Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Integrating AI/ML-Enabled Sonochemical Bioprocessing for Sustainable Valorization of Sweet Potato Leaves Waste into Functional Food Ingredients: Structural, Nutritional, and Metabolic Insights

Md Shadman Khan
Hajee Mohammad Danesh Science & Technology University
Md Atikul Islam
Hajee Mohammad Danesh Science & Technology University
Maruf Ahmed
maruf@hstu.ac.bd

Hajee Mohammad Danesh Science & Technology University

Research Article

Keywords: Artificial intelligence, bioprocessing, functional foods, waste valorization; metabolic insights

Posted Date: November 11th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11304656/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Integrating AI/ML-Enabled Sonochemical Bioprocessing for Sustainable Valorization of Sweet Potato Leaves Waste into Functional Food Ingredients: Structural, Nutritional, and Metabolic Insights

Md Shadman Khan[1], Md Atikul Islam[2], Maruf Ahmed [1*]

[1] Department of Food Processing and Preservation, Hajee Mohammad Danesh Science & Technology University, Dinajpur, 5200, Bangladesh
[2] Department of Chemistry, Hajee Mohammad Danesh Science & Technology University, Dinajpur, 5200, Bangladesh

Abstract

Sweet potato leaves are agro-waste produced in substantial amounts during the post-harvest period. In this study, sweet potato leaves were converted into functional food ingredients to achieve zero waste, using AI/ML in the process system, and their metabolic potential was evaluated. The protein- and fiber-rich fractions were extracted using single-stage sonication, and the fiber-rich residue (OPSF) was subsequently treated by dual-stage citric acid-assisted sonication (TPSF). ANN, Bayesian, and NSGA-II models were independently coupled with the extraction process to identify optimal processing conditions for protein fractions, with training R^2 values of 0.980, 0.998, and 0.942 and cross-validation values of 0.854, 0.872, and 0.807, respectively. The protein extract exhibited a diverse HPLC amino acid profile. ATR-FTIR from 4000–400 cm^{-1} was employed to evaluate the effects of sonication on the protein extract and the introduction of C=O, indicating enhanced hydrophilicity of the citric acid-treated fiber fraction (TPSF). TPSF showed a 19.4% increase in glucose adsorption capacity compared to OPSF. TPSF also increased cholesterol adsorption capacity by 10.3% at pH 2.0, and at pH 7.0 it increased from 25.03 ± 0.17 to 26.15 ± 0.41 mg/g compared to OPSF, reflecting its potential as a functional food ingredient for improving metabolic function.

Keywords: Artificial intelligence, bioprocessing, functional foods, waste valorization; metabolic insights.

*Corresponding Author: Professor Dr. Maruf Ahmed,
Department of Food Processing and Preservation,
Hajee Mohammad Danesh Science and Technology University, Dinajpur 5200, Bangladesh
Email: maruf@hstu.ac.bd
Mobile: +8801750777899

1. Introduction

Sweet potato leaves are nutrient-rich edible leafy vegetables consumed in several regions and typically prepared through cooking or processing, however, despite their considerable nutritional and functional potential, remain largely underutilized and are commonly discarded as agricultural waste (Alam et al., 2025; Yi et al., 2024). Global sweet potato production reached approximately 93.52 million tons in 2023 across an estimated 7.57 million hectares**,** corresponding to an average yield of 12.35 tons per hectare (Sheepo et al., 2025). However, according to the Bangladesh Bureau of Statistics (BBS) report for the 2024-25 fiscal year, Bangladesh cultivated 67,690.73 acres, yielding 303,272.58 metric tons, underscoring its acceptance, resilience, and contribution to economic growth. Across different sweet potato genotypes, the leaves exhibit crude protein 16.69–31.08, crude fiber 9.15–14.26, crude fat 2.08–5.28, carbohydrate 42.03–61.36, and ash 7.39–14.66 g/100 g dry weight, respectively (Sun et al., 2014)**.** In recent years, functional foods have gained increasing attention in food science owing to their potential to confer physiological benefits and support the prevention and management of non-communicable diseases (Topolska et al., 2021). The protein- and fiber-rich composition of sweet potato leaves may provide a promising basis for developing functional food ingredients with potential health-promoting properties. Plant protein may serve complementary essential amino acids that improve the nutritional density of products (Mäkinen et al., 2016). Concomitantly, insoluble dietary fiber may contribute to the intestinal microbiota, with implications for glucose homeostasis, lipid metabolism, intestinal integrity, and immune function (Liu et al., 2024; Ma et al., 2025; G. Zhang et al., 2024). Organic acid hydrolysis of insoluble dietary fiber offers a comparatively mild and potentially more sustainable approach to chemical treatment, while maintaining effective modification efficiency (Yu et al., 2025). Previous studies have reported that citric acid modification can enhance the hydration, lipid-retention, and adsorption properties of kiwifruit insoluble dietary fiber, particularly its affinity for water, oil, bile acids, nitrite ions, and glucose (K. Wang et al., 2021).

Hybrid data-driven AI, particularly classical machine learning (ML) and deep learning (DL), provides powerful tools for extracting complex relationships from multidimensional food datasets, supporting advanced analysis in food chemistry and processing, while opening new avenues for addressing emerging challenges, including the characterization, extraction, and prediction of four-dimensional food structures (Ekrami et al., 2025; Gu et al., 2026; Wu et al., 2025). However,

ultrasound-assisted extraction is a sustainable approach for recovering plant components (Bernardi et al., 2021; Ochoa-Rivas et al., 2017). Ultrasound-assisted extraction can efficiently increase extraction yield, preserve native structure, reduce extraction time, lower solvent requirements, and minimize ecological impact (Dabbour et al., 2018; Sert et al., 2022). Artificial neural networks (ANN) models can capture intricate nonlinear interactions and provide predictive representations of experimental systems, while Bayesian optimization (BO) enables efficient exploration of experimental conditions toward optimal responses (Ervilha et al., 2025; G. E. Lee et al., 2022). In parallel, non-dominated sorting genetic algorithm II (NSGA-II) facilitates multi-objective optimization by identifying a set of Pareto-optimal solutions when multiple, potentially conflicting process objectives must be considered. Collectively, these approaches can strengthen experimental design, model development, process understanding, and optimization of process-variable effects on target responses (Wei et al., 2023).

Artificial intelligence (AI) is rapidly emerging as a powerful tool for functional food research (Alkalbani et al., 2026). Although ultrasound-assisted extraction and chemical modification have been investigated for other plant-derived materials, the application of these approaches to sweet potato leaves waste have not been explored. Furthermore, artificial intelligence (AI) and machine learning (ML), evaluation of the functional and metabolic potential of the resulting functional food ingredients, have received limited attention, particularly within a sequential zero-waste valorization strategy (Su & Cavaco-Paulo, 2021; Tang, 2024; Eldiehy et al., 2022; Ademosun et al., 2026). In this study, a comparative hybrid modeling framework incorporating artificial neural networks (ANN), Gaussian process (GP)-based Bayesian optimization, and cubic B-spline–Ridge modeling coupled with non-dominated sorting genetic algorithm II (NSGA-II) was developed to optimize protein-rich extract recovery from sweet potato leaves, followed by the development of fiber-rich extracts through single-stage and dual-stage citric acid-coupled ultrasonic treatment of the residual biomass. The optimized protein fractions were characterized by ATR-FTIR to elucidate amplitude-dependent structural changes, while the single-stage and dual-stage citric acid-coupled ultrasonic fiber fractions were comparatively assessed for their structural and physicochemical properties. The protein fractions were further analyzed for amino acid composition using HPLC, whereas the fiber fractions were subjected to comparative in vitro adsorption assays to evaluate their functional properties and potential metabolic effects.

2. Materials & methods

2.1. Sample preparation

Sweet potato leaves were collected as post-harvest waste from riverside fields in Gaibandha, Bangladesh. The collected leaves were washed thoroughly to remove extraneous matter and dried in a food dehydrator (LT-31-17, China) at 60 °C for 12 h. The dried leaves were pulverized into a fine powder using a Panasonic MX-AV425 1500W Super Mixer Grinder (Japan) and stored at 4 °C to maintain stability prior to analysis. The dried sweet potato leaf powder contained 8.50 ± 0.07% moisture, measured with a moisture analyzer (Xing Yun XY-105MW, China), and 25.29% crude protein (Aguirre, 2023).

2.2. Bayesian optimization

Protein extraction was optimized using a Bayesian optimization model with adaptive sampling. The model was developed using Python (version 3.11.9) and implemented in VS Code (version 1.132.1), considering solvent-to-solid ratios (1:15, 1:25, 1:35), amplitudes (50%, 70%, 100%) at 40 kHz, and extraction times (5, 10, 15 min). Using Python, a Gaussian process (SingleTaskGP) was trained on (torch.tensor) experimental inputs in BoTorch, serving as a surrogate model for protein extraction yield. The surrogate model was integrated and fitted using the Kernel Hyperparameter Prior exact marginal log-likelihood. Gamma Prior (3.0, 6.0) was assigned to the Kernel Hyperparameters to facilitate initial probabilistic belief before the posterior was performed. A Bayesian optimization model was constructed based on the following theoretical framework, reported by Ervilha et al. (2025), which encompasses the mathematical principles of the surrogate method, Gaussian process regression, model updating, and the expected improvement (EI) acquisition function. To determine optimized experimental extraction conditions, the q-Expected Improvement (qEI) acquisition function**,** a Monte Carlo-based criterion, was employed using SobolQMCNormalSampler to guide the next suggested experiment. `The` expected improvement (EI) (q=1, in each iteration) was measured considering the best observed response, and it's given by:

$$\alpha_{\mathrm{EI}}(\mathbf{x}) = \mathbb{E}\left[\max(0, f(\mathbf{x}) - f(\mathbf{x}^{+}))\right] \quad (1)$$

where $f(\mathbf{x}^+)$ is the best observed response. Bayesian optimization was employed with a surrogate model to estimate the objective function $f(\mathbf{x})$, where the GP was mathematically described by the mean function $\mu(\mathbf{x})$ and covariance function:

$$k(\mathbf{x}, \mathbf{x}'): f(\mathbf{x}) \sim \mathcal{GP}(\mu(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')) \quad (2)$$

For a new point $\mathbf{x}$, the posterior distribution of the Gaussian process was expressed as:

$$\mu_{n+1}(\mathbf{x}) = \mu(\mathbf{x}) + k(\mathbf{x}, \mathbf{X})K^{-1}(\mathbf{y} - \mu(\mathbf{X})) \quad (3)$$

$$\sigma^2{}_{n+1}(\mathbf{x}) = k(\mathbf{x}, \mathbf{x}) - k(\mathbf{x}, \mathbf{X})K^{-1}k(\mathbf{X}, \mathbf{x}) \quad (4)$$

where $\mathbf{X}$, $\mathbf{y}$, and K represent the matrix of observed data points, the vector of observed responses, and the covariance matrix of the observed data points, respectively.

2.3 Ultrasound-assisted extraction (UAE)

2.3.1. Protein and fiber isolation using single-phase Ultrasonic treatment

The solid-liquid matrix was prepared by incorporating 10 g of the sample into 50 mM Tris-HCl buffer at solvent-to-solid ratios of 1:15, 1:25, and 1:35 for each experimental condition (Bianco et al., 2022). Alkalinity was adjusted to pH 8.0 and stirred at 750 rpm for 10 min at 25 ± 2 °C. Sonication (Ultrasonic cleaner TUC-32, China) was performed at 40 kHz with amplitudes of 50%, 70%, and 100% for 5, 10, or 15 min at 25 ± 2 °C. The supernatant was collected by vacuum filtration, adjusted to pH 4.7 to induce isoelectric precipitation, and then centrifuged at 7500 × g for 25 min to collect the extract. The extract was dried in a dark desiccator with silica gel for 14 h, finely ground, and stored at 4 °C. The residue was dried at 60 °C for 12 hours and ground to a fine powder. For One-phase ultrasonic treatment of fiber isolation, protein was separated using sonication, maintaining 40 kHz with an amplitude of 50%, a solvent-to-solid ratio of 1:25, and a sonication time of 15 min. The mixture was filtered after sonication and dried at 60 °C for 12 hours and ground to the fine powder. The residue after protein extraction contained 35% crude fiber, which denoted as (OPSF).

2.3.2. Dual-phase Ultrasonic treatment of fiber extract

After the first-phase ultrasonic treatment, the remaining residue was sonicated again using a 1% (w/v) citric acid solution, and the treatment was conducted at 40 kHz with an amplitude of 50%, a solvent-to-solid ratio of 1:25, and for 15 min. Hot water was used as the medium, and a temperature of 80 °C ± 2 was maintained throughout the sonication. After sonication, the residue was filtered and washed thoroughly with distilled water to neutral pH, then dried at 60 °C for 12 hours and ground to develop a fine powder (TPSF).

2.4. *Artificial neural network*

In this study, a single-hidden-layer multilayer perceptron (MLP) was employed to assess the nonlinear relationship between the ultrasonic process variables and protein extraction yield. Table 1 represents the experimental runs of ultrasound extraction of protein. The range of the inputs and the output were placed between −1 to +1 prior to training using min–max scaling, and scalers were re-fit exclusively on the training partition within every resampling iteration to prevent information leakage from validation observations into the preprocessing step. Network weights were optimized with the limited-memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) quasi-Newton solver, which shows favourable environment for limited dataset. Which severed a fully connected network with, for example, 10 hidden neurons would carry approximately 51 trainable parameters (10×3 input weights + 10 hidden biases + 10 output weights + 1 output bias) against only 17 observations, the hyper-parameter search space was deliberately constrained to reduce the risk of over parameterization: hidden-layer size ∈ {2, 3, 4, 5}, activation function ∈ {hyperbolic tangent, logistic sigmoid, rectified linear unit}, and L2 weight-decay coefficient $\alpha \in \{10^{-3}, 10^{-2}, 10^{-1}, 1\}$, giving 48 candidate configurations. The model architectures are deliberately restricted rather than exhaustively searched to avoid the model memorizing run-to-run noise (Xue et al., 2022).The trained ANN ensemble was then used as a surrogate objective function for constrained numerical optimization via the differential evolution (DE) algorithm (strategy: best1bin; population size: 15; maximum 300 generations; tolerance: 10^{-8}; polishing enabled), with the search restricted to the experimentally investigated ultrasound factor ranges (amplitude 50 to 100%; time 5 to 15 min; ratio 1:15 to 1:35) to avoid extrapolation beyond the design space.

2.5. Non-dominated Sorting Genetic Algorithm II (NSGA-II)

The 17 experimental observations of ultrasound extractions were used to develop a cubic B-spline–Ridge response model for predicting ultrasound protein extraction yield. The model was optimized by cross-validation and subsequently refitted using the complete experimental dataset. The validated response model was then integrated with the Non-dominated Sorting Genetic Algorithm II (NSGA-II) to determine optimal trade-offs among process performance and resource requirements. The optimization simultaneously maximized predicted protein extraction yield, minimized extraction time, and minimized solid-to-solvent ratio. NSGA-II generated an initial population of feasible solutions, predicted the corresponding yield using the response model, ranked solutions through non-dominated sorting and crowding distance, and generated successive populations through selection, simulated binary crossover ($P_c = 0.90$, $\eta_c = 20$), and polynomial mutation ($\eta_m = 25$). The procedure was repeated for 600 generations with a population size of 300, using a fixed random seed (42) to ensure reproducibility.

2.6. ATR-FTIR spectrometer

To assess the structural characteristics of the samples as a function of sonication amplitude and treatment behavior, ATR-FTIR spectroscopy was performed using a PerkinElmer Spectra II ATR-FTIR spectrometer equipped with a $LiTaO_3$ detector (Waltham, MA). Data were acquired from 4000 to 400 cm^{-1} at 4 cm^{-1} resolution within 32 accumulated scans per sample. Initial data tune-up and baseline correction were performed using PerkinElmer Spectrum IR software (version 10.6.0.893).

2.7. High-performance liquid chromatography (HPLC)

Following AOAC Official Method 982.30**,** the composition of amino acids in sweet potato leaf protein extract was determined after hydrolysis of 0.20 to 1.0 g samples with 6 M HCl containing 0.1% phenol at 110 °C for 24 hours to perform acidic hydrolysis. Following hydrolysis, the pH was adjusted to 2.9- 3.1 with NaOH solution; the mixture was filtered through a 0.25 μm syringe membrane, and the filtrate was loaded into the autosampler. Amino acid separation and quantification were performed using the Sykam Amino Acid Analyzer (Model: S433, Sykam, Germany) via ion-exchange chromatography with post-column ninhydrin-based derivatization.

Amino acid Separation was conducted on a 150 mm × 4.26 mm capillary column with an auto-injector at 57°C column temperature and a pressure of 39 bar. The mobile phase was operated at a constant flow rate of 0.45 ml/min, resulting in a total run time of 63 min. Amino acid detection was conducted using a dual-channel photometer-integrated system over the wavelength range of 440 to 570 nm, and amino acid standards were subsequently used to quantify amino acid concentrations in the protein samples.

2.8. Determination of water-holding capacity

Assessment of water-holding capacity of the extracts were carried out using the method reported by (Chen et al. 1984), with minor adjustments. A pre-weighed 0.5 g protein extract was dissolved in 25 mL of distilled water and vortexed for 1 min. The hydrated mixture was then incubated in a water bath for 30 min and centrifuged (Labwe BT16R, China) at 4500 x g for 15 min to decant the supernatant. The mass of hydrated residue was weighed, and water-holding capacity was calculated using the following equation:

$$\text{Water-Holding Capacity} = \frac{\text{Residue}}{\text{Total weigh of extract}} \text{x } 100\% \quad (5)$$

2.9. Determination of oil holding capacity

Using the method of Saha & Deka (2017), with minor adjustments, 25 mL of vegetable oil was added to 0.5 g of sample to evaluate the lipid-binding ability of extracts. After vortexing, the samples were incubated for 25 min at room temperature, and then centrifuged at 4500 x g for 15 min. The supernatant was removed, and the Falcon tube was kept at 45°C for 10 min. The residue was weighed, and oil-holding capacity was determined using the following equation:

$$\text{Oil-Holding Capacity} = \frac{\text{Residue} - \text{Total weight of extract}}{\text{Total weigh of extract}} \text{x } 100\% \quad (6)$$

2.10. Determination of solubility index & swelling capacity

To assess solubility and swelling characteristics, we applied the analytical approach of Nadiha et al. (2010) with minor modifications. 0.5 g of samples were introduced into 50 mL of distilled water

at pH 7.0. The vortex-agitated solid-liquid matrix was incubated at 40 °C for 30 min and then centrifuged at 5000 x g for 25 min. The retained hydrated residue was quantified by evaporating the aqueous extract at 105 °C overnight, whereas the following equations were employed to determine the solubility index and swelling capacity:

$$\text{Solubility index} = \frac{\text{Weigh of oven dried extract}}{\text{Total weigh of extract}} \text{ x } 100\% \quad (7)$$

$$\text{Swelling capacity} = \frac{\text{Weigh of residue}}{\text{Total weigh of extract X } \left(\frac{1-ws\%}{100}\right)} \quad (8)$$

2.11. In vitro biochemical assays

2.11.1. Cholesterol adsorption capacity

To assess the cholesterol adsorption capacity of single and dual phase sonicated fiber rich powder samples, the analytical protocol was followed with minor adjustments (Lai et al., 2022). 15 g of egg yolk was weighed and incorporated into 500 mL distilled water to prepare initial cholesterol latex and the mixture was homogenized properly to prepare uniform solution. 0.200 g of single and dual phase sonicated fiber rich powder samples were mixed individually with 10.0 mL of cholesterol latex, while pH was adjusted to 2.0 and 7.0, using 1 M HCl and 1 M NaOH to maintain gastric and intestinal phase, respectively. The samples then shifted to water bath 37 °C and 120 rpm for 2 hours and centrifuged at 4000 X g for 20 minutes after incubation period. O-phthalaldehyde-sulfuric acid method was employed to determine the concentration of the samples. 5 mL of supernatant was incorporated with 15 mL of O-phthalaldehyde solution and 10 mL of 98.3% sulfuric acid, and kept for 15 minutes to develop a stable color. The absorbance was determined at 550 nm using a UV-Vis spectrophotometer (Shaanxi UV-1800, China). The cholesterol adsorption capacity of single and dual phase sonicated fiber rich powder samples were determined using the cholesterol standard calibration curve and following equation:

$$\text{Cholesterol adsorption capacity (mg/g)} = \frac{(C_0 - C_1)\ X\ V}{m} \quad (9)$$

Where, C_0 is the initial concentration of the cholesterol latex (mg/mL); C_1 is the cholesterol concentration of single and dual phase sonicated fiber rich powder samples (mg/mL); V is the volume of the cholesterol latex added in the system; m is the initial weight of the sample (g).

2.11.2. Glucose adsorption capacity

The Glucose adsorption capacity of single- and dual-phase sonicated fiber-rich powder samples was measured using the procedure reported by Jiang et al. 2023), with minor modifications. Single-phase (OPSF) and dual-phase sonicated fiber-rich powder samples (TPSF) were incorporated into 50 mL of glucose solution (10 mmol/L), maintaining the sample/solvent ratio at 1:50. The incubation period was 2 hours, maintaining 37 °C in a shaking water bath. After the incubation period, samples were centrifuged at 4000 X g for 20 minutes to separate the supernatant. 1 mL of supernatant was mixed with 02 ml of DNS reagent and incubated in boiling water for 5 minutes. Using distilled water, the supernatant-DNS mixture was adjusted to 15 mL and kept for 20 minutes to develop color. To measure the Glucose adsorption capacity of Single (OPSF) and dual-phase sonicated fiber-rich powder samples (TPSF), a UV-Vis spectrophotometer was used to measure the absorbance at 540 nm, and the following equation was employed to calculate GAC:

$$\text{Glucose adsorption capacity (mg/g)} = \frac{(C_0 - C_1)\ X\ V}{m} \quad (10)$$

Where C_0 is the glucose concentration without dietary fiber (mg/mL); C_1 is the glucose concentration of Single (OPSF) and dual-phase sonicated fiber-rich powder samples (TPSF) samples (mg/mL); V is the volume of the supernatant; m is the initial weight of the sample (g).

2.12. Statistical analysis

All data from the experimental study were reported as mean ± standard deviation SD and three independent replications were obtained. We performed one-way analysis of variance (one-way ANOVA) in OriginPro 2021 (version 9.8.0.200), and determined significant differences using Tukey's multiple comparison test; $p < 0.05$ was considered statistically significant.

3. Results and discussions

3.1. Bayesian optimization and correlation analysis of sonication effects

3.1.1. Effect of amplitude on protein extraction

Ultrasonic intensity strongly influences ultrasound efficiency and cavitation dynamics. Elevated sonication intensity can disrupt structural integrity (Dabbour et al., 2018). A frequency of 40 kHz is expected to generate strong cavitation, which may contribute to adverse effects such as aggregation, protein unfolding, and reduced yield (Liu et al., 2019). In the surrogate model Figure 1 (a), the GP mean curve followed the experimental response behavior across the intensity range, whereas the qEI curve (Figure 1 b) exhibited the dominant peak at 50% amplitude. The experimental observations in Table 1 indicate that applying a 70% amplitude for protein extraction yields the lowest yields (11.77–17.73%), suggesting that excessive amplitude at 40 kHz and intensified cavitation may induce protein unfolding, aggregation, and instability**.** The clustering of high yield (23.09–30.94%) was observed around 50% amplitude, suggesting that substantially lower ultrasonic amplitude created optimal cavitation conditions and enhanced extraction yield; a similar effect of intensity was reported by (Sharma et al., 2023).

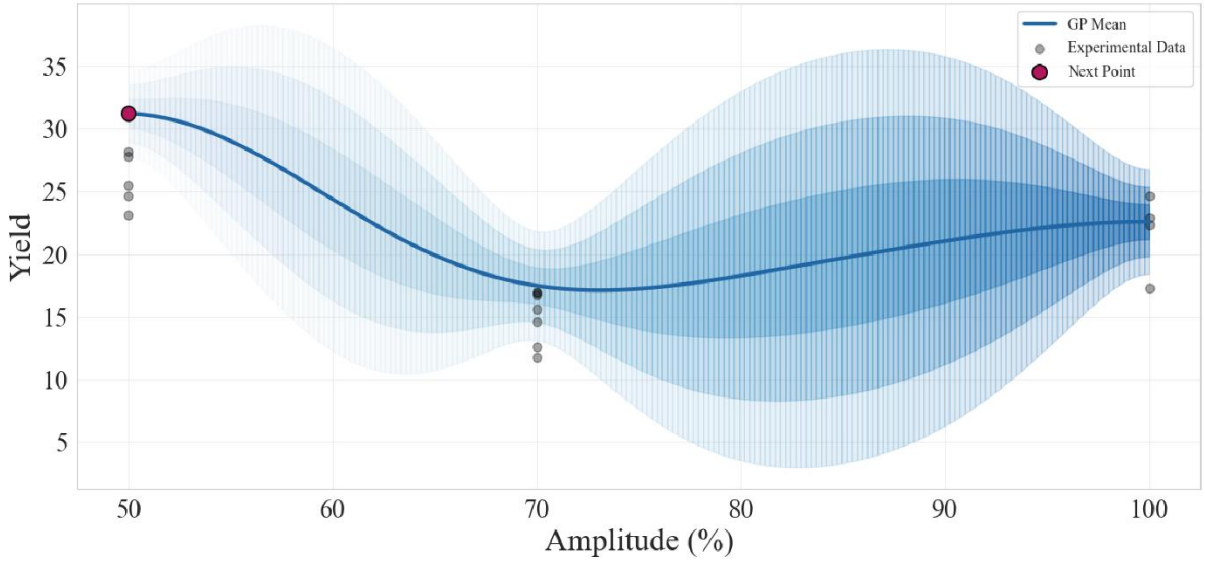


Figure 1(a). Bayesian optimization surrogate model of protein yield (%) versus amplitude (%)

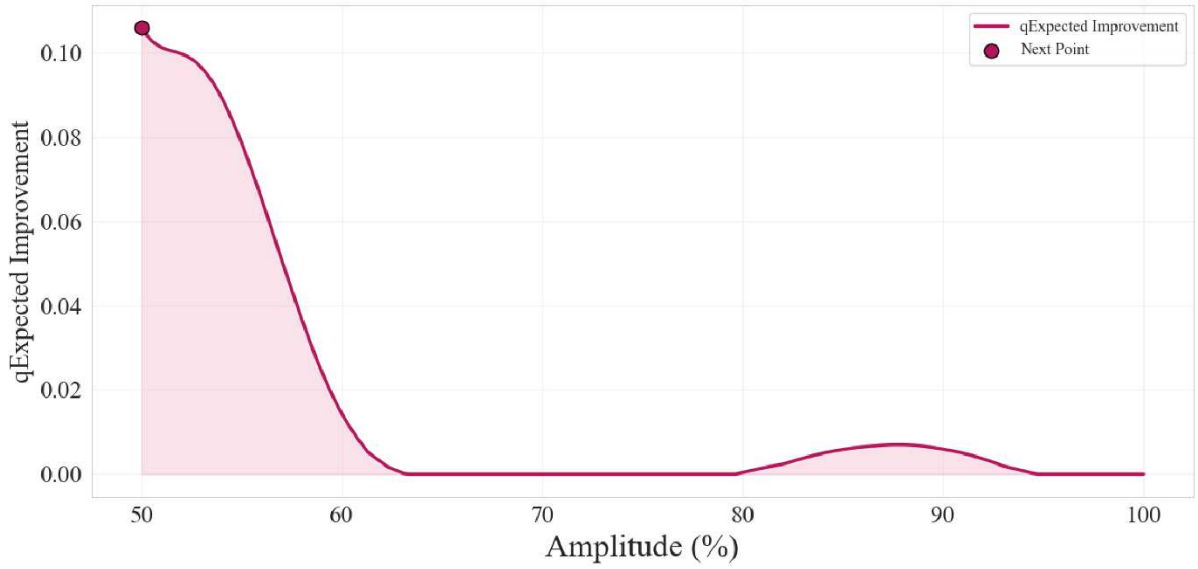


Figure 1(b). Bayesian optimization acquisition function (expected improvement) versus amplitude (%)

Table 1. Effect of ultrasound amplitude, extraction time, and solid-to-solvent ratio on the yield percentage of sweet potato leaf protein extracts.

Run No.	Amplitude (%)	Extraction time (min)	Sample/solvent ratio (g/mL)	Yield (%)
1	50	5	1:25	27.78
2	100	5	1:25	22.32
3	50	15	1:25	30.94
4	50	15	1:23	28.19
5	100	15	1:25	24.64
6	50	10	1:15	23.09
7	100	10	1:15	22.86
8	50	15	1:35	25.45
9	50	10	1:35	24.62
10	100	10	1:35	17.28
11	70	5	1:15	11.77
12	70	15	1:15	12.60
13	70	5	1:35	14.62
14	70	15	1:35	15.60
15	70	10	1:25	17.01
16	70	10	1:25	16.88
17	70	10	1:25	17.73

3.1.2. Effect of extraction time on protein extraction

Sonication time significantly influenced protein recovery during ultrasound-assisted extraction; the experiment was conducted over a range of 5- 15 min. Figure 1(c), suggested that longer sonication generally enhanced mass transfer between the solvent and the plant matrix, although the optimal extraction time varies with substrate properties and extraction duration (Tu et al., 2015). However, excessive ultrasonic exposure may affect the structural stability of protein and release undesired components (Das et al., 2022). As illustrated by the qEI acquisition function in Figure 1 (d), the predicted yield gradually increased with increasing sonication time. The experimental results in Table 1 confirmed this observation, with a 30.94% yield achieved at a 15 min extraction time, indicating a promising condition for yield enrichment.

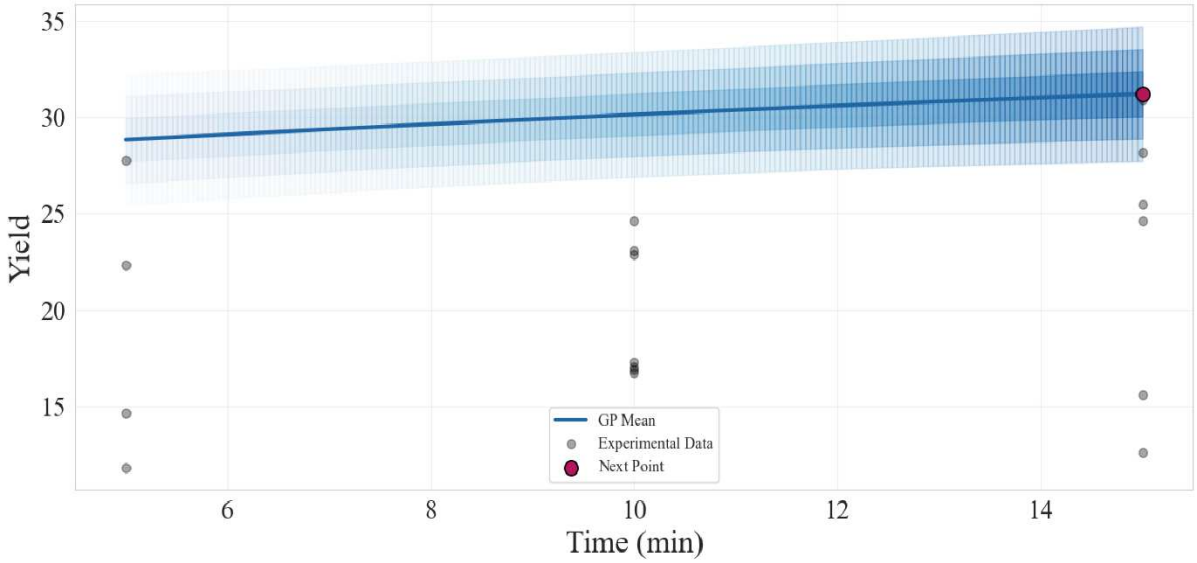


Figure 1(c). Bayesian optimization surrogate model of protein yield (%) versus time (min)

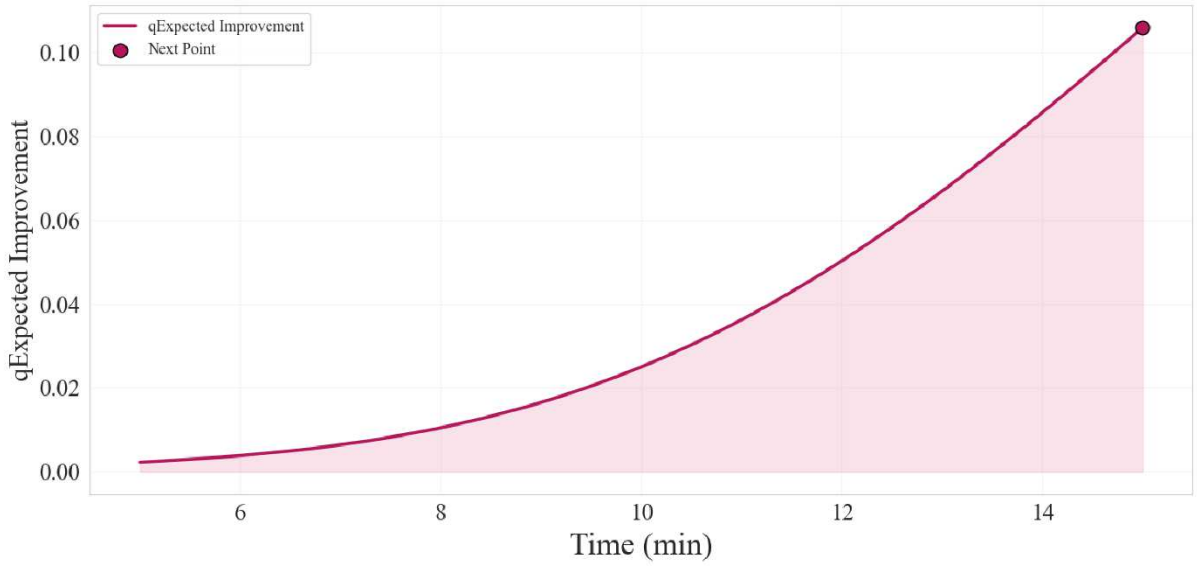


Figure 1(d). Bayesian optimization acquisition function (expected improvement) versus time (min)

3.1.3. Effect of solid-solvent ratio on protein extraction

Table 1 summarizes the 17 experimental runs. The surrogate model provided a detailed nonlinear relationship between the solid-solvent ratio and extraction yield. The predicted response curve indicated that the yield increased steadily as the solid-to-solvent ratio increased from 1:15 (g/mL), reaching a maximum of 24-30% at ratios of 1:25 to 1:30 (g/mL). Increasing solvent volume enhances mass-transfer kinetics and cavitation disintegration of plant tissues, thereby facilitating the release of target components (Zhang et al., 2011). However, in figure 1(e), increasing the ratio from 1:30 to 1:35 (g/mL) slightly reduces the extraction yield, which may be attributed to dilution effects and reduced cavitation intensity. The qEI function, in figure 1(f) indicated that a sample-to-buffer (50 mM Tris-HCl) ratio of 1:25 to 1:28 (g/mL) is the most favorable condition for achieving higher extraction yields.

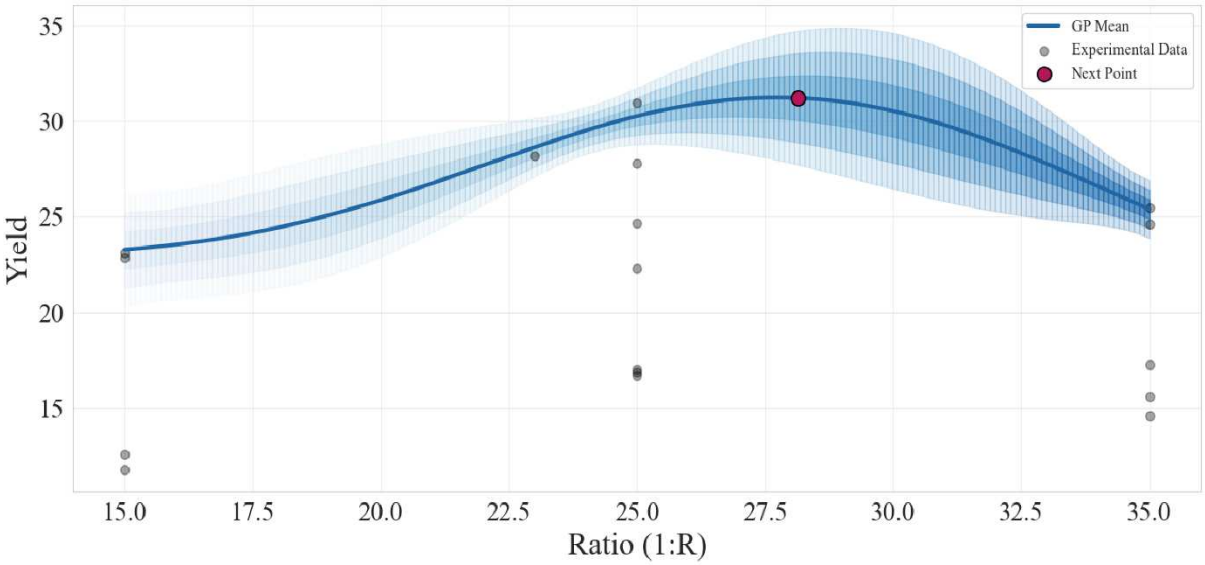


Figure 1(e). Bayesian optimization surrogate model of protein yield (%) versus solid/solvent ratio (g/mL)

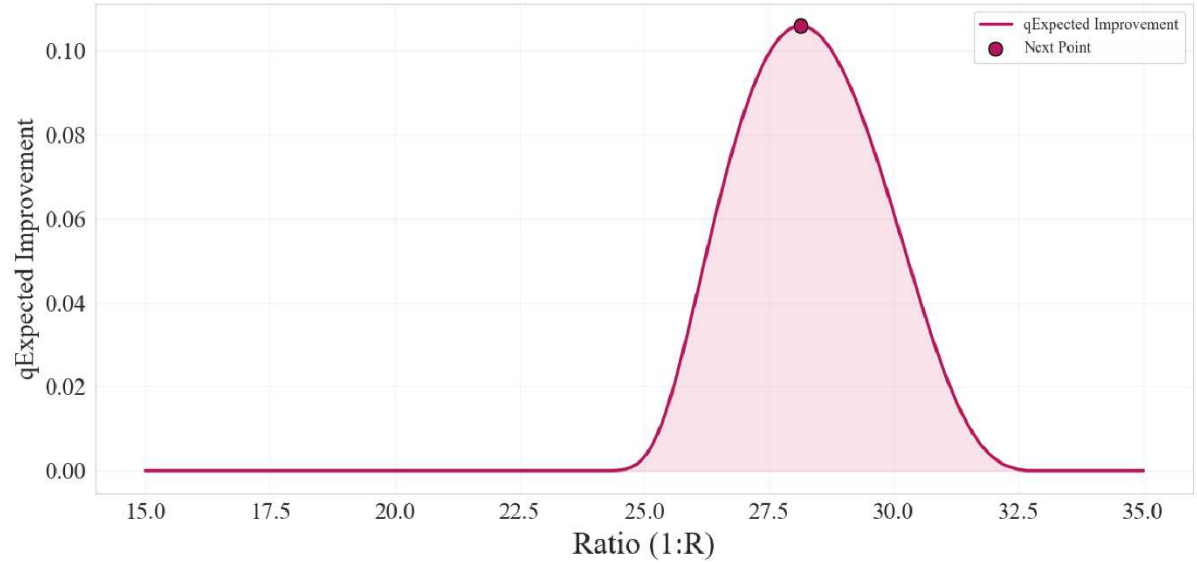


Figure 1(f). Bayesian optimization acquisition function (expected improvement) versus solid/solvent ratio (g/mL)

3.1.4. Bayesian Optimal experimental conditions

The optimal parameters for ultrasound-assisted extraction were effectively delineated via Bayesian optimization (BO), thereby enhancing protein recovery from sweet potato leaves. The predictive yield of 31.22% within the resolved extraction regime-ultrasound amplitude 50% at 40 kHz, solid-to-solvent ratio 1:28.14, extraction duration 15 min- was closely aligned with the experimental

yield of 29.92%, confirming the BO model's robustness in navigating the nonlinear response. The ultrasound-assisted recovered extract exhibited a crude protein content of 47.17%, as quantified using the Kjeldahl method. The temperature and pH were maintained at 25 ± 2 °C and 8.0, respectively, throughout extraction. The model exhibited strong predictive performance, with R^2 = 0.998 on the training dataset, an RMSE of 0.224, a 95% confidence interval of [28.93, 33.50], and a standard deviation (uncertainty) of 1.17, indicating robustness and predictive stability.

3.2. Analysis of Predictive Performance and ANN-Integrated Process Optimization

The Artificial Neural Network (ANN) model effectively captured the nonlinear relationship between ultrasound amplitude, extraction time, sample-to-solvent ratio, and protein extraction yield (Table 2). The model achieved a training R^2 of 0.980, while cross-validation yielded an R^2 of 0.854, with RMSE and MAE of 2.140 and 1.560%, respectively. Comparison of experimental and ANN-predicted yield of sweet potato leaf protein extract is illustrated in Figure 2. Across the 17 runs, the absolute prediction error ranged from 0.02 to 4.30%, with particularly close agreement for several runs and the center-point replicates. The prediction profile indicates that the ANN reliably represented the experimental response within the investigated design space. ANN optimization identified 50.0% ultrasound amplitude, 15.0 min extraction time, and a sample-to-solvent ratio of 1:26.23 as the optimal combination, with a predicted protein extraction yield of 28.82%**.** The experimental yield was 28.39% under these optimization conditions. The low ensemble seed SD (0.12%) indicates good stability of the predicted solution across independent network initializations. The optimum at the lowest amplitude combined with the longest extraction time suggests that prolonged exposure to moderate acoustic intensity favored protein solubilization, whereas increasing amplitude alone did not improve recovery. This behavior may reflect a balance between cavitation-induced matrix disruption and potential limitations associated with excessive acoustic intensity. The intermediate solvent ratio further suggests that enhanced solvent accessibility contributed to protein transfer without requiring the maximum solvent level.

Table 2. Experimental design matrix and ANN predictions of protein extraction yield.

Run	Observed yield (%)	Predicted yield (%)	Residual (%)	Absolute error (%)	APE (%)
1	27.78	27.76	0.02	0.02	0.09
2	22.32	22.63	-0.31	0.31	1.37
3	30.94	28.58	2.36	2.36	7.64
4	28.19	28.47	-0.28	0.28	0.99
5	24.64	23.00	1.64	1.64	6.66
6	23.09	22.84	0.25	0.25	1.07
7	22.86	18.70	4.16	4.16	18.19
8	25.45	26.70	-1.25	1.25	4.90
9	24.62	25.95	-1.33	1.33	5.39
10	17.28	21.58	-4.30	4.30	24.91
11	11.77	14.38	-2.61	2.61	22.15
12	12.60	16.82	-4.22	4.22	33.52
13	14.62	14.40	0.22	0.22	1.50
14	15.60	17.84	-2.24	2.24	14.34
15	17.01	17.42	-0.41	0.41	2.39
16	16.88	17.18	-0.30	0.30	1.75
17	17.73	17.17	0.56	0.56	3.18

APE, absolute percentage error. Predicted yield values are the ensemble mean across five random-initialization seeds (2026, 7, 42, 123, 1001). Residual = observed − predicted.

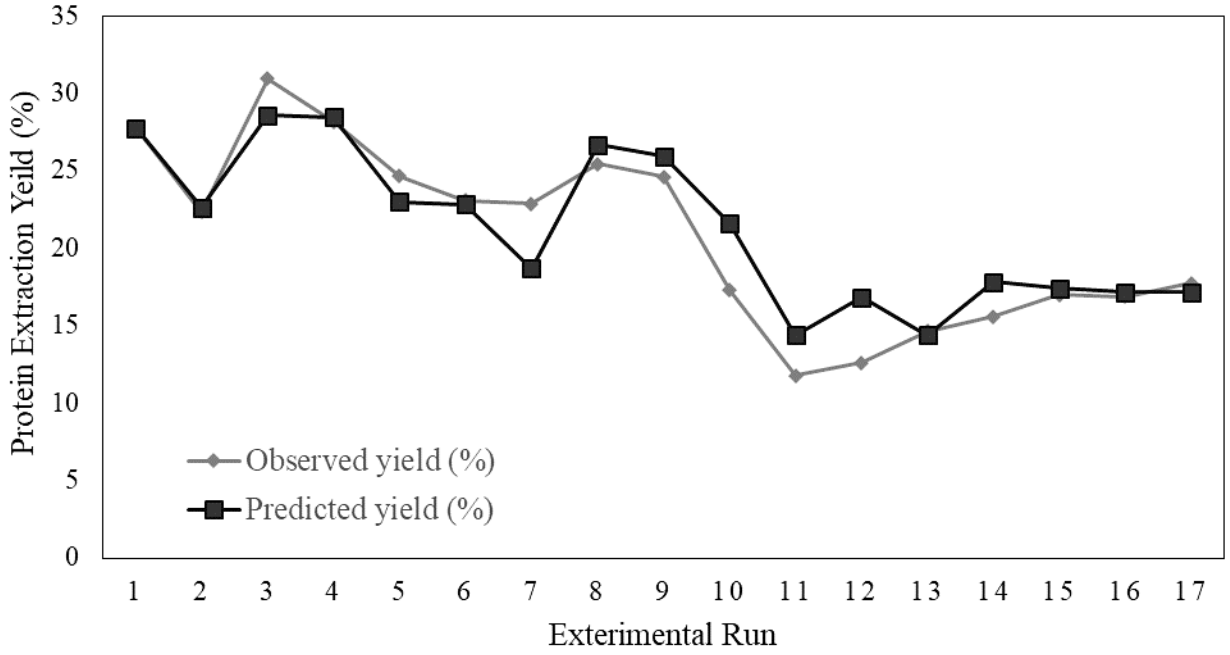


Figure 2. Comparison of experimental and ANN-predicted yield of sweet potato leaf protein extract.

3.3. NSGA-II-based multi-objective optimization and Pareto trade-offs

In Figure 3. 3D Pareto-Front Projection of Yield, Solid-to-Solvent Ratio, and Extraction Time Obtained by NSGA-II is illustrated. The validated cubic B-spline–Ridge model was coupled with NSGA-II to resolve the competing objectives of maximizing protein extraction yield while minimizing extraction time and solid-to-solvent ratio (Figure 4). The model showed satisfactory predictive performance, with (R^2) values of 0.942 for training and 0.807 under cross-validation, and RMSE and MAE of 2.465% and 1.747%, respectively. The comparative predictive performance of Gaussian Process (BO), ANN, and NSGA-II models for ultrasound-assisted protein extraction represents in Table 3. All representative Pareto solutions converged at a minimum amplitude of 50%, suggesting that increasing amplitude above this level provided limited benefit when yield was considered alongside processing requirements. The maximum-yield Pareto solution **(**50%, 15.00 min, 1:24.76**)** predicted 29.05% yield, closely matching the experimentally highest yield of 30.94% obtained at 50%, 15 min, and 1:25. Thus, the NSGA-II solution was located within the same favorable operating region identified experimentally, although the predicted value did not exceed the experimental maximum. More importantly, the Pareto front showed the potential to improve process efficiency with relatively small yield losses.

Reducing the extraction time and ratio to 13.65 min and 1:24.12**,** respectively, decreased predicted yield only marginally to 28.82%**,** representing a more resource-efficient alternative to the yield-oriented solution. In contrast, a stronger reduction in processing requirements **(**6.09 min; 1:19.85**)** reduced predicted yield to 26.20%**,** indicating a more pronounced recovery penalty as extraction intensity decreased. Collectively, the Pareto solutions demonstrate a clear yield–process efficiency trade-off**.** The 29.05% solution was preferable when maximizing protein recovery is prioritized, whereas the 28.82% solution provides a more balanced operating condition with only a 0.23 percentage-point reduction in predicted yield but lower time and solvent requirements. This comparative behavior highlights the advantage of NSGA-II over single-response optimization by identifying a range of technically meaningful operating conditions rather than a single optimum.

Table 3. Comparative predictive performance of Gaussian Process (BO), ANN, and NSGA-II models for ultrasound-assisted protein extraction.

Model	Dataset	R^2	RMSE (%)	MAE (%)	MAPE (%)
Gaussian Process (BO surrogate)	Training	0.998	0.224	0.150	0.655
Gaussian Process (BO surrogate)	Cross-Validation	0.872	2.022	1.484	7.087
ANN	Training	0.980	0.800	0.590	3.040
ANN	Cross-Validation	0.854	2.140	1.560	8.830
NSGA-II	Training	0.942	1.354	0.988	5.364
NSGA-II	Cross-Validation	0.807	2.465	1.747	9.598

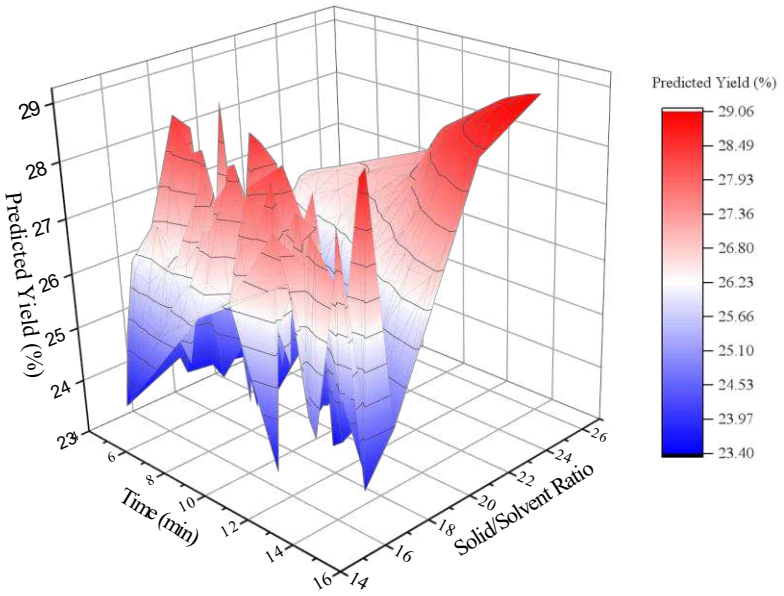


Figure 3. 3D Pareto-Front Projection of Yield, Solid-to-Solvent Ratio, and Extraction Time Obtained by NSGA-II

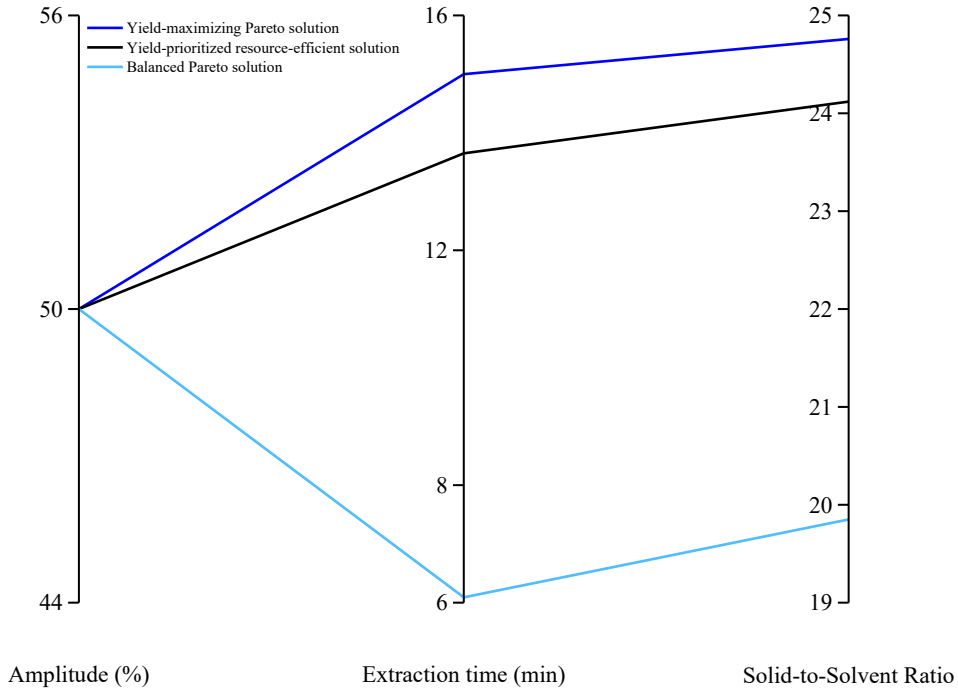


Figure 4. Parallel coordinates plot illustrating the trade-offs among ultrasound amplitude, extraction time, solid-to-solvent ratio, and predicted protein extraction yield for representative Pareto-optimal solutions obtained using NSGA-II

3.4. Characterization of functional groups of sweet potato leaf protein extract

ATR-FTIR provides structural insights and characterizes changes in protein functional groups. The amide I region (1700–1600 cm^{-1}) is attributed to C–O stretching vibration of the peptide backbone, whereas the amide II band (1600–1500 cm^{-1}) is associated with –NH bending along with C–N stretching, representing the effects of amplitude on protein secondary structure (Hu & Xiong, 2022). Figures 5(a) and 5(b) show that ultrasound treatment at 40 kHz induces significant amplitude-dependent changes in the conformation of sweet potato leaf protein over the wavenumber range of 400 to 4000 cm^{-1}. The most substantial spectral variations occurred in the amide I region. The amide I region (1700–1600 cm^{-1}) and amide II band (1600–1500 cm^{-1}) exhibited markedly broader profiles and higher intensity at 100% amplitude compared to 50% amplitude. Such peak broadening is generally indicative of structural disorder, protein denaturation, and disrupted hydrogen-bonding interactions. At 50% amplitude, reducing cavitation intensity at 40 kHz preserved the protein's native structure. Notably, the 70% amplitude at 40 kHz showed the weakest amide band spectra, suggesting that exposure of hydrophobic amino acids may have facilitated intermolecular interactions, leading to aggregation and partial unfolding, which also aligned with the lower extraction yield. Interestingly, under alkaline extraction conditions, covalent bonds such as C–N could form between polyphenol and protein (X. Wang et al., 2023). In Figure 5(a), the spectral region between 950–1200 cm^{-1} showed prominent C–O stretching and glycosidic C–O–C vibrations. These are generally attributed to carbohydrate-rich compounds such as polysaccharides, glycoproteins, pectin, and residual cellulose derivatives (Guo et al., 2018). At 100% amplitude, 40 kHz produced an exponentially intense peak, resulting in excessive ultrasound-induced cavitation of polysaccharide-rich cell wall structures. Consequently, this led to higher extraction of carbohydrate-associated compounds, reflecting decreased selectivity. At 50% amplitude, the peak showed reduced intensity, indicating a more balanced extraction condition, in which sweet potato leaf proteins were extracted without excessive carbohydrate-rich compounds. Conversely, the weak spectral intensity at 70% amplitude suggests insufficient disruption and lower mass-transfer efficiency. High sonication intensity, such as 100% amplitude at 40 kHz, broadened the peak between 3200–3500 cm^{-1} for O-H and amine N-H stretching, resulting in strong hydrogen bonding and maintaining an unfolded state. The stretching vibration of aliphatic groups in proteins was observed at 2927-2928 cm^{-1} and exhibited

pronounced broadening at 100% amplitude, indicating enhanced conformational freedom and structural disruption of protein aliphatic chains (Chen et al., 2016).

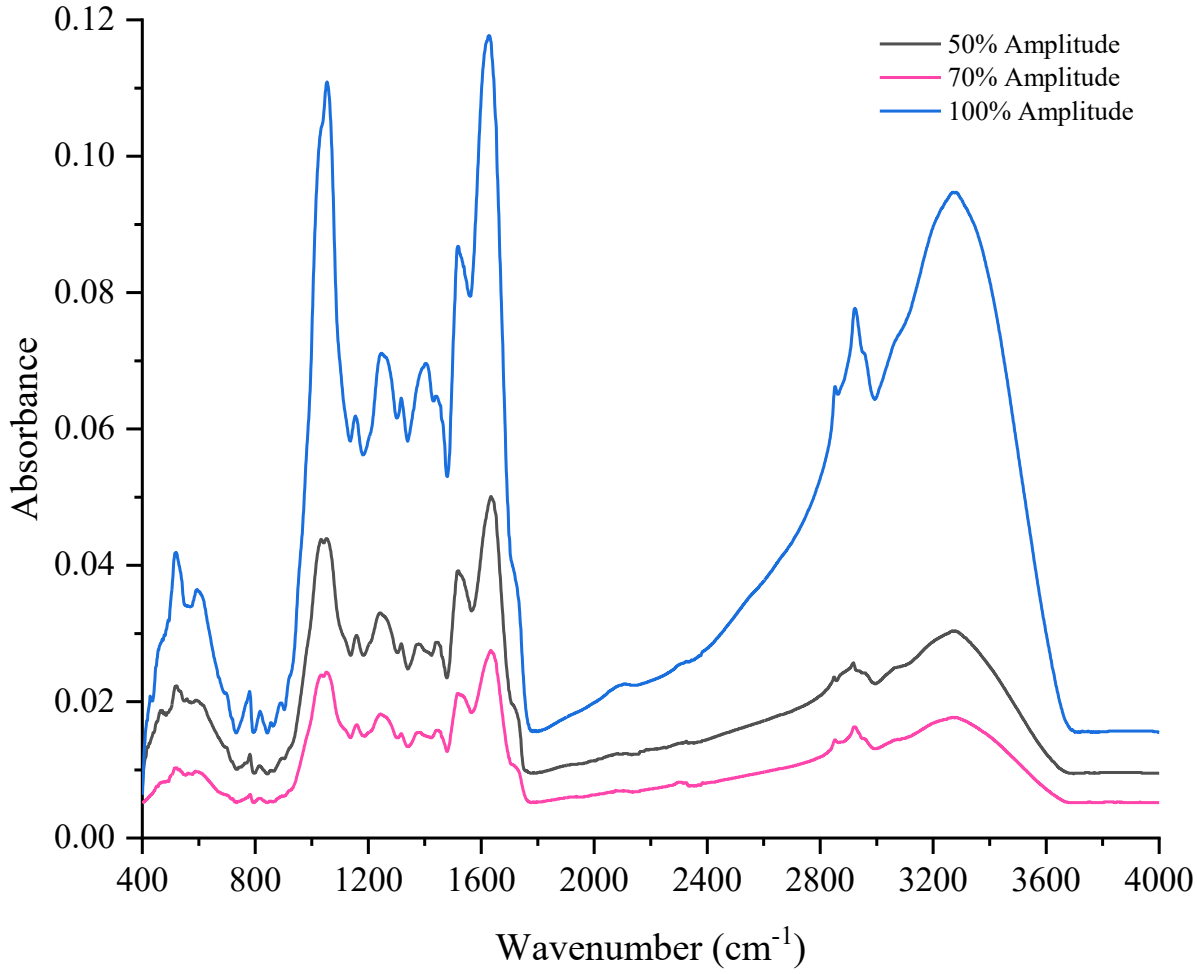


Figure 5(a). Full-range ATR-FTIR spectra (400–4000 cm^{-1}) of sweet potato leaf protein extracts obtained at ultrasound amplitudes of 50%, 70%, and 100%.

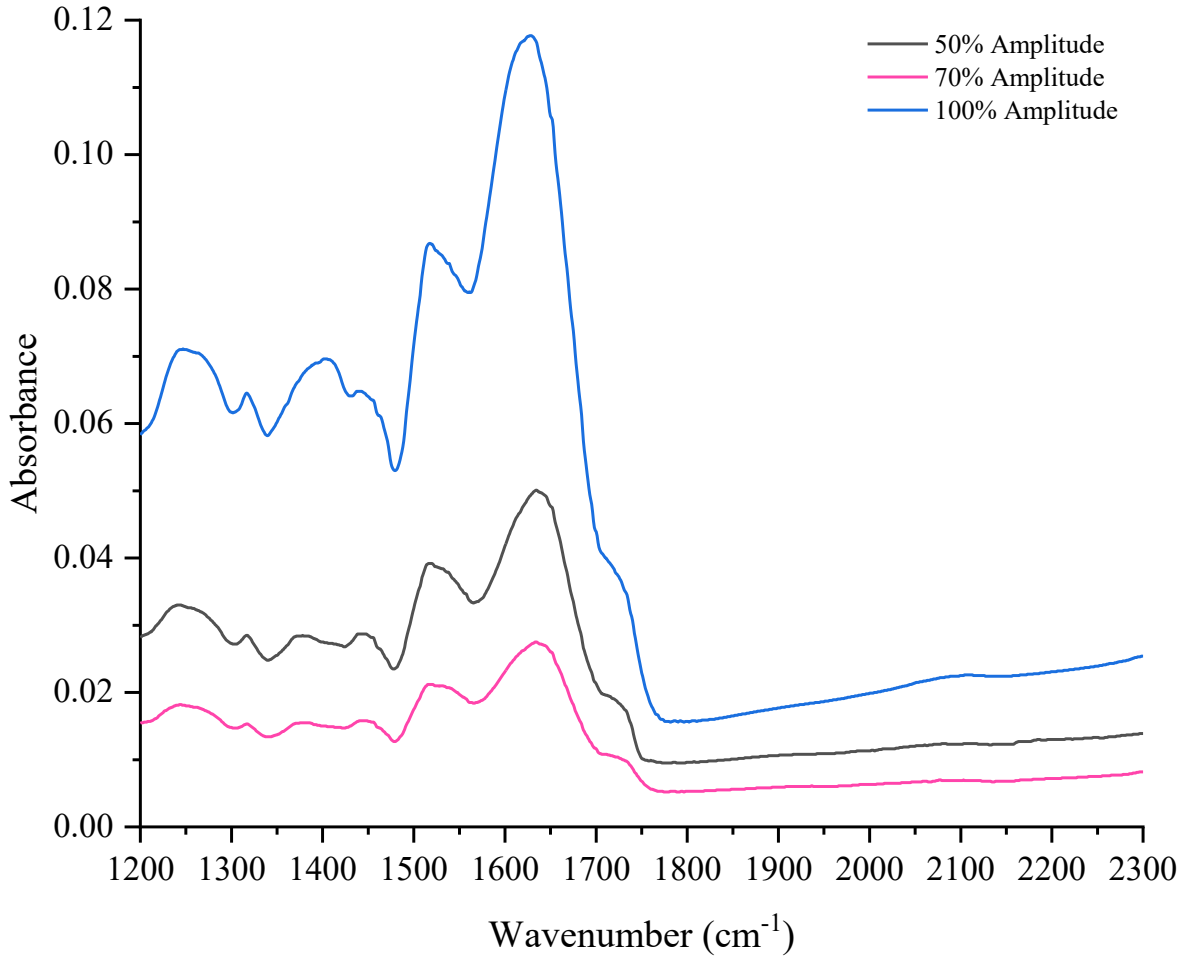


Figure 5(b). ATR-FTIR spectral range (1200–2300 cm^{-1}) of sweet potato leaf protein extracts obtained at ultrasound amplitudes of 50%, 70%, and 100%.

3.5. Water-holding and oil-holding capacity of optimized protein extracts

Water-holding capacity (WHC) provides indirect insights into the cellular microstructure of food matrices, reflecting the protein-polysaccharide matrix's ability to retain water within a three-dimensional network (Mosibo et al., 2022). Table 4. represents comparative physicochemical of optimized protein extract samples obtained using different modeling approaches. In contrast, oil-holding capacity (OHC) indicates overall hydrophobic affinity and potential lipid-protein interaction. The combined presence of water- and oil-holding capacities enhances functional properties, including aqueous dispersion stability and mouthfeel, in formulated food systems (Webb et al., 2023). The BO, ANN, and NSGA-II-optimized sweet potato leaf protein extracts exhibited water-holding capacity values of 3.17 ± 0.15 (g/g), 3.34 ± 0.40 (g/g), and 3.42 ± 0.33

(g/g). Although the values obtained from the three optimization approaches did not differ significantly ($p > 0.05$). The relatively high WHC may also enhance hydration stability and textural properties in food applications. Meanwhile, the oil-holding capacity of BO, ANN, and NSGA-II-optimized sweet potato leaf protein extracts were 1.35±0.09 (g/g), 1.09±0.18 (g/g), and 1.24±0.46 (g/g), indicating moderate hydrophobic interactions and partial exposure to nonpolar amino acid residues. Comparable water- and oil-holding capacities have been reported for other plant-derived protein sources (Shen et al., 2022). Collectively, these properties highlight balanced amphiphilic characteristics in food applications.

Table 4. Comparative physicochemical of optimized protein extract samples obtained using different modeling approaches.

Sample	WHC (g/g)	OHC (g/g)	SWC (g/g)	Solubility (%)
P1 (BO)	3.17±0.15[a]	1.35±0.09[a]	6.32 ± 0.34[a]	20.04± 0.27[a]
P2 (ANN)	3.34± 0.40[a]	1.09±0.18[a]	6.16 ± 0.74[a]	20.58± 0.51[a]
P3 (NSGA-II)	3.42± 0.33[a]	1.24±0.46[a]	5.88 ± 0.58[a]	20.11± 0.5[a]

Values are presented as mean ± SD (n = 3). Different lowercase letters (a, b) within the same column indicate significant differences among samples according to Tukey's multiple comparison test ($p < 0.05$). P1 = protein extract obtained under Bayesian optimization (BO)-optimized conditions; P2 = protein extract obtained under artificial neural network (ANN)-optimized conditions; P3 = protein extract obtained under non-dominated sorting genetic algorithm II (NSGA-II)-optimized conditions.

3.6. Solubility and Swelling Properties of optimized protein extracts

The solubility of BO, ANN, and NSGA-II-optimized sweet potato leaf protein extracts was 20.04±0.27 %, 20.58±0.51%, and 20.11±0.5%, respectively, at pH 7.0, suggesting moderate hydration functionality. The relatively high solubility likely reflects ultrasound-assisted extraction at pH 8.0, whereas pH 4.7 was maintained for isoelectric precipitation. A similar trend has been reported for other ultrasound-assisted extractions for protein recovery from other plant sources, suggesting that protein solubility is significantly enhanced. pH shifted toward acidic to neutral or slightly alkaline conditions (Yalçın et al., 2008). Ultrasonic extraction may also enhance solubility by inducing conformational changes, allowing hydrophilic amino acids to become more hydrated in aqueous dispersion (Yildiz et al., 2017). The swelling capacity was 6.32 ± 0.34 (g/g), 6.16 ± 0.74 (g/g), and 5.88 ± 0.58 (g/g), respectively, for BO, ANN, and NSGA-II-optimized sweet potato leaf protein extracts, indicating a moderate capacity for water uptake and retention. This

characteristic may be attributed to favorable hydration properties, which are closely associated with enhanced product stability, texture, structural integrity, and overall sensory qualities (Xiao et al., 2025).

3.7. Analysis of Physicochemical properties of Single and Dual-Stage Sonicated Fiber Extracts

Single-stage sonication and sequential double-stage ultrasound-assisted citric acid treatment at 80 °C produced distinct changes in the physicochemical properties of the fiber-rich residues (Table 5). The water-holding capacity of TPSF increased significantly from 5.11 ± 0.26 (g/g) to 6.13 ± 0.24 (g/g) compared to OPSF, indicating that the second treatment substantially modified the water capacity of the residue. The oil-holding capacity of TPSF increased from 1.86 ± 0.08 to 2.03 ± 0.37 g/g; however, it was not statistically significant ($p > 0.05$). Similarly, swelling capacity decreased slightly, with values of 6.81 ± 0.07 and 6.70 ± 0.26 g/g for OPSF and TPSF, respectively. Solubility was significantly reduced after the second-stage treatment, declining from 12.77 ± 0.63% in the one-stage sonicated fiber (OPSF) to 8.66 ± 0.54% in the two-stage sonicated fiber (TPSF) ($p < 0.05$), corresponding to a 32.2% decrease relative to OPSF. TPSF underwent an additional ultrasound–citric acid treatment, this reduction may reflect that citric-acid-induced cross-linking alters the inter and intramolecular interactions among polysaccharide chains, restricting their mobility (Karma et al., 2022). The dual-stage sonicated citric-acid-induced structural cross-linking may enhance the thermal stability of TPSF by reducing its solubility and swelling characteristics (Arachchi et al., 2025).

Table 5. Comparative physicochemical properties of single and dual phase ultrasonic isolated fiber extract.

Sample	WHC (g/g)	OHC (g/g)	SWC (g/g)	Solubility (%)
OPSF	5.11 ± 0.26[a]	1.80 ± 0.48[a]	6.81 ± 0.07[a]	12.77 ± 0.63[a]
TPSF	6.13 ± 0.24[b]	2.03 ± 0.37[a]	6.70 ± 0.26[a]	8.66 ± 0.54[b]

Values are presented as mean ± SD (n = 3). Different lowercase letters (a, b) within the same column indicate significant differences among samples according to Tukey's multiple comparison test ($p < 0.05$). OPSF = one-stage sonicated protein-extraction residue-derived fiber; TPSF = two-stage sonicated protein-extraction residue-derived fiber.

3.8. Structural and Chemical Group Characterization of Single and Dual-Stage Sonicated Fiber Extracts

The ATR-FTIR spectra of OPSF and TPSF showed the characteristic vibrational bands of the polysaccharide-rich fiber matrix, with distinct changes following the second-stage ultrasound-assisted citric acid treatment at 80 °C (Figure 6). The broad absorption at 3300–3500 cm^{-1} was assigned to O–H stretching vibrations, reflecting extensive inter- and intramolecular hydrogen bonding among hydroxyl-containing polysaccharides, including the galacturonic acid-rich pectic fraction of the fiber matrix (Arana et al., 2026). Its altered intensity and breadth in TPSF indicate changes in hydrogen bonding and hydroxyl-group accessibility. The band at 2926 cm^{-1} was attributed to C–H stretching of methylene groups, while increased intensity in TPSF suggested greater exposure of aliphatic C–H groups, potentially resulting from treatment-induced disruption and rearrangement of the polysaccharide matrix (Jiang et al., 2024). TPSF showed a distinct absorption around 1730 cm^{-1}, corresponding to the C=O stretching vibration of esterified carboxyl groups**.** The emergence of this band provides evidence of citric-acid-mediated esterification of accessible hydroxyl groups within the polysaccharide matrix, indicating successful chemical modification of the fiber structure (Gamonpilas et al., 2011). The fingerprint region exhibited enhanced absorptions at 1152, 1080, and 925 cm^{-1} in TPSF relative to OPSF. The band at 1152 cm^{-1} was associated with asymmetric C–O–C stretching, while the 1080 cm^{-1} band corresponded primarily to C–O stretching of carbohydrate moieties. The absorption near 925 cm^{-1} was related to glycosidic and carbohydrate ring-associated vibrations (Ye et al., 2019). The increased intensity of these bands following ultrasound-assisted citric-acid treatment suggests reorganization of the residual polysaccharide matrix, accompanied by greater exposure and accessibility of C–O–C, C–O, and glycosidic functionalities.

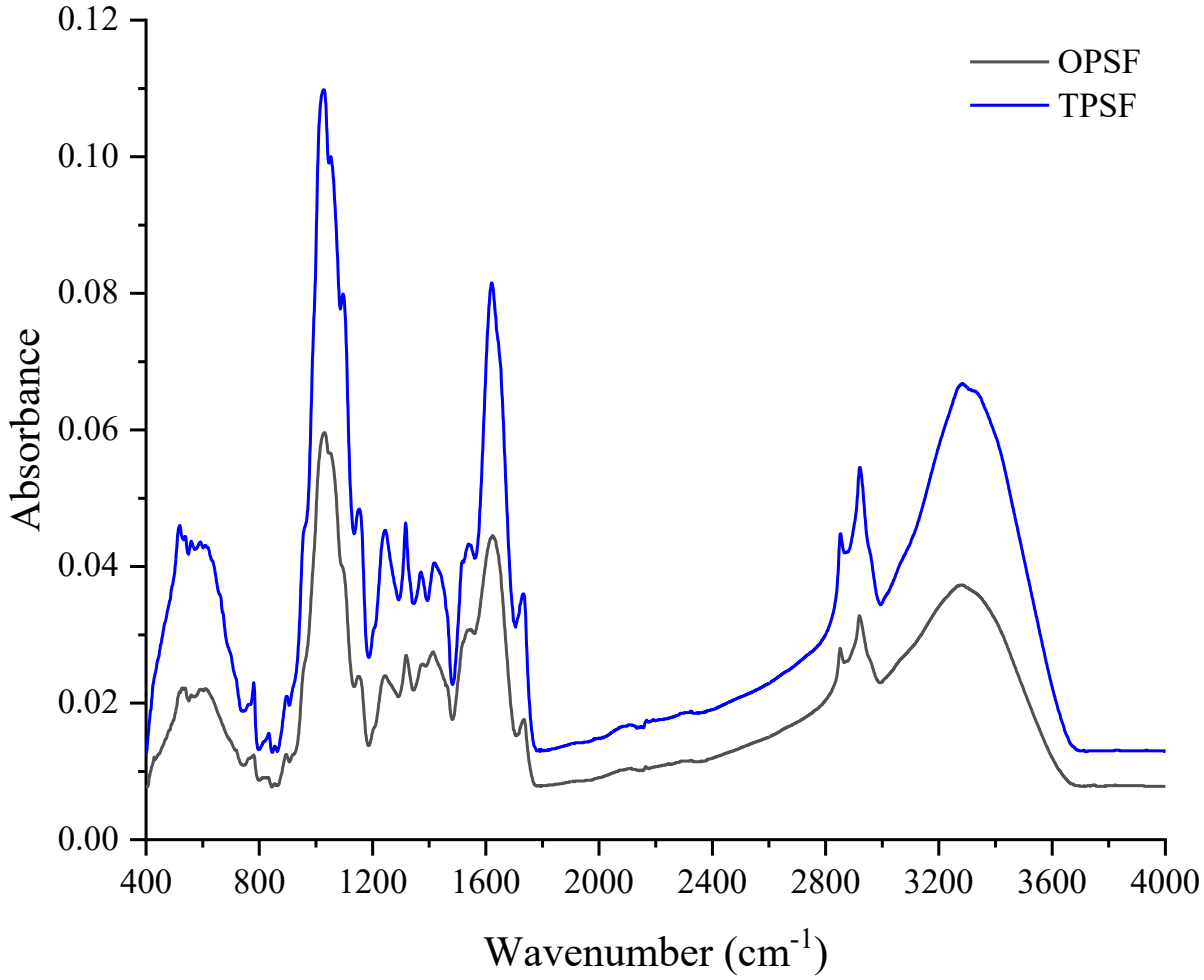


Figure 6. ATR-FTIR spectral range (400–4000 cm^{-1}) of single-phase sonicated fiber (OPSF) and dual-phase treated fiber (TPSF)

3.9. Amino acid composition and nutritional evaluation of sweet potato leaf protein extract

The amino acid profile of the sweet potato leaf protein extract showed substantial diversity, including essential and non-essential amino acids, reflecting its sustainability as an alternative plant protein source. Plant-based protein sources, including cereal seeds such as rice and wheat, as well as leafy vegetables, typically contain lower amino acid concentrations than animal-derived sources (Grácio et al., 2023). Table 6. represents the amino acid composition of sweet potato leaf protein extract. Among all quantified amino acids, aspartic acid and glutamic acid dominated the profile, with concentrations of 19.01 (mg/g) and 17.08 (mg/g), respectively. Leucine exhibited the highest value among essential amino acids, with a concentration of 12.41 (mg/g), followed by lysine (8.60 mg/g), histidine (7.56 mg/g), threonine (7.32 mg/g), valine (6.08 mg/g), and isoleucine

(5.05 mg/g), respectively. In the amino acid profile, the aromatic amino acids phenylalanine (8.88 mg/g) and tyrosine (6.25 mg/g) were also present in notable quantities. Similarly, histidine content (7.56 mg/g extract) was present in notable amounts and has been shown to stimulate muscle growth and increase protein deposition (Zeng et al., 2024). The amino acid profile of sweet potato leaf protein extract also showed considerable levels of arginine (7.62 mg/g), serine (7.00 mg/g), glycine (6.94 mg/g), alanine (5.90 mg/g), and proline (5.67 mg/g), respectively. Although several amino acids were below the WHO/FAO/UNU reference pattern, the sweet potato leaf crude protein extract exhibited a broad profile of essential and non-essential amino acids. The relatively lower concentrations may partly reflect the crude nature of the extract, where residual non-protein constituents can affect amino-acid values expressed per unit extract mass. Nevertheless, its diverse amino-acid composition supports its potential as a supplementary plant protein ingredient, particularly combined with complementary proteins to improve the overall amino-acid balance.

Table 6. Essential and non-essential amino acid composition of sweet potato leaf protein extract.

Amino Acid	Concentration (mg/g)
Aspartic acid	19.01 ± 1.11
Threonine	7.32 ± 0.60
Serine	7.00 ± 0.03
Glutamic acid	17.08 ± 0.46
Glycine	6.94 ± 0.98
Alanine	5.90 ± 0.72
Valine	6.08 ± 0.26
Methionine	2.00 ± 0.58
Isoleucine	5.05 ± 0.61
Leucine	12.41 ± 1.37
Tyrosine	6.25 ± 0.54
Phenylalanine	8.88 ± 0.93
Histidine	7.56 ± 1.59
Lysine	8.60 ± 1.16
Arginine	7.62 ± 0.88
Proline	5.67 ± 0.83

Values are presented as mean ± standard deviation (SD) (n = 2). Amino acid concentrations are expressed as mg/g of sample.

3.10. Glucose adsorption capacity of single and dual-stage sonicated fiber extracts

Glucose adsorption capacity was significantly enhanced by the second-stage treatment, increasing from 4.95 ± 0.03 g/g in OPSF to 5.91 ± 0.05 g/g in TPSF ($p < 0.05$), representing a 19.4% increase (Table 7). The approximately 19.4% increase following dual-stage sonication indicates enhanced glucose-binding functionality of TPSF. This improvement may be associated with sonication-induced structural rearrangement of the fiber matrix, which can increase the accessibility of hydroxyl and other hydrophilic functional groups and thereby provide more accessible sites for glucose interaction. The glucose adsorption capacity of insoluble dietary fiber may be influenced by the initial glucose concentration of the solution, as increasing glucose availability can alter the concentration gradient and the extent of glucose–fiber interactions (Zhang et al., 2023). The enhanced glucose adsorption capacity of TPSF suggests greater potential to delay glucose availability in the gastrointestinal environment and in vitro hypoglycemic activity. Previous studies have shown that citric acid treatment can increase the proportion of slowly digestible carbohydrate retained within the residual fiber matrix, potentially through acid-mediated structural modification (Lee et al., 2019).

Table 7. Comparative in vitro glucose and cholesterol adsorption capacities of one-phase sonicated fiber (OPSF) and dual-phase treated fiber (TPSF)

Sample	Glucose adsorption capacity (g/g)	Cholesterol adsorption capacity (g/g) (pH 2)	Cholesterol adsorption capacity (g/g) pH 7
OPSF	4.95 ± 0.03[a]	15.37± 0.63[a]	25.03 ± 0.17[a]
TPSF	5.91 ± 0.05[b]	16.95± 0.24[b]	26.15 ± 0.41[b]

Values are presented as mean ± SD (n = 3). Different lowercase letters within the same column indicate significant differences among samples according to Tukey's multiple comparison test ($p < 0.05$). OPSF = one-stage sonicated protein-extraction residue-derived fiber; TPSF = two-stage sonicated protein-extraction residue-derived fiber;

3.11. Cholesterol adsorption capacity of single and dual-stage sonicated fiber extracts

Dual-stage sonication significantly enhanced cholesterol adsorption capacity at both pH conditions ($p < 0.05$). At pH 2, the capacity increased from 15.37 ± 0.63 mg/g in OPSF to 16.95 ± 0.24 mg/g in TPSF, representing a 10.3% increase relative to OPSF. Similarly, at pH 7, TPSF demonstrated significantly greater cholesterol adsorption than OPSF, reaching 26.15 ± 0.41 mg/g compared with

25.03 ± 0.17 mg/g for OPSF ($p < 0.05$). Consistently, both fiber fractions showed greater cholesterol adsorption at pH 7 than at pH 2, indicating that the physicochemical environment of the fiber matrix influenced adsorption. The enhanced performance of TPSF may be associated with sonication-induced disruption and reorganization of the fiber structure, potentially exposing additional accessible binding domains and increasing fiber–cholesterol interactions. The cholesterol adsorption capacities of both OPSF and TPSF were markedly higher than those reported for several previously investigated dietary-fiber sources. Reported cholesterol adsorption capacities for wheat bran, soybean hull, apple peel, and mixed fiber were 2.17–10.75 mg/g at pH 2.0 and 3.48–11.34 mg/g at pH 7.0, respectively (Zhang et al., 2011). This substantially greater adsorption performance highlights the strong cholesterol-binding potential of the sonicated sweet potato leaf fiber fractions.

4. Conclusions

This study integrated AI/ML to develop functional food ingredients from agro-waste sweet potato leaves. The comparative study across models, validated with ATR-FTIR, provided meaningful insights into resolving process-system complexity, while each model showed strong predictive performance. The functional, glucose-, and cholesterol-binding capacities of the fiber extracts provided significant insights into their potential to lower cholesterol under both phases and modulate glucose levels. The findings further confirmed that sequential dual-stage treatment enhanced the fiber extract's binding capacities. The study also characterized the single- and dual-phase sonication behaviors of the fiber matrix. The diverse presence of crude amino acids also showed potential as supplements and for integration with other protein sources, while purity enhancement may further improve its nutritional profile. The water-holding, oil-holding, and swelling capacities also suggested considerable techno-functional properties for both protein and fiber extracts. Additionally, the sequential conversion of sweet potato leaf waste into complementary protein- and fiber-based functional ingredients demonstrates a novel integrated pathway for achieving zero-agro waste and its potential for metabolic-health-oriented food applications.

Funding Declaration: This research received no external funding.

Author contribution: Md Shadman Khan: conceptualization, methodology, Data curation, formal analysis, Visualization, investigation and writing-original draft; **Md Atikul Islam:** writing-review and editing; **Maruf Ahmed:** supervision, investigation, Validation, writing-review and editing, Funding acquisition.

Data availability

Data materials are available on request.

Declaration of Competing Interest

The authors confirm that they have no conflicts of interest with respect to the work described in the manuscript.

Ethics, Consent to Participate, and Consent to Publish declarations: not applicable.

References

Ademosun, A. O., Ademosun, M. T., Fasehun, O. I., & Oboh, G. (2026). Artificial intelligence as the new transformative driver in functional food development. *Food Research International: Reports*, *1*, 100010. https://doi.org/10.1016/j.frinr.2026.100010

Aguirre, J. (2023). The Kjeldahl Method. In J. Aguirre, *The Kjeldahl Method: 140 Years* (pp. 53–78). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-31458-2_4

Alam, M. J., Mahmud, A.-A., Rana, M. S., Hossain Molla, M. S., Alam, M. R., Anwar, M. M., Mahmud, T., & Hossain, A. (2025). Optimizing sweet potato planting time to enhance productivity in the riverine islands of Bangladesh. *Heliyon*, *11*(6), e42972. https://doi.org/10.1016/j.heliyon.2025.e42972

Alkalbani, N., Shahin, L., Benzeghiba, H., Obaid, R. S., Osaili, T. M., Cheikh Ismail, L., Al Qasssimi, G., Rauf, M., Abdulrahim, K., Almashgouni, A., Ashuweihi, F., & AL-Fuqaha, D. (2026). Artificial intelligence in functional food innovation: Bioactive enhancement and formulation optimization: A quasi-systematic review. *Food Chemistry: X*, *34*, 103628. https://doi.org/10.1016/j.fochx.2026.103628

Arachchi, D. M., Halim, A., Fadimu, G., Farahnaky, A., & Majzoobi, M. (2025). Green Starch Modification Using Citric Acid: Quinoa, Chickpea, and Cassava Starches. *Foods*, *14*(2), 164. https://doi.org/10.3390/foods14020164

Bernardi, S., Lupatini-Menegotto, A. L., Kalschne, D. L., Moraes Flores, É. L., Bittencourt, P. R. S., Colla, E., & Canan, C. (2021). Ultrasound: A suitable technology to improve the extraction and techno-functional properties of vegetable food proteins. *Plant Foods for Human Nutrition*, *76*(1), 1–11. https://doi.org/10.1007/s11130-021-00884-w

Bianco, M., Calvano, C. D., Ventura, G., Losito, I., & Cataldi, T. R. I. (2022). Determination of hidden milk allergens in meat-based foodstuffs by liquid chromatography coupled to electrospray ionization and high-resolution tandem mass spectrometry. *Food Control*, *131*, 108443. https://doi.org/10.1016/j.foodcont.2021.108443

Chen, Y.-C., Chen, W.-H., Lin, B.-J., Chang, J.-S., & Ong, H. C. (2016). Impact of torrefaction on the composition, structure and reactivity of a microalga residue. *Applied Energy*, *181*, 110–119. https://doi.org/10.1016/j.apenergy.2016.07.130

Dabbour, M., He, R., Ma, H., & Musa, A. (2018). Optimization of ultrasound assisted extraction of protein from sunflower meal and its physicochemical and functional properties. *Journal of Food Process Engineering*, *41*(5), e12799. https://doi.org/10.1111/jfpe.12799

Das, D., Panesar, P. S., & Saini, C. S. (2022). pH shifting treatment of ultrasonically extracted soybean meal protein isolate: Effect on functional, structural, morphological and thermal properties. *Process Biochemistry*, *120*, 227–238. https://doi.org/10.1016/j.procbio.2022.06.015

Ekrami, M., Shokrollahi Yancheshmeh, B., Roshani-Dehlaghi, N., Mobahi, N., Emam-Djomeh, Z., & Mohammadifar, M. (2025). Next-generation smart and safe foods: Artificial intelligence -driven strategies for 4D food pre-printing challenges. *Trends in Food Science & Technology*, *165*, 105317. https://doi.org/10.1016/j.tifs.2025.105317

Eldiehy, K. S. H., Daimary, N., Borah, D., Sarmah, D., Bora, U., Mandal, M., & Deka, D. (2022). Towards biodiesel sustainability: Waste sweet potato leaves as a green heterogeneous catalyst for biodiesel production using microalgal oil and waste cooking oil. *Industrial Crops and Products*, *187*, 115467. https://doi.org/10.1016/j.indcrop.2022.115467

Ervilha, R., Campos, D., Da Silva Macêdo, B., Bertuol, D. A., Tanabe, E. H., Yaseen, Z. M., Bodini, M., Saporetti, C. M., & Goliatt, L. (2025). Bayesian optimization-guided machine learning models for bio-oil yield prediction from lignocellulosic biomass pyrolysis. *Biomass and Bioenergy*, *203*, 108280. https://doi.org/10.1016/j.biombioe.2025.108280

Gamonpilas, C., Pongjaruvat, W., Fuongfuchat, A., Methacanon, P., Seetapan, N., & Thamjedsada, N. (2011). Physicochemical and rheological characteristics of commercial chili sauces as thickened by modified starch or modified starch/xanthan mixture. *Journal of Food Engineering*, *105*(2), 233–240. https://doi.org/10.1016/j.jfoodeng.2011.02.024

Grácio, M., Oliveira, S., Lima, A., & Boavida Ferreira, R. (2023). RuBisCO as a protein source for potential food applications: A review. *Food Chemistry*, *419*, 135993. https://doi.org/10.1016/j.foodchem.2023.135993

Gu, C., Wang, G., Zhuang, W., Hu, J., He, X., Zhang, L., Du, Z., Xu, X., Yin, M., Yao, Y., Sun, X., & Hu, W. (Walter). (2026). Artificial intelligence-enabled analysis methods and their applications in food chemistry. *Critical Reviews in Food Science and Nutrition*, *66*(1), 206–227. https://doi.org/10.1080/10408398.2025.2521648

Guo, Q., Ai, L., & Cui, S. W. (2018). Fourier Transform Infrared Spectroscopy (FTIR) for Carbohydrate Analysis. In Q. Guo, L. Ai, & S. Cui, *Methodology for Structural Analysis of Polysaccharides* (pp. 69–71). Springer International Publishing. https://doi.org/10.1007/978-3-319-96370-9_9

Hu, C., & Xiong, H. (2022). Structure, interfacial adsorption and emulsifying properties of potato protein isolate modified by chitosan. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, *638*, 128314. https://doi.org/10.1016/j.colsurfa.2022.128314

Jiang, C., Zeng, X., Wei, X., Liu, X., Wang, J., & Zheng, X. (2024). Improvement of the functional properties of insoluble dietary fiber from corn bran by ultrasonic-microwave synergistic modification. *Ultrasonics Sonochemistry*, *104*, 106817. https://doi.org/10.1016/j.ultsonch.2024.106817

Karma, V., Gupta, A. D., Yadav, D. K., Singh, A. A., Verma, M., & Singh, H. (2022). Recent Developments in Starch Modification by Organic Acids: A Review. *Starch - Stärke*, *74*(9–10), 2200025. https://doi.org/10.1002/star.202200025

Lai, S., Chen, Z., Zhang, Y., Li, G., Wang, Y., & Cui, Q. (2022). Micronization Effects on Structural, Functional, and Antioxidant Properties of Wheat Bran. *Foods*, *12*(1), 98. https://doi.org/10.3390/foods12010098

Lee, G. E., Kim, R. H., Lim, T., Kim, J., Kim, S., Kim, H.-G., & Hwang, K. T. (2022). Optimization of accelerated solvent extraction of ellagitannins in black raspberry seeds using artificial neural network coupled with genetic algorithm. *Food Chemistry*, *396*, 133712. https://doi.org/10.1016/j.foodchem.2022.133712

Lee, N., Seo, J.-M., Kim, H.-S., Seo, D.-H., Kim, J., Choi, H., & Lee, B.-H. (2019). Citric-acid treatment during rice processing increases the level of slowly digestible starch with a potential to regulate the post-prandial blood glucose level. *Journal of Cereal Science*, *89*, 102821. https://doi.org/10.1016/j.jcs.2019.102821

Liu, T., Lei, H., Zhen, X., Liu, J., Xie, W., Tang, Q., Gou, D., & Zhao, J. (2024). Advancements in modifying insoluble dietary fiber: Exploring the microstructure, physicochemical properties, biological activity, and applications in food industry—A review. *Food Chemistry*, *458*, 140154. https://doi.org/10.1016/j.foodchem.2024.140154

Ma, Y., Ke, J., Wang, Y., Zhou, Y., Gao, X., Wang, X., & Shen, Q. (2025). The improvement effect of insoluble dietary fiber of Polygonatum sibiricum on hyperlipidemia in high-fat diet

mice via gut microbiota and metabolites. *Frontiers in Nutrition*, *12*, 1601867. https://doi.org/10.3389/fnut.2025.1601867

Mäkinen, O. E., Wanhalinna, V., Zannini, E., & Arendt, E. K. (2016). Foods for Special Dietary Needs: Non-dairy Plant-based Milk Substitutes and Fermented Dairy-type Products. *Critical Reviews in Food Science and Nutrition*, *56*(3), 339–349. https://doi.org/10.1080/10408398.2012.761950

Mosibo, O. K., Ferrentino, G., Alam, M. R., Morozova, K., & Scampicchio, M. (2022). Extrusion cooking of protein-based products: Potentials and challenges. *Critical Reviews in Food Science and Nutrition*, *62*(9), 2526–2547. https://doi.org/10.1080/10408398.2020.1854674

Ochoa-Rivas, A., Nava-Valdez, Y., Serna-Saldívar, S. O., & Chuck-Hernández, C. (2017). Microwave and Ultrasound to Enhance Protein Extraction from Peanut Flour under Alkaline Conditions: Effects in Yield and Functional Properties of Protein Isolates. *Food and Bioprocess Technology*, *10*(3), 543–555. https://doi.org/10.1007/s11947-016-1838-3

Sert, D., Rohm, H., & Struck, S. (2022). Ultrasound-Assisted Extraction of Protein from Pumpkin Seed Press Cake: Impact on Protein Yield and Techno-Functionality. *Foods*, *11*(24), 4029. https://doi.org/10.3390/foods11244029

Sharma, N., Sahil, Madhumita, M., Kumar, Y., & Prabhakar, P. K. (2023). Ultrasonic modulated rice bran protein concentrate: Induced effects on morphological, functional, rheological, and thermal characteristics. *Innovative Food Science & Emerging Technologies*, *85*, 103332. https://doi.org/10.1016/j.ifset.2023.103332

Sheepo, M. L., Thungo, Z. G., & Sibiya, J. (2025). Socio-economic factors and cropping systems in sweet potato (Ipomoea batatas) cultivation among north-central regions of Namibia. *Journal of Agriculture and Food Research*, *21*, 101879. https://doi.org/10.1016/j.jafr.2025.101879

Shen, Y., Hong, S., Du, Z., Chao, M., O'Quinn, T., & Li, Y. (2022). Effect of adding modified pea protein as functional extender on the physical and sensory properties of beef patties. *LWT*, *154*, 112774. https://doi.org/10.1016/j.lwt.2021.112774

Su, J., & Cavaco-Paulo, A. (2021). Effect of ultrasound on protein functionality. *Ultrasonics Sonochemistry*, *76*, 105653. https://doi.org/10.1016/j.ultsonch.2021.105653

Sun, H., Mu, T., Xi, L., Zhang, M., & Chen, J. (2014). Sweet potato (Ipomoea batatas L.) leaves as nutritional and functional foods. *Food Chemistry*, *156*, 380–389. https://doi.org/10.1016/j.foodchem.2014.01.079

Tang, J. (2024). Recent progress in plant-based proteins: From extraction and modification methods to applications in the food industry. *Food Chemistry*. https://doi.org/10.1016/j.fochx.2024.101540

Topolska, K., Florkiewicz, A., & Filipiak-Florkiewicz, A. (2021). Functional Food—Consumer Motivations and Expectations. *International Journal of Environmental Research and Public Health*, *18*(10), 5327. https://doi.org/10.3390/ijerph18105327

Tu, G. L., (2015) Comparison of Enzymatic and Ultrasonic Extraction of Albumin from Defatted Pumpkin (Cucurbita pepo) Seed Powder. *Food Technology and Biotechnology*, *53*. https://doi.org/10.17113/ftb.53.04.15.4159

Valdez-Arana, J. D. C., Ramos, A. H., Orué, B. S., Diego, R. J., Pósito Baca, M. A., & Delgado-Soriano, V. (2026). Optimization of pectin extraction from quince peels (Cydonia oblonga) using ultrasound-assisted extraction with citric acid. *Applied Food Research*, *6*(2), 102461. https://doi.org/10.1016/j.afres.2026.102461

Wang, K., Li, M., Wang, Y., Liu, Z., & Ni, Y. (2021). Effects of extraction methods on the structural characteristics and functional properties of dietary fiber extracted from kiwifruit (Actinidia deliciosa). *Food Hydrocolloids*, *110*, 106162. https://doi.org/10.1016/j.foodhyd.2020.106162

Wang, X., Wang, W., Luo, S., Wang, N., Wang, L., Zhang, N., & Yu, D. (2023). Evaluation of ohmic heating modified soybean protein isolate structure and antioxidant film under different catechin concentrations. *LWT*, *186*, 115224. https://doi.org/10.1016/j.lwt.2023.115224

Webb, D., Dogan, H., Li, Y., & Alavi, S. (2023). Physico-Chemical Properties and Texturization of Pea, Wheat and Soy Proteins Using Extrusion and Their Application in Plant-Based Meat. *Foods*, *12*(8), 1586. https://doi.org/10.3390/foods12081586

Wei, X., Chen, T., Shen, G., & Ding, F. (2023). Alkali-activated slag backfill grout materials optimization: Statistical RSM modelling and NSGA-II performance control. *Journal of Materials Research and Technology*, *24*, 5613–5627. https://doi.org/10.1016/j.jmrt.2023.04.182

Wu, M., Xu, G., Chen, Y., Xu, X., Zhu, S., Shen, M., Zhang, T., & Shi, S. (2025). Artificial neural network model-based optimization of Polygonum perfoliatum L. Polysaccharide ultrasonic-assisted extraction, structural characterization, and anti-inflammatory activity. *Ultrasonics Sonochemistry*, *123*, 107654. https://doi.org/10.1016/j.ultsonch.2025.107654

Xiao, R., Flory, J., Alavi, S., & Li, Y. (2025). Physicochemical and functional properties of plant proteins before and after extrusion texturization. *Food Hydrocolloids*, *163*, 111119. https://doi.org/10.1016/j.foodhyd.2025.111119

Xue, H., Xu, J., Zhang, J., Wei, Y., Cai, X., & Tan, J. (2022). Modeling, optimization, purification, and characterization of polysaccharides from Lilium lancifolium Thunb. *LWT*, *162*, 113491. https://doi.org/10.1016/j.lwt.2022.113491

Yalçın, E., Çelik, S., & İbanoğlu, E. (2008). Foaming properties of barley protein isolates and hydrolysates. *European Food Research and Technology*, *226*(5), 967–974. https://doi.org/10.1007/s00217-007-0618-8

Ye, J., Luo, S., Huang, A., Chen, J., Liu, C., & McClements, D. J. (2019). Synthesis and characterization of citric acid esterified rice starch by reactive extrusion: A new method of producing resistant starch. *Food Hydrocolloids*, *92*, 135–142. https://doi.org/10.1016/j.foodhyd.2019.01.064

Yi, J., Li, L., Li, X., Duan, X., Wang, J., Han, Y., & Gao, Y. (2024). Valorisation of sweet potato leaves as a potential agri-food resource: Assessing nutritional and nutraceutical values altered by food processing—A review. *International Journal of Food Science and Technology*, *59*(11), 8825–8836. https://doi.org/10.1111/ijfs.17014

Yildiz, G., Andrade, J., Engeseth, N. E., & Feng, H. (2017). Functionalizing soy protein nano-aggregates with pH-shifting and mano-thermo-sonication. *Journal of Colloid and Interface Science*, *505*, 836–846. https://doi.org/10.1016/j.jcis.2017.06.088

Yu, Z., Yang, Y., Zhang, Y., Zhang, R., Lyu, B., Wang, S., Jing, W., & Yu, H. (2025). Citric acid modification of insoluble dietary fiber from soybean residue: Microstructural reconstruction, interfacial activity enhancement and stability by pickering emulsions. *Food Research International*, *221*, 117465. https://doi.org/10.1016/j.foodres.2025.117465

Zeng, X., Zhou, X.-Q., Jiang, W.-D., Wu, P., Liu, Y., Ma, Y.-B., Tang, L., Li, S.-W., Kuang, S.-Y., & Feng, L. (2024). Histidine is effective in promoting muscle growth and protein

deposition: In vivo and in vitro experimental models of grass carp (Ctenopharyngodon idella). *Food Bioscience*, *62*, 105537. https://doi.org/10.1016/j.fbio.2024.105537

Zhang, G., Wang, D., Ding, Y., Zhang, J., Ding, Y., & Lyu, F. (2024). Effect and mechanism of insoluble dietary fiber on postprandial blood sugar regulation. *Trends in Food Science & Technology*, *146*, 104354. https://doi.org/10.1016/j.tifs.2024.104354

Zhang, H., Claver, I. P., Zhu, K.-X., & Zhou, H. (2011). The Effect of Ultrasound on the Functional Properties of Wheat Gluten. *Molecules*, *16*(5), 4231–4240. https://doi.org/10.3390/molecules16054231

Zhang, N., Huang, C., & Ou, S. (2011). In vitro binding capacities of three dietary fibers and their mixture for four toxic elements, cholesterol, and bile acid. *Journal of Hazardous Materials*, *186*(1), 236–239. https://doi.org/10.1016/j.jhazmat.2010.10.120

Zhang, X., Zeng, Y., Liu, J., Men, Y., & Sun, Y. (2023). Effects of three extraction methods on the structural and functional properties of insoluble dietary fibers from mycoprotein. *Food Chemistry Advances*, *2*, 100299. https://doi.org/10.1016/j.focha.2023.100299

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Suplimentaryfile.docx