A Robotic Approach for Accelerating Discoveries in Food Science Through Automated Preparation, Assessment and Optimization

Stefan Ilić, Edgar Chávez Montes, Constantijn Sanders, Cécile Gehin-Delval, Giulia Marchesini, Christoph Hartmann, Josie Hughes

***Abstract*—Addressing the challenges of a growing global population and the agricultural impacts of climate change requires transitioning to sustainable food solutions that enhance food security, affordability and nutrition. Plant-based food products are among the most promising options, but the introduction of novel ingredients and processing methods in this emerging design space poses significant challenges. To accelerate the design of such food products, we propose autonomous online experimentation and optimization by combining robotic preparation and assessment with active learning. We focus on minimizing the viscosity of a model plant-based food system consisting of water, flour, protein, salt and enzyme, which undergoes multi-step processing, including mixing and enzymatic hydrolysis. We demonstrate that robotic automation not only automates the experiment, but also improves precision compared to humans. Using Bayesian Optimization, we identify the lowest viscosity formulation in just 20 experiments. Constructed Gaussian process models enable viscosity predictions for untested formulations with an error as low as 6%. Multi-objective optimization further expands the set of optimal solutions by balancing cost and viscosity, enabling the discovery of formulations optimized for viscosity, cost, nutrition and trade-offs. These findings highlight the potential of robotic and learning-based-approach to accelerate food product development across diverse food science applications.**

***Note to practitioners*—To meet the growing demand for sustainable and nutritious food products, the food industry is increasingly focusing on plant-based food alternatives. Developing these products involves diverse ingredients and extensive laboratory trials, which are both time-consuming and labor-intensive. We propose an integrated robotic and active learning system to accelerate this process. By focusing on minimization of viscosity in a model plant-based food system composed of water, flour, protein, salt and enzyme, robotic automation is used to automate experiments, increase throughput, improve precision and save researcher's time. Bayesian Optimization in just 20 experiments minimizes viscosity of such food system, while Thompson Sampling balances competing objectives such as cost and viscosity, to identify set of optimal solutions across the design space and constructing the Pareto Front. Experimental data is further used to build predictive models, providing a valuable tool for food scientists to quickly estimate viscosity for untested formulations. Expansion of this work will focus on optimizing stability of the model food system, incorporating oil and high-pressure homogenization as an additional processing step.**

***Index terms*—Robotic automation, Online Optimization, Food Science.**

Stefan Ilić and Josie Hughes are with École polytechnique fédérale de Lausanne (EPFL), School of Engineering (email: stefan.ilic@epfl.ch; josie.hughes@epfl.ch).

Edgar Chávez Montes and Cécile Gehin-Delval are with Nestlé Product Technology Centre Coffee in Orbe, Switzerland (email: brunoedgar.chavezmontes@rd.nestle.com; cecile.gehin-delval@rdor.nestle.com).

Constantijn Sanders is with Nestlé Product Technology Center in Konolfingen, Switzerland (email: constantijn.sanders@rdko.nestle.com).

Giulia Marchesini and Christoph Hartmann are with Nestlé Research in Lausanne, Switzerland (email: giulia.marchesini1@rd.nestle.com; christoph.hartmann@rdls.nestle.com).

I. INTRODUCTION

The food industry supplies economically viable, tasty, and nutritionally balanced food to meet the global population's nutritional needs [1]. However, this is becoming increasingly difficult due to growing global population and deteriorating agricultural conditions caused by climate change [2]. Furthermore, food production is also part of the problem as it has substantial environmental impact with high resource use and significant greenhouse gas emissions [3]. It is therefore critical that we radically move towards more sustainable food solutions that improve food security, cost and nutrition [4]. One viable path which targets both sustainability and nutrition is the adoption of plant-based ingredients [5], [6]. This shift can offer environmental benefits such as reduced greenhouse gas emissions and land use [7], alongside health benefits, including a lower risk of chronic diseases, often linked to plant-based diets [8], [9]. However, the development of such food products introduces many new ingredients and processes for which we have limited expertise [10], [11], necessitating active exploration to discover new formulations. Currently, addressing this challenge requires extensive laboratory trials, which are resource intensive, time-consuming and often fail to provide the solutions needed to develop plant-based foods with the desired attributes. Therefore, new methodologies are required for the discovery of better plant-based food products.

Robotics and machine learning methods present a compelling solution to rapidly accelerate the exploration and discovery of optimal food formulations and processing conditions for new food systems [12]. Automation of the experiments offers precise control of process parameters, motions and timing, enabling continuous process monitoring for error detection and data capture [13], [14]. Such approach has been already applied to several scientific domains, including biotechnology [15], [16], chemistry [17], [18], and even for engineering sciences [19], [20]. These approaches have demonstrated the improved throughput, accuracy and precision enabled by the robots [21], [22]. However, the environments in which these robots operate are typically well structured and use fully automated analysis equipment [23], [24]. In contrast, food science environments require robots capable of robust operation with diverse and often non-automation friendly equipment and measurement devices, while also handling solid and liquid ingredients to formulate or prepare different recipes.

A second opportunity enabled by the robotic automation, is to combine the improved throughput with machine learning for automated optimization and discovery [25]. This has been shown to be a powerful approach for accelerating scientific discoveries. For example, optimization of a photocatalyst for

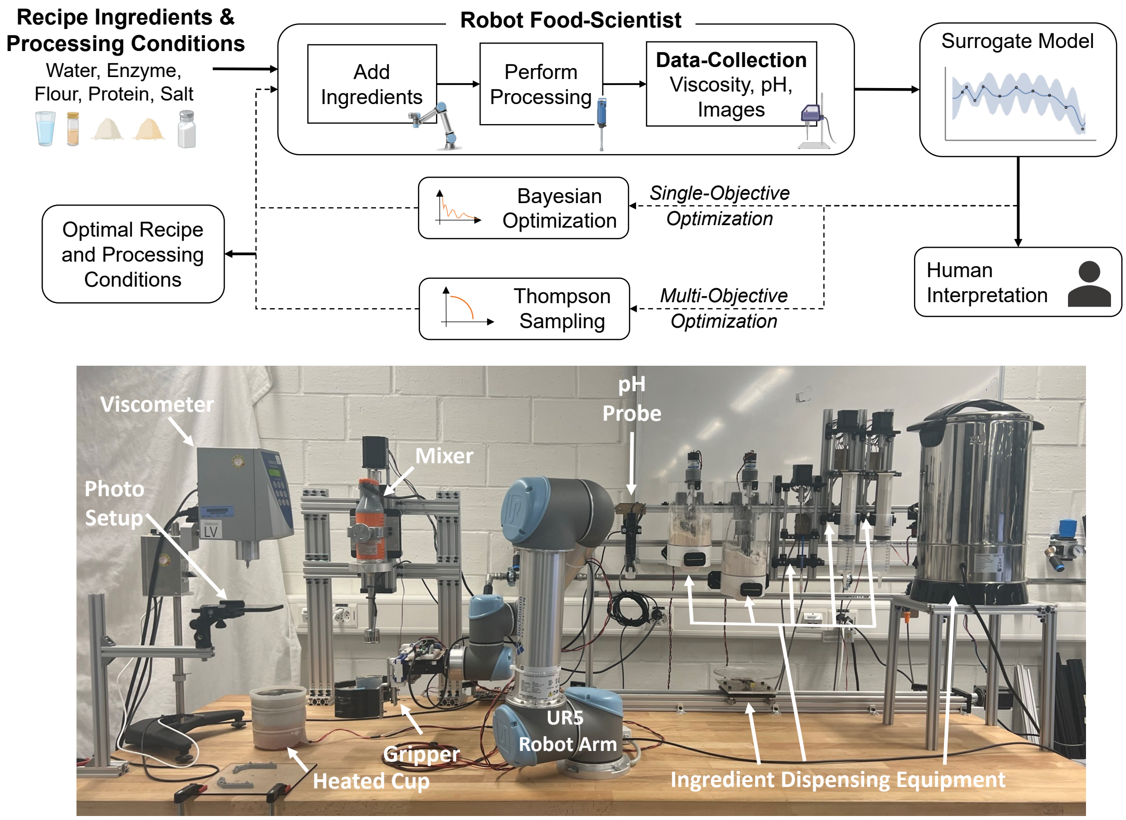


Fig. 1. The robotic system used for the preparation, processing, analysis and optimization of a liquid plant-based food systems. The robotic system consists of the UR5 robotic arm with mounted custom gripper used for grasping of the heated cup in which powder and liquid ingredients are dosed using the dispensing system. The plant-based model food system is processed using the mixer and is analyzed using the viscometer, pH probe and photo capturing setup.

hydrogen production has been performed with a mobile robotic platform and Bayesian Optimization [26], the use of automated microfluidics and linear discriminant analysis (LDA) has been applied to construct a model of the associated chemical space, while neural networks alongside a reconfigurable continuous-flow robotic platform have been used for exploration of chemical synthesis [27]. Within the domain of food science, the time scale of experiments and the variety of processing conditions make online optimization more challenging. However, a robot 'cook' has been combined with Bayesian optimization to identify optimal recipes and processing methods for robotic cooking of scrambled eggs [28]. Additionally, robotic pouring and computer vision have been coupled with Artificial Neural Networks to assess and classify the foaming behavior of beer and sparkling wines [29], [30]. While existing research highlights the potential of combining robotics with active learning, applications in food science remain scarce.

The design and optimization of food products typically requires discovery of optimal formulations and processing conditions. These are evaluated through quantitative metrics measured by scientific instruments and qualitative assessments performed by humans. For liquid food systems, which are the focus of this paper, this typically involves ingredient selection and processing steps such as mixing, homogenization and heat treatment. This creates a vast design space, with formulations ranging from a few to a dozen ingredients in varying quantities and processing times ranging from minutes to hours under diverse mixing and temperature conditions. Traditional manual methods struggle to handle many possible combinations, making it challenging to discover optimal food formulations. To accelerate the discovery and optimization of new food systems, we propose combining full robotic automation of the experiments with active learning. Specifically, we focus on developing a novel plant-based food concentrate which forms the basis of several food products, including plant-based dairy alternatives. The formulation of this food base requires addition and incorporation of five ingredients which then undergo mixing and heating processes, with single experiment lasting up to one hour. We propose a robotic approach for automating this process, combining robotic manipulation with automation of the lab environment to achieve the required flexibility, robustness and adaptability for diverse experimental workflows. This enables full control over timing, motions and the overall process, allowing experiments to be conducted in reliable and repeatable conditions with continuous data-capture. Compared to manual methods, this approach improves accuracy and repeatability. For the discovery of optimal food formulations, the robotic system is coupled with single and multi-objective optimization methods, creating a closed-loop optimization framework to balance competing objectives, such as cost and processing duration.

The developed robotic system (Figure 1) utilizes a 6 degree of freedom (DoF) robotic arm, a custom gripper, temperature-controlled cup and automated laboratory infrastructure. It is capable of dosing and combining ingredients, performing processing tasks (mixing and temperature control) and collecting data (viscosity, pH and imaging data). By integrating robotic automation with Bayesian Optimization and Thomp-

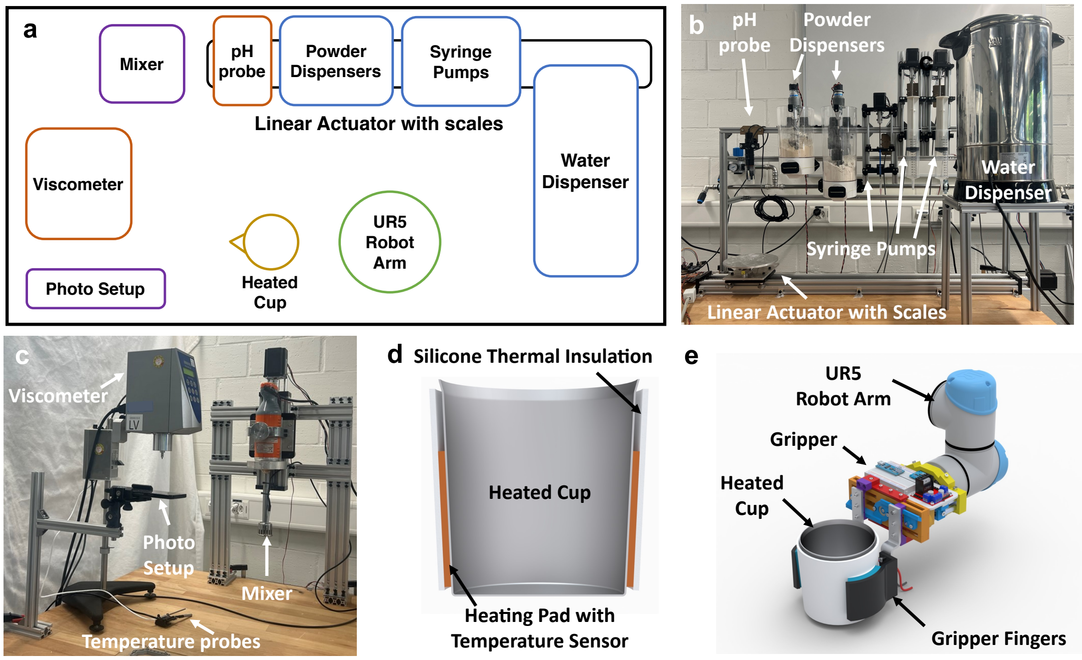


Fig. 2. (a) Layout of the robotic setup. (b) pH probe and the equipment used for dispensing of ingredients including the linear actuator with load-cell based scales. (c) High shear mixer used for the food system processing, viscometer used for viscosity measurements alongside the temperature probes and photo setup used for food system appearance capture. (d) The heated cup with externally attached heating pad with integrated temperature sensor and silicone thermal insulation. (e) The gripper used for the heated cup manipulation.

son Sampling, the robotic system enables discovery of new formulations that balance multiple objectives, such as cost, sensory experience, processability and nutrition. We demonstrate the applicability of these simple optimization methods for food science applications, highlighting their ability to optimize complex food interactions that cannot be mathematically expressed. By reusing data across optimization methods, we reduced experimental requirements, identifying optimal formulations and building predictive viscosity models with as few as 30 experiments for a given dry matter concentration. By developing the first (to our knowledge) fully automated food optimization robotic system, we demonstrate the potential of combining robotics with active learning for the development of plant-based and other novel food products.

In the rest of the paper, we first introduce the analyzed plant-based food system and its experimental process, and provide details on the robotic system and the applied formulation optimization techniques. We then benchmark the efficiency, throughput and precision of the robotic system against human-driven experimentation. Furthermore, we demonstrate the system's ability to minimize and predict viscosity, as well as to identify optimal food formulations by balancing competing objectives. We conclude by discussing the applicability of the robotic system to diverse food products and outlining future directions for both automation and recipe optimization.

II. METHODS

The specific food science problem we address is the viscosity optimization of a concentrated liquid food system prepared with plant-based ingredients. This proof-of-concept food system is composed of flour, protein, enzyme, salt and water, and is prepared using mixing and heating processes. During processing, the enzyme hydrolyzes the starch in the flour under the influence of heat and mixing, reducing the food system's viscosity. The resulting liquid food system serves as the foundation of many final products. The diverse range of possible ingredients and processing steps creates a vast design space in terms of formulations and processes, making final product properties difficult to predict. Additionally, optimization must account for numerous (and often competing) requirements and constraints, including processability, nutritional value, and cost. In our case-study, the primary objective is to minimize the viscosity, which affects processability and mouthfeel, while also reducing the use of costly ingredients. The robotic system has been developed to automate preparation and analysis, enhancing repeatability, reliability and accuracy of experiments. By coupling the robotic system with single and multi-objective optimization methods, we can discover optimized formulations, identify tradeoffs between competing objectives and build predictive models.

A. The Liquid Plant-Based Model Food System

The explored proof-of-concept concentrated liquid food system is formed by combining plant-based flour, protein, salt, enzyme and water. The mixture is heated to 65-75°C to activate the enzyme, while continuous high-shear mixing enables the enzyme to perform the hydrolysis of starch into sugars, reducing the viscosity of the food system [31]. pH regulation is performed using salt, keeping it within acceptable range. This process represents an initial development phase of the food product, prior to scaling to factory-level production.

The analyzed food system recipe is defined by ingredient and process parameters given in Table I. The bounds of the F/P Ratio are determined by protein levels (12-24%), balancing nutritional value and cost-effectiveness, while mixing time and temperature are selected based on enzyme properties. The dry

Parameters	Range
Flour/Protein (F/P) Ratio	3-7
Enzyme	0.05-0.15%
Mixing Time	9-24 min.
Temperature	65-75 °C
Dry Matter (DM)	25-35%

TABLE I
THE INGREDIENT AND PROCESS PARAMETERS USED TO PARAMETRIZE THE ANALYZED PLANT-BASED MODEL FOOD SYSTEM

matter (DM) parameter, representing the ratio of solid to wet ingredients in the mix, is calculated using Equation 1. For the analyzed beverage, a DM range of 25-35% is relevant, with two specific ratios, 27% and 30% being explored. These broad parameter limits were intentionally chosen to extend beyond typical ranges, enabling the discovery of novel formulations.

$$\text{Dry matter (\%)} = \frac{\text{Total food system weight} - \text{Water weight}}{\text{Total food system weight}} \quad (1)$$

B. Robotic Automation & Data Capture

The robotic setup exploits both robotic manipulation and automation of the lab equipment to enable flexible automation of all processes. The setup shown on Figure 1 uses a 6 degree of freedom robot arm (UR5) with a custom gripper, high shear mixer (DynamiX 160), rotational viscometer (Brookfield DV-II+ Pro LV), pH probe (DFRobot Meter Pro Kit V2), custom dispensing system and a heated cup. The 6 DoF robotic arm was chosen primarily for its compact footprint and ability to perform versatile motions, making the robotic system highly adaptable to changes in the experimental process or equipment layouts. Temperature monitoring is performed by three DS18B20 probes positioned at the mixer head, pH probe and attached to a cup during viscosity measurement. The robotic setup is designed for human-robot collaboration, allowing task switching between humans and robots for improved throughput. The robotic system is controlled by a centralized Python-based controller hosted on a PC, with the experimental process fully pre-programmed. This controller manages and synchronizes all equipment, including the robotic arm, which employs inverse kinematics to execute pre-set, parametrized paths. Optimization methods interact with the controller to exclusively exchange formulation parameters. Human interaction with the system is minimal and limited to the viscometer, which is operated manually due to compatibility constraints.

The gripper shown on Figure 2 e) has a DC motor linked to a rack-and-pinion mechanism to enable finger motion. Motor current feedback is used to detect when an object is held, while a linear potentiometer provides position control. The fingers are padded with sponge and silicone to enhance gripping performance. Mounted on the robot arm, the gripper is used to grasp and move the heated cup around the workspace.

The custom heated cup shown on Figure 2 d) is a 0.7-liter stainless steel container with an externally mounted 150 W heating pad featuring an integrated temperature sensor. To minimize heat loss and ensure safe, damage-free robotic handling, the cup is insulated with a silicone sheet. Temperature is regulated using a bang-bang controller.

The dispensing system shown on Figure 2 b), enables precise dosing of both dry and liquid ingredients. During the dispensing process, the heated cup is placed on a custom load-cell-based scales, moved by a one-meter-long lead screw linear actuator. By integrating real-time weight measurements with dispenser control, the system achieves closed-loop feedback control for precise ingredient dosing. Water is first dispensed in bulk using a heated liquid dispenser equipped with a tap, automated using a servo motor for opening and closing. Precise dosing of water, enzyme and saline solution (for salt) is achieved using three custom syringe pumps driven by stepper motors. Water and saline solution are dispensed using a 150 ml syringe with ±0.5 ml precision, while enzyme is dispensed using a 1 ml syringe. Powder ingredients, such as flour and protein, are dispensed using two automated powder dispensers which use a servo motor to pull a string to open or close the dispenser outlet. To prevent clogging, powders are continuously stirred during dispensing.

A calibrated rotational viscometer shown in Figure 2 c), is used for viscosity measurements. Equipped with RV6 spindle, it covers the full viscosity range of the analyzed food system, with measurements falling within 15-45% torque and a tolerance of ±1% at 20 RPM [32]. The viscometer is operated using Brookfield Rheocalc32 software.

C. Experimental Process

The experimental process, shown on Figure 3, begins with the robot picking up the heated cup and placing it on a custom scale mounted on a linear guide. The heated cup's temperature controller is activated at the start of the experiment and remains active through the process. The linear guide moves the cup between dispensers, starting with the hot liquid dispenser for bulk water dosing. This step targets a weight 30 g below the total water amount due to the dispenser's lower precision. The remaining water is then accurately dosed using a syringe pump. Following water dispensing, the heated cup is moved to the enzyme syringe pump for dosing. Once complete, the linear guide moves the cup to the robot pickup position. The robot arm then transfers the cup to the mixer, where the enzyme and water are mixed for 10 s at 3000 RPM, ensuring enzyme dispersion and uniform enzymatic reaction.

In the next stage, flour and protein are added incrementally to ensure thorough incorporation. The process begins with the robot placing the cup on the linear guide, which moves it to the corresponding dispensers. After dispensing, the robot moves the cup to the mixer for ingredient incorporation. Flour is added in three stages, each followed by mixing at 3000 RPM for 10 s, then protein is dispensed in two stages, each followed by mixing at 5000 RPM for 45 s. Salt is added as a saline solution with a 25% salt-to-water weight ratio using a syringe pump. To minimize foaming during the ingredient incorporation process, the robot arm moves the cup around the mixer head at 0.24 m/s in a clockwise direction.

Following the ingredient incorporation process, the liquid food system undergoes extensive mixing for an additional 9-24 minutes. During this time, the robot arm continuously moves the cup around the mixer head at a speed of 0.08 m/s in

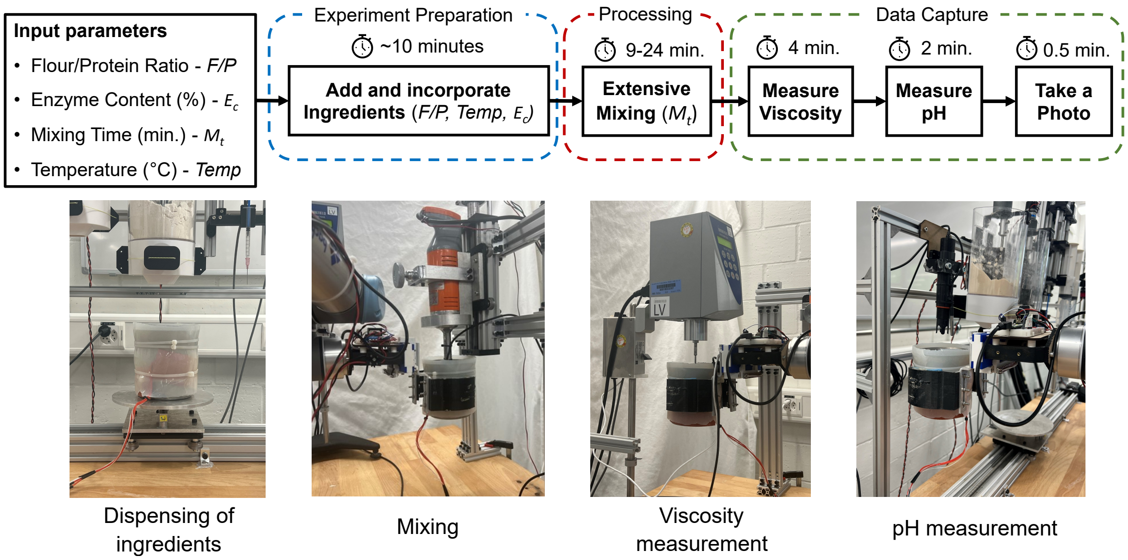


Fig. 3. Illustration of the experimental laboratory process used for optimizing liquid plant-based food systems. The controllable input parameters consisting of two ingredient and two process parameters, influence the measured metrics, primarily viscosity. Each experiment lasts between 35-50 minutes, followed by cleaning, with the duration largely dependent on the mixing time.

a clockwise direction. The mixing begins at 5000 RPM for the first 9 minutes, after which the mixer speed is reduced to 3000 RPM due to a decrease in the viscosity of the food system. After mixing, the robot transfers the heated cup to the viscometer, where viscosity is measured according to a predefined procedure. Viscosity measurements are taken at the center of the liquid every 15 s for a total duration of 150 s at a spindle speed of 20 RPM. Next, the cup is moved to the pH probe for a 60 s pH measurement at the center of the liquid. An initial waiting period of 15 s is included to minimize temperature disparity between the pH probe and the liquid food system. In the final step, the heated cup is transferred to a camera, which captures an image of the sample. Upon completion of all experimental steps, the cup is released at a predefined location, requiring manual cleaning of the equipment to prepare for the next experiment.

Temperature control is performed using a bang-bang controller, which maintains the temperature within ± 0.5 °C of the setpoint, and compensates for the 1-2 °C rise caused by the mixing process. However, during viscosity and pH measurements, when the liquid is stationary, a temperature difference of 3-5 °C develops between the center and the edge of the liquid due to heat transfer from the cup's sides. This temperature disparity cannot be corrected by the temperature controller alone and requires additional stirring to achieve uniformity. In the experimental process, however, no additional stirring was performed during viscosity and pH measurements to avoid further enzymatic reaction.

Although the entire experimental process is automated, it is often quicker and more convenient to manually weigh and dispense the ingredients, as this process step tends to be slow. However, the enzyme must always be dosed robotically to ensure high accuracy, as it has a significant impact on viscosity. Powdered ingredients are dispensed in multiple steps due to their lower density compared to water, which creates a surface powder layer that hinders ingredient incorporation.

Compared to fully manual experimentation, the robotic system follows the same process steps but performs them sequentially. In the manual approach, many tasks can be done in parallel, such as combining ingredient incorporation with mixing or conducting pH and viscosity measurements simultaneously using separate vessels. Equipment layout also differs, as in the robotic system all equipment is centralized, while in the manual approach it is often dispersed across different labs, impacting measurement conditions and experiment duration.

D. Bayesian Optimization

Bayesian optimization is used to identify formulations with the lowest viscosity in the analyzed food system, prepared using a predefined experimental process. Bayesian Optimization combines active learning with surrogate model creation to iteratively sample the design space, aiming to minimize the objective function while maximizing the certainty of the surrogate model [33]. The surrogate model is based on Gaussian Processes, which estimate both the mean and uncertainty at any point in the design space. This enables Bayesian optimization to iteratively explore the design space, focusing on areas with high potential reward and a strong likelihood of identifying an optimal solution. The next sampling point is determined using the 'Expected Improvement' acquisition function, which offers a balanced approach between improving the model and identifying high-performing recipes [33].

Bayesian optimization is particularly well suited for problems involving costly experiments like ours, as it can identify effective solutions with relatively few observations by relying on surrogate models rather than requiring an explicit functional form for the objective function [33]. Unlike other optimization methods such as Genetic Algorithms or Simulated Annealing, it also quantifies uncertainty, enabling interpretation of the surrogate model in the context of its accuracy [34]. Additionally, Bayesian Optimization is highly sample-efficient for expensive evaluations [35], requiring only a minimal initial dataset and effectively balancing exploration and exploitation, making it especially appropriate for our application. Bayesian

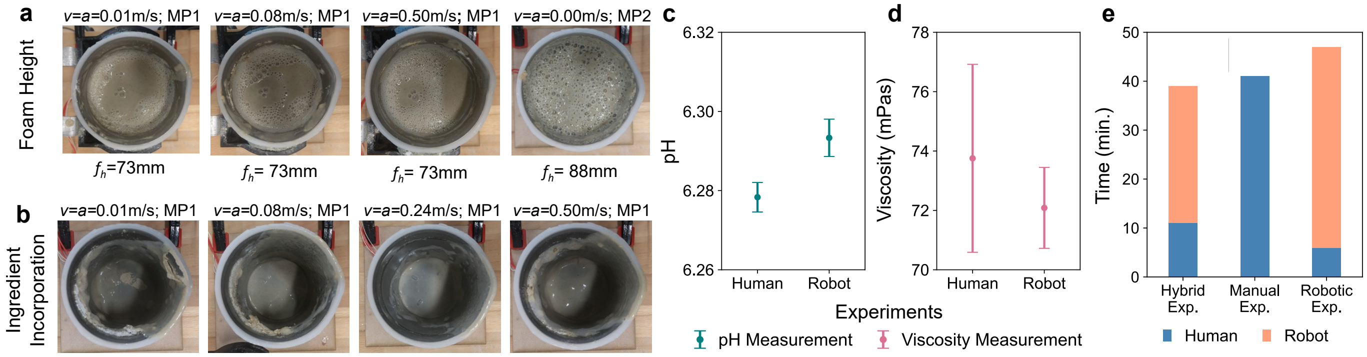


Fig. 4. (a) The minimization of foam height (f_h) by varying robot arm speeds; MP1 (Mixing Pattern 1) – the cup is continuously moved clockwise by a robot arm around the static mixer head; MP2 (Mixing Pattern 2) – the cup is moved in a clockwise direction between one of the four points every 3 minutes. (b) The improvement in ingredient incorporation achieved by varying robot arm speeds. (c) The comparison of robot and human performance in pH measurement (d) and viscosity measurement tasks. (e) Time required to perform an experiment by a robot-only, human-only and hybrid human-robot system.

Optimization was implemented using the MATLAB Statistics and Machine Learning Toolbox. The algorithm takes variables and their bounds as an input, along with data from seed experiments which were in our case defined using a random search. Given the large design space and time required to perform the experiment, exploration-exploitation ratio was set to 0.25 favoring early exploitation. The objective function, as shown in the equation below, aims to minimize viscosity as a function of formulation parameters $P_1, P_2, \dots, P_n$:

$$\begin{aligned} &\text{minimize Viscosity}(P_1, P_2, \dots, P_n) \\ &lb_i \le P_i \le ub_i,\ i = 1, \dots, n \end{aligned} \tag{2}$$

where lb_i and ub_i represent lower and upper bounds for the formulation parameters P_i, with n being the maximum number of such parameters. Stopping criteria is outlined in the equation below:

$$\frac{\min\limits_{i-5\to i} \text{Viscosity}_i(P)}{\min\limits_{1\to i} \text{Viscosity}_i(P)} \ge 0.95 \quad \text{or} \quad i \le 15 \tag{3}$$

where $\min\limits_{1\to i} \text{Viscosity}_i(P)$ represents the lowest overall measured viscosity and $\min\limits_{i-5\to i} \text{Viscosity}_i(P)$ represents the lowest viscosity in the last five experiments. Optimization stops when the viscosity improvement over the last five experiments is below 5% of the global minimum, or when a total of 15 experiments have been completed.

E. Gaussian Process Surrogate Model

The surrogate model built as a part of the Bayesian Optimization process can be utilized to make predictions for untested formulations [36]. By querying the model, an estimation of the objective function (i.e. the viscosity) and the corresponding uncertainty can be obtained. The accuracy and certainty of the model relies on the distribution and number of data points used for model construction.

F. Multi-Objective Pareto Front Optimization

The Bayesian Optimization and food system surrogate model offer a good starting point for minimizing viscosity and understanding the design space. However, they are limited in their ability to optimize multiple competing objectives. Multi-objective optimization addresses this challenge by balancing competing objectives and focusing on the set of most optimal solutions, a concept known as the Pareto Front [37]. To construct the Pareto Front for expensive black-box functions such as viscosity optimization, surrogate-based multi-objective optimization algorithms are often used [38], [39]. These algorithms fit cost-effective surrogate models to a limited number of samples of the expensive objective [40], achieving high sample efficiency without requiring the objective to have an explicit functional form.

For viscosity optimization, the Thompson sampling efficient multi-objective optimization (TSEMO) algorithm has been applied [41]. It uses seed data to train the Gaussian Process models for each objective. The resulting objective function shown in the equation below, aims to minimize both viscosity and P_c, the parameter with highest cost impact:

$$\begin{aligned} &\text{minimize Viscosity}(P), P_c \\ &P = (P_1, P_2, \dots, P_n), P_c \in P \\ &lb_i \le P_i \le ub_i,\ i = 1, \dots, n \end{aligned} \tag{4}$$

where $P = (P_1, P_2, \dots, P_n)$ represents the set of formulation parameters, where lb_i and ub_i denote the lower and upper bounds for each parameter P_i. The value n represents the total number of formulation parameters.

These Gaussian Process models are sampled using spectral sampling to construct individual functions, which are used in conjunction with the NSGA-II algorithm to construct the Pareto Front. Additionally, Thompson Sampling is used to identify the next sampling point, using hypervolume criterion to assess both convergence and diversity metrics [42]. The proposed sampling point is selected based on the highest hypervolume value. By providing the viscosity at the proposed location in the design space, the Pareto Front is updated, and a new sampling point is suggested using Thompson Sampling and hypervolume criteria. This iterative process continues until either no further improvement is observed or experimentation capacity is reached. The viscosity of each formulation proposed by TSEMO can be determined either experimentally or by using the surrogate model constructed during the Bayesian Optimization. Constructed Pareto Fronts can be compared by calculating the area under the front using Trapezoidal rule [43].

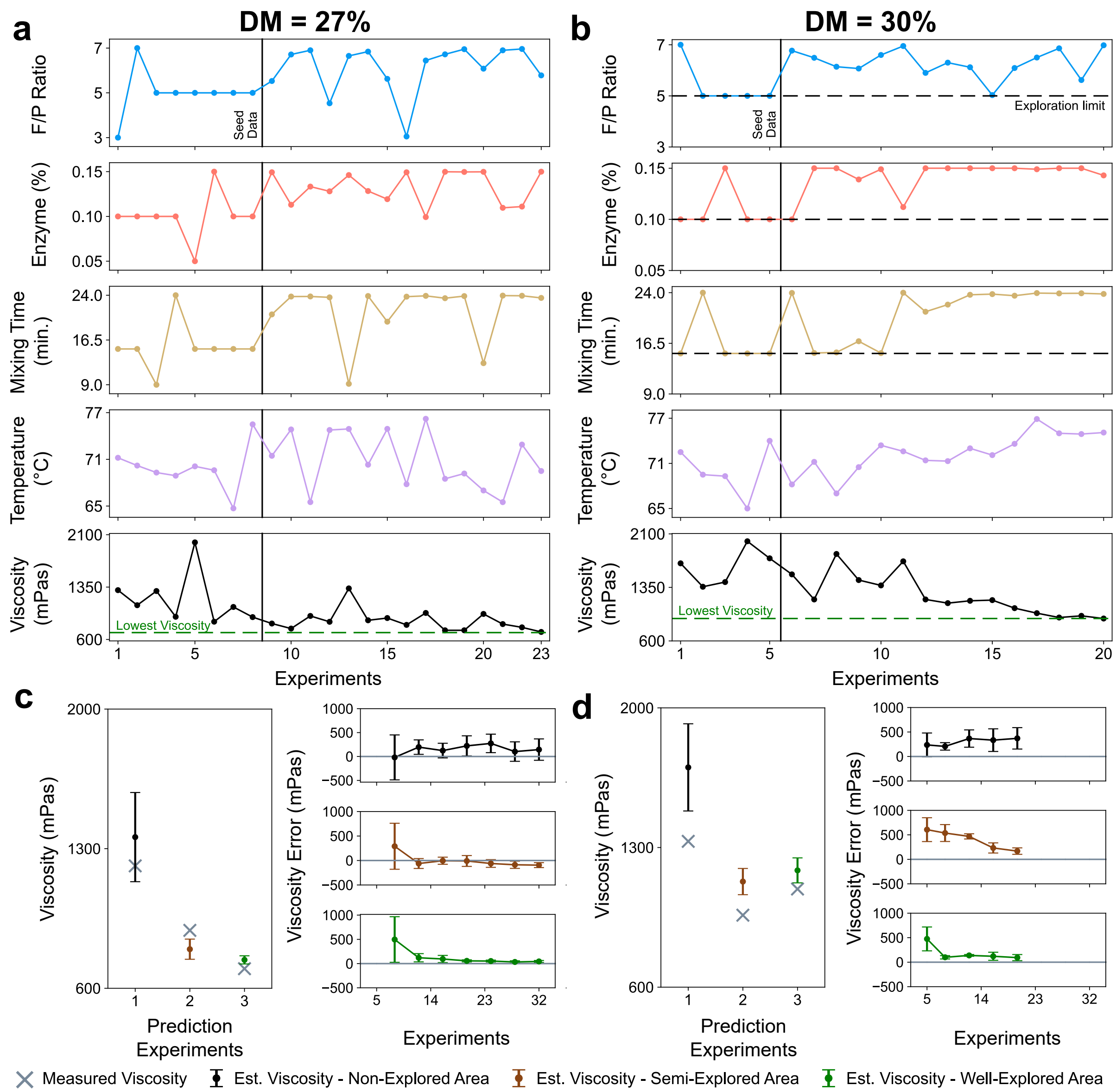


Fig. 5. (a) The minimization of viscosity using the Bayesian optimization algorithm by varying 4 input parameters for the dry matter DM=27% and (b) DM=30% food systems. (c) The estimation of viscosity and its uncertainty using the surrogate model based on Gaussian Process models, including a comparison of estimated and measured viscosities and the evaluation of error based on the amount of training data for the DM=27% and (d) DM=30%.

III. EXPERIMENTAL RESULTS

A. Robotic Automation

Automating the formulation and assessment of the model food system enables the exploration of how mixing motion affects the preparation process and supports identification of optimal motion patterns and speeds. This is particularly important, as inadequate mixing motion or poor ingredient incorporation can result in foaming and incomplete mixing.

To investigate the effect of robot motion on foaming and ingredient incorporation during the mixing phase, various speeds and motion patterns (Supplementary Figure 1) were tested on identical recipes. When examining foam height (Figure 4 a), significant foaming was observed in the absence of cup motion during mixing. However, as the arm motion speed increased (0.01, 0.08 and 0.50 m/s), foaming was significantly reduced and remained approximately constant. The impact of robot speed and motion on ingredient incorporation was also evaluated. Poor ingredient incorporation was observed at robot speeds of 0.01, 0.08 and 0.50 m/s, while the best incorporation occurred at 0.24 m/s, as shown on Figure 4 b). At low speeds, powder ingredients lacked sufficient velocity for effective incorporation, while at the highest speeds, incorporation time became the limiting factor. These findings highlight the sensitivity of the food system to small process variations and illustrate why robotic control is highly advantageous.

To compare measurement precision between humans and robots, pH and viscosity were measured in a stable beverage prepared from a water-powdered milk mix at room temperature. For five repeated measurements, humans and robotic system exhibited identical precision in pH measurements as shown on Figure 4 c). However, the robotic system demonstrated higher precision in viscosity measurements with a 1.9% deviation compared to 4.3% for a human while showing identical accuracy, as shown on Figure 4 d).

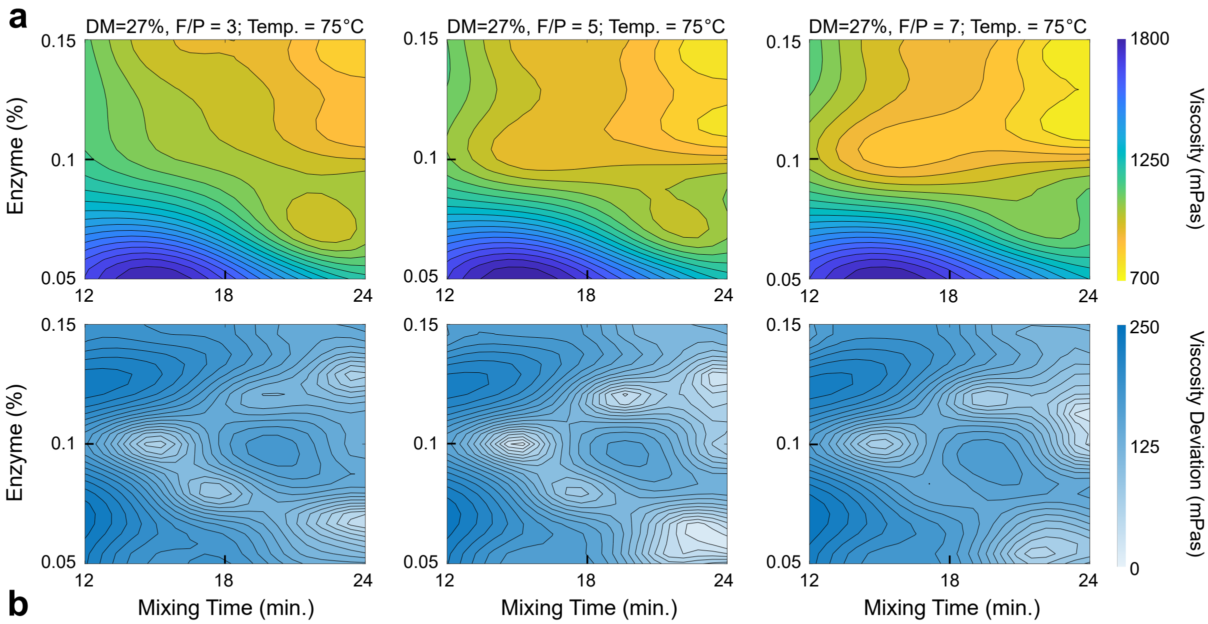


Fig. 6. (a) The model of the analyzed plant-based food system over the discretized design space, built using Gaussian Process models and trained on all the available experimental data for DM=27%. (b) The corresponding standard deviation model for DM=27%.

Considering efficiency, a comparison of human and robotic approaches, as seen in Table II shows that, although the fully robotic approach is 15% slower than a fully manual one, it reduces human intervention by over 85% and saves an average of 30 minutes of researcher time per experiment. Limited human involvement remains necessary, primarily for cleaning tasks that are difficult to automate. By leveraging both humans and robots, and manually handling the slow, multi-step ingredient incorporation process, the hybrid approach achieves the highest throughput, completing up to 12 experiments within an 8-hour day, with much of the process running unattended.

B. Single Objective Optimization: Minimizing Viscosity for Processability

Minimizing the viscosity of liquid food systems is essential for enhancing processability, taste and mouthfeel. To achieve this, we use single-objective Bayesian Optimization to identify the optimal formulation, parametrized by four factors: flour-to-protein ratio (F/P ratio), enzyme content, mixing time and processing temperature. The analyzed model food system was optimized at two dry matter (DM) concentrations, 27% and 30%. This approach was used to demonstrate broad applicability of Bayesian optimization across a range of viscosities [44], [45].

An initial set of seed experiments was conducted, followed by optimization until convergence. Upper and lower bounds (Table I) were defined for each optimizable parameter to ensure they remained within the robot's operational limits. These bounds were intentionally kept broad to minimize bias and enable exploration of unconventional or unexpected formulations. The results of the Bayesian Optimization process are shown in Figure 5, highlighting the evolution of viscosity and recipe parameters for two different DM concentrations.

For the DM=27% food system (Figure 5 a), viscosity gradually decreases with each iteration, reaching a point of minimal further improvement around the 20th experiment. All parameters except temperature begin to converge by the 18th experiment, resulting in one of the lowest measured viscosities. Exploration focuses primarily on the upper regions of the F/P ratio, enzyme and mixing time, while temperature is more broadly explored. This suggests that temperature has the least impact on viscosity, whereas the other parameters play a more significant role. Building on the results from the DM=27% food system, which showed that the upper range of each process parameter yielded lower viscosity, the design space for the DM=30% food system was restricted to the upper half of the range for all parameters except temperature (Figure 5 b). For this food system, viscosity begins to decrease decisively from the 12th experiment onwards, with subsequent iterations achieving progressively lower viscosity and plateauing towards 20th experiment. The enzyme and mixing time converge rapidly by the 12th experiment, while the F/P ratio continues to be explored, and temperature remains in the higher range.

Metrics	Manual Exp.	Hybrid Exp.	Robotic Exp.
Total Experiment Duration (min.)	41	39	47
Automated Tasks	1/6	4/6	5/6
Manual Interventions	50+	15-20	5-10
Researcher Hands-On Time	100% (41 min.)	30% (11 min.)	13% (6 min.)
Throughput (exp./8h)	11.7	12.3	10.2

TABLE II
COMPARISON OF EXPERIMENTAL EFFICIENCY METRICS ACROSS MANUAL, HYBRID AND ROBOTIC APPROACHES

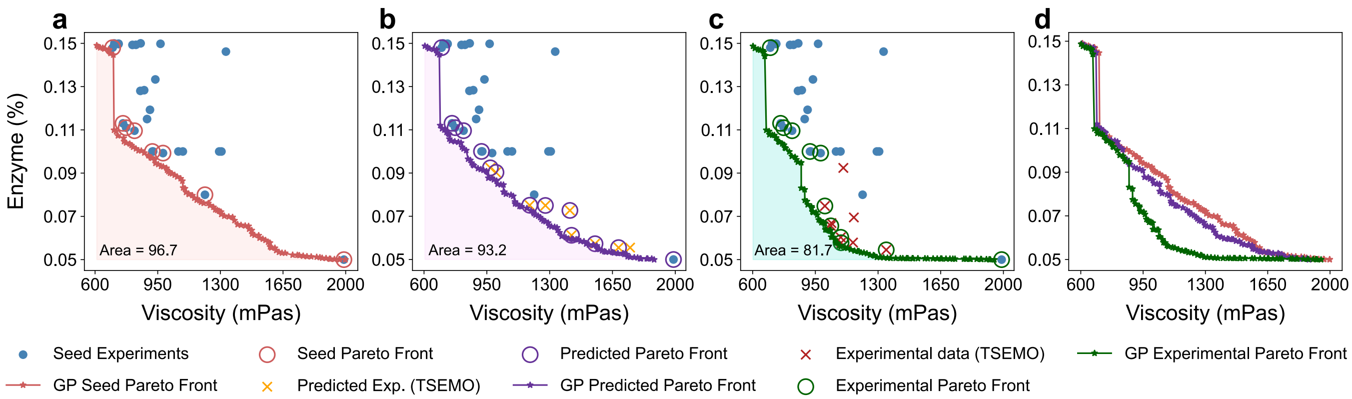


Fig. 7. (a) Pareto Front constructed by the TSEMO algorithm using seed experimental data from all prior experiments for the DM=27% food system. (b) Pareto Front constructed with 9 additional experiments proposed by the TSEMO algorithm, where viscosity is provided by the surrogate model built during the single-objective optimization process. (c) Pareto Front constructed with viscosity data provided experimentally. (d) Comparison of Pareto Fronts constructed using the TSEMO algorithm.

The lowest viscosity formulation, identified as F1 in Table III, was achieved at the maximum values for all parameters except temperature for the DM=27% food system. However, for the DM=30% food system, the lowest viscosity was reached at the maximum values for all parameters, suggesting the need for additional exploration in the high temperature range for the DM=27% food system. The viscosity observed for the DM=30% is higher than that of the DM=27% food system. For both systems, enzyme content and mixing time have the most significant impact on reducing viscosity, as demonstrated by their rapid convergence. The F/P ratio also has a measurable impact on viscosity, with recipes containing more flour and less protein (high F/P) exhibiting lower viscosity. Furthermore, the interaction between temperature and enzyme was unexpected, as no slowdown in enzymatic activity was observed with increase in temperature.

C. Model-Based Predictions of Viscosity

The data from Bayesian Optimization experiments is used to create a surrogate model based on Gaussian Processes. This model enables the estimation of viscosity for untested recipes, offering a valuable tool for food scientists. The surrogate model for the DM=27% food system, shown in Figure 5 c), is constructed using Bayesian optimization data (23 data points) and Pareto Front data (9 data points) from a separate experiment described later in the paper. In contrast, the surrogate model for DM=30% food system, shown in Figure 5 d), is created solely from Bayesian Optimization data (20 data points). To evaluate the surrogate model, three test points were selected, representing untested parameter combinations from well-explored, semi-explored and unexplored regions of the design space. Ground truth data for model evaluation was obtained by conducting experiments at these exact parameter combinations. The accuracy of these models is assessed based on the amount of training data used for their construction.

The surrogate model for the DM=27% food system, built from a larger dataset, is more accurate, with measured viscosity largely falling within the model's uncertainty. Both error and model uncertainty are significantly lower in well-explored areas, with an absolute error as low as 6%, but increasing to 12% in unexplored regions. The error analysis indicates that the first 20 data points contribute the most to reducing model error. For the DM=30% food system, the model is constructed from fewer data points, resulting in a higher overall error. In well explored areas, the error is as low as 9% and remains close to the model's uncertainty, whereas in unexplored regions, it increases to 28%. Importantly, the model demonstrates that the error continues to decrease with additional data, suggesting that acquiring more data would further enhance accuracy. In both cases, data gathered from Bayesian Optimization enables viscosity predictions with higher accuracy in low-viscosity regions while also supporting meaningful analysis in areas with higher model uncertainty.

D. Visualization of the Surrogate model

Visualization of the surrogate model provides a more comprehensive understanding of the design space. To visualize the analyzed food system, viscosity and standard deviation have been estimated from the surrogate model across a discretized design space of 25 points for the range of each parameter

Exp. Description	F/P Ratio	Enzyme (%)	Mix. Time (min.)	Temp °C	Viscosity (mPas)	Exp. Source
F1 - Lowest viscosity	5.78	0.150	23.55	69.5	707.7	Bayes. Opt.
F2 - Best tradeoff	6.96	0.111	23.91	72.9	773.3	Bayes. Opt.
F3 - High protein	3.05	0.150	23.76	67.8	808.7	Bayes. Opt.
F4 - Lowest processing cost	6.08	0.150	12.62	67.0	965.4	Bayes. Opt.
F5 - Lowest overall cost	5.78	0.055	21.84	74.9	1345.7	Pareto Front
F6 - Median viscosity	6.44	0.099	23.9	76.2	979.5	Bayes. Opt.

TABLE III
THE OPTIMAL FORMULATIONS OF THE ANALYZED DM=27% FOOD SYSTEM

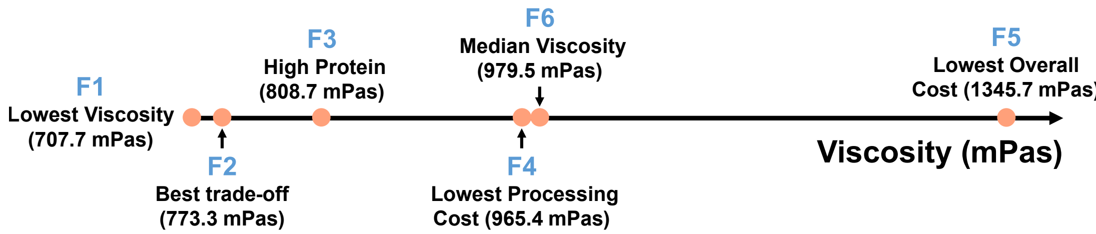


Fig. 8. Viscosity visualization of optimal formulations shown in Table III, derived from over 30 experiments conducted to optimize the DM=27% food system.

(Figure 6). By visualizing the design space for the DM=27% food system as a function of enzyme content and mixing time, the two most dominant parameters, it is observed that viscosity decreases with an increase in F/P ratio. The lowest viscosity is achieved in regions of highest enzyme and mixing time, while the highest viscosity is observed on the opposite side of the design space. Additionally, the food system model reveals the relationship between mixing time and enzyme content, indicating that the area of highest enzyme content and shortest mixing time has lower viscosity than the opposite region of the design space. This suggests that the enzyme content is the most dominant parameter in the experiment. The certainty of the developed models is illustrated by standard deviation plot (Figure 6 b) which indicates the highest model accuracy in regions with highest enzyme and mixing time.

E. Pareto-Optimal Solutions

While single-objective optimization offers valuable insights into minimizing viscosity, practical food system design must consider multiple, often competing objectives. In this context, enzyme content serves as an important control parameter, as minimizing enzyme levels has the greatest impact on overall process costs in our model food system. Multi-objective optimization enables balancing of competing goals, minimizing viscosity and enzyme concentration by identifying Pareto-optimal solutions and constructing the Pareto Front, a frontier beyond which no further improvements are possible [37].

Starting from the existing data for the DM=27% food system (26 data points), a Pareto Front is constructed as seen in Figure 7 a), however, it lacks experimental data beyond a viscosity of 950 mPas. To address this, the TSEMO algorithm is applied, employing Thompson sampling to identify the next sampling point by balancing maximizing information gain and expanding the feasible region of the Pareto Front [41]. The TSEMO Pareto Front, driven by a surrogate model from single-objective optimization (Figure 7 b), reduces the space of unfeasible solutions by only 3.6% compared to the Pareto Front constructed from the seed data (Figure 7 a). In contrast, experimentally driven TSEMO Pareto Front (Figure 7 c) achieves a 15.5% reduction, highlighting the limitations of surrogate model due to insufficient training data. Comparing all three Pareto Fronts (Figure 7 d) reveals significant improvements in discovered recipe formulations within the region sampled by TSEMO, suggesting further improvements with an increased number of experiments. Focusing on the experimental Pareto Front (Figure 7 c), a reduction in enzyme content from 0.15% to 0.11% corresponds to a 25% decrease in enzyme levels with only an 10% increase in viscosity. Furthermore, reducing the enzyme level to 0.066% leads to a 48% increase in the viscosity. However, decreasing enzyme content below 0.06%, causes a rapid rise in viscosity, signifying a region to be avoided.

F. Discovered Food Formulations

Leveraging over 30 experiments conducted to optimize DM=27% food system, we identified a range of food formulations which seek to address specific objectives, as seen in Table III. Formulation F1 achieves the lowest viscosity but comes with the maximum enzyme content and extended mixing time. In contrast, formulation F2 with a modest 10% increase in viscosity, offers the best trade-off between viscosity and affordability, achieving the lowest cost in terms of flour-to-protein ratio while maintaining moderate enzyme levels. However, processing expenses remain high due to elevated temperature and long mixing time. Formulation F3 prioritizes nutrition, delivering the highest protein content with a 15% higher viscosity along with high ingredient and processing cost. A comparison between formulations F4 (lowest processing cost) and F5 (lowest overall cost), reveals that reducing enzyme content significantly increases viscosity, making this trade-off suboptimal. When compared to a median-viscosity of the baseline formulation (F6), as illustrated on Figure 8, all optimized formulations, except F5, perform better, highlighting the effectiveness of the employed optimization methods. Furthermore, the diversity and performance of the optimal formulations underscore the effectiveness of our combined robotic and recipe optimization approach, demonstrating its high throughput, accuracy, and ability to identify both specific and optimal food formulations.

IV. DISCUSSIONS & CONCLUSIONS

We have introduced the first robotic system for discovery of optimized formulations and processing conditions of plant-based food systems. The proposed robotic system is capable of dosing ingredients, processing the food system through mixing and heating, and measuring viscosity, pH, as well as capturing appearance using a camera. Our observations indicate that distinct motion speeds of the heated cup around the mixer head are necessary, one to enhance ingredient incorporation and the other to minimize the effects of foaming during extensive mixing. By applying the Bayesian Optimization algorithm for viscosity minimization, we demonstrate that the lowest viscosity is achieved at the highest values of input parameters, including F/P ratio, Enzyme content, Mixing Time and Temperature. Furthermore, a surrogate model based on Gaussian Processes and trained on experimental data was used to estimate the viscosity for unknown parameter combinations, achieving a prediction error of 6% in well-explored areas of the design space. Leveraging the capabilities of the surrogate model across the parameterized design space, we visualize the food system model to gain further insights into the relationship between the parameters, identifying enzyme content as the most dominant formulation parameter. We conclude our findings by constructing the Enzyme-Viscosity Pareto Front, demonstrating a set of optimal solutions for the most dominant input parameter. The experimentally driven Pareto Front reduces the area of unfeasible solutions by 15.5% compared to the seed data and surrogate model-driven Pareto Fronts, which exhibit similar performance. From an efficiency perspective, robotic experimentation reduced human effort by 85%, while the hybrid human-robot approach achieved the highest throughput, completing 12 experiments in 8 hours.

By coupling robotic automation with active learning methods, we provide an alternative to the largely manual approach used in food science laboratories. Through the application of robotic automation to food processing, we can achieve higher experiment throughput while improving accuracy and precision through precise parameter control, consistent motions and timing. The ability to rapidly explore novel formulations and processes can enable the discovery of new food systems optimized for various objectives. Our novel approach can now be extended to include additional processing steps, for example, high pressure homogenization, oil addition or dilution [46], [47]. These additional processing steps would enable us to also analyze stability (phase separation) of the food system. This would expand the number of possible ingredients and process parameters, providing a wider design space in which we can discover the most affordable, nutritious and stable formulations. To support this modified experimental process, we would also explore scaling up the existing system using industrial Cartesian automation systems. By benchmarking 6-DoF robotic arm, industrial systems and hybrid approaches, our goal is to identify the optimal automation strategy for specific tasks and overall experimental process.

The integrated robotic and machine learning approach can be extended to other food systems featuring diverse ingredients, processing methods and objectives. However, substantial modifications to the existing hardware, software and data-processing pipelines may be required to accommodate differences in preparation process, ingredient interactions and the physical properties of various food matrices. The current robotic system is well-suited for investigating effects of minor formulation changes, such as introducing new ingredients or adjusting process parameters. Optimization of food systems involving such minor changes could be accelerated using transfer learning methodologies, allowing past experimental data to be reused to identify optimal formulations with a minimal number of additional trials. Similar strategies could help manage variability in ingredient quality across different batches, a challenge inherent to agriculture and food production that can lead up to a 30% difference in measured viscosity. Enhancing predictive models with a few targeted experiments could help account for batch-to-batch variability, ensuring consistent food system quality with minimal formulation adjustments.

ACKNOWLEDGMENT

This project was supported by the EU's Horizon 2020 research and innovation program under the Marie Skłodowska Curie grant agreement No 945363.

REFERENCES

[1] R. Capone, H. E. Bilali, P. Debs, G. Cardone, and N. Driouech, "Food system sustainability and food security: connecting the dots," *Journal of Food Security*, vol. 2, no. 1, pp. 13–22, 2014.

[2] U. A. Schneider, P. Havlík, E. Schmid, H. Valin, A. Mosnier, M. Obersteiner, H. Böttcher, R. Skalskỳ, J. Balkovič, T. Sauer *et al.*, "Impacts of population growth, economic development, and technical change on global food production and consumption," *Agricultural Systems*, vol. 104, no. 2, pp. 204–215, 2011.

[3] L. Baroni, L. Cenci, M. Tettamanti, and M. Berati, "Evaluating the environmental impact of various dietary patterns combined with different food production systems," *European journal of clinical nutrition*, vol. 61, no. 2, pp. 279–286, 2007.

[4] J. Poore and T. Nemecek, "Reducing food's environmental impacts through producers and consumers," *Science*, vol. 360, no. 6392, pp. 987–992, 2018.

[5] B. C. Chai, J. R. Van Der Voort, K. Grofelnik, H. G. Eliasdottir, I. Klöss, and F. J. Perez-Cueto, "Which diet has the least environmental impact on our planet? a systematic review of vegan, vegetarian and omnivorous diets," *Sustainability*, vol. 11, no. 15, p. 4110, 2019.

[6] J. Sabaté and S. Soret, "Sustainability of plant-based diets: back to the future," *The American journal of clinical nutrition*, vol. 100, pp. 476S–482S, 2014.

[7] E. Stehfest, L. Bouwman, D. P. Van Vuuren, M. G. Den Elzen, B. Eickhout, and P. Kabat, "Climate benefits of changing diet," *Climatic change*, vol. 95, no. 1, pp. 83–102, 2009.

[8] F. B. Hu, "Plant-based foods and prevention of cardiovascular disease: an overview," *The American journal of clinical nutrition*, vol. 78, no. 3, pp. 544S–551S, 2003.

[9] P. J. Tuso, M. H. Ismail, B. P. Ha, and C. Bartolotto, "Nutritional update for physicians: plant-based diets," *The Permanente Journal*, vol. 17, no. 2, p. 61, 2013.

[10] A. V. Cardello, F. Llobell, D. Giacalone, C. M. Roigard, and S. R. Jaeger, "Plant-based alternatives vs dairy milk: Consumer segments and their sensory, emotional, cognitive and situational use responses to tasted products," *Food Quality and Preference*, vol. 100, p. 104599, 2022.

[11] V. Caputo, G. Sogari, and E. J. Van Loo, "Do plant-based and blend meat alternatives taste like meat? a combined sensory and choice experiment study," *Applied Economic Perspectives and Policy*, vol. 45, no. 1, pp. 86–105, 2023.

[12] G. Lippi and G. Da Rin, "Advantages and limitations of total laboratory automation: a personal overview," *Clinical Chemistry and Laboratory Medicine (CCLM)*, vol. 57, no. 6, pp. 802–811, 2019.

[13] J. Iqbal, Z. H. Khan, and A. Khalid, "Prospects of robotics in food industry," *Food Science and Technology*, vol. 37, pp. 159–165, 2017.

[14] W. H. Wang, X. Y. Liu, and Y. Sun, "High-throughput automated injection of individual biological cells," *IEEE Transactions on automation science and engineering*, vol. 6, no. 2, pp. 209–219, 2008.

[15] M. Dörr, M. P. Fibinger, D. Last, S. Schmidt, J. Santos-Aberturas, D. Böttcher, A. Hummel, C. Vickers, M. Voss, and U. T. Bornscheuer, "Fully automatized high-throughput enzyme library screening using a robotic platform," *Biotechnology and Bioengineering*, vol. 113, no. 7, pp. 1421–1432, 2016.

[16] M. Hornsby, M. Paduch, S. Miersch, A. Sääf, T. Matsuguchi, B. Lee, K. Wypisniak, A. Doak, D. King, S. Usatyuk *et al.*, "A high through-put platform for recombinant antibodies to folded proteins," *Molecular & Cellular Proteomics*, vol. 14, no. 10, pp. 2833–2847, 2015.

[17] S. Asche, G. J. Cooper, G. Keenan, C. Mathis, and L. Cronin, "A robotic prebiotic chemist probes long term reactions of complexifying mixtures," *Nature Communications*, vol. 12, no. 1, p. 3547, 2021.

[18] A.-C. Bédard, A. Adamo, K. C. Aroh, M. G. Russell, A. A. Bedermann, J. Torosian, B. Yue, K. F. Jensen, and T. F. Jamison, "Reconfigurable system for automated optimization of diverse chemical reactions," *Science*, vol. 361, no. 6408, pp. 1220–1225, 2018.

[19] N. Obayashi, K. Junge, S. Ilić, and J. Hughes, "Robotic automation and unsupervised cluster assisted modeling for solving the forward and reverse design problem of paper airplanes," *Scientific Reports*, vol. 13, no. 1, p. 4212, 2023.

[20] T. Czimmermann, M. Chiurazzi, M. Milazzo, S. Roccella, M. Barbieri, P. Dario, C. M. Oddo, and G. Ciuti, "An autonomous robotic platform for manipulation and inspection of metallic surfaces in industry 4.0," *IEEE Transactions on Automation Science and Engineering*, vol. 19, no. 3, pp. 1691–1706, 2021.

[21] H. Zhao, W. Chen, H. Huang, Z. Sun, Z. Chen, L. Wu, B. Zhang, F. Lai, Z. Wang, M. L. Adam *et al.*, "A robotic platform for the synthesis of colloidal nanocrystals," *Nature Synthesis*, vol. 2, no. 6, pp. 505–514, 2023.

[22] S. Steiner, J. Wolf, S. Glatzel, A. Andreou, J. M. Granda, G. Keenan, T. Hinkley, G. Aragon-Camarasa, P. J. Kitson, D. Angelone *et al.*, "Organic synthesis in a modular robotic system driven by a chemical programming language," *Science*, vol. 363, no. 6423, p. eaav2211, 2019.

[23] M. Beetz, U. Klank, I. Kresse, A. Maldonado, L. Mösenlechner, D. Pangercic, T. Rühr, and M. Tenorth, "Robotic roommates making pancakes," in *2011 11th IEEE-RAS International Conference on Humanoid Robots*. IEEE, 2011, pp. 529–536.

[24] M. Abolhasani and E. Kumacheva, "The rise of self-driving labs in chemical and materials sciences," *Nature Synthesis*, vol. 2, no. 6, pp. 483–492, 2023.

[25] G. Schneider, "Automating drug discovery," *Nature reviews drug discovery*, vol. 17, no. 2, pp. 97–113, 2018.

[26] B. Burger, P. M. Maffettone, V. V. Gusev, C. M. Aitchison, Y. Bai, X. Wang, X. Li, B. M. Alston, B. Li, R. Clowes *et al.*, "A mobile robotic chemist," *Nature*, vol. 583, no. 7815, pp. 237–241, 2020.

[27] J. M. Granda, L. Donina, V. Dragone, D.-L. Long, and L. Cronin, "Controlling an organic synthesis robot with machine learning to search for new reactivity," *Nature*, vol. 559, no. 7714, pp. 377–381, 2018.

[28] K. Junge, J. Hughes, T. G. Thuruthel, and F. Iida, "Improving robotic cooking using batch bayesian optimization," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 760–765, 2020.

[29] C. G. Viejo, S. Fuentes, G. Li, R. Collmann, B. Condé, and D. Torrico, "Development of a robotic pourer constructed with ubiquitous materials, open hardware and sensors to assess beer foam quality using computer vision and pattern recognition algorithms: Robobeer," *Food Research International*, vol. 89, pp. 504–513, 2016.

[30] B. C. Condé, S. Fuentes, M. Caron, D. Xiao, R. Collmann, and K. S. Howell, "Development of a robotic and computer vision method to assess foam quality in sparkling wines," *Food Control*, vol. 71, pp. 383–392, 2017.

[31] T. Palmer and P. L. R. Bonner, *Enzymes: biochemistry, biotechnology and clinical chemistry, Second edition*. Cazarin, Cinthia Baú Betim, 2011.

[32] *Brookfield DV-II+ Pro Programmable Viscometer, Operating Instructions.*

[33] E. Brochu, V. M. Cora, and N. De Freitas, "A tutorial on bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning," *arXiv preprint arXiv:1012.2599*, 2010.

[34] P. Papis, C. Siettos, T. Dasaklis *et al.*, "Fuzzy sets, systems and applications," in *Encyclopedia of Operations Research and Management Science, Third edition*, 2013.

[35] F. Hutter, L. Kotthoff, and J. Vanschoren, *Automated machine learning: methods, systems, challenges*. Springer Nature, 2019.

[36] E. Schulz, M. Speekenbrink, and A. Krause, "A tutorial on gaussian process regression: Modelling, exploring, and exploiting functions," *Journal of Mathematical Psychology*, vol. 85, pp. 1–16, 2018.

[37] M. T. Emmerich and A. H. Deutz, "A tutorial on multiobjective optimization: fundamentals and evolutionary methods," *Natural computing*, vol. 17, pp. 585–609, 2018.

[38] J. Knowles, "Parego: A hybrid algorithm with on-line landscape approximation for expensive multiobjective optimization problems," *IEEE transactions on evolutionary computation*, vol. 10, no. 1, pp. 50–66, 2006.

[39] A. M. Schweidtmann, A. D. Clayton, N. Holmes, E. Bradford, R. A. Bourne, and A. A. Lapkin, "Machine learning meets continuous flow chemistry: Automated optimization towards the pareto front of multiple objectives," *Chemical Engineering Journal*, vol. 352, pp. 277–282, 2018.

[40] D. Becker, C. Schmitt, L. Bovetto, C. Rauh, C. McHardy, and C. Hartmann, "Optimization of complex food formulations using robotics and active learning," *Innovative Food Science & Emerging Technologies*, vol. 83, p. 103232, 2023.

[41] E. Bradford, A. M. Schweidtmann, and A. Lapkin, "Efficient multiobjective optimization employing gaussian processes, spectral sampling and a genetic algorithm," *Journal of global optimization*, vol. 71, no. 2, pp. 407–438, 2018.

[42] Y. Cao, B. J. Smucker, and T. J. Robinson, "On using the hypervolume indicator to compare pareto fronts: Applications to multi-criteria optimal experimental design," *Journal of Statistical Planning and Inference*, vol. 160, pp. 60–74, 2015.

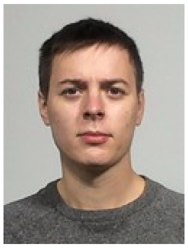
[43] B. Fornberg, "Improving the accuracy of the trapezoidal rule," *SIAM Review*, vol. 63, no. 1, pp. 167–180, 2021.

[44] S. Nielsen, "Food analysis (food science texts series)," *Springer Science+ Business Media, LLC*, 2010.

[45] W. Routray and H. N. Mishra, "Scientific and technical aspects of yogurt aroma and taste: a review," *Comprehensive Reviews in Food Science and Food Safety*, vol. 10, no. 4, pp. 208–220, 2011.

[46] O. E. Mäkinen, V. Wanhalinna, E. Zannini, and E. K. Arendt, "Foods for special dietary needs: Non-dairy plant-based milk substitutes and fermented dairy-type products," *Critical reviews in food science and nutrition*, vol. 56, no. 3, pp. 339–349, 2016.

[47] D. J. McClements, "Development of next-generation nutritionally fortified plant-based milk substitutes: Structural design principles," *Foods*, vol. 9, no. 4, p. 421, 2020.



Stefan Ilic Stefan Ilic received his M.Sc. degree in Mechanical Engineering from the University of Kragujevac, Serbia. He is currently pursuing a Ph.D. in Robotics at the CREATE Lab, EPFL. His research focuses on robotic laboratory automation and formulation optimization, particularly in food science and the development of novel food products. Previously, he worked for five years in various roles across the automotive, medical, and CAD consulting industries.

Edgar Chávez Montes Edgar Chávez Montes completed his undergraduate studies in Food Science at the National Autonomous University of Mexico (UNAM) and subsequently earned a PhD in Food Process Engineering from Nancy University in France in 2002. Following several years of research on food powder processing at Unilever R&D in the United Kingdom and the Netherlands, he has been leading research projects at Nestlé R&D in Switzerland for the past 17 years in the fields of food powder science, technology, and analytics. His work also explores the integration of digital technologies and robotics within these areas to facilitate the development of new food products.

Constantijn Sanders Constantijn Sanders completed undergraduate studies in Chemistry and Chemical Engineering at the University of Groningen. His PhD project about pharmaceutical powder processes was at The University of Sheffield with Prof. Mike Hounslow, followed by a post-doc at UC Santa Barbara in the process control team with Prof. Frank Doyle. Since 2012 he had various roles at Nestle R&D in food science and lab automation.

Cécile Gehin-Delval Cécile Gehin-Delval holds both an M.Sc. degree and a PhD in Physical Chemistry from the University of Franche-Comté, France. Following a two-year tenure as an assistant professor at the University of Geneva, Dr. Gehin-Delval joined the Nestlé Research Center in Lausanne, Switzerland. With 20 years of expertise in colloid science, she has specialized in emulsions, foams, and delivery systems. Currently, she is engaged in developing digital technologies and robotic systems to enhance coffee beverage formulations and expedite innovation.

Christoph Hartmann Christoph Hartmann holds a diploma in mechanical engineering and a PhD in Computational Fluid Dynamics. His habilitation is in the area of biomechanics. Today, he leads the Research Alliances and Expertise Development function for Nestlé R&D. Since 2020, he acts as Honorary Professor for Food Process Engineering at Technische Universität Berlin. In this role, Christoph is active as lecturer and scientist and pursues own research projects in the area of food materials and digitalization. Since March 1, 2025, Christoph works as Professor of Practice at Ecole Polytechnique Federale de Lausanne (EPFL) on a part-time basis.

Giulia Marchesini Giulia Marchesini holds an MSc Degree in Food Science and Technology and has completed a PhD in Food Science at the University of Udine, Italy. She began her career with the R&D team at Chiquita before joining Nestlé R&D in Switzerland in 2012 as a scientist. She has been working in the Protein Science field for several years and subsequently transitioned to various leadership roles within R&D. In recent years, alongside her teams, she has been exploring the potential of using robotics and data analytics to accelerate discoveries and innovation in the food industry.

Josie Hughes Josie Hughes is an Assistant Professor at EPFL where she established the CREATE in 2021. She undertook her undergraduate, masters and PhD studies at the University of Cambridge, joining the Bio-inspired Robotics Lab (BIRL). Following this, she worked as a postdoctoral associate at the CSAIL, MIT in the Distributed Robotics Lab. Her research focuses on developing novel design paradigms for designing robot structures that exploit their physicality and interactions with the environment. This includes the development of robotic hands, soft manipulators and locomoting robots, and automation systems for applications focused on sustainability and science.