Journal of Food Engineering 393 (2025) 112467

Contents lists available at ScienceDirect

Journal of Food Engineering

journal homepage: www.elsevier.com/locate/jfoodeng

# Digital twin for predicting and controlling food fermentation: A case study of kombucha fermentation

Songguang Zhao [a], Tianhui Jiao [b], Selorm Yao-Say Solomon Adade [b], Zhen Wang [a], Qin Ouyang [a,**], Quansheng Chen [a,b,*]

[a] School of Food and Biological Engineering, Jiangsu University, Zhenjiang 212013, PR China
[b] College of Food and Biological Engineering, Jimei University, Xiamen 361021, PR China

ARTICLE INFO

*Keywords:*
Digital twin
Artificial intelligence
Food fermentation
Process prediction
Optimal control

ABSTRACT

In the era of rapid advancements in computing and the Internet of Things, the food fermentation sector is undergoing a digital and intelligent transformation. This research developed a food fermentation prediction and control system based on digital twin technology. The system employs multi-scale feature extraction and convolution feature fusion to establish partial least squares (PLS) prediction models for C source and bacterial concentration. The results showed that the PLS prediction models of C source and bacterial concentration exhibited excellent performance, with RMSEP of 0.5538 mg/mL and 0.0558 (Au), and RPD of 5.63 and 6.52, respectively. An optimal control system for the fermentation process was constructed by integrating the prediction models with a genetic algorithm (GA), yielding satisfactory simulation and testing outcomes. The study showed that the proposed digital twin-based fermentation prediction and control system offers superior robustness and reliability, advancing the digital and intelligent development of the food fermentation industry.

## 1. Introduction

In the midst of exploring the modernization of traditional fermented foods, the field of food processing is actively seeking to incorporate cutting-edge technologies such as Artificial Intelligence (AI) and the Internet of Things (IoT), aiming to optimize the quality, safety, and unique flavor characteristics during food processing (Zhang et al., 2023). Among these innovative technologies, Digital Twin (DT) technology stands out as one of the most promising tools (Liu et al., 2023; Qi et al., 2021). By creating a virtual mirror image of the physical processing system, DT technology can simulate and analyze complex processing dynamics in real-time, providing crucial feedback information and laying a solid foundation for formulating optimal strategies (Liu, 2024). Coupled with new hardware devices, such as IoT sensors, these sensors enable real-time connectivity between the physical processing environment and the DT model, thereby facilitating continuous improvement in food processing lines and final products (Yao et al., 2022; Zhang et al., 2022).

Given the popularity of kombucha fermentation products due to their unique flavor components and health benefits (Barakat et al., 2022), a thorough exploration of the scientific basis of their fermentation mechanism and flavor formation is particularly crucial for advancing new fermentation technologies and optimizing traditional fermentation processes (Vinestock et al., 2024; Yu et al., 2022). Currently, a significant trend in food fermentation science is the application of DT technology for real-time monitoring and controlling of fermentation processes (Cabeza-Gil et al., 2023; Zhang et al., 2023). This technology can effectively analyze the process variability caused by differences in fermentation production, thereby providing the possibility for highly personalized fermentation control (Tancredi et al., 2023).

Recent researches have ventured into applying DT technology for developing pH prediction model in cheese fermentation processes and control strategies for fed-batch cultivation of *Saccharomyces cerevisiae* (Appl et al., 2020; Li et al., 2020), demonstrating its immense potential in enhancing product quality and efficiency. However, in food fermentation processes, the establishment of fermentation process models and multi-parameter decoupling control have emerged as pressing challenges, given their heavy reliance on the complex nature of fermentation

This article is part of a special issue entitled: Digital Twins published in Journal of Food Engineering.

* Corresponding author. School of Food and Biological Engineering, Jiangsu University, Zhenjiang 212013, PR China.
** Corresponding author.
*E-mail addresses:* oyqyf@ujs.edu.cn (Q. Ouyang), qschen@ujs.edu.cn (Q. Chen).

https://doi.org/10.1016/j.jfoodeng.2025.112467
Received 10 September 2024; Received in revised form 9 November 2024; Accepted 6 January 2025
Available online 25 January 2025

itself (González-Hernández et al., 2022; Rydal et al., 2024). Currently, the five-dimensional DT model is the mainstream approach for constructing the DT framework, encompassing the physical system, virtual system, twin data, business, and connectivity (Duan et al., 2021; Hu et al., 2022; Wang et al., 2023). Among these, the virtual system and twin data can provide solutions for process modeling and decoupling control, respectively (Liu, 2024; Mahanty, 2023). Therefore, in this study, when constructing the five-dimensional DT model for food fermentation, the focus is on the virtual system and twin data.

This study aims to develop a DT model for monitoring and controlling the fermentation process of Kombucha (Scheme 1). Specifically, (1) the fermentation system integrates various IoT sensors for real-time monitoring of process parameters such as bacterial concentration, C source consumption rate, pH, dissolved oxygen, and others; (2) a process model is established based on the deep features extracted by the proposed multi-scale convolutional filters; (3) a multi-parameter decoupling control system is developed based on the process model and multivariate optimization algorithms; (4) validation experiments are conducted based on the developed fermentation DT model.

## 2. Materials and methods

The fermentation process data is collected through IoT sensors to monitor key parameters of the physical entity, including bacterial concentration (BC), C source (CS) concentration, total acid (TA) concentration, pH, dissolved oxygen (DO), and L-Theanine (LT). By monitoring this information in real-time, changes in key indicators at the next moment can be predicted, forming the basis for optimal control of the food fermentation process. This study focuses on the prediction model and optimal control strategy of kombucha fermentation, which correspond to the virtual system and twin data of DT, respectively.

### 2.1. Kombucha fermentation

The experiment was carried out in a 5 L bioreactor (Purchased from Green Bio-engineering Co., LTD., Zhenjiang, China) in the laboratory with the temperature set at 30 °C and the stirring speed set at 150 rpm (Aung et al., 2022), referred to Fig. S1. The specific details of the fermentation experiment process are provided in the experimental introduction below Fig. S1. Kombucha culture medium is made by mixing 30 g of instant tea powder (supplyed from Nanjing Rongdian Food Technology Co., LTD., Nanjing, China), 60 g of sucrose (purchased from Xi'an Musen Bioengineering Co., LTD., Xi'an, China), and 3000 mL of sterile water. The fermentation cycle lasted for 72 h, with a sampling interval of 2 h, resulting in 37 samples collected per batch. The process variables selected were current fermentation time (t), BC, CS, TA, DO, pH and LT. Data from 5 batches under normal production conditions were collected, with 4 batches used as training samples and the remaining as test samples to validate the process model.

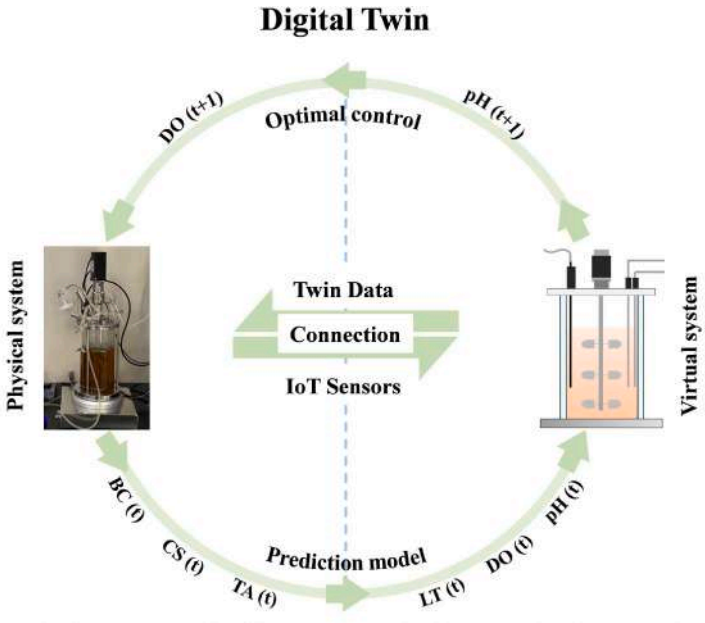
### 2.2. Experimental materials and physicochemical value determination

In the determination of physicochemical values, the main reagents utilized include anthranone, sulfuric acid (98%), LT standard (purity≥99%), acetonitrile (chromatographic grade), sodium hydroxide solution (1 mol/L). All chemicals were purchased from Youchen (Xiamen) Instrument Equipment Co., LTD. (Xiamen, China). Ultrapure water (18.35 MΩ) is applied to dispense reagents.

An ultraviolet spectrophotometer (UV1920, Shanghai Lengguang Technology Co., LTD., Shanghai, China) was employed to measure the optical density (OD) value of bacteria at 600 nm (Krepelka et al., 2020), and a high performance liquid chromatography (HPLC) (EasySep-3030, Shanghai Unimicro Technologies Co., LTD., Shanghai, China) was used to determine LT (Aung et al., 2022). The content of CS was determined by anthrone sulfuric acid method and TA was determined by potentiometric titration method. DO and pH were measured using the sensor electrodes contained in the bioreactor. Additionally, an on-line visible near-infrared (vis-NIR) monitoring system was developed for real-time detection of BC, CS, TA and LT (Mayr et al., 2021; Zhao et al., 2023). The vis-NIR monitoring system was described in detail in Fig. S2 and Table S1. Based on this, the IoT system of kombucha fermentation was built (described in Fig. 1), and the real-time transmission of BC, CS, TA, DO, pH and LT data was realized.

### 2.3. Construction of the prediction model

Prediction models were developed for the concentration of BC and CS, offering rational insights into the optimal control of the fermentation




**Scheme 1.** The schematic diagram of this study.

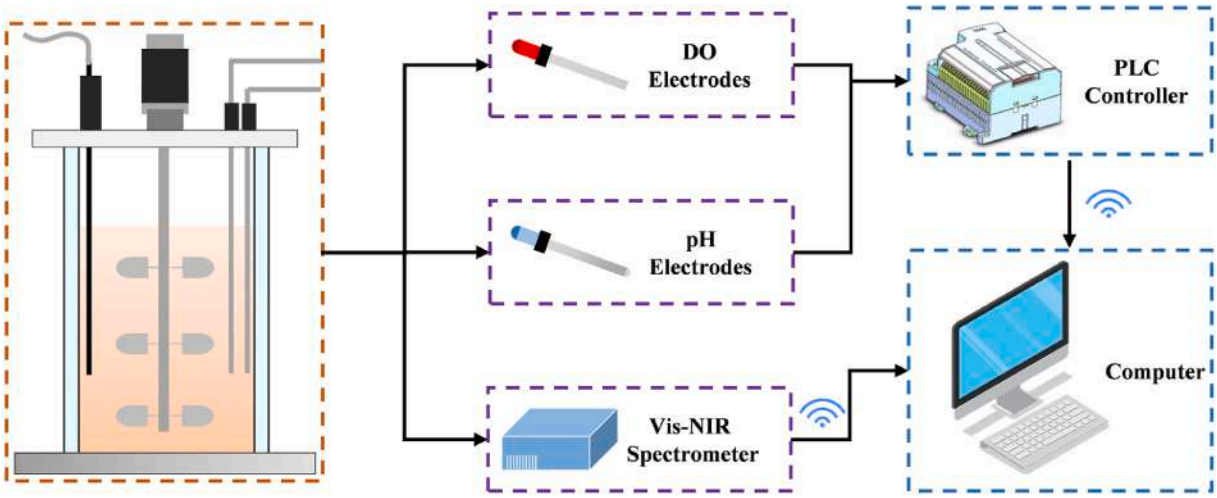



**Fig. 1.** The IoT system of kombucha fermentation.

process. Subsequently, a multi-scale filter was designed to efficiently extract and hybrid relevant process data features. Using these fused features, prediction models for the target values were constructed.

### 2.3.1. The construction of process data sets

The process data of the first 36 samples in each batch corresponded to the target value at the next moment, as illustrated in Fig. 1. Each batch yielded 36 groups of process data, resulting in 144 groups across four batches. When establishing CS prediction model, t, BC,TA, DO, pH and LT were used as input variables. When building the BC prediction model, t, CS,TA, DO, pH and LT were used as input variables. The functional form of the prediction models of CS and BC were respectively as follows:

$$f_{CS}(t+\Delta t) = f(t, BC(t), TA(t), DO(t), pH(t), LT(t)) \quad (1)$$

$$f_{BC}(t+\Delta t) = f(t, CS(t), TA(t), DO(t), pH(t), LT(t)) \quad (2)$$

where t is the current sampling time, $\Delta t$ is the next sampling time.

Due to differing units of process data, normalization was applied (Kammoun et al., 2024). Max-min normalization was used, and the calculation is shown in equation (1):

$$X_n = \frac{(X_o\text{-MIN}(X))}{(\text{MAX}(X)\text{-MIN}(X))} \quad (3)$$

where $X_o$ and $X_n$ are the original and normalized values, respectively, and MAX (X) and MIN (X) represent the maximum and minimum values of each dimension.

The SPXY method was used to divide the training and test sets in a 7:3 ratio (Sun et al., 2021). The training set contained 100 samples, while the test set included 44.

### 2.3.2. Design of single-scale filter

The historical data in the fermentation process contained rich information, and features could be extracted with strong correlation to the target. Therefore, mining historical data features has become the premise and research focus for establishing reliable and stable prediction models. In recent years, convolutional neural networks (CNN) have been widely used for their powerful feature extraction capabilities, such as one-dimensional CNN for spectral analysis and two-dimensional CNN for computer vision (Meng et al., 2022; Sun et al., 2023). Inspired by this, a feature filter was designed using convolutional modules to extract key features of historical data of the food fermentation process. As shown in Fig. 2, the filter was designed to incorporate 5 convolution modules of the same size by taking advantage of the CNN local connection and weight-sharing properties (Song et al., 2023). Each

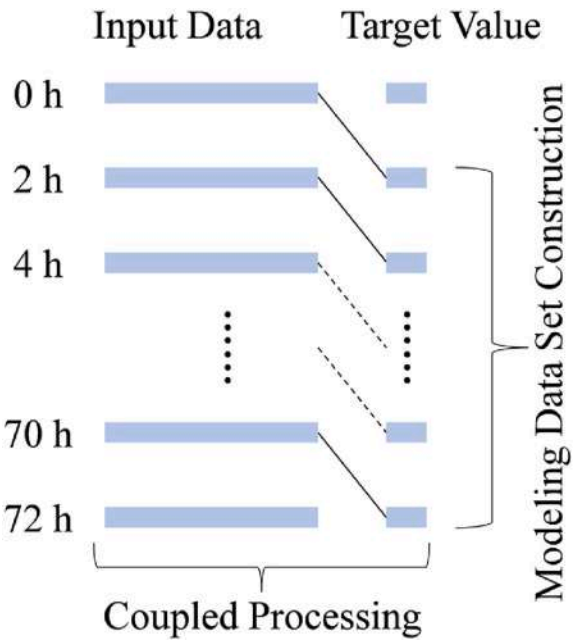



**Fig. 2.** The construction of process data.

convolution module contained convolution (Conv), batch normalization (BN), and maximum pooling (MaxPool) operations. Because the 5 convolution modules had the same size and described the features at the same scale, the filter was called single-scale filters. Subsequently, the features of the fermentation process data extracted through 5 convolutional modules were flattened and used as the output of a single-scale filter. The calculation process is shown in equation (2):

$$\text{Output} = \bigvee_{i=1}^{5} \text{MaxPool}_i(\text{BN}_i(\text{Conv}_i(X))) \quad (4)$$

Where i can be 1, 2, 3, 4, and 5, X represents the input vector, $\bigvee$ stands for the flattening loop operation.

### 2.3.3. Design of multi-scale filter

The features extracted by the single-scale filter were relatively rudimentary, prompting the design of a multi-scale filter to address this limitation (Feng et al., 2024). Single-scale filters were equivalent to longitudinal extraction of the process data feature, while multi-scale

filters were equivalent to horizontal extraction, forming a two-dimensional mesh filter. As depicted in Fig. 3, in this study, single-scale filters with convolution kernels [2,1], [3,1], and [4,1] were designed and assembled into three-scale filters. Although the five convolution kernels of a single-scale filter were of the same size, the weights of each convolution kernel were different, achieving feature extraction of process data from multiple perspectives. The addition of multi-scale changed the feature extraction method from line to surface, increasing the diversity of extracted features and enabling the deep mining of process data features.

### 2.3.4. Multi-scale feature fusion

The features extracted by the multi-scale filter need to be fused to obtain the model input data. As displayed in Fig. 4, this study investigated three fusion strategies: average fusion (Ahmad and Khan, 2021), attention-weighted average fusion (Li et al., 2024), and convolution fusion (Zhang et al., 2021), to achieve the fusion of each filter feature map data. Average fusion refers to the average value of features under three scales, and the calculation method can be seen in equation (3):

$$X\_average\_fusion = \frac{\sum X_{[i,1]}}{3} \tag{5}$$

Where, $X_{[i,1]}$ is the feature map vector at different scales, and $i = 2, 3, 4$;

Attention-weighted average fusion means that the attention value of each scale output feature map is extracted by average pooling first, then an attention factor is added, and finally, an activation function is used to activate. The calculation method can be seen equation (4):

$$\begin{cases} Value\_attention(i) = AveragePooling\left(X_{[i,1]}\right) \\ Sigmoid(X) = \dfrac{1}{1+e^{-X}} \\ Attention(i) = Sigmoid(\alpha * Value\_attention(i)) \\ X\_attention\_average\_fusion = \dfrac{\sum Attention(i) * X_{[i,1]}}{3} \end{cases} \tag{6}$$

Where $X_{[i,1]}$ is the feature map vector at different scales, α s the attention factor with a value of 0.8, and $i = 2, 3, 4$.

Convolution fusion referes to the use of convolution method to map the features of three scales to a new space. The convolution kernel was set to size [3,1], step size 1, and filled with '0'. The calculation method can be seen in Fig. 5 or equation (5):

$$X\_Conv\_fusion = \prod_{i=1}^{3} Conv_i(X_i) \tag{7}$$

Where $i = 1, 2, 3$, $X_i$ represents the feature map at the ith scale, $\prod$ stands for flattening operation.

### 2.3.5. Construction and evaluation of predictive model

Partial Least Squares (PLS) is a popular data-driven soft sensor development method because it can handle high-dimensional and multicollinear data (Fourie et al., 2020; Jiang et al., 2020). This study used fusion feature data as input to PLS to train prediction models of CS and BC. Although the fermentation process exhibits typical non-linear, time-delay, and strong-coupling characteristics (Mohan et al., 2022), the PLS model is inherently linear, which may not adequately capture these complexities. Therefore, employing a multi-scale feature filter becomes crucial for achieving non-linear mapping of process data (Meng et al., 2022), serving as a non-linear complement to linear modeling methods.

Several indices were used to evaluate the model's performance (Huang et al., 2021). The coefficient of determination ($R^2$) assessed the correlation between input and output data, with the $R_C^2$ for the calibration and $R_P^2$ for the prediction sets. Root Mean Square Error (RMSE) evaluated the model's prediction error, with RMSEC and RMSEP representing the calibration and prediction sets, respectively. The Mean Absolute Error (MAE) and the Ratio of Performance to Deviation (RPD) further evaluated the error and reliability of the model. The formulas for each indicator are as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_{i,actual} - y_{i,predicted}\right)^2}{\sum_{i=1}^{n}\left(y_{i,actual} - \bar{y}_{actual}\right)^2} \tag{8}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(y_{i,actual} - y_{i,predicted}\right)^2}{n\text{-}1}} \tag{9}$$

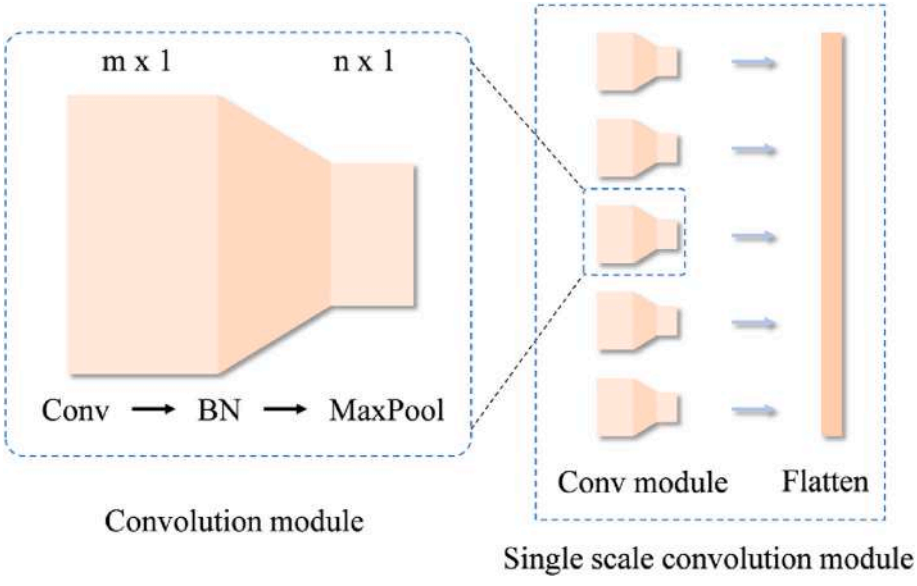



**Fig. 3.** The single-scale filter design.

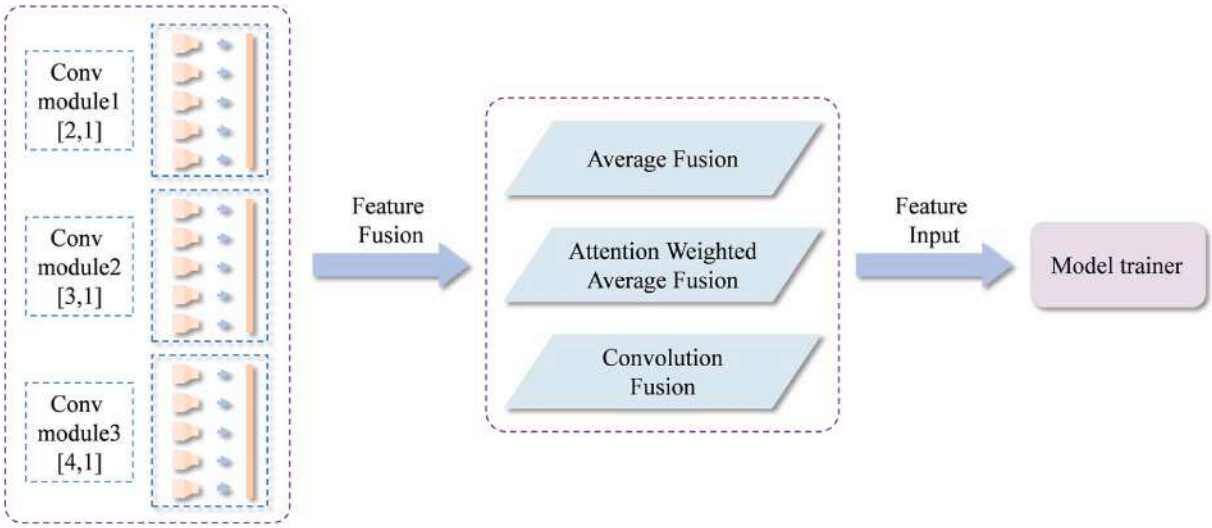



**Fig. 4.** Design of multi-scale filters and feature fusion.

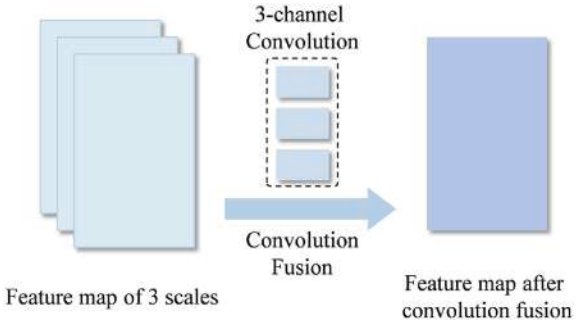



**Fig. 5.** Schematic diagram of convolution fusion.

$$RPD = \frac{SD_V}{\sqrt{\frac{\sum_{i=1}^{n}\left(y_{i,actual}\text{-}y_{i,predicted}\right)^2}{n\text{-}1}}} \tag{10}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_{i,actual}\text{-}y_{i,predicted}\right| \tag{11}$$

Where $y_{i,actual}$ is the actual value, $y_{i,predicted}$ is the predicted value, $\overline{y}_{actual}$ is the mean of actual values, n is the number of samples, and $SD_V$ is the standard deviation of the validation set.

### 2.4. The construction of optimal control logic framework

Kombucha fermentation is a complex microbial process characterized by uncertainty (Augustin et al., 2023). The process parameters exhibit a high degree of coupling, such that any adjustment to one parameter inevitably influences others, thereby complicating the overall process control. To mitigate this issue, a method is imperative for simultaneously adjusting multiple parameters, thereby minimizing the coupling effect. In the context of Kombucha fermentation, pH and DO play pivotal roles in microbial growth. Consequently, the objective of this study was to devise a system that optimizes the regulation of both pH and DO. Recently, meta-heuristic algorithms have been widely used in engineering for their optimization capabilities (Ansari and Daxini, 2021). This study employed the Genetic Algorithm (GA), a classical meta-heuristic approach, as central control algorithm in control strategy (Li, 2023). GA obtains the optimal value of the fitness function by seeking the optimal combination of control actions. The GA used a population of 20 individuals, a maximum of 100 iterations, and an individual boundary of [−5, 5]. The pH and DO were encoded as an individual in the GA, and the optimal combination was determined when the optimal individual was identified. In contrast to traditional control methods, this study introduced a control strategy akin to an expert recommendation system, designed to supplant manual labor. However, this strategy necessitated collaboration with the actuator to complete the entire control action. The actuator can be realized by combining automation elements and classical control algorithms, such as Proportional-Integral-Derivative (PID).

#### 2.4.1. Individual encoding and decoding

As shown in Fig. 6A, to increase the possibility of individual variation and better find the optimal solution, individuals were first encoded as decimal numbers in the range [−5, 5], and then converted to binary numbers. The individual dimension is 16, with DO and pH each represented by eight-bit binary numbers. As shown in Fig. 6B, to obtain the true DO and pH values, decode the solved optimal individual by converting the eight-bit binary numbers to decimal numbers between [0, 255].

#### 2.4.2. Design of fitness function

The fitness function guides the search algorithm in a better direction, and a well-designed fitness function can improve the performance and efficiency of the algorithm (Hormozi et al., 2022). In kombucha fermentation, higher BC and consumption of CS indicate favorable progress. Therefore, the growth rate of bacteria and the consumption rate of CS were used to guide the optimization control process. The algorithm flow chart of the fitness function is depicted in Fig. S3.

#### 2.4.3. Construction of optimized control system

As illustrated in Fig. 7, an advanced control logic framework tailored for the kombucha fermentation process had been constructed. This framework integrated prediction models for CS and BC, along with a GA optimization technique. The aim of control was to enhance both the efficiency and quality of the fermentation process. Specifically, fermentation state parameters, which were captured by IoT sensors, were utilized as inputs into the prediction models to derive estimates of CS and BC availability. Subsequently, the GA optimization algorithm was employed to identify the optimal combination of fermentation state parameters that maximized CS consumption and BC concentration. Based on the outcomes of this optimization, the fermentation system underwent adjustments in DO and pH levels, thereby achieving a decoupling optimization control across multiple parameters.

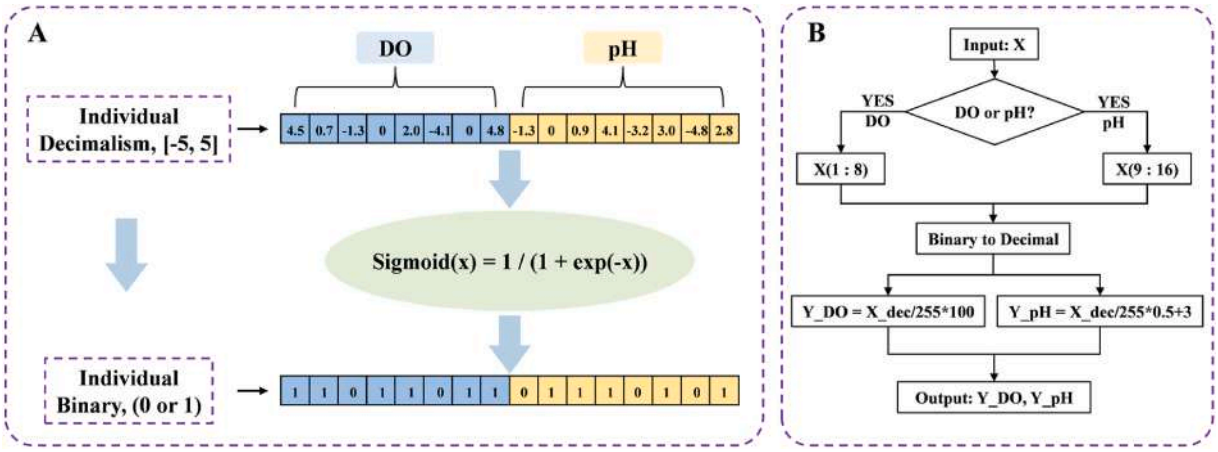



**Fig. 6.** Individual encoding and decoding of GA algorithm.

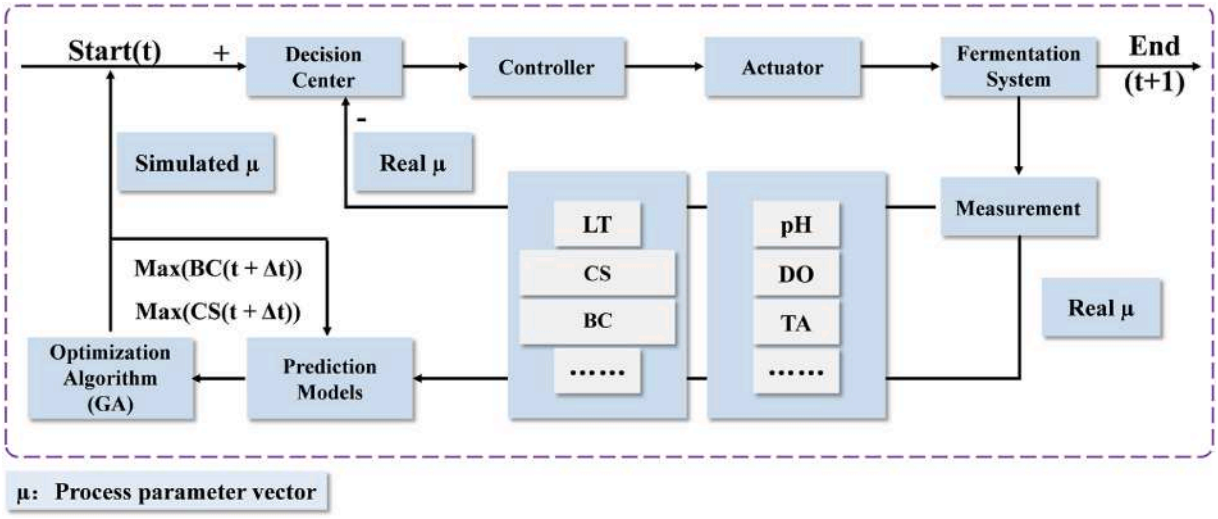



**Fig. 7.** The optimal control logic framework of kombucha fermentation process.

*2.4.4. Simulation experiment*

To verify the effectiveness of the proposed method, a batch of fermentation experiments was simulated using the developed optimized control system. Three types of simulation experiments were conducted with different regulatory objectives: maximum CS consumption, maximum bacteria number, and segmented control. Segmented control aimed for maximum bacteria numbers in the early stage of fermentation and maximum CS consumption in the late stage.

*2.5. Testing experiment*

A comprehensive testing experiment was carried out to evaluate the performance of the developed DT. This involved conducting two sets of fermentation experiments: one without control and another with control implemented. The primary objective of the testing experiment was to optimize the control of CS consumption. This was achieved by utilizing the DT to calculate the optimal DO and pH levels at 12 h, 24 h, 36 h, 48 h, and 60 h, respectively. Subsequently, these calculated values were employed to regulate and control the fermentation system.

*2.6. Statistical analysis and tools*

The data used in the experiment were average values from three measurements. Origin 2018 (OriginLab Corporation, USA) was used to create charts, while MATLAB R2014a (MathWorks Inc., USA) was employed for calculations and analyses.

## 3. Results and discussion

*3.1. The changing trend of process parameters*

During the fermentation process, the process data, including CS, TA, BC, DO, pH, and LT, were measured by utilizing IoT sensors, as depicted in Fig. 8. Given that the kombucha liquid inoculum was self-cultivated within the laboratory, there existed certain variability in both the CS concentration and pH value. Consequently, the initial CS and pH levels at the commencement of fermentation exhibit fluctuations, which precisely increased the diversity of the fermentation and facilitated the enhancement of the generalization ability of the model. After approximately 15 h, the oxygen had been nearly depleted (Fig. 8D). By roughly 40 h, the consumption of the CS had stabilized (Fig. 8A), aligning closely with the microbial growth trend (Fig. 8C). At approximately 25 h, the TA content attained its peak, which corresponded with the observed pH shift (Fig. 8E). Notably, the level of LT began to decrease rapidly around 25 h (Fig. 8F), potentially due to the significant accumulation of metabolic products (Júnior et al., 2022).

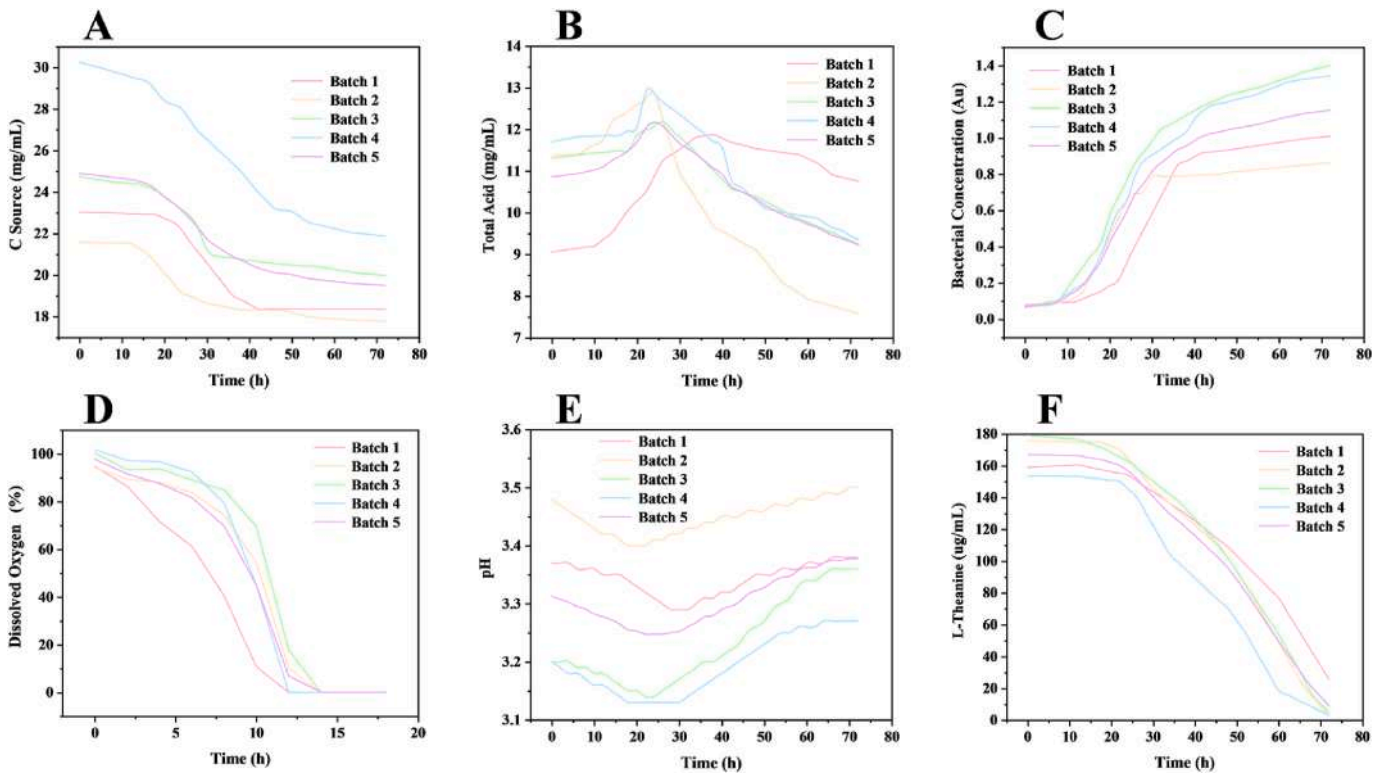



**Fig. 8.** The changing trend of CS (A), TA (B), BC (C), DO (D), pH (E), and LT (F) during kombucha fermentation.

### 3.2. The results of CS and BC prediction models

To evaluate the investigate effectiveness of the designed filter, the performance of the PLS models with 7 different input features were compared. The included the original data, single-scale features with convolution kernel [2,1] (Scale1-PLS), single-scale features with convolution kernel [3,1] (Scale2-PLS), single-scale features with convolution kernel [4,1] (Scale3-PLS), average features of three scales (Ave_PLS), attention weighted average features of three scales (Atten_PLS), and convolution fusion features of three scales (Conv_PLS).

As shown in Table 1, the performance of the 7 PLS models for the CS were compared. Most of the features extracted by both single-scale and multi-scale filters improved the PLS model performance. Additionally, the performance of the best PLS models based on multi-scale features was generally better than that of single-scale. The Conv_PLS model achieved the best performance, with $R_P^2$ of 0.9677, RMSEP of 0.5538 mg/mL, MAE of 0.4435 mg/mL, and RPD of 5.63. Compared to PLS, the Conv_PLS model demonstrated a significant improvement, with a 7% increase in correlation and a 42% reduction in error.

As shown in Table 2, the performance of the 7 PLS models for BC was compared. Both Scale3-PLS and Conv_PLS achieved excellent model performance, with RPD values exceeding 6. Compared to Scale1-PLS and Scale2-PLS, Scale3-PLS showed better performance, possibly due to the stronger feature extraction capability of the convolution kernel [4,1] in Scale3-PLS. Furthermore, Conv_PLS outperformed Scale3-PLS with an $R_P^2$ of 0.9759, RMSEP of 0.0558 (Au), MAE of 0.0435 (Au), RPD of 6.52, as the multiple scales of features provided a more comprehensive description of the process data (Feng et al., 2024). Compared to PLS, the

**Table 1**
Performance comparison of 7 PLS models of C source.

| Modeling Methods | PC[a] | Train set | | Test set | | MAE | RPD |
|---|---|---|---|---|---|---|---|
| | | $R_C^2$ | RMSEC | $R_P^2$ | RMSEP | | |
| PLS | 6 | 0.9171 | 0.9144 | 0.9028 | 0.9604 | 0.7880 | 3.24 |
| Scale1-PLS[b] | 10 | 0.9189 | 0.9188 | 0.8745 | 1.0912 | 0.7880 | 2.86 |
| Scale2-PLS[c] | 10 | 0.9546 | 0.6880 | 0.9462 | 0.7147 | 0.5730 | 4.36 |
| Scale3-PLS[d] | 10 | 0.9711 | 0.5484 | 0.9616 | 0.6035 | 0.4796 | 5.16 |
| Ave_PLS[e] | 10 | 0.9236 | 0.8922 | 0.8979 | 0.9843 | 0.7718 | 3.17 |
| Atten_PLS[f] | 10 | 0.9229 | 0.8963 | 0.8972 | 0.9879 | 0.7775 | 3.15 |
| **Conv_PLS[g]** | **10** | **0.9720** | **0.5399** | **0.9677** | **0.5538** | **0.4435** | **5.63** |

$R_c^2$ = determination coefficient of calibration; RMSEC = root mean square error of calibration; $R_P^2$ = determination coefficient of prediction; RMSEP = root mean square error of prediction; MAE = Mean absolute error; RPD = The ratio of the standard deviation of validation set to RMSE of prediction; Data in bold are the best result for each method.

[a] PC = The principal component.
[b] Scale1-PLS = A single scale filter with convolution kernel [2,1].
[c] Scale2-PLS = A single scale filter with convolution kernel [3,1].
[d] Scale3-PLS = A single scale filter with convolution kernel [4,1].
[e] Ave-PLS = The average of three scale features.
[f] Atten-PLS = The average of three scale features weighted by attention.
[g] Conv-PLS = Fusion of three scale features by convolution.

**Table 2**
Performance comparison of 7 PLS models of bacterial concentration.

| Modeling Methods | PC[a] | Train set | | Test set | | MAE | RPD |
|---|---|---|---|---|---|---|---|
| | | $R_C^2$ | RMSEC | $R_P^2$ | RMSEP | | |
| PLS | 6 | 0.9092 | 0.1275 | 0.8172 | 0.1536 | 0.1355 | 2.37 |
| Scale1-PLS[b] | 10 | 0.9534 | 0.0913 | 0.9279 | 0.0965 | 0.0757 | 3.77 |
| Scale2-PLS[c] | 10 | 0.9689 | 0.0746 | 0.9458 | 0.0837 | 0.0706 | 4.34 |
| Scale3-PLS[d] | 10 | 0.9764 | 0.0650 | 0.9717 | 0.0606 | 0.0492 | 6.01 |
| Ave_PLS[e] | 10 | 0.9628 | 0.0816 | 0.9593 | 0.0725 | 0.0582 | 5.01 |
| Atten_PLS[f] | 10 | 0.9629 | 0.0815 | 0.9588 | 0.0729 | 0.0587 | 4.98 |
| **Conv_PLS[g]** | **10** | **0.9858** | **0.0504** | **0.9759** | **0.0558** | **0.0435** | **6.52** |

$R_c^2$ = determination coefficient of calibration; RMSEC = root mean square error of calibration; $R_p^2$ = determination coefficient of prediction; RMSEP = root mean square error of prediction; MAE = Mean absolute error; RPD = The ratio of the standard deviation of validation set to RMSE of prediction; Data in bold are the best result for each method.

[a] PC = The principal component.
[b] Scale1-PLS = A single scale filter with convolution kernel of [2,1].
[c] Scale2-PLS = A single scale filter with convolution kernel of [3,1].
[d] Scale3-PLS = A single scale filter with convolution kernel of [4,1].
[e] Ave-PLS = The average of three scale features.
[f] Atten-PLS = The average of three scale features weighted by attention.
[g] Conv-PLS = Fusion of three scale features by convolution.

Conv_PLS model showed a significant improvement, with a 19% increase in correlation and a 64% reduction in error.

### 3.3. Validation of CS and BC prediction models

The prediction models for CS and BC were validated using batch 5 fermentation data to assess their reliability and generalization ability. As presented in Fig. 9A, the predicted value for the CS prediction model closely matched the true value, with an average error of only 0.2681 mg/mL, and the predicted trend was consistent with the true value. Similarly, Fig. 9B illustrates the predicted values for BC were very close to the true value, with an average error of 0.06 (Au), effectively simulating the trend of BC.

### 3.4. The effect of IoT sensor noise on DT

In this work, random perturbations (10% of the range of change) were added to the input data, as shown in Table 3. In the context of error perturbation, the RMSE of CS was observed to be less than 2.19, with the MAE falling below 1.84. The RMSE of BC was less than 0.09, and the MAE did not exceed 0.07. It was noteworthy that variations in DO leaded to prediction errors in CS, potentially due to the intimate correlation between yeast aerobic respiration and DO levels. Furthermore, perturbations in pH impaired the predictive accuracy of BC, likely because bacterial growth was reliant on the pH of their growth environment.

### 3.5. The result of simulation experiment

Fig. 10A illustrates the optimal control of the fermentation process, which was guided by the consumption of the CS. Throughout the process, the consumption of the CS was consistently higher than that of natural fermentation. As shown in Fig. 10B, the pH was maintained at a level of about 3.4 during this time. In contrast, under natural conditions, the DO was depleted after approximately 24 h, causing the fermentation system to enter an oxygen-free mode. However, when the process was guided by CS consumption, oxygenation was required after 24 h (Fig. 10C), likely due to the yeast's increased oxygen requirements for aerobic respiration and reproduction. These results were obtained using the CS process model and fitness function. As shown in Fig. 10D, the control of the fermentation process was guided by the number of bacteria. In the early stage of fermentation, the increase in the number of bacteria was not significant, but it changed substantially in the later stage. Overall, the pH was maintained at around 3.1 (Fig. 10E). After 24 h, oxygen was also required for the yeast's aerobic respiration and reproduction (Fig. 10F).

Fig. 11 illustrates a two-stage fermentation control: early stage (8–24 h) governed by bacterial count, and late stage (24–50 h) by CS consumption. Fig. 11A and B demonstrate increased bacterial numbers

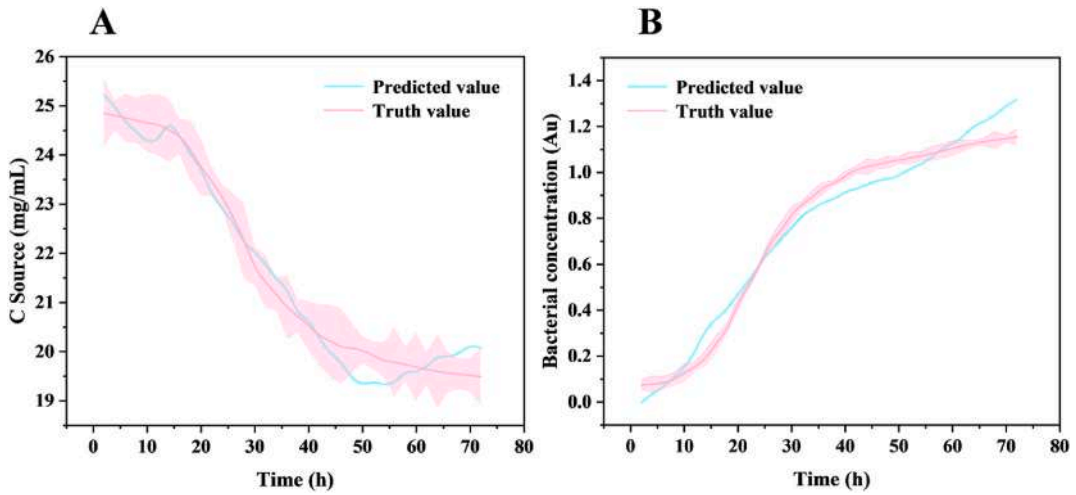



**Fig. 9.** The verification experiment results of C sources (A) and bacterial concentration (B).

**Table 3**
The impact of IoT sensor noise on prediction models of DT.

| IoT sensor input | Disturbance range | CS (mg/mL) | | BC (Au) | |
|---|---|---|---|---|---|
| | | RMSE | MAE | RMSE | MAE |
| – | – | 1.9492 ± 0.8161 | 1.6845 ± 0.6808 | 0.0514 ± 0.0094 | 0.0417 ± 0.0069 |
| pH | [-0.02, 0.02] | 1.9560 ± 0.8127 | 1.6435 ± 0.6566 | 0.0817 ± 0.0068 | 0.0649 ± 0.0043 |
| DO | [-5, 5] (%) | 2.1412 ± 0.8681 | 1.8426 ± 0.7351 | 0.0541 ± 0.0074 | 0.0427 ± 0.0059 |
| LT | [-9, 9] (μg/mL) | 1.9423 ± 0.8202 | 1.6874 ± 0.6699 | 0.0550 ± 0.0103 | 0.0429 ± 0.0076 |
| TA | [-1.5, 1.5] (mg/mL) | 2.1847 ± 0.7697 | 1.8232 ± 0.7727 | 0.0850 ± 0.0155 | 0.0651 ± 0.0085 |

RMSE = Root mean square error; MAE = Mean absolute error.

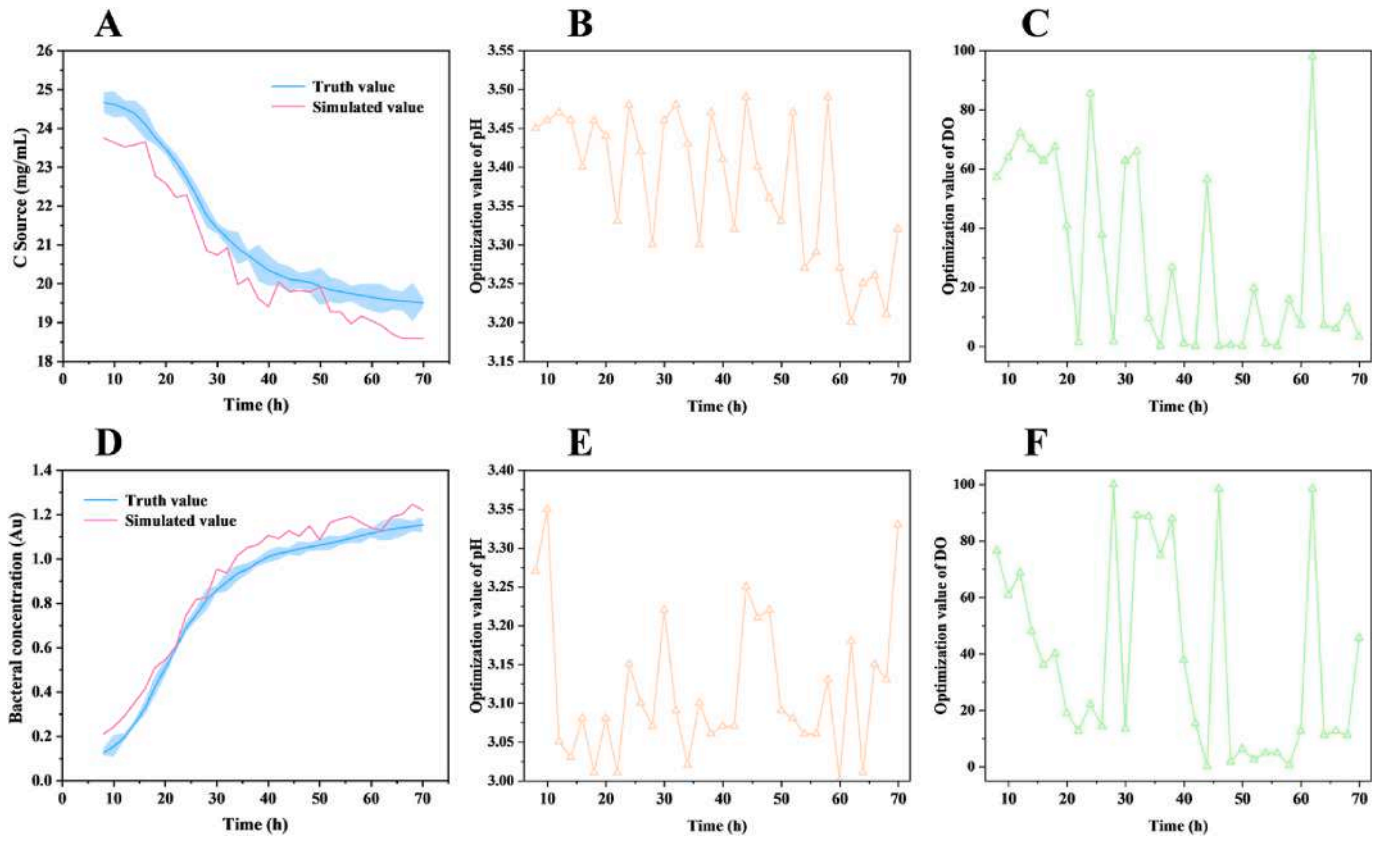



**Fig. 10.** C source optimization control simulation results (A) and pH (B) and dissolved oxygen (C) that need to be regulated; The bacterial concentration optimization control simulation results (D) and pH (E) and dissolved oxygen (F) that need to be regulated.

early on, followed by heightened CS consumption later. Fig. 11C and D reveal that post-25 h, pH adjustment to 3.4 and oxygenation were required, aligning with single-factor optimization control findings.

### 3.6. The result of testing experimen

After implementing optimized control measures, a significant increase in the consumption of CS was observed, as depicted in Fig. 12A. The most pronounced effects were achieved at the 12 h and 24 h, potentially attributed to the enhanced aerobic respiration of yeast and acetic acid bacteria. The increase of pH favored the growth of yeast, and in addition the passage of oxygen favored the rapid multiplication of yeast and acetic acid bacteria (Jayabalan et al., 2014). This was consistent with the results of the DT reasoning, which increased the pH (Fig. 12B) and DO (Fig. 12C) of the fermentation broth.

### 3.7. Discussion and perspective

The five-dimensional model is one of the most popular methods to for constructing DT, particularly in the context of food fermentation. The construction of virtual systems and twin data is relatively complex, making it a research focus in food fermentation DT (Cabeza-Gil et al., 2023; Duan et al., 2021; Tancredi et al., 2023). Using kombucha fermentation as an example, this study proposes a method for constructing virtual systems and twin data for food fermentation DT. This involves the process model construction and optimization control method in the kombucha fermentation process, effectively building an expert system for fermentation.

The food fermentation process was quite complicated, making it crucial to clean and filter fermentation process data. Features extracted from deep mining process data are essential for building robust and reliable prediction models. According to Tables 1 and 2, models based on extracted features generally performed better than the original models. Among the three multi-scale feature fusion methods, convolution fusion achieved the best performance. These results indicate that effective feature extraction and fusion can enhance the reliability and confidence of process prediction models.

As shown in Fig. 5A and B, DO and pH were encoded using 8-bit binary numbers, extending the search space of the optimal individual to 16 dimensions, thereby increasing the likelihood of finding the

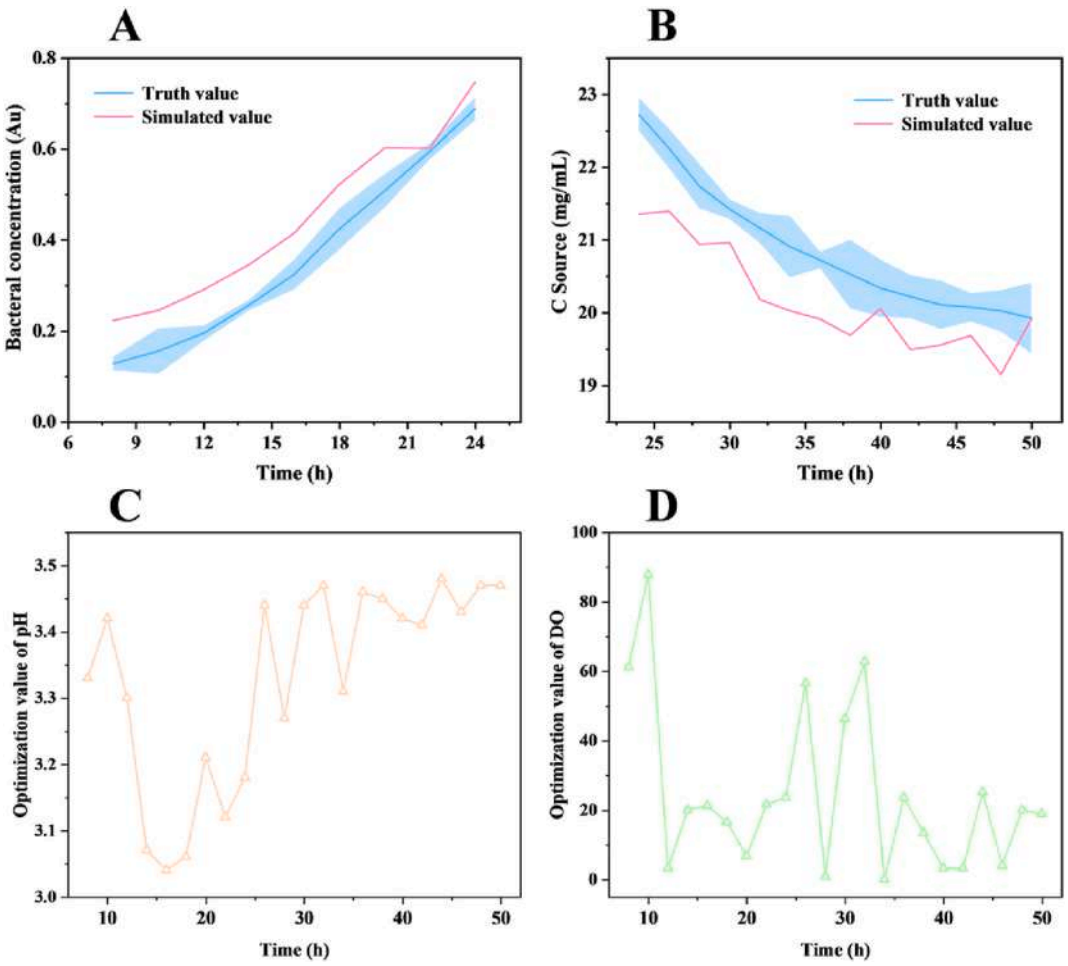



**Fig. 11.** Simulated values of bacterial concentration (A) and source C concentration (B) in the stage control process, as well as pH values (C) and dissolved oxygen (D) that need to be regulated throughout the fermentation process.

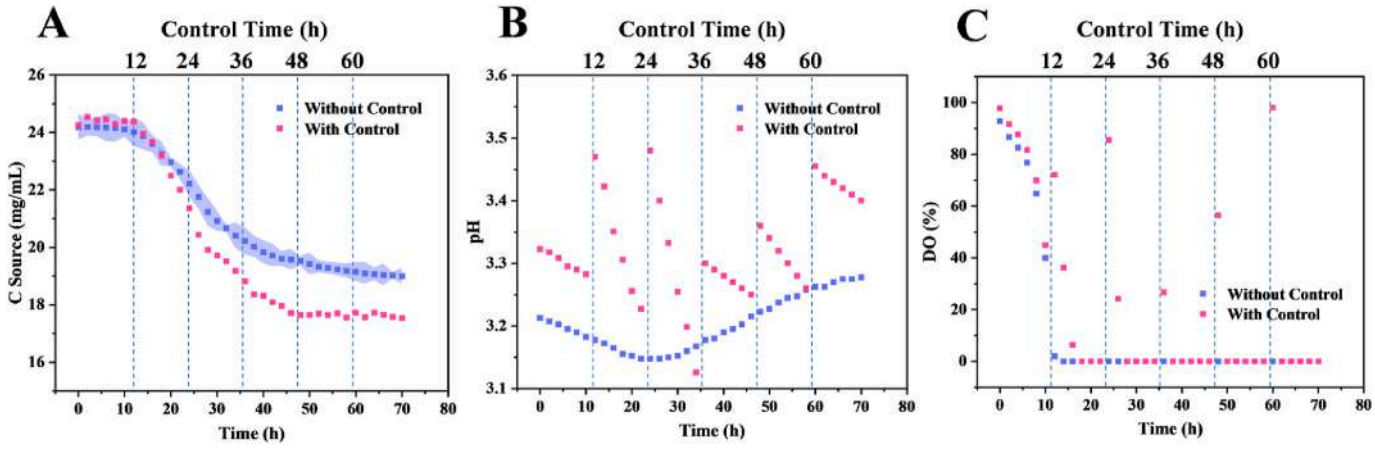



**Fig. 12.** The optimal control results for the purpose of C source consumption (A), and the regulation of pH (B) and DO (C) in the optimization process.

optimal solution. Figs. 8 and 9 demonstrate that the three optimal control simulation experiments yielded different regulatory outcomes due to varying fitness functions or fermentation guidance schemes. Therefore, in production, the fitness function should be designed based on production demands and expert experience. In addition, we have carried out a test experiment for the purpose of CS consumption and obtained satisfactory results.

The dataset constructed in this study only analyzed the relationship between the before and after moments (Fig. 1), and did not consider the global time series. Food fermentation processes typically exhibit significant time delays, so constructing a dataset from a time series perspective should better serve the prediction of future indicators, which can be explored in future discussions (Hua et al., 2023). Additionally, this study's model was constructed based on a data-driven approach, which lacked the explanatory power of fermentation mechanisms compared to knowledge-driven models. Therefore, it is necessary

to compare data-driven, knowledge-driven, and hybrid approaches to more comprehensively explain the fermentation process and build a more scientifically sound virtual system for DT applications (Kurniawan et al., 2023). Furthermore, generative models have gradually become popular tools in AI development in recent years, such as Generative Adversarial Networks (GANs) and diffusion models (Bond-Taylor et al., 2022). Using these models to generate process parameters that need to be regulated to achieve optimal control of food fermentation and build twin data for DT applications appears to be a promising solution.

## 4. Conclusion

A food fermentation prediction and control system based on DT technology was developed to realize real-time prediction and optimization of fermentation process parameters. A multi-scale feature filter was constructed to extract the deep features of fermentation process data. Three feature fusion strategies were investigated to achieve multi-scale feature fusion. Based on the fused features, PLS prediction models of CS and BC were constructed to achieve future time predictions. The experimental results showed that the Conv_PLS models of CS and BC had perfect performance and excellent verification effects, with RMSEP of 0.5538 mg/mL and 0.0558 (Au) and RPD of 5.63 and 6.52, respectively. Based on the prediction model and GA algorithm, a fermentation process optimization control system was built, and satisfactory simulation results were obtained. The results showed that the proposed fermentation prediction and control system based on DT had superior robustness and reliability, which promoted the digitization and intelligent development of the food fermentation industry.

## CRediT authorship contribution statement

**Songguang Zhao:** Writing – review & editing, Writing – original draft, Software, Methodology, Conceptualization. **Tianhui Jiao:** Methodology, Data curation, Conceptualization. **Selorm Yao-Say Solomon Adade:** Validation, Formal analysis. **Zhen Wang:** Methodology, Investigation, Formal analysis. **Qin Ouyang:** Methodology, Investigation, Data curation. **Quansheng Chen:** Supervision, Funding acquisition, Conceptualization.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

This work was financially supported by the Scientific and Technological Projects of Fujian Province (2022N5017).

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jfoodeng.2025.112467.

## Data availability

Data will be made available on request.

## References

Ahmad, Z., Khan, N., 2021. CNN-based multistage gated average fusion (MGAF) for human action recognition using depth and inertial sensors. IEEE Sens. J. 21 (3), 3623–3634. https://doi.org/10.1109/jsen.2020.3028561.

Ansari, Z.N., Daxini, S.D., 2021. A state-of-the-art review on meta-heuristics application in remanufacturing. Arch. Comput. Methods Eng. 29 (1), 427–470. https://doi.org/10.1007/s11831-021-09580-z.

Appl, C., Moser, A., Baganz, F., Hass, V.C., 2020. Digital twins for bioprocess control strategy development and realisation. Digital Twins: APPLICATIONS to the DESIGN and OPTIMIZATION of BIOPROCESSES, pp. 63–94. https://doi.org/10.1007/10_2020_151.

Augustin, M.A., Hartley, C.J., Maloney, G., Tyndall, S., 2023. Innovation in precision fermentation for food ingredients. Crit. Rev. Food Sci. Nutr. 64 (18), 6218–6238. https://doi.org/10.1080/10408398.2023.2166014.

Aung, T., Lee, W.-H., Eun, J.-B., 2022. Metabolite profiling and pathway prediction of laver (Porphyra dentata) kombucha during fermentation at different temperatures. Food Chem. 397, 133636. https://doi.org/10.1016/j.foodchem.2022.133636.

Barakat, N., Beaufort, S., Rizk, Z., Bouajila, J., Taillandier, P., El Rayess, Y., 2022. Kombucha analogues around the world: a review. Crit. Rev. Food Sci. Nutr. 63 (29), 10105–10129. https://doi.org/10.1080/10408398.2022.2069673.

Bond-Taylor, S., Leach, A., Long, Y., Willcocks, C.G., 2022. Deep generative modelling: a comparative review of VAEs, GANs, normalizing flows, energy-based and autoregressive models. IEEE Trans. Pattern Anal. Mach. Intell. 44 (11), 7327–7347. https://doi.org/10.1109/tpami.2021.3116668.

Cabeza-Gil, I., Ríos-Ruiz, I., Martínez, M.Á., Calvo, B., Grasa, J., 2023. Digital twins for monitoring and predicting the cooking of food products: a case study for a French crêpe. J. Food Eng. 359, 111697. https://doi.org/10.1016/j.jfoodeng.2023.111697.

Duan, J.-G., Ma, T.-Y., Zhang, Q.-L., Liu, Z., Qin, J.-Y., 2021. Design and application of digital twin system for the blade-rotor test rig. J. Intell. Manuf. 34 (2), 753–769. https://doi.org/10.1007/s10845-021-01824-w.

Feng, Y., Cong, Y., Xing, S., Wang, H., Ren, Z., Zhang, X., 2024. GCFormer: multi-scale feature plays a crucial role in medical images segmentation. Knowledge-Based Syst. 300, 112170. https://doi.org/10.1016/j.knosys.2024.112170.

Fourie, E., Aleixandre-Tudo, J.L., Mihnea, M., du Toit, W., 2020. Partial least squares calibrations and batch statistical process control to monitor phenolic extraction in red wine fermentations under different maceration conditions. Food Control 115, 107303. https://doi.org/10.1016/j.foodcont.2020.107303.

González-Hernández, Y., Michiels, E., Perré, P., 2022. A comprehensive mechanistic yeast model able to switch metabolism according to growth conditions. Fermentation 8 (12), 710. https://doi.org/10.3390/fermentation8120710.

Hormozi, E., Hu, S., Ding, Z., Tian, Y.-C., Wang, Y.-G., Yu, Z.-G., Zhang, W., 2022. Energy-efficient virtual machine placement in data centres via an accelerated Genetic Algorithm with improved fitness computation. Energy 252, 123884. https://doi.org/10.1016/j.energy.2022.123884.

Hu, F., Qiu, X., Jing, G., Tang, J., Zhu, Y., 2022. Digital twin-based decision making paradigm of raise boring method. J. Intell. Manuf. 34 (5), 2387–2405. https://doi.org/10.1007/s10845-022-01941-0.

Hua, L., Zhang, C., Sun, W., Li, Y., Xiong, J., Nazir, M.S., 2023. An evolutionary deep learning soft sensor model based on random forest feature selection technique for penicillin fermentation process. ISA Trans. 136, 139–151. https://doi.org/10.1016/j.isatra.2022.10.044.

Huang, X., Wang, J., Huang, B., 2021. Two novel hybrid linear and nonlinear models for wind speed forecasting. Energy Convers. Manage. 238, 114162. https://doi.org/10.1016/j.enconman.2021.114162.

Jayabalan, R., Malbaša, R.V., Lončar, E.S., Vitas, J.S., Sathishkumar, M., 2014. A review on kombucha tea—microbiology, composition, fermentation, beneficial effects, toxicity, and tea fungus. Compr. Rev. Food Sci. Food Saf. 13 (4), 538–550. https://doi.org/10.1111/1541-4337.12073.

Jiang, Q., Yan, X., Yi, H., Gao, F., 2020. Data-driven batch-end quality modeling and monitoring based on optimized sparse partial least squares. IEEE Trans. Ind. Electron. 67 (5), 4098–4107. https://doi.org/10.1109/tie.2019.2922941.

Júnior, J.C.d.S., Meireles Mafaldo, Í., de Lima Brito, I., Tribuzy de Magalhães Cordeiro, A.M., 2022. Kombucha: formulation, chemical composition, and therapeutic potentialities. Curr. Res. Food Sci. 5, 360–365. https://doi.org/10.1016/j.crfs.2022.01.023.

Kammoun, A., Ravier, P., Buttelli, O., 2024. Impact of PCA pre-normalization methods on ground reaction force estimation accuracy. Sensors 24 (4), 1137. https://doi.org/10.3390/s24041137.

Krepelka, P., Bolívar, A., Pérez-Rodríguez, F., 2020. Two-dimensional mid and near infrared correlation spectroscopy for bacterial identification. J. Near Infrared Spectrosc. 29 (2), 63–72. https://doi.org/10.1177/0967033520974518.

Kurniawan, W., Vikromvarasiri, N., Pisutpaisal, N., Koyama, M., Nakasaki, K., 2023. Utilization of first-order mass action kinetic models to elucidate the mechanism of methane fermentation of glycerol. Biochem. Eng. J. 200, 109099. https://doi.org/10.1016/j.bej.2023.109099.

Li, B., Lin, Y., Yu, W., Wilson, D.I., Young, B.R., 2020. Application of mechanistic modelling and machine learning for cream cheese fermentation pH prediction. J. Chem. Technol. Biotechnol. 96 (1), 125–133. https://doi.org/10.1002/jctb.6517.

Li, N., 2023. Comparison of the characteristics of the control strategies based on artificial neural network and genetic algorithm for air conditioning systems. J. Build. Eng. 66, 105830. https://doi.org/10.1016/j.jobe.2023.105830.

Li, X., Yao, X., Liu, Y., 2024. Combining swin transformer and attention-weighted fusion for scene text detection. Neural Process. Lett. 56 (2), 52. https://doi.org/10.1007/s11063-024-11501-7.

Liu, G.-P., 2024. Control strategies for digital twin systems. IEEE/CAA J. Autom. Sin. 11 (1), 170–180. https://doi.org/10.1109/jas.2023.123834.

Liu, X., Jiang, D., Tao, B., Xiang, F., Jiang, G., Sun, Y., Kong, J., Li, G., 2023. A systematic review of digital twin about physical entities, virtual models, twin data, and applications. Adv. Eng. Inf. 55, 101876. https://doi.org/10.1016/j.aei.2023.101876.

Mahanty, B., 2023. Hybrid modeling in bioprocess dynamics: structural variabilities, implementation strategies, and practical challenges. Biotechnol. Bioeng. 120 (8), 2072–2091. https://doi.org/10.1002/bit.28503.

Mayr, S., Schmelzer, J., Kirchler, C.G., Pezzei, C.K., Beć, K.B., Grabska, J., Huck, C.W., 2021. Theae nigrae folium: comparing the analytical performance of benchtop and handheld near-infrared spectrometers. Talanta 221, 121165. https://doi.org/10.1016/j.talanta.2020.121165.

Meng, Y., Ma, Z., Ji, Z., Gao, R., Su, Z., 2022. Fine hyperspectral classification of rice varieties based on attention module 3D-2DCNN. Comput. Electron. Agric. 203, 107474. https://doi.org/10.1016/j.compag.2022.107474.

Mohan, V., Pachauri, N., Panjwani, B., Kamath, D.V., 2022. A novel cascaded fractional fuzzy approach for control of fermentation process. Bioresour. Technol. 357, 127377. https://doi.org/10.1016/j.biortech.2022.127377.

Qi, Q., Tao, F., Hu, T., Anwer, N., Liu, A., Wei, Y., Wang, L., Nee, A.Y.C., 2021. Enabling technologies and tools for digital twin. J. Manuf. Syst. 58, 3–21. https://doi.org/10.1016/j.jmsy.2019.10.001.

Rydal, T., Frandsen, J., Nadal-Rey, G., Albæk, M.O., Ramin, P., 2024. Bringing a scalable adaptive hybrid modeling framework closer to industrial use: application on a multiscale fungal fermentation. Biotechnol. Bioeng. 121 (5), 1609–1625. https://doi.org/10.1002/bit.28670.

Song, S., Jeon, B., Kim, M., Kim, J.-J., 2023. Efficient convolutional processing of spiking neural network with weight-sharing filters. IEEE Electron. Device Lett. 44 (6), 1007–1010. https://doi.org/10.1109/led.2023.3265065.

Sun, D., Zhou, C., Hu, J., Li, L., Ye, H., 2023. Off-flavor profiling of cultured salmonids using hyperspectral imaging combined with machine learning. Food Chem. 408, 135166. https://doi.org/10.1016/j.foodchem.2022.135166.

Sun, Y., Yuan, M., Liu, X., Su, M., Wang, L., Zeng, Y., Zang, H., Nie, L., 2021. A sample selection method specific to unknown test samples for calibration and validation sets based on spectra similarity. Spectrochim. Acta, Part A 258, 119870. https://doi.org/10.1016/j.saa.2021.119870.

Tancredi, G.P.C., Bottani, E., Vignali, G., 2023. Digital twin-enabled process control in the food industry: proposal of a framework based on two case studies. Int. J. Prod. Res. 62 (12), 4331–4348. https://doi.org/10.1080/00207543.2023.2260495.

Vinestock, T., Short, M., Ward, K., Guo, M., 2024. Computer-aided chemical engineering research advances in precision fermentation. Curr. Opin. Food Sci. 58, 101196. https://doi.org/10.1016/j.cofs.2024.101196.

Wang, H., Liu, Y., Mu, Z., Xiang, J., Li, J., 2023. Real-time precision reliability prediction for the worm drive system supported by digital twins. Reliab. Eng. Syst. Saf. 240, 109589. https://doi.org/10.1016/j.ress.2023.109589.

Yao, Z., Zhu, Y., Wu, Q., Xu, Y., 2022. Challenges and perspectives of quantitative microbiome profiling in food fermentations. Crit. Rev. Food Sci. Nutr. 64 (15), 4995–5015. https://doi.org/10.1080/10408398.2022.2147899.

Yu, H., Liu, S., Qin, H., Zhou, Z., Zhao, H., Zhang, S., Mao, J., 2022. Artificial intelligence-based approaches for traditional fermented alcoholic beverages' development: review and prospect. Crit. Rev. Food Sci. Nutr. 64 (10), 2879–2889. https://doi.org/10.1080/10408398.2022.2128034.

Zhang, D., Jia, C., Sun, D., Gao, C., Fu, D., Cai, P., Hu, Q.-N., 2023. Data-driven prediction of molecular biotransformations in food fermentation. J. Agric. Food Chem. 71 (22), 8488–8496. https://doi.org/10.1021/acs.jafc.3c01172.

Zhang, R., Wang, F., Cai, J., Wang, Y., Guo, H., Zheng, J., 2022. Digital twin and its applications: a survey. Int. J. Adv. Des. Manuf. Technol. 123 (11–12), 4123–4136. https://doi.org/10.1007/s00170-022-10445-3.

Zhang, X., Liu, S., Li, L., Lei, J., Chang, G., 2021. Multiscale holospectrum convolutional neural network-based fault diagnosis of rolling bearings with variable operating conditions. Meas. Sci. Technol. 32 (10), 105027. https://doi.org/10.1088/1361-6501/ac05f8.

Zhao, S., Jiao, T., Wang, Z., Adade, S.Y.-S.S., Wu, X., Ouyang, Q., Chen, Q., 2023. On-line detecting soluble sugar, total acids, and bacterial concentration during kombucha fermentation based on the visible/near infrared combined meta-heuristic algorithm. J. Food Compos. Anal. 123, 105653. https://doi.org/10.1016/j.jfca.2023.105653.