Sustainable Food Technology

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal | View Issue

Cite this: *Sustainable Food Technol.*, 2026, **4**, 947

A comparative study of theoretical model-based optimization and experimental design approaches for functional beverage formulation

Regina Navarrete-González, Aurelio López-Malo, Enrique Palou and Nelly Ramírez-Corona *

The demand for functional beverages offering specific health benefits beyond hydration has grown, while advancements in computer-aided food formulation introduce new approaches to reduce innovation costs in the food industry. Two case studies, a juice mixture and a plant-based beverage, were optimized using two approaches: one based on theoretical models (TMOs) and the other on experimental design (DoE). In both approaches, the objective function focused on maximizing the target property; in the first case study, it maximizes antioxidant content, whereas in the second case study, it maximizes protein content. The optimal juice formula based on the TMO approximation consisted of 14% apple, 44% grape, and 42% cranberry, whereas the DoE approach's formula comprised 28.5% apple, 32.2% grape, and 39.3% cranberry. Validation showed that TMO had a lower error rate of 2.0% in phenolic content compared to 13.7% from DoE. For the plant-based beverage, TMO estimated 74% rice, 16% peas, and 10% almonds, compared to DoE's 60%, 28%, and 12%, respectively. Total protein estimation errors were 14.5% for DoE and 4.2% for TMO. Overall, water activity estimation was most accurate for both cases (0.6% and 0.1%, respectively). Larger errors were observed in estimations of pH and acidity (20–24%) for the juices and in viscosity (22%) for the milk analogue. Sensory tests found no significant difference ($p > 0.05$) in consumer acceptance between the two approaches, with mean scores of 7.5 ± 1.2 (mix design) and 7.7 ± 1.9 (theoretical) for juices. Similarly, for the milk analogue, the values were 6.2 ± 2.5 (mix design) and 6.3 ± 2.4 (theoretical). Thus, although theoretical estimation has limitations in accuracy, it can produce acceptable, cost-effective formulations that consumers accept, saving time and resources.

Received 15th August 2025
Accepted 3rd November 2025

DOI: 10.1039/d5fb00480b

rsc.li/susfoodtech

Sustainability spotlight

Advances in computer-aided food formulation are introducing new methods to cut innovation costs in the food industry. As computer-aided food processing improves, the food industry is getting closer to technologies that allow for optimizing formulations using theoretical methods, paired with experimental designs that can greatly reduce innovation costs in the food industry. This proposed approach could enhance product formulation, enabling the optimization of ingredient selection and processing conditions, minimizing waste and resource use, which translates into a more efficient, cost-effective, and sustainable food design.

1. Introduction

The increase of the global population necessitates food design centered on optimizing available resources, enhancing process efficiency, reducing costs, and minimizing design times and energy consumption, all while maintaining nutritional and sensory quality. Nowadays, there are an infinite number of possible combinations between ingredients, concentrations, and process unit operations, which suggests that developing quality-oriented food formulation is an opportunity area for responding to the needs demanded by consumers.[1,2]

Moreover, consumers are becoming increasingly aware of the significant role that foods and beverages play in their diet in terms of health, as they aid in preventing or slowing degenerative diseases caused by oxidative stress or protein deficiency. A functional beverage is a drink that provides a specific health benefit beyond hydration and contains relevant ingredients (such as vitamins, minerals, probiotics or antioxidants) designed to support or enhance a particular function related to human wellness. Commercial functional beverages include sports and energy drinks, vitamin-enhanced waters, functional teas, juices, smoothies, and fortified drinks. These various types of functional beverages can be categorized into seven major

Departamento de Ingeniería Química, Alimentos y Ambiental. Universidad de las Américas Puebla, San Andrés Cholula, Puebla, 72810, Mexico. E-mail: nelly.ramirez@udlap.mx

groups: energy, performance enhancers, weight management, beverages for digestive health, immunity, cardiovascular health, and cognitive health.[3–5] Among all functional beverages, many consumers today prefer fruit juice blends and dairy-free alternatives due to their nutritional, sensory, and sustainability benefits, as well as dietary restrictions. Fruit juice blends are naturally high in vitamins, minerals, fiber, and bioactive phytochemicals like polyphenols and flavonoids, which are linked to antioxidant, anti-inflammatory, and metabolic health benefits.[6–8] Likewise, dairy-free alternatives are increasingly popular, not only for accommodating lactose intolerance, milk protein allergies, or vegan diets, but also because they typically have a smaller environmental footprint than traditional dairy.[8] Driven by these factors, the global demand for dairy-free options is increasing rapidly.

Nowadays, the functional beverage category stands out, with a market value of $208.13 billion dollars according to 2024 market reports, highlighting a Compound Annual Growth Rate (CAGR) of $7.5 million dollars growth for 2022–2027.[9] Particularly, in 2024, the global plant-based milk market was valued at approximately USD 20.93 billion and is expected to reach USD 43.63 billion by 2034.[10] Similarly, the broader market for dairy alternative beverages, including almond, soy, oat, and other non-dairy milks, was valued at around USD 26.06 billion in 2024 and is projected to expand to USD 69.59 billion by 2033.[11]

Given this growing trend regarding the formulation of functional beverages, technological challenges are focused on reducing required design, development, and production times, focusing mainly on fortification with essential nutrients and nutraceuticals without compromising sensory attributes.[12,13] Food formulation has typically been evaluated by designing mixtures, optimizing the characteristics of each ingredient, and maximizing their values through statistical optimization.

Some examples of this approach can be found in Ogundele *et al.*,[14] who evaluated the effects of blends of pineapple, orange juice, carrot, and *Hibiscus sabdariffa* extracts on the antioxidant properties of juice formulations, using a response surface methodology (DoE) to find a combination with the best antioxidant properties. Kim *et al.*[15] optimized the mixing ratio of broccoli, cabbage, and carrot powders to develop juice powders with high phenolic content, strong antioxidants, and good sensory qualities using a mixture design. The DoE mixture design approach has also been applied to the formulation of vegetable-based fermented products, evaluating their fermentability and the resulting physicochemical parameters, texture, microbiological quality, and viability of the lactic acid bacteria during shelf life.[16] Seifu *et al.*[17] formulated a milk analogue from peanuts, oats, and chickpeas for adults, middle-aged individuals, and the elderly. The formulated vegetal beverage was optimized using a mixture design in terms of its nutritional value, considering protein and fat content, as well as its mineral composition.

On the other hand, available computational tools have enabled several authors to develop product design using theoretical equations as models for product formulation. Although computer-aided design has been successfully implemented in the petrochemical, personal care, pharmaceutical, and cosmetic industries,[1,18] this methodology has rarely been applied in the food industry. In this regard, Erdogdu *et al.*[19] discussed the importance of virtualization in food processing, emphasizing that food companies must remain competitive, cost-effective, and capable of producing high-quality products while meeting market demands. For new formulations, they also observed that food materials are quite complex, and formulations face constraints on functional, nutritional, and organoleptic properties. Musina *et al.*[2] reviewed various computational tools for food formulation and development, most of which are based on experimental designs of mixtures. A key goal in food design is to find an optimal mix ratio within specified limits, where many raw materials can serve as ingredients with often variable nutritional content, performance, and availability. For instance, Yeboah *et al.*[20] developed a theoretical approach for ice cream formulations using jaggery as a sugar substitute. They converted this problem into mathematical models for multivariable and multi-objective optimization of food products. The designed ice cream was experimentally tested for density, overrun, viscosity, pH, texture, freezing point, and melting rate, and evaluated through sensory assessment. Results indicate that computer-designed ice creams are enjoyable, with jaggery not affecting sensory perception.

Advances in computer-aided food processing bring us closer to technologies that allow us to optimize formulations through theoretical approaches, matching them with experimental designs that significantly reduce the costs of innovation associated with the food industry (*i.e.*, time, materials, and personnel).[2,21]

As noted, theoretical food system models based on thermophysical property estimation are still in development due to the complexity of food matrices. Therefore, the present study aims to propose model formulation for food mixtures by estimating physicochemical properties. Two beverage case studies are presented, namely a fruit-juice mixture and a plant-based beverage. The consumption of juice blends and plant-based drinks is on the rise, and both beverage types can be characterized by the physical and functional properties of their components, including density, viscosity, soluble solids, pH, antioxidants, and proteins. In this work, beverage formulation was optimized using two approaches: one based on theoretical models and the other based on experimental design.

2. Methodology

2.1 Definition of case studies

The proposed model was applied in two case studies. The first study utilized a blend of fruit juices (grape, apple, and cranberry) to maximize antioxidant content. The second study focused on developing a drink formulation that maximizes protein content, incorporating a blend of high-protein vegetable sources.

For both case studies, the first step is to define the product in terms of the target properties and desired outcome attributes. Next, the product requirements were translated into physical properties, specifying target values to establish a list of attributes that align with the formulation of a beverage meeting

Table 1 Heuristics dashboard for case study formulations

Consumer assessments		Constraints	Measure unit
Juice formulation	Juice blend	$x_{apple} + x_{grape} + x_{cran} = 1$	
	Antioxidant activity	Total phenolics > 500	mg GAE per mL
		% scavenging > 70%	% inhibition
	Flavor profile	10 < solid content < 15	°Brix
		Acidity <2%; 0.05 mg acid per mL	—
		2.5 < pH < 6	—
	Consistency	0.980 < a_w < 0.990	—
		20 < viscosity < 26	Centipoise
		1.00 < density < 1.04	g mL^{-1}
Milk analog formulation	Protein mass	$x_{rice} + x_{pea} + x_{almond} = 1$	—
	Flavor profile	$x_{Albumin} + x_{Globulin} + x_{Glutenin} + x_{Prolamin} = 1$	—
		0.995 < a_w < 0.998	—
		100 < total protein < 200	mg mL^{-1}
		6 < pH < 7	—
		65 < viscosity < 75	Centipoise
		1.00 < density < 1.09	g mL^{-1}

a flavor profile, an ingredient balance, texture, and the target nutrient to be optimized. The first case includes a fruit-based beverage, which is described in terms of properties such as balance of solids, titratable acidity, pH, density, and viscosity. The second case describes the plant-based beverage as a milk analog, highlighting the protein content and distribution as a proportional ratio of albumin, globulin, glutenin, and prolamin content. The beverage formulation was defined by translating consumer preferences (flavor, color, and appearance) into physical properties (acid–sweetness balance, viscosity, density, *etc.*) that can be estimated using empirical or semi-empirical models, as in Calvo *et al.*[22] The constraint boards describe the physical properties, including viscosity, density, solubility, and visual attributes such as color and particle size. The physicochemical attributes are presented in a constraint board format outlining the main restrictions in Table 1.

Once the properties' dashboards have been defined, a set of theoretical models is required to describe each attribute in terms of physicochemical constraints. It is important to highlight that the targeted properties were defined only if they can be measured experimentally to validate the theoretical estimation accuracy.

2.2 Product design optimization

2.2.1 Theoretical estimation of physicochemical properties for product formulation. Theoretical modeling includes the following equations:

First, the sum of each mass fraction must be equal to 1 to comply with the mass balance constraint.

$$\sum_{i=1}^{n} x_i = 1 \tag{1}$$

where x_i is the mass fraction of each mixture component, and n is the number of components.

Several colligative properties in food systems depend on the activity coefficient of the mixture, typically expressed as water activity (a_w) for food matrices. Among the various food properties, a_w is one of the most important parameters, directly related to the amount of water available in the food material for physical, chemical, and biochemical reactions. For ideal aqueous solutions, a_w is defined as the ratio of the partial pressure of water present in the solution (p_w) to the vapor pressure of pure water (p_s), and can be related to the molar fraction of the solute in the solution.[23] In aqueous sugar solutions, water activity is affected by water–water, water–sugar, and sugar–sugar interactions, which are concentration and temperature-dependent. The Ross equation provides a straightforward method for estimating water activity (a_w) in mixtures containing multiple solutes, under the assumption of ideal solution behavior. It has been proven effective for systems with high a_w values (up to 0.95). This work estimated a_w using the Norrish and Ross equations.[24,25] The Norrish equation (eqn (2)) estimates the water activity for binary solutions that do not involve non-electrolyte solutes. For the aqueous solution containing a mixture of solutes, the Ross equation (eqn (3)) was implemented; this equation assumes that the interaction among solutes is negligible, and the law of the mixture can be obtained by multiplying the water activity of binary systems.

$$a_{ws} = \frac{p_w}{p_s} = \gamma_w \bar{x}_w = \bar{x}_w e^{\left(-k\left(x_s^2\right)\right)} \tag{2}$$

$$a_{wmix} = \prod_{s=1}^{n} a_{ws} \tag{3}$$

where γ_w is the water activity coefficient, $\bar{x}_w$ is the mole fraction of water, and x_s is the mole fraction of solutes in the binary mixtures. The Norrish constants (k) for each solute are: 6.47 for sucrose, 2.25 for fructose and glucose, 2.52 for β-alanine, and 2.59 for α-aminoacid-*n*-butyric acid.[26]

Texture is one of the primary sensory properties of food, which is related to the size, shape, and structure of food molecules. For liquid foods, this sensory response can be related to rheological parameters. Viscosity is a measure of internal fluid friction, and for a mixture of liquids, the dependence of viscosity on composition can be nearly linear for ideal systems. The Grunberg and Nissan equation (eqn (4) and (5))

enables the theoretical calculation of viscosity in mixtures using the group contribution method and a simple mixing rule for systems at low temperatures, wherein the viscosity is expressed in centipoises per mL.[27,28]

$$\mathrm{In}(\eta)_{\mathrm{mix}} = \sum_{i=1}^{n} x_i \mathrm{In}(\eta_i) + \frac{1}{2}\sum_{i=1}^{n} x_i G_{ij} \quad (4)$$

$$G_{ij} = \sum \Delta_i - \sum \Delta_j + W;\ W = 0 \quad (5)$$

where x_i is the mass fraction for each component in the mixture, η_i is the viscosity contribution of each mixture component, and G_{ij} is the interaction parameter (Tables S1 and S2 in SI data).

The proper estimation of density is crucial for characterizing food systems, as it is a key property required to determine other variables, such as viscosity. The density of liquid foods, like juices or plant-based beverages, depends on several factors, including composition, soluble solid content, and processing conditions like temperature and pressure.[29] In general, it can be accurately estimated using the well-known densities of the juice constituents at a reference temperature. For the mixture design considered here, the effects of temperature and pressure can be neglected, and the mixture density (g mL^{-1}) was estimated using eqn (6), which considers the density contribution of each component in the mixture.[30]

$$\rho_{\mathrm{mix}} = \frac{1}{\sum x_i/\rho_i} \quad (6)$$

where x_i is the mass fraction in the aqueous solution and ρ_i is the density contribution of each component in the mixture (g mL^{-1}, Table S3 in SI data).

Various models in the scientific literature have attempted to explain food choice based on physical–chemical factors such as nutrients, sweet–acid levels, pH, and antioxidant content.[27] In this regard, estimating these characteristics is less straightforward than the other properties described in previous sections. To accurately determine how properties such as pH, total phenol content, and protein content depend on composition, calibration curves must be established to approximate values within the formulation range. In the theoretical estimation of acidity and the pH_{mix} in the mixture, an equation similar to Ross's is used to assume that the pH_i of each component contributes to the mixture based on its mass fraction (eqn (7)), which is obtained by multiplying the pH_i of the ternary mixture. The pH_i is estimated by the Henderson–Hasselbalch equation (eqn (7)), where [A−] represents the concentration of undissociated acid. [HA+] represents the concentration of its conjugated base, in this case NaOH solution (0.1 N concentration was used).

$$\mathrm{pH}_{\mathrm{mix}} = \sum_{i=1}^{n} f_i \mathrm{pH}_i;\ f_i = \frac{C_i}{\sum_{i:1}^{n} C_i} \quad (7)$$

$$\mathrm{pH}_i = \mathrm{p}K_{a1} - \log\frac{[\mathrm{A}-]}{[\mathrm{HA}+]} \quad (8)$$

pK_a is the pH_i at which equal quantities of undissociated acid [A−] and conjugated base [HA+] are present.[31] In case study 1, the pK_{a1} of malic acid in apple juice was 3.45; the $pK_{a1,2}$ of tartaric acid in grape juice was $0.35 \times pK_{a1}$: 3.03 and $0.65 \times pK_{a2}$ 4.36; and the $pK_{a1,2}$ of citric acid in cranberry juice was 3.08 and 4.75. In case 2, the rice, pea, and almond extract mixtures were based on phytic acid with pK_{a1} of 5.40 and pK_{a2} equal to 4.21, as reference data on lactic acid analogues (data for pH estimation in Tables S4 and S5 in SI data).

The concentration of total polyphenols in each mixture followed an additive rule of component fraction and binary interaction (eqn (9)). The coefficients C_i and β_i were obtained from a multiple linear regression, based on the calibration curves and the experimental values for different mixtures ($R^2 > 0.9$).

$$\mathrm{TP\ content(mix)} = \sum_{i:1}^{3} x_i C_i + \beta_1 x_1 x_2 + \beta_2 x_2 x_3 + \beta_3 x_3 x_1 \quad (9)$$

where x_i is the mass fraction of each sample at the dilution of mixture formulation. The interaction coefficients β, in case 2 were: β_1: 33.90, β_2: 197.52, β_3: −241.47. The coefficients for each pure ingredient C_i were obtained from the calibration curve: C_{apple}: 224.62, C_{grape}: 606.16, $C_{\mathrm{cranberry}}$: 600.54 mg GAE per mL juice.

For the protein content estimation, similarly to eqn (9) and (10), an additive model equation (eqn (10)) with interaction coefficients was considered, where coefficients C_i and β_i were also obtained from a multiple linear regression, from experimental values for different mixtures and calibration curves ($R^2 > 0.9$).

$$\mathrm{Protein\ con(mix)} = \sum_{i:1}^{3} x_i C_i + \beta_1 x_1 x_2 + \beta_2 x_2 x_3 + \beta_3 x_3 x_1 \quad (10)$$

where x_i is the mass fraction of each sample at the dilution of mixture formulation. The interaction coefficients β in this case were: β_1: 13.2, β_2: 30.15, β_3: −79.4, the coefficients for each pure ingredient C_i being C_{rice}: 221.8, C_{pea}: 202.6, C_{almond}: 141.9.

2.2.2 Experimental mixture design. Various experimental designs are available to optimize the development of food products. Fractional factorial designs, particularly Taguchi methods and simple lattice designs, have proven to be effective design tools in various areas. Techniques such as response surface methodology and mixture designs are effective in optimizing formulations. This work used a simplex lattice mixture design. In the case of formulations composed of several ingredients that sum to a constant (such as systems considered here), a simplex lattice design is particularly suitable because it systematically explores all possible combinations of component proportions within the mixture. For a 3-component system, this design enables the estimation of linear and nonlinear blending effects among ingredients, providing a comprehensive understanding of how different proportions influence product characteristics and allowing efficient optimization.[32–34]

Tables 2 and 3 show the mixture design for fruit juice blends and the plant-based beverage, respectively. Each point of the design was prepared in triplicate in the indicated proportions

Table 2 Experimental design for juice blend formulation using the simplex lattice mixture design

Component 1	Component 2	Component 3
Apple (%)	Grape (%)	Cranberry (%)
20	20	60
20	60	20
25	25	50
25	50	25
50	25	25
60	20	20

Table 3 Experimental design for plant-based beverage formulation using the simplex lattice

Component 1	Component 2	Component 3
Rice (%)	Pea (%)	Almond (%)
100	0	0
0	100	0
0	0	100
50	50	0
0	50	50
50	0	50
25	25	50
25	50	25
50	25	25
33	33	34

and analyzed for the following parameters: total soluble solids (Brix), water activity, pH, titratable acidity, antioxidants (total phenolics), antioxidant capacity (DPPH), viscosity, and density. The polynomial models describing all responses were developed, incorporating the linear effects of each component and their interactions; simple lattice designs permit the comparison of k^2 varieties in blocks of size k. The response polynomial is represented in eqn (11) where b_0, $b_1 \ldots b_7$ are the regression coefficients and u_1, u_2, u_3 are the mass fractions of each component in the mixture.

$$\text{Response} = b_1u_1 + b_2u_2 + b_3u_3 + b_4u_1u_2 + b_5u_1u_3 + b_6u_2u_3 \quad (11)$$

The obtained polynomial expressions allow the development of a three-dimensional surface for each response (*e.g.*, water activity, solids, density, viscosity). Each term in the polynomial model was statistically significant ($p < 0.05$). The correlation coefficients and the RMSE were determined for each response.

2.2.3 Formulation of the optimization problem. Optimizing ingredient choice and product formulation is based on two approaches: (i) theoretical equations and (ii) experimental design using a simple lattice mixture design. For the theoretical approximation, the equations described in Section 2.2.1 were used, whereas for the experimental design approach, polynomial models derived from regression analysis were implemented. The optimization problem was defined as a nonlinear optimization problem for both approaches:

$$\max Z = f(x)$$

where $f(x)$ was defined as maximizing (a) antioxidant content for case study 1 and (b) maximizing protein content for case study 2; s.t., $h(x) = 0$, mass balance equality constraints; $g(x) \leq 0$, target inequality constraints defined in the property's dashboards; $x \in X$, composition variables.

All the estimated properties under the optimal conditions were experimentally validated. The optimization problem for the theoretical model was solved using the Generalized Reduced Gradient (GRG) non-linear solver available in Microsoft Excel tools. The surface response model was developed using Minitab software to optimize experimental design. The optimal solutions obtained were experimentally validated using the methods described in Section 2.5.

2.3 Sample preparation and experimental validation

2.3.1 Sample preparation. The juice samples (*Vaccinium macrocarpon* and *Malus domestica*) were obtained from Ingredion commercial brand (Guadalajara, Mexico). Grape juice concentrate (*Vitis vinifera*, cv. Victoria) was obtained from Casa Leal vineyard (Aguascalientes, Mexico). All juice concentrations exceeded 69.1 ± 1.0 °Brix, then all samples were diluted to 10, 15, and 20 °Brix with distilled water.

The second case was based on powders of vegetable protein concentrates from rice and peas (70% d.b. protein concentrate) and almonds (30% d.b. protein concentrate) obtained in a commercial store (Future Foods). The vegetable powders and powder mixtures were rehydrated to 15% solids, suspended with 1% xanthan gum, and homogenized until stable. Xanthan gum was chosen for its high water-binding capacity and ability to form stable, shear-thinning dispersions that prevent sedimentation and phase separation in plant-based matrices.[35] The concentration of 1% was determined from preliminary tests and supported by previous studies, which reported that ≈1.0–2.0% xanthan gum provides optimal viscosity and sensory balance in dairy-alternative beverages.[36]

2.3.2 Physical and chemical characterization for experimental validation. Water activity (Aqua-Lab, Decagon Devices Inc, Washington, USA) and viscosity (Brookfield viscometer, DV3TRVTJ0, AMETEK Inc., Devon-Berwyn, PA, USA) were determined for each mixture. A portable refractometer was used to determine the total soluble solids of the juices; in °Brix; the pH at 25 °C was measured with a HANNA Instruments potentiometer (model HI5521, Providence, USA). A spectrophotometer (Shimadzu UV-1900i) was used to carry out color and turbidity measurements of all mixture samples, according to Camelo-Méndez.[37] The visible spectrum (380–770 nm) was recorded at constant intervals (wavelength interval = 2 nm), using 2 mm path-length glass cells and distilled water as a reference. The CIELAB parameters (L^*, a^*, b^*) were determined using the software Croma Lab (Konica Minolta, Osaka, Japan). The total acidity of the samples was measured by

titration, using 5 mL of each sample and transferred to a 250 mL beaker, adding 50 mL of distilled water and mixing (with 0.1 N sodium hydroxide standard, pH of 8.2). Eqn (12) was used to determine the total acidity value.

$$\text{Acidity}w/w(\%) = \frac{v \cdot 0.1 \cdot 0.64}{S \times 10} \quad (12)$$

where v = NaOH volume used for each sample titration and S = initial volume of the sample, 0.1 N indicates NaOH concentration, 0.64 is the acid constant (malic and tartaric acid reference) and 10 the dilution order. The determination of total phenolic content was carried out according to the Folin–Ciocalteu method, where the calibration curve was prepared using gallic acid (0–100 mg mL^{-1}), and the absorbance was linearly correlated ($R^2 = 0.98$) with GAE over the studied range. One milliliter of diluted samples (1/10) or gallic acid was mixed with 5 mL of FCR (10%). Then 4 mL of saturated sodium carbonate was added to the solution and kept for 30 min. The absorbance of the samples was determined at 765 nm using an ultraviolet-visible (UV-vis) spectrophotometer (UV-1900i, Shimadzu, Japan), and the total phenolic compounds (TP) were estimated as mg of gallic acid equivalent (GAE) per 100 mL of sample from the standard curve.

The antioxidant activity (%) was evaluated by DPPH radical scavenging assay, as in Sahraee *et al.*;[38] 0.1 mL of diluted samples (1/10) was mixed with 3.9 mL of DPPH solution (100 M). After shaking the solution, it was kept in the dark at 25 °C for 30 min. Ethanol (96%) was used instead of the sample in the above procedure as blank. The antioxidant activity of the samples was reported as the inhibition percentage of the DPPH radical according to eqn (13)

$$\text{Inhibition } (\%) = \frac{A_c - A_s}{A_c} \quad (13)$$

where A_c is the absorbance of the control (methanol and DPPH solution) and A_s is the absorbance of the sample solution (sample and DPPH solution), the sample's absorbance was determined at 517 and 734 nm using an ultraviolet-visible (UV-vis) spectrophotometer (UV-1900i, Shimadzu, Japan).

2.3.3 Protein characterization. The protein content of the solutions was determined according to the Bradford method.[39] The standard curve was generated with different concentrations of bovine serum albumin (0–400 mg mL^{-1}) that were linearly correlated ($R^2 = 0.96$) with the absorbance at 565 nm (UV-vis spectrophotometer UV-1900i, Shimadzu, Japan). Samples were dissolved/extracted by adding 25 mL of distilled water (adjusted to the corresponding pH value) to 500 mg of each powder. All extractions were conducted in triplicate. After stirring for 30 min at room temperature, the suspensions were centrifuged at 3900 × *g* for 30 min at 5 °C. Aliquots of the supernatant were analyzed.[40]

The solubility profile of the proteins from the samples was determined by varying the pH values (distilled water from 1 to 12).

2.4 Sensory evaluation for comparing the optimal formulations

The sensory evaluation was conducted with 30 untrained panelists (54% women and 46% men) aged 18–60 years (median age: 25–27), recruited through an institutional email invitation. Participants were regular consumers of fruit-based or milk-alternative beverages and reported no allergies. All participants provided informed consent in accordance with the ethics committee of the Doctoral Program in Food Science at Universidad de las Americas Puebla (UDLAP). Sensory evaluations were carried out in individual booths under controlled white lighting and ambient temperature (22 ± 2 °C) in the Sensory Evaluation Laboratory of the Department of Chemical, Food and Environmental Engineering at UDLAP, following ISO 8589:2023 guidelines.[41] Each panelist received two 20 mL samples per evaluation—one corresponding to the theoretically optimal formulation and the other to the mixture design optimization—served at room temperature in odor-free plastic cups labeled with randomized three-digit codes to ensure blinding. The order of presentation was balanced and randomized to minimize order effects. A 9-point hedonic scale (1 = dislike very much to 9 = like very much) was used to evaluate the selected attributes (color, odor, flavor, texture and appearance, general acceptability). Also, panelists were asked to choose a word from the suggested descriptors and give a score to identify a taste profile. The sample size ($n = 30$) was consistent with previous sensory studies on beverage formulations and considered adequate to detect moderate differences in hedonic perception among formulations. Mean scores were analyzed using one-way ANOVA, followed by Tukey's post hoc test ($p < 0.05$) to determine significant differences among samples.

2.5 Statistical analysis

All physicochemical and instrumental measurements were performed in triplicate for each mixture, and the results are expressed as mean ± standard deviation. One-way analysis of variance (ANOVA) followed by Tukey's test ($p < 0.05$) was used to determine significant differences among mixtures. The same statistical approach was applied to the sensory evaluation data to compare mean hedonic scores for texture and taste profiles among formulations. All data were processed using Minitab *v.20.3.2021 statistical software (Minitab Inc., State College, PA, USA)*. Additionally, a regression model based on the mixture design was developed and analyzed using the same software.

3. Results and discussion

3.1 Formulation optimization by the theoretical approach

Although the application of semi-empirical equations in designing foods and beverages, such as juice blends and vegetable blends, addresses limitations related to insufficient data, during food formulation,[42] these models may offer a first approximation and reduce experimental effort times. The constraints set in the dashboard (Table 1) allow for delimiting the optimization problem and finding the objective function.

The optimal solution for case study 1 indicates that the mixture that maximizes antioxidant content should consist of 14% apple juice, 44% grape juice, and 42% cranberry juice. The estimation of the optimal mixture was performed by means of a mathematical solver, given the properties estimated by

Table 4 Optimal solution for juice mixture by the theoretical approach and experimental validation

Constraints	Measurement unit	Theorical approach	Experimental validation	Error (%)
$x_{apple} + x_{grape} + x_{cran} = 1$	—	14/44/42%	—	
Total phenolics > 500	mg GAE per mL	564.52	553 ± 16.83	2.0
% scavenging > 85%	% inhibition	89.71	77 ± 5.4	14.2
10 < solid content < 15	°Brix	11	12 ± 0.4	9.1
Acidity < 2%; 0.05 mg	—	0.045	0.054 ± 0.02	20.0
2.5 < pH < 6	—	4.65	5.8 ± 0.2	24.7
0.980 < a_w < 0.990	—	0.986	0.992 ± 0.03	0.6
20 < η < 26	Centipoise	21.67	24 ± 1.8	10.8
1.00 < ρ < 1.04	g mL^{-1}	1.02	1.15 ± 0.1	12.7

Table 5 Optimal solution for the milk analogue by the theoretical approach and experimental validation

Constraints	Measurement unit	Theorical approach	Experimental validation	Error (%)
$X_{rice} + X_{pea} + X_{almond} = 1$	Mass percentage	74/16/10%	—	
$X_{Albumin} + X_{Globulin} + X_{Glutenin} + X_{Prolamin} = 100$	Proportion	12/22/6/60	—	
a_w > 0.995	—	0.996	0.997 ± 0.01	0.1
Total protein > 150	mg mL^{-1}	200	171 ± 11.4	14.5
6 < pH < 7.5	—	6.94	6.8 ± 0.4	2.0
40< η < 80	Centipoise	70.5	55 ± 15.78	22.0
ρ > 1.00	g mL^{-1}	1.07	1.072 ± 0.05	0.2

equations in Section 2.2, subjected to the constraints in Table 1. This combination yields a 564.53 mg of GAE per mL concentration, corresponding to an 89.71% scavenging percentage. As shown in Table 4, all the established formulation constraints are satisfactorily met.

The estimated sugar content for this mixture is 11 °Bx, which aligns with the standard sugar content for juices and nectars under Mexican regulations.[43] Regarding the properties associated with the flavor profile, the mixture contained 0.045 mg of malic/tartaric acid in 3 : 2 proportion, yielding a pH of 4.6. The acidity of the juice can be considered from the range of reported commercial juices with minimum % acidity values in terms of malic/tartaric acid from 0.053 mg mL^{-1} to a maximum of 0.075 mg mL^{-1} in cranberry, apple, and cranberry blends.[44]

The model estimates a water activity of 0.986, a parameter closely linked to the content and type of sugars, which can impact properties such as freezing temperatures. Regarding texture-related variables, the estimated viscosity and density values for this optimal mixture were 21.67 cP and 1.02 g mL^{-1}, corresponding to a Newtonian behavior fluid, being a low solids juice, according to the quality standard of 20–30 cP as viscosity measure.[45,46]

It is important to note that this formulation and its thermophysical property values were theoretically estimated. Therefore, subsequent experimental validation of this optimal solution was conducted with all properties measured. The experimental results showed an estimation error ranging from 0.6% to 24% to the theoretical estimations. Water activity showed the most accurate theoretical estimation, with an error margin of only 0.6%. Viscosity and density displayed higher errors, with discrepancies of 10.8% and 12.7%, respectively, between theoretical estimations and experimental measurements; although the percentage error appears large, it is important to note that for the order of magnitude of these properties, and considering the standard deviation from experimental validation, the difference is less than one cps for viscosity and less than 0.1 g L^{-1} for density. The largest estimation errors were seen in pH and acidity, which ranged from 20% to 24.7%. The lack of accuracy in pH and acidity estimations can arise from several sources that have not been considered due to the simplified model and assumptions, such as the ionic interactions or the fact that dissociation constants are only valid under certain conditions. Additionally, in juices, there is a mixture of acids that can compete for protons. From a mathematical perspective, pH is a logarithmic function, so small changes in acidity may result in large changes in pH. In contrast, the total phenolic content exhibited a good estimation, demonstrating only a 2% error between theoretical and experimental values, while the percentage of inhibition of DPPH showed a 14.2% error. The observed error may be attributed to the variability of the data and the concentration limits of each juice selected; although the samples were measured under standardized conditions, the measurement of phenolic compounds may differ as the gradients are in parts per million, making this response variable very sensitive to changes by reagents or oxidation for the optimal solution reached.

For study case 2, the optimization problem is described in Table 5. The optimal theorical formula that satisfies the defined restrictions and maximizes the vegetable protein content in the drink (200.8 mg mL^{-1}) should be formulated with 74% rice extract, 16% pea extract, and 10% almond extract, corresponding to a theoretical distribution of vegetable proteins as:

Table 6 Coefficients of the regression polynomial model for the studied juice mixture

Terms	Total soluble solids (°Brix)	a_w	Acidity (%)	Total phenol content (mg GAE per mL)	Free radical scavenging (DPPH %)	Viscosity (cP)	Density (g mL^{-1})
Apple, %	0.0882	0.01046	0.00049	18.0824	2.3654	0.1685	0.00938
Grape, %	0.1883	0.01078	−0.00020	37.9699	0.6657	0.0059	0.01040
Cranberry, %	0.1637	0.00888	0.00103	−19.1003	0.4351	0.2709	0.01060
Apple *grape (%)	0.0001	−0.00006	0.00003	−1.8713	−0.0531	0.0093	0.00003
Apple *cranberry (%)	0.0005	0.00003	−0.00003	0.9585	−0.0428	−0.0044	0.00002
Grape *cranberry (%)	−0.0045	0.00001	0.00000	0.1497	0.0496	0.0026	−0.00003
R-sq (%)	71.06	79.52	70.11	89.31	79.2	99.91	34.05
RMSE	0.2639	0.0017	0.0013	44.5363	3.1073	0.0246	0.0049

12% albumins, 60% glutenins, 22% globulins and 6% prolamins (Fig. S2 in the SI).

The contribution of the rice extract yields a higher proportion of globulin proteins, as these proteins are abundant in cereal. At the same time, the albumin and prolamin content is mainly associated with the pea extract. Regarding the almond extract, this component contributes less to the protein balance but plays an important role in the fat profile.[47,48]

Regarding the properties related to the consistency of this plant-based drink, the mixture has a viscosity of 70.5 cP, a density of 1.07 g mL^{-1}, and 0.996 water activity. This blend formulation is considered a low solid content liquid (around 15%). Presumably, the milk analogs with high viscosities (2.2 to 48 mPa s^{-1}) contain thickening agents to inhibit gravitational separation, such as pea, which acts as a natural thickening agent. They also provide a creamy mouthfeel because they do not contain many oil bodies or fat droplets.[12]

For this case study, the property estimation shows a standard error from 0.1% for water activity to 22% for viscosity when compared to experimental determinations. The error in viscosity estimation for the plant-based beverage can be attributed to treating the components as a mixture of solutions, when they are actually aqueous suspensions of oil bodies or oil-in-water,[49] so that the group contribution methods fail to estimate that behavior accurately.

Water activity, pH, and density are the properties with high estimation accuracy. Regarding the protein content estimation, the error was around 14.5%. This could be due to the interaction between the extracts and their rheological variability, and the solids' contribution in the blend.[50]

In general, even for those properties with larger estimation errors, a strong determination coefficient between theoretical estimations and experimental data was observed ($R^2 > 0.9$). For illustrative purposes, Fig. S1 in the SI shows the correlations for viscosity and a_w.

As noted, although theoretical estimation has some drawbacks related to estimation accuracy, this approach offers feasible solutions that help reduce the number of potential data and ingredients to be tested within a reasonable time frame. This theoretical estimate is based on the chemical composition of the main ingredients.[51]

3.2 Formulation optimization by the experimental mixture design approach

In the experimental design, the initial case centered on a fruit juice formula aimed at maximizing antioxidant capacity. Table 6 presents the coefficients for the polynomial models (eqn

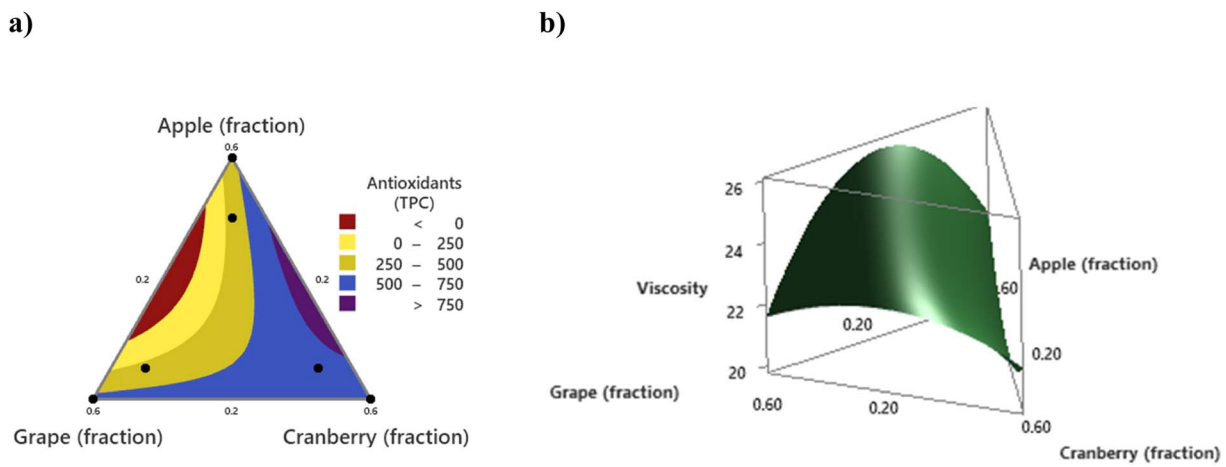

Fig. 1 (a) Contour plot of antioxidant activity measured as total phenol content (TPC) in mg mL^{-1} of juice mixture from apple, grape, and cranberry blends (purple zone indicates optimal response). (b) Surface plot showing viscosity response as a function of blending ratio for the same juice mixture.

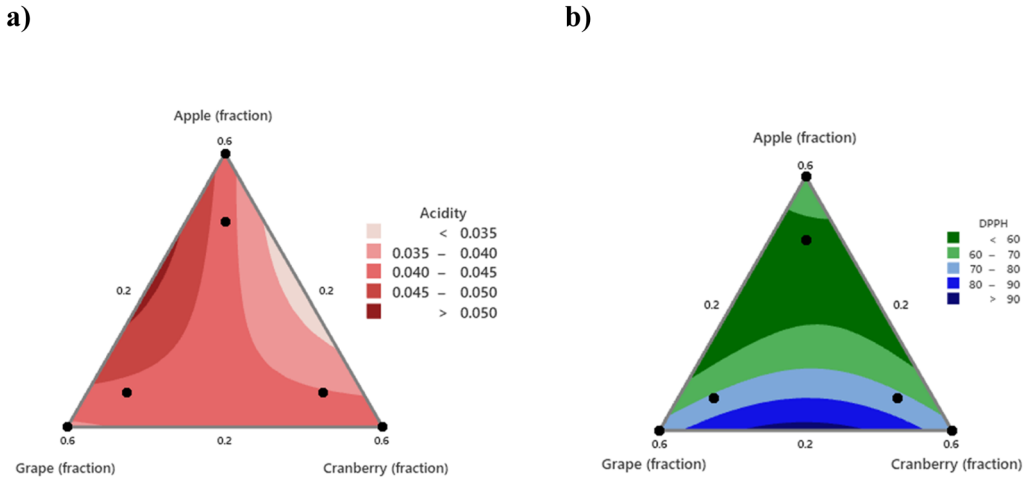

Fig. 2 (a) Contour plot displaying the acidity (%) response for the juice mixture at various proportions of apple, grape, and cranberry. The region with a strong red color indicates the highest value in mg citric/malic acid per mL of juice. (b) Contour plot illustrates the antioxidant response (DPPH inhibition %) based on the blending ratio for the same juice mixture.

Table 7 Optimal juice mixture obtained by the mixture experimental design approach

Constraints	Measurement unit	Mixture design approach	Experimental validation	Error (%)
$x_{apple} + x_{grape} + x_{cran} = 1$	(–)	28.5/32.2/39.3%	—	
Total phenolics > 500	mg GAE per mL	532	605 ± 32.1	13.7
% scavenging > 85%	% inhibition	75	72 ± 3.1	4.0
10 < solid content < 15	°Brix	10.4	11 ± 0.4	5.8
Acid < 2%; 0.05 mg acid per mL	mg acid per mL	0.041	0.058 ± 0.01	41.5
2.5 < pH < 6.5	(–)	4.7	5.7 ± 1.3	21.3
0.900 < a_w < 0.990	(–)	0.982	0.981 ± 0.03	0.1
η > 22	Centipoise	22.5	24.1 ± 0.9	7.1
ρ > 1.00	g mL^{-1}	1.02	1.10 ± 0.2	7.8

(11)), derived from data regression based on the experimental design for all responses measured in Case Study 1.

The contour plot in Fig. 1a, obtained using a polynomial regression model (Table 6), illustrates the blend optimization, highlighting the area with the highest concentration of polyphenols, indicated in purple; the black dots denote the experimental data points. The findings suggest that the best formulation should include 23% apple juice, 38% grape juice, and 39% cranberry juice, yielding a polyphenol value of 532 mg per mL of juice. It has been reported that apple, grape, and cranberry are valuable sources of antioxidant components expressed as chlorogenic acid, caffeic acid, hydroxycinnamic acids, and flavanols (catechin and epicatechin), with levels between 200 and 500 mg mL^{-1}.[52,53]

Fig. 1b shows the surface plot of viscosity for the juice mixtures. A maximum value in viscosity is observed as the proportion of grape juice blend increases, suggesting complex interactions between the juice components; according to Wolfe,[54] this behavior may be related to differences in sugar content (fructose, glucose, and sucrose) between the fruit juices and the formation of intermolecular interactions related to sugar and acids in the beverage.

Table 8 Coefficients of the regression polynomial model for the studied plant-based beverage

Responses/terms	a_w	Protein content (mg mL^{-1})	pH	Particle size (mm)	Color (delta hue)	Viscosity (cP)	Density (g mL^{-1})
Rice (%)	0.99198	102.37	7.8005	41	9.336	31.23	1.0796
Pea (%)	0.99498	96.71	7.0504	37.7	8.724	82.58	1.0783
Almond (%)	0.99648	42.85	6.1654	25.6	8.864	40.88	1.0243
Rice *pea (%)	0.0174	−51.9	−6.559	244	−0.9	−172.9	−0.704
Rice *alm (%)	0.0194	−87.6	−2.779	−52	−0.02	367.2	0.422
pea *alm (%)	−0.0094	59.7	6.799	223	1.14	−71.1	0.201
R-sq (%)	95.45	99.52	99.96	91.64	82.82	98.81	91.95
RSME	0.002	15.361	0.525	11.066	0.126	34.879	0.078

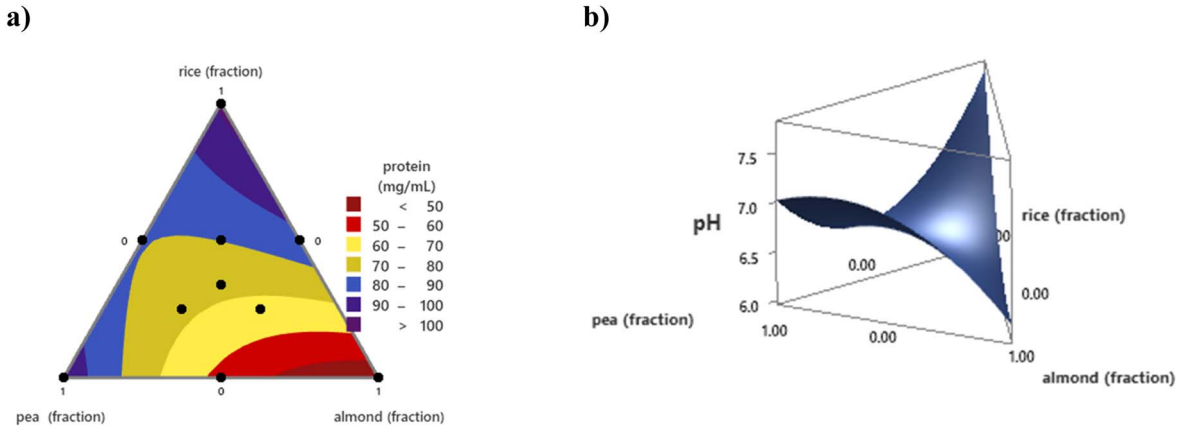

Fig. 3 (a) Contour plot showing the protein content across various almond, pea, and rice blend ratios for the beverage formulation. (b) Surface plot depicting the pH response of the studied plant-based beverage at different ingredient proportions.

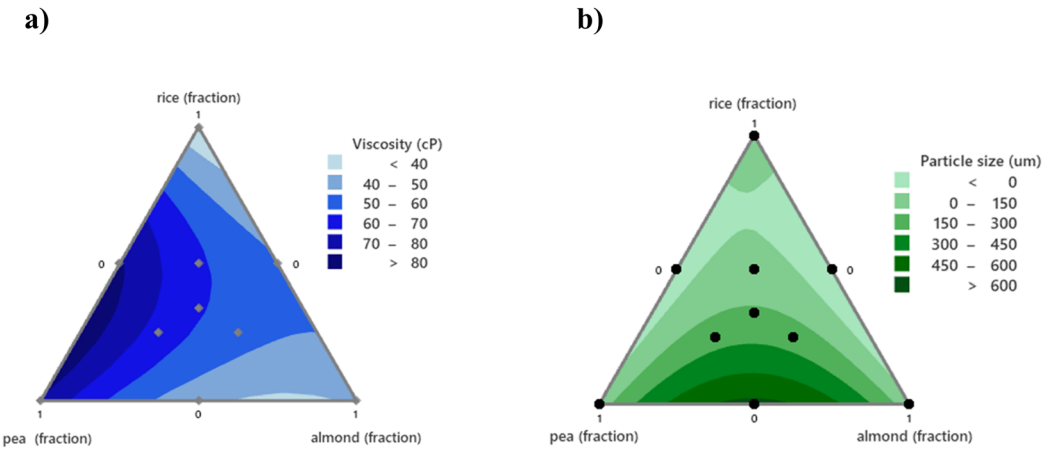

Fig. 4 (a) Contour plot of viscosity (cP) response of the blend of almond, pea, and rice contribution to beverage formulation, and (b) contour plot of particle size (μm) for the studied plant-based beverage as a function of the formula composition.

In Fig. 2a, the acidity of juice blends shows minimal variation due to juice proportions. This behavior could be related to the nature of the organic acids and their concentration in the fruit juices.[44,55] Fig. 2b shows how the DPPH radical behaves in the juice mixture (related to its role in oxidative inhibition). The obtained contour indicates that the main sources of DPPH radical are grape–cranberry blends.

Table 7 displays the optimum formula obtained by the experimental design approach including the experimental validation. The polynomial model approaches a phenolics content of 532 mg mL^{-1} and an inhibition percentage of the DPPH radical % of 75% for the optimum formula.

In this optimization using polynomial regression, the error estimation for water activity, percentage inhibition, solid content, viscosity, and density is below 8%. However, like the theoretical estimates, pH and acidity show larger deviations, with errors of 21.3% and 41.5%, respectively. In this case, the estimation of total phenolic content has an error of 13.7%.

For case study 2, a vegetable drink resembling dairy was formulated to showcase the protein contributions of almonds, rice, and peas. Table 8 presents the significant coefficients for polynomial models for each measured response.

The obtained contour plot in Fig. 3a, shows the optimized mixture of plant-based beverages, with the purple area indicating the highest total protein area, corresponding to the lowest almond contribution to the mixture. The black markers indicate the experimentally tested mixtures.

One of the most important nutrients of interest in milk substitutes is their dietary protein quality,[32,56] defined in terms of the protein's ability to be digested and absorbed by the body, which depends on the essential amino acid composition. A mixture of proteins from different plant sources maximizes the quality of the protein in a dairy analog beverage compared to single-source extracts, even from maximum legume-sources such as soy milk. Fig. 3b shows that the pH of the vegetable drink mixture fluctuates within a narrow range, resulting in minimal changes in pH levels. The pH is a crucial factor affecting the texture and taste of the formulation, primarily influencing the stability and solubility of the beverage dispersion.

Table 9 Optimal plant-based beverage obtained by the mixture experimental design approach

Constraints	Measurement unit	Mixture design approach	Experimental validation	Error (%)
$X_{rice} + X_{pea} + X_{almond} = 1$		60/28/12%	—	
$X_{Albumin} + X_{Globulin} + X_{Glutenin} + X_{Prolamin} = 100$		$0.15x + 0.28x + 0.05x + 0.50x$	—	
$a_w > 0.995$		0.996	0.997 ± 0.01	0.1
Total protein > 150	mg mL^{-1}	161.2	168 ± 7.90	4.2
6 < pH < 7.5		6.32	6.73 ± 0.17	6.5
40 < η < 80	Centipoise	41.77	64.92 ± 19.23	55.4
1.00 < ρ > 1.1	g mL^{-1}	1.00	1.07 ± 0.02	7.0

Table 10 Sensory evaluation of the formulated beverages[a]

	Juice blend formulation		Plant-based beverage	
Attributes	TMO-approach	DoE-approach	TMO-approach	DoE-approach
Color	8.3 ± 0.8^{A}	7.5 ± 1.9^{A}	4.1 ± 1.3^{B}	5.1 ± 1.4^{B}
Odor	6.4 ± 1.5^{A}	6.7 ± 1.4^{A}	5.5 ± 1.2^{B}	5.3 ± 1.6^{B}
Flavor	5.4 ± 1.8^{A}	5.2 ± 1.1^{A}	6.3 ± 1.2^{B}	5.2 ± 1.6^{B}
Texture and appearance	7.2 ± 1.8^{A}	6.6 ± 1.6^{A}	5.8 ± 1.3^{B}	4.1 ± 1.3^{B}
General acceptability	7.7 ± 1.9^{A}	7.5 ± 1.2^{A}	6.2 ± 2.5^{B}	6.3 ± 2.4^{B}

[a] Different capital letters for each attribute in each case (juice or plant-based beverage) indicate significant differences ($p < 0.05$).

The viscosity response shown in Fig. 4a indicates that the pea and rice mixtures contribute differently compared to the almond mixture. The dark blue zone represents the maximum viscosity value influenced by the contribution of the pea extract. This behavior could be related to interactions between components, such as the formation of intermolecular structures depending on amino acid composition and peptide bonds in proteins from different plant sources (see Fig. S2 in SI data). As estimated for the optimized formula (Table 9), the larger presence of rice indicates that there would be a higher percentage of glutenin.[50] Regarding the particle size, Fig. 4b shows a non-linear variation of this property as a function of the components' proportion. A maximum particle size peak is identified, which could be related to differences in the molecular structure or molecular weight of the protein diversity. The observed variability in the mixture could be related to the most abundant protein types, glutenins, and albumins, contributed by the rice–peas.[50] Table 9 presents the beverage optimization results obtained through the experimental design approach and their experimental validation. The highest error in estimation was observed in viscosity, with a deviation of 55.4%. In contrast, water activity exhibited the most accurate estimation, with an error of just 0.1%. Density and pH had estimation deviations of 7% and 6.5%, respectively. The error in estimating total protein content was 4.2%.

3.3 Comparison among the optimal drink formulae obtained by both implemented approaches

Finally, when comparing both the theoretical approach and the experimental mixture design one can observe that, in the case of juice formulation, the mixture design approach results in a more balanced formulation containing 28.5% of apple, 32.2% of grape and 39.3% of cranberry juices, compared with theoretical approximation that yielded a mixture containing 14%, 44%, and 42%, respectively. It is important to notice that some key parameters in the theoretical model were estimated by considering only components such as water and different sugars.

Although the proposed models allow the estimation of formulations, they do not necessarily provide a practical solution. Therefore, both mixtures underwent a sensory evaluation. Overall, panelists did not perceive a significant difference ($p > 0.05$) between the two formulae (Table 10), with a mean value of 7.5 ± 1.2 for the mixture design approach and 7.7 ± 1.9 for the theoretical estimation. Therefore, the difference between the two formulations does not alter the perception of flavor or its intensity.

In general, flavor attributes denote acid perception, grape note presence, and cranberry notes. The panel preferred the sample illustrated in the red line, indicating an affinity for the theoretical approach with this mixture (14% apple, 44% grape, 42% cranberry), identifying flavor descriptors such as grape, bitter notes, and traces of wine flavor (Fig. 5).

For case study 2 (plant-based beverage), the theoretical approach optimizes the formulation to achieve a higher protein content and a more balanced theorical protein distribution, including 40% glutenins, 33% globulins, 29% prolamins, and 22% albumins.

According to the sensory evaluation (Table 10), panelists indicated that they did not perceive a significant difference ($p > 0.05$) between the two formulae, with an average value of 6.2 ± 2.5 for the mix design approach and 6.3 ± 2.4 for the theoretical

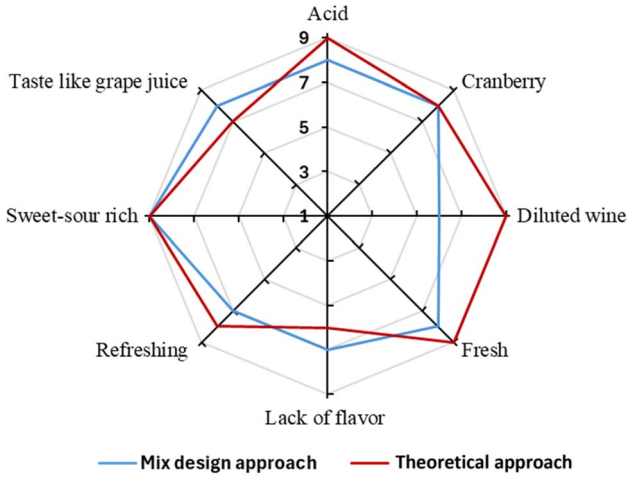

Fig. 5 Sensorial evaluation of optimal juice formulations (TMO and DoE, final samples). The blue line represents the mix design approach (DoE), while the red line indicates the theoretical approach (TMO). Panelist descriptors are shown around the graph.

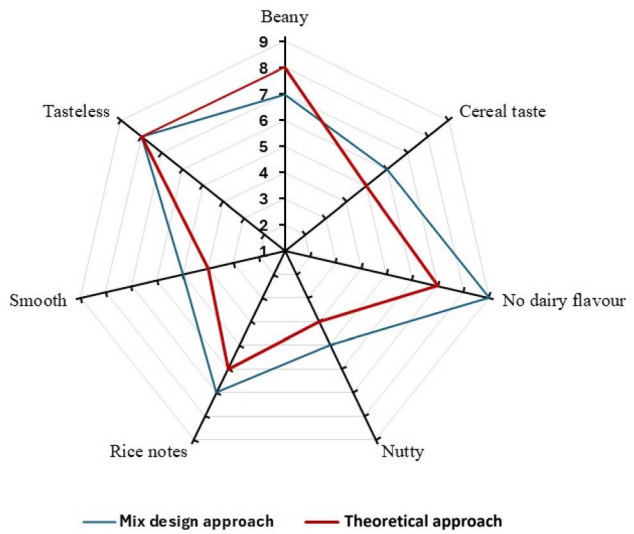

Fig. 6 Sensorial evaluation of optimal plant-based beverage formulations (TMO and DoE, final samples). The blue line shows the DoE approach, while the red solid line indicates TMO. Descriptors from panelists are displayed around the graph.

approach. They rated the plant-based beverage as less acceptable, noting characteristics such as poor flavor and bitter aftertaste. The mixture design formula demonstrates a balance of notes, including cereal, intensity, and smoothie, with a distinct rice note standing out (Fig. 6). Conversely, the theoretical formula features a simple, dairy-free beany note. Overall, the panelists prefer the mixture design approach (60% rice, 28% pea, and 12% almond), highlighting the legume notes as predominant.

As noted, the findings indicate that consumers did not notice a significant difference between formulae based on theoretical and experimental methods. This suggests that the proposed approach could reduce the need for extensive experimental testing. Theoretical modeling can significantly reduce experimental efforts, resulting in less raw material use, faster formulation processes, lower energy consumption, and reduced waste, making it a sustainable alternative for beverage development.

4. Conclusions

Two case studies focused on mixture optimization were compared, highlighting the use of theoretical and experimental designs for beverage formulation development. Mixture design facilitated the evaluation of physicochemical interactions at varying proportions, allowing for the estimation of optimal blends. The theoretical framework helped define the constraint space, reducing the search area for feasible formulations and requiring minimal experimental trials. Additionally, sensory evaluations provided qualitative insights into the similarities and differences between formulae derived from both theoretical and experimental approaches. The proposed theoretical approach may enhance product formulation while promoting agile, cost-effective, and sustainable food design. Future work could involve enhancing the theoretical approach by improving the property estimation, incorporating additional properties, and utilizing larger databases. The methodology could also be expanded and refined to handle more complex cases, such as dispersion systems like emulsions, solid foams, and gels. Furthermore, food formula development is an open field for the implementation of artificial intelligence tools and big data analytics, broadening the scope of computational-aided food design.

Author contributions

Project conceptualization and supervision NRC, ALM, and EPG; RNG and ALM performed formal analysis and data curation; RNG and EPG wrote the original draft; ALM and NRC reviewed and edited the main manuscript. All authors reviewed the manuscript.

Conflicts of interest

The authors declare that they have no conflict of interest.

Data availability

All data have been included in the manuscript and supplementary information (SI). Supplementary information: model parameters for property estimation. See DOI: **https://doi.org/10.1039/d5fb00480b**.

Acknowledgements

The author, Navarrete-González, gratefully acknowledges the financial support for her PhD studies from the Universidad de las Américas Puebla (UDLAP) and the Secretariat of Science, Humanities, Technology, and Innovation (SECIHTI) of Mexico.

References

1 J. A. Arrieta-Escobar, M. Camargo, L. Morel, F. P. Bernardo, A. Orjuela and L. Wendling, Design of formulated products integrating heuristic knowledge and consumer assessment, *AIChE J.*, 2021, **67**(4), e17117.
2 O. Musina, P. Putnik, M. Koubaa, F. J. Barba, R. Greiner, D. Granato, *et al.*, Application of modern computer algebra systems in food formulations and development: A case study, *Trends Food Sci. Technol.*, 2017, **64**, 48–59.
3 A. Gupta, N. Sanwal, M. A. Bareen, S. Barua, N. Sharma, O. Joshua Olatunji, *et al.*, Trends in functional beverages: Functional ingredients, processing technologies, stability, health benefits, and consumer perspective, *Food Res. Int.*, 2023, **170**, 113046.
4 G. KS and J. A. John, Functional beverages: Special focus on anti-diabetic potential, *J. Food Process. Preserv.*, 2021, **45**(11), e15974.
5 S. Li, S. Chu, J. Lu, P. Wang and M. Ma, Molecular and structural properties of three major protein components from almond kernel, *J. Food Process. Preserv.*, 2018, **42**(3), e13536.
6 M. Butu and S. Rodino, 11 – Fruit and Vegetable-Based Beverages—Nutritional Properties and Health Benefits, in *Natural Beverages*, ed. A. M. Grumezescu and A. M. Holban, Academic Press, 2019, pp. 303–38, [cited 2025 Oct 8], available from: **https://www.sciencedirect.com/science/article/pii/B9780128166895000110**.
7 B. Hidalgo-Fuentes, E. de Jesús-José, A. D. J. Cabrera-Hidalgo, O. Sandoval-Castilla, T. Espinosa-Solares, R. M. González-Reza, *et al.*, Plant-Based Fermented Beverages: Nutritional Composition, Sensory Properties, and Health Benefits, *Foods*, 2024, **13**(6), 844.
8 A. Xie, Y. Dong, Z. Liu, Z. Li, J. Shao, M. Li, *et al.*, A Review of Plant-Based Drinks Addressing Nutrients, Flavor, and Processing Technologies, *Foods*, 2023, **12**(21), 3952.
9 T. Prabhu, S. Shenbagavalli, R. Rajamanickam, M. Manivannan, A. Baskaran and J. Rajangam, Unlocking the full potential of functional beverages in medicinal and aromatic plants, *Ann. Phytomed. An. Int. J.*, 2024, **13**, 337–346.
10 A. Chopra, J. Jagose and A. Pandey, Health is wealth-eating for tomorrow: factors influencing purchase intention of plant-based diets in India, *Cogent Bus. Manag.*, 2025, **12**(1), 2443567.
11 Global Growth Insights|Market Research Report & Consulting, [Internet]. [cited 2025 Oct 8], available from: **https://www.globalgrowthinsights.com/**.
12 D. J. McClements, Development of Next-Generation Nutritionally Fortified Plant-Based Milk Substitutes: Structural Design Principles, *Foods*, 2020, **9**(4), 421.
13 S. Sethi, S. K. Tyagi and R. K. Anurag, Plant-based milk alternatives an emerging segment of functional beverages: a review, *J. Food Sci. Technol.*, 2016, **53**(9), 3408–3423.
14 O. M. A. Ogundele, O. O. Awolu, A. A. Badejo, I. D. Nwachukwu and T. N. Fagbemi, Development of functional beverages from blends of Hibiscus sabdariffa extract and selected fruit juices for optimal antioxidant properties, *Food Sci. Nutr.*, 2016, **4**(5), 679–685.
15 M. B. Kim, J. Y. Ko and S. B. Lim, Formulation Optimization of Antioxidant-Rich Juice Powders Based on Experimental Mixture Design, *J. Food Process. Preserv.*, 2017, **41**(3), e12897.
16 A. A. Chekdid, C. J. F. Kahn, E. Prévot, M. Ferrières, B. Lemois, C. Choquet, *et al.*, Mixture design applied for formulation and characterization of vegetal-based fermented products, *LWT–Food Sci. Technol.*, 2021, **146**, 111336.
17 Y. Seifu, T. Birhanu, B. Abera and R. Duraisamy, Peanut, oat, and chickpea composite plant-based milk alternatives: Formulation, optimization, and evaluation of nutritional, microbial, and shelf-life characteristics, *Appl. Food Res.*, 2025, **5**(1), 101035.
18 J. D. Gomez, D. Pradilla and O. Alvarez, A Multiscale Approach to the Design and Manipulation of Oil-in-Water Emulsion-Based Products, *Int. J. Chem. Eng.*, 2021, **2021**, e8897983.
19 F. Erdogdu, O. Altin, F. Marra and T. F. Bedane, A computational study to design process conditions in industrial radio-frequency tempering/thawing process, *J. Food Eng.*, 2017, **213**, 99–112.
20 J. Yeboah, A. M. Santoro, J. A. Arrieta-Escobar, I. M. Caballero, A. Orjuela, C. F. Novoa, *et al.*, Heuristic-based computer-aided design of ice creams and validation by using jaggery as refined sugar substitute, *Chem. Eng. Res. Des.*, 2022, **184**, 256–266.
21 A. K. Datta, Toward computer-aided food engineering: Mechanistic frameworks for evolution of product, quality and safety during processing, *J. Food Eng.*, 2016, **176**, 9–27.
22 F. Calvo, O. Álvarez, J. M. Gómez and L. Ricardez-Sandoval, Optimal design and experimental validation of emulsified cosmetic products: A multiscale approach, *Chem. Eng. Process. Process Intensif.*, 2024, **197**, 109711.
23 S. Ghosh, A. Khuntia and A. Kadival, Thermodynamic and mathematical background of water activity and its significance from microbiological point of view for food products, in *Trends & Prospects in Food Technology, Processing and Preservation*, Today & Tomorrow's Printers and Publishers, Delhi, India, 2018, pp. 1–12.
24 A. J. Maneffa, R. Stenner, A. S. Matharu, J. H. Clark, N. Matubayasi and S. Shimizu, Water activity in liquid food systems: A molecular scale interpretation, *Food Chem.*, 2017, **237**, 1133–1138.
25 R. S. Norrish and G. H. Twigg, Single-Constant Linear Equation for Vapor-Liquid Equilibria in Binary Systems, *Ind. Eng. Chem.*, 1954, **46**(1), 201–204.
26 J. Chirife and G. J. Favetto, Some physico-chemical basis of food preservation by combined methods, *Food Res. Int.*, 1992, **25**(5), 389–396.
27 M. Bourne, *Food Texture and Viscosity: Concept and Measurement*, Elsevier, 2002, p. 446.
28 W. Marczak, N. Adamczyk and M. Łężniak, Viscosity of Associated Mixtures Approximated by the Grunberg-Nissan Model, *Int. J. Thermophys.*, 2012, **33**(4), 680–691.

29 M. A. Magerramov, A. I. Abdulagatov, N. D. Azizov and I. M. Abdulagatov, Pressure- and Temperature-Dependent Density Change of Juices During Concentration, *Food Bioprocess Technol.*, 2008, **1**(3), 254–269.
30 D. M. Phinney, J. C. Frelka and D. R. Heldman, Composition-Based Prediction of Temperature-Dependent Thermophysical Food Properties: Reevaluating Component Groups and Prediction Models, *J. Food Sci.*, 2017, **82**(1), 6–15.
31 C. Tyl and G. D. Sadler, pH and Titratable Acidity, in *Food Analysis*, ed. S. S. Nielsen, Springer International Publishing, Cham, 2017, pp. 389–406, [cited 2025 Feb 28], available from: DOI: **10.1007/978-3-319-45776-5_22**.
32 A. Cosson, D. Blumenthal, N. Descamps, I. Souchon and A. Saint-Eve, Using a mixture design and fraction-based formulation to better understand perceptions of plant-protein-based solutions, *Food Res. Int.*, 2021, **141**, 110151.
33 P. Yu, M. Y. Low and W. Zhou, Design of experiments and regression modelling in food flavour and sensory analysis: A review, *Trends Food Sci. Technol.*, 2018, **71**, 202–215.
34 G. E. Arteaga, E. Li-Chan, M. C. Vazquez-Arteaga and S. Nakai, Systematic experimental designs for product formula optimization, *Trends Food Sci. Technol.*, 1994, **5**(8), 243–254.
35 D. S. Taghian, Y. Zhang, B. Vardhanabhuti and A. Jan van der Goot, Enhancing textural properties in plant-based meat alternatives: The impact of hydrocolloids and salts on soy protein-based products, *Curr. Res. Food Sci.*, 2023, **7**, 100571.
36 H. Kim, H. I. Hwang, K. W. Song and J. Lee, Sensory and rheological characteristics of thickened liquids differing concentrations of a xanthan gum-based thickener, *J. Texture Stud.*, 2017, **48**(6), 571–585.
37 G. A. Camelo-Méndez, P. E. Vanegas-Espinoza, M. L. Escudero-Gilete, F. J. Heredia, O. Paredes-López and A. A. Del Villar-Martínez, Colorimetric Analysis of Hibiscus Beverages and their Potential Antioxidant Properties, *Plant Foods Hum. Nutr.*, 2018, **73**(3), 247–252.
38 S. Sahraee, B. Ghanbarzadeh and P. M. Falcone, Application of mixture design methodology for development of high antioxidant fruity functional beverage, *Food Sci. Nutr.*, 2022, **10**(7), 2245–2254.
39 M. M. Bradford, A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Anal. Biochem.*, 1976, **72**(1), 248–254.
40 J. M. Lynch and D. M. Barbano, Kjeldahl Nitrogen Analysis as a Reference Method for Protein Determination in Dairy Products, *J. AOAC Int.*, 1999, **82**(6), 1389–1398.
41 ISO[Internet], [cited 2025 Oct 8], ISO 8586:2023. available from: **https://www.iso.org/standard/76667.html**.
42 C. Wrigley and R. Ramsey, Emotional food design: From designing food products to designing food systems, *Int. J. Food Des.*, 2016, **1**(1), 11–28.
43 DOF – Diario Oficial de la Federación, [Internet], [cited 2025 May 30]. available from: **https://www.dof.gob.mx/nota_detalle_popup.php?codigo=5643986**.
44 G. Main, M. Faupel, J. Morris and R. Mcnew, Quality and Stability of Blueberry Juice Blended with Apple, Grape and Cranberry Juice, *J. Food Qual.*, 2001, **24**(2), 111–125.
45 L. Diamante and M. Umemoto, Rheological Properties of Fruits and Vegetables: A Review, *Int. J. Food Prop.*, 2015, **18**(6), 1191–1210.
46 F. Salehi, Physicochemical characteristics and rheological behaviour of some fruit juices and their concentrates, *J. Food Meas. Char.*, 2020, **14**(5), 2472–2488.
47 L. M. Jiménez-Munoz, G. M. Tavares and M. Corredig, Design future foods using plant protein blends for best nutritional and technological functionality, *Trends Food Sci. Technol.*, 2021, **113**, 139–150.
48 R. Karoui and I. Bouaicha, A review on nutritional quality of animal and plant-based milk alternatives: a focus on protein, *Front. Nutr.*, 2024, **11**, 1378556.
49 D. J. McClements, E. Newman and I. F. McClements, Plant-based Milks: A Review of the Science Underpinning Their Design, Fabrication, and Performance, *Compr. Rev. Food Sci. Food Saf.*, 2019, **18**(6), 2047–2067.
50 M. Kumar, M. Tomar, J. Potkule, P. S. Reetu, J. Dhakane-Lad, *et al.*, Functional characterization of plant-based protein to determine its quality for food applications, *Food Hydrocoll.*, 2022, **123**, 106986.
51 A. Schmidt, L. Briceno-Mena, S. Rajagopalan, K. Ma, B. Reiner and B. Braun, Introduction to Formulation Optimization, in *The Digital Transformation of Product Formulation*, CRC Press, 2024.
52 A. Sabra, T. Netticadan and C. Wijekoon, Grape bioactive molecules, and the potential health benefits in reducing the risk of heart diseases, *Food Chem.: X*, 2021, **12**, 100149.
53 M. B. Vallée, M. Verheyde, S. Pomerleau, A. Doyen and C. Couillard, Health Benefits of Apple Juice Consumption: A Review of Interventional Trials on Humans, *Nutrients*, 2022, **14**(4), 821.
54 K. Wolfe, X. Wu and R. H. Liu, Antioxidant Activity of Apple Peels, *J. Agric. Food Chem.*, 2003, **51**(3), 609–614.
55 E. Pappas and K. M. Schaich, Phytochemicals of Cranberries and Cranberry Products: Characterization, Potential Health Effects, and Processing Stability, *Crit. Rev. Food Sci. Nutr.*, 2009, **49**(9), 741–781.
56 G. Sarwar and R. W. Peace, Comparisons between True Digestibility of Total Nitrogen and Limiting Amino Acids in Vegetable Proteins Fed to Rats, *J. Nutr.*, 1986, **116**(7), 1172–1184.