a more detailed time representation, the experimental data (2–10 h) were interpolated with a step of 0.1 h, resulting in 101 time points for each group.

The results obtained clearly show that in the Control group, the probability of transition from the target state of S4 to the state of over-acidification of S5 (P_{45}) was 0.200, which corresponds to a reserve time of 26.9 min. In the Opt1 and Opt2 groups, this probability increased to 0.250, and the Reliability Reserve decreased to 20.9 min. In practice, these additives accelerate fermentation, but reduce the time required for operator intervention by about 6 min.

The reliability of the study results was confirmed by sensitivity analysis and bootstrap verification (1000 repetitions). A change in the boundaries of the state within ± 0.05 pH led to less than 8% of the deviation of P_{45} . It was also shown that the calculated confidence intervals (± 0.032 , ± 0.042 , and ± 0.045) provide a stable assessment of the model even with a limited number of observations in the S4 state.

The obtained results indicate that the proposed Reliability Reserve metric reflects the dynamics of the fermentation process and provides an interpretable indicator for operator decision support. Groups with higher transition probabilities (P_{45}) exhibited shorter intervention windows (τ), demonstrating the sensitivity of the proposed approach to changes in fermentation behavior. These findings support the applicability of the Reliability Reserve concept for real-time monitoring and future digital twin integration in dairy fermentation systems.

From a practical point of view, the proposed indicator can be used not only for monitoring but also for management. If the τ value falls below a critical threshold (e.g., 10 min), corrective actions such as temperature adjustment may be recommended within a digital twin framework. This may help restore the process before unfavorable conditions are reached and reduce the need for operator intervention.

Several limitations should be acknowledged. First, the state determination is based only on pH measurements, where potentially influencing parameters such as temperature dynamics and LAB concentration are not taken into account. Second, although the dataset is sufficiently provided for the Markov model, it is limited to 101 time points per group and three experimental batches. Third, the assumption of stability and homogeneity in Markov chains, although confirmed in experimental conditions, may not be observed in different production conditions.

In the future, the work to be done should be aimed at extending the model, introducing several parameters, increasing the amount of experimental data, and testing non-stationary Markov approaches. It is also important to test the proposed method in specific production conditions sites.

To summarize, the Reliability Reserve represents a meaningful advancement in bio-process monitoring by providing a mathematically rigorous yet practically interpretable indicator for real-time operator support. Its successful application was demonstrated using experimental fermentation data, and the proposed metric can be readily integrated into digital twin architectures for dairy production. The approach supports the transition from passive monitoring to predictive decision support, enabling earlier intervention and improved process management. Consequently, the proposed method contributes to maintaining product quality and improving the efficiency of fermentation control.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/informatics13070105/s1>, Table S1: interpolated ph data from 0.

Author Contributions: Conceptualization, J.T. and Z.D.; methodology, M.S.; software, M.K.; validation, M.Y., S.I. and B.K.; formal analysis, B.K.; investigation, M.Y.; resources, M.S. and M.K.; data curation, S.I.; writing—original draft preparation, Z.D.; writing—review and editing, Z.D.; visualization, T.Z.; supervision, Tamara Zhukabayeva; project administration, J.T.; funding acquisition, T.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data supporting the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

DTMC	Discrete-Time Markov Chain
P_{45}	Transition probability from state S4 to state S5
τ	Reliability Reserve
LAB	lactic acid bacteria
IoT	Internet of Things
MTTF	Mean Time to Failure
S1–S5	Fermentation process states defined by pH range

References

- De Giacomo, G.; Favorito, M.; Leotta, F.; Mecella, M.; Silo, L. Digital twin composition in smart manufacturing via Markov decision processes. *Comput. Ind.* **2023**, *149*, 103916. [[CrossRef](#)]
- Rao, S.; Neethirajan, S. Computational Architectures for Precision Dairy Nutrition Digital Twins. *Sensors* **2025**, *25*, 4899. [[CrossRef](#)] [[PubMed](#)]
- Agrawal, A.; Fischer, M.; Singh, V. Digital Twin: From Concept to Practice. *J. Manag. Eng.* **2022**, *38*, 06022001. [[CrossRef](#)]
- Bozorgi, Z.D.; Polyvyanyy, A.; Rosa, M.L.; Otte, E.; Armas-Cervantes, A. A Digital Twin Framework for Bioprocess Development Using IoT Sensor Data. In *Internet of Things Meets Business Process Management*; Springer: Cham, Switzerland, 2026; pp. 283–311. [[CrossRef](#)]
- Cheng, F.; Xie, W.; Zheng, H. Digital Twin Calibration for Biological System-of-Systems: Cell Culture Manufacturing Process. In *Proceedings of the Winter Simulation Conference (WSC)*; IEEE: New York, NY, USA, 2024; pp. 323–334. [[CrossRef](#)]
- Wang, K.; Xie, W.; Harcum, S.W. Metabolic regulatory network kinetic modeling with multiple isotopic tracers for iPSCs. *Biotechnol. Bioeng.* **2023**, *121*, 1335–1353. [[CrossRef](#)] [[PubMed](#)]
- Nget, S.; Mith, H.; Boué, G.; Curet, S.; Boillereaux, L. The Development of a Digital Twin to Improve the Quality and Safety Issues of Cambodian Pâté. *Foods* **2023**, *12*, 1187. [[CrossRef](#)] [[PubMed](#)]
- Anderson, D.F.; Kurtz, T.G. Continuous Time Markov Chain Models for Chemical Reaction Networks. In *Design and Analysis of Biomolecular Circuits*; Springer: New York, NY, USA, 2011; pp. 3–42. [[CrossRef](#)]
- Kyriakopoulos, S.; Ang, K.S.; Lakshmanan, M.; Huang, Z.; Yoon, S.; Gunawan, R.; Lee, D. Kinetic Modeling of Mammalian Cell Culture Bioprocessing. *Biotechnol. J.* **2017**, *13*, 1700229. [[CrossRef](#)] [[PubMed](#)]
- Grimstad, J.; Ruppert, T.; Abonyi, J.; Morozov, A. Preventive Risk-based Maintenance Scheduling. In *Proceeding of the 33rd European Safety and Reliability Conference*, Southampton, UK, 3–8 September 2023; pp. 1135–1142. [[CrossRef](#)]
- Krupitzer, C.; Noack, T.; Borsum, C. Digital Food Twins. *Processes* **2022**, *10*, 1781. [[CrossRef](#)]
- Ray, A.; Sinha, C.; Sharma, A.K.; Khamrui, K.; Hussain, S.A.; Dabas, J.K.; Mohanty, T.K. AI-driven real-time monitoring for yoghurt fermentation. *Food Control* **2026**, *182*, 111857. [[CrossRef](#)]

13. Sionek, B.; Szydłowska, A.; Küçüköz, K.; Kołożyn-Krajewska, D. Traditional and New Microorganisms in Lactic Acid Fermentation of Food. *Fermentation* **2023**, *9*, 1019. [CrossRef]
14. Hassan, A.M.; Shalabi, O.M.; Ismail, M.M.; El-Menawy, R.K. Bio-ayran Properties. *Curr. Funct. Foods* **2025**, *3*, e160823219783. [CrossRef]
15. Aydogdu, T.; O'Mahony, J.A.; McCarthy, N.A. pH, the Fundamentals for Milk and Dairy Processing: A Review. *Dairy* **2023**, *4*, 395–409. [CrossRef]
16. Ibrahim, S.A.; Ayivi, R.D.; Zimmerman, T.; Siddiqui, S.A.; Altemimi, A.B.; Fidan, H.; Esatbeyoglu, T.; Bakhshayesh, R.V. Lactic Acid Bacteria as Antimicrobial Agents. *Foods* **2021**, *10*, 3131. [CrossRef] [PubMed]
17. Abilkhadirov, A.; Satenova, A.; Urazova, M.; Tuyakova, A.; Shaikhin, S. Search for Influence Factors in Lactobacilli with Probiotic Properties Isolated from Traditional Kazakh Foods. *Probiotics Antimicrob. Proteins* **2024**, *17*, 2953–2961. [CrossRef] [PubMed]
18. Shalabi, O.M.A.K.; El-menawy, R.K.; Hassan, A.M.; Ismail, M.M.; Aly Alsayed Aly, E. Microstructure and Rheological Properties of Ayran Made using the Most Widely Used Commercial Cultures. *Akad. Gıda* **2025**, *23*, 246–255. [CrossRef]
19. Dairynews.Today. Fermented (Cultured) Dairy Products: 2022–2023 Production Statistics. Dairynews.Today. 2024. Available online: <https://dairynews.today> (accessed on 19 June 2026).
20. Vera.kz. Dairy Industry of Kazakhstan: Production Growth, Prices, and Consumption. Vera.kz. 2025. Available online: <https://vera.kz> (accessed on 19 June 2026).
21. Tengemonitor.kz. Can Kazakhstan's Dairy Industry Withstand Competition? Tengemonitor.kazgazeta.kz. 2023. Available online: <https://tengemonitor.kazgazeta.kz> (accessed on 19 June 2026).
22. Jpra-kazniipk.kz. Dairy Subcomplex of Kazakhstan: Current State and Measures to Improve Efficiency. Problemy Agrynk. 2024. Available online: <https://www.jpka-kazniipk.kz> (accessed on 19 June 2026).
23. Turanews.kz. Kazakhstan Faces Annual \$800 Million Loss Due to Meat and Dairy Imports. Turanews.kz. 2024. Available online: <https://turaneews.kz> (accessed on 19 June 2026).
24. Ulysmedia.kz. Money Down the Drain, or Why Kazakhstan's Own Dairy Products Are Not to Everyone's Taste. Ulysmedia.kz. 2023. Available online: <https://ulymedia.kz> (accessed on 19 June 2026).
25. An, M.; Zhao, Z.; Zhao, L.; Yang, J.; Gao, H.; Zhang, L.; Zhao, G.; Hou, B.; He, J.; Hung, W.-L.; et al. Yogurt Fermentation Kinetics. *Foods* **2025**, *14*, 2935. [CrossRef] [PubMed]
26. Robles-Rodríguez, C.E.; Szymańska, E.; Huppertz, T.; Özkan, L. Dynamic modeling of milk acidification. *Food Bioprod. Process.* **2021**, *128*, 41–51. [CrossRef]
27. Kennedy, M.C.; O'Hagan, A. Bayesian Calibration of Computer Models. *J. R. Stat. Soc. B* **2001**, *63*, 425–464. [CrossRef]
28. Tuo, R.; Wu, C.F.J. Calibration in Computer Models. *SIAM/ASA J. Uncertain. Quantif.* **2016**, *4*, 767–795. [CrossRef]
29. Kingma, D.P.; Ba, J. Adam: A Method for Stochastic Optimization. *arXiv* **2014**, arXiv:1412.6980.
30. Rasmussen, E.; Williams, C.K.I. *Gaussian Processes for Machine Learning*; MIT Press: Cambridge, MA, USA, 2005. [CrossRef]
31. Arango, O.; Trujillo, A.J.; Castillo, M. Inline control of yoghurt fermentation process. *J. Food Eng.* **2020**, *277*, 109885. [CrossRef]
32. Muncan, J.; Tei, K.; Tsenkova, R. Real-Time Monitoring of Yogurt Fermentation. *Sensors* **2020**, *21*, 177. [CrossRef] [PubMed]
33. Fearnhead, P.; Giagos, V.; Sherlock, C. Inference for reaction networks using the linear noise approximation. *Biometrics* **2014**, *70*, 457–466. [CrossRef] [PubMed]
34. Saro, J.; Ducháček, J.; Brožová, H.; Stádník, L.; Bláhová, P.; Horáková, T.; Hlavatý, R. Discrete Markov Chains for Dairy Cow Diseases. *Animals* **2024**, *14*, 2542. [CrossRef] [PubMed]
35. Zin, Y.; Htet, Y.; Lwin, T.C.; Tin, P. Markov-Dependent stochastic approach to lactation curves. *Smart Agric. Technol.* **2023**, *6*, 100335. [CrossRef]
36. Maw, S.Z.; Zin, T.T.; Tin, P.; Kobayashi, I.; Horii, Y. Absorbing Markov Chain Model for Dairy Cow Calving Time. *Sensors* **2021**, *21*, 6490. [CrossRef] [PubMed]
37. Freire, P.; Freire, D.; Licon, C.C. Machine learning in dairy products. *Crit. Rev. Food Sci. Nutr.* **2024**, *65*, 1878–1893. [CrossRef] [PubMed]
38. Yee, C.S.; Zahia-Azizan, N.A.; Abd Rahim, M.H.; Mohd Zaini, N.A.; Raja-Razali, R.B.; Ushidee-Radzi, M.A.; Ilham, Z.; Wan-Mohtar, W.A.A.Q.I. Smart Fermentation Technologies. *Fermentation* **2025**, *11*, 323. [CrossRef]
39. TR CU 033/2013; Technical Regulation of the Customs Union “On Safety of Milk and Dairy Products” (Adopted by Decision of the Council of the Eurasian Economic Commission No. 67 of 9 October 2013). Eurasian Economic Commission: Moscow, Russia, 2013.
40. Khanashyam, A.C.; Jagtap, S.; Agrawal, T.K.; Thorakkattu, P.; Malav, O.P.; Trollman, H.; Hassoun, A.; Ramesh, B.; Manoj, V.; Rathnakumar, K.; et al. Applications of artificial intelligence in the dairy Industry. *Comput. Electron. Agric.* **2025**, *238*, 110879. [CrossRef]

41. Ghorbaniaghdam, A.; Chen, J.; Henry, O.; Jolicoeur, M. Analyzing Clonal Variation of Monoclonal Antibody-Producing CHO Cell Lines. *PLoS ONE* **2014**, *9*, e90832. [[CrossRef](#)] [[PubMed](#)]
42. Corlu, C.G.; Akcay, A.; Xie, W. Stochastic simulation under input uncertainty. *Oper. Res. Perspect.* **2020**, *7*, 100162. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.