Scifood
vol. 20, 2026, p. 135-145
https://doi.org/10.5219/scifood.101
ISSN: 2989-4034 online
https://scifood.eu
© 2026 Authors, License: CC BY-NC-ND 4.0
Received: 8.1.2026
Revised: 15.2.2026
Accepted: 24.2.2026
Published: 25.2.2026

Tofu processing: selected parameters from the seeds to the final products and by-products

Daniela Lukáčová, Martina Fikselová, Radoslav Židek, Jozef Čapla, Peter Zajác

ABSTRACT

The efficiency of soy food production, particularly tofu, depends on a complex of physicochemical properties of soybean seeds (*Glycine max*). This study analyzes 187 soybean samples to evaluate the primary determinants of tofu yield/production. The main objective was to analyze the relationships among selected physicochemical properties of soybeans and tofu production. The methods applied were Near-Infrared Spectroscopy (NIR), physical parameters (hydration coefficient, thousand seed weight [TSW]), and correlation and regression analyses to identify relationships among hydration potential, seed properties, dry matter content, and the amino acid profile. Our results show that the hydration potential is strongly positively associated with seed dry matter content, indicating that dry matter is a critical predictor of water absorption efficiency. Raw tofu yield demonstrated a clear positive correlation with the physical properties of the seeds, including TSW and the hydration coefficient. An original contribution of this work is the identification of a negative amino-acid signature associated with reduced tofu yield. Aspartic acid, histidine, alanine, and glycine showed significant negative correlations with tofu raw yield, whereas proline content showed a slight positive association.

Keywords: soybean, tofu, yield, amino acids, hydration

INTRODUCTION

Soybean *(Glycine max)* represents an important commodity in the global food system due to its high protein and oil content across both food and feed value chains. Global soybean production in 2024/25 exceeded 400 Mt, reflecting increased expansion and reinforcing soybeans as a strategic raw material for protein supply **[1]**. Tofu remains one of the most important soy-based products, and its processing performance depends strongly on raw-material behavior during early unit operations - particularly soaking and hydration. It was demonstrated that higher soaking temperatures accelerated water absorption (hydration coefficient), but increased solids loss **[2]**. Another study showed that higher water absorption in soybean seeds significantly increased extraction yields of protein (and other components) into soy extract, establishing hydration as a gateway step that can affect downstream processing efficiency **[3]**. It was also identified that soybean water uptake during soaking is affected by measurable factors (including seed-related properties and conditions), supporting the concept that hydration behavior is not random noise but a technologically relevant trait **[4]**. In controlled soaking experiments, the water absorption rate of soybean seeds increased with higher soaking temperature and pH, while the equilibrium value was comparatively stable **[3]**. Hydration performance is also sensitive to pre-treatments in food technology to stabilize processing and reduce safety risks. Results on soybean soaking showed that pre-blanching had been especially beneficial at lower soaking temperatures (25–50 °C) by decreasing initial microbial counts and shortening the soaking time required, illustrating how unit operations can be tuned for both, safety and performance **[5]**. Storage under high temperature and humidity is associated with the "hard-to-cook" (HTC) phenomenon. These conditions have been experimentally linked to the increased lipid peroxidation, supporting the idea that storage-induced biochemical changes can co-occur with functional deterioration relevant to hydration and processing quality **[6]**. Downstream of soaking/extraction, tofu yield is strongly affected by the thermal history and curd handling. Mechanical dewatering process further modulates yield, pressing trials demonstrated that increasing the applied pressure decreased yield as water drained from tofu, while simultaneously shifting texture (e.g., harder gels under higher loads) **[7]**. Bhardwaj et al. (1999) found that tofu yield was significantly correlated with seed size, seed oil, and the rate of water absorption at 1 h and 16 h of soaking. They also reported a strong negative correlation between tofu yield and seed protein content and identified seed protein content, seed size, and soy extract solids as significant determinants of tofu yield in a stepwise regression model ($R^2 \approx 0.80$) **[8]**. Protein composition, beyond total protein content, has demonstrable effects on tofu gel properties. Tofu-making experiments using crude storage-protein components showed clear functional differences between gels prepared from crude 11S and 7S fractions: gels derived from crude 11S were markedly harder than those from crude 7S. These findings further suggest that varietal variation in the 11S/7S proportion can translate into measurable differences in tofu gel physical properties **[9]**. Quantitative mass-balance data from tofu processing show that a substantial fraction of soybean solids and protein remains in the water-insoluble residue (okara). Across five fractions generated during manufacture (soak water, okara, soy extract, tofu, and whey), the residue can account for approximately 30% of bean solids and 20% of bean protein, indicating that incomplete extraction and separation directly reduce the amount of protein available for curd formation. These observations support the interpretation that improving hydration and extraction efficiency should reduce protein retention in okara, thereby increasing tofu yield **[10]**. Rapid analytical technologies are increasingly used to screen incoming lots and support data-driven process control. Near-infrared spectroscopy modeling on a large soybean sample set achieved high predictive performance for moisture, fat, and protein, supporting NIR as a feasible tool for rapid compositional assessment relevant to processing decisions **[11]**. Building on established links among hydration behavior, compositional status, and processing performance, this model study analyzes soybean sample data, focusing on hydration metrics and yield outcomes to identify practical predictors of soaking efficiency and raw yield under standardized conditions.

Scientific Hypothesis

We hypothesize that selected bean parameters, including dry matter content and 1000-seed weight (TSW), affect hydration kinetics, and that a specific amino acid profile serves as a biochemical marker for predicting raw tofu yield. In particular, variation in the amount/portion of amino acids is expected to correlate positively or negatively with tofu yield.

Objectives

The main objective of this model study is to analyze the relationships between the physicochemical properties of soybeans and the technological parameters of tofu production, using 187 bean samples. Specific objectives include:

1. To quantify the relationship between the hydration coefficient and the dry matter content of the seeds.
2. To determine the correlation between the raw yield of tofu and physical properties such as hydration coefficient and 1000-seed weight.
3. To assess the amino acid profile of the seeds and to determine which specific amino acids exhibit positive/ negative associations with processing performance.

MATERIAL AND METHODS

Samples description: The analysis included 187 samples (1 sample = 1 kg of seeds) of soybean *(Glycine max),* comprising both conventional and organic production, originating from the Danube region (Donau Soya certified).

Sample collection: Soybean seeds were collected from commercial deliveries during the 2024/2025 harvest and stored at room temperature prior to analysis (max. 24 hours).

Samples preparation: Before analysis, samples were cleaned of impurities and foreign bodies. For NIR analysis, samples were used as whole grains.

Number of samples analysed: A total of 187 samples of bean seeds were analyzed and used for the tofu production.

Chemicals

Food-grade coagulants were used for soy extract coagulation: calcium chloride ($CaCl_2$) and magnesium chloride ($MgCl_2$) in a 5:1 Ca:Mg ratio. Cation-exchanged water was used for soaking.

Animals, Plants and Biological Materials

soybean (*Glycine max*) seeds.

Instruments

Determinations of the dry matter and amino acid content were performed using NIR (Near-Infrared Spectroscopy) at the TANGO spectrometer (Bruker Optik GmbH, Germany).

Physical measurements (thousand seed weight -TSW) were performed using a Class II precision scale Kern 572-35 (Kern & Sohn GmbH, Germany). Range: 0,2 – 2400 g; d= 0,01 g.

Laboratory Methods

NIR Spectroscopy: the sample was placed in a Petri dish and placed on the instrument's measuring head. Spectrum measurement was initiated; the TANGO instrument illuminated the sample in the NIR region and measured the reflected light intensity as a function of wavelength. Prediction (Calculation): collected spectral data were digitally transmitted to the Evonik AMINONIR® analytical platform. The platform processes the spectra using a robust global calibration model based on extensive reference data. The system then returns predicted values for parameters: dry matter content and the complete amino acid profile.

Hydration coefficient determination: 250 g soybean sample was soaked in water at ambient temperature (approx. 20-22 °C) for 7 hours. The coefficient was calculated as the ratio of the weight of soaked beans to the weight of dry beans.

Thousand seed weight (TSW): was determined by manually counting the seeds, followed by weighing, using the Class II precision scale Kern 572-35 (Kern & Sohn GmbH, Germany).

Tofu production: the process involves grinding the soaked beans with water, cooking the slurry, filtering to obtain soy extract, and coagulation. The curd was pressed in a standardized mold. Raw yield of tofu was calculated as the ratio of fresh tofu weight / dry soybean weight (soybean weight of seeds used for hydration was 250 g).

Description of the Experiment

Study flow: The experimental design was divided into three operational phases spanning the workflow from raw material to finished product data. In Phase 1 (raw material characterization), samples were cleaned and prepared before undergoing non-destructive chemical analysis using a TANGO Bruker NIR spectrometer to establish baseline-independent variables, including dry matter content and the quantitative amino acid profile. 17 individual amino acids plus 1 combined variable: methionine (MET), cysteine (CYS), methionine+cysteine (MET_CYS), lysine (LYS), threonine (THR), tryptophan (TRP), arginine (ARG), isoleucine (ILE), leucine (LEU), valine (VAL), histidine (HIS), phenylalanine (PHE), glycine (GLY), serine (SER), proline (PRO), alanine (ALA), aspartic acid/aspartate (ASP), and glutamic acid/glutamate (GLU) were observed. At the same time, selected physical properties of the seeds were determined. In Phase 2 (standardized processing trial), fixed input mass 250 g of each previously characterized sample was processed using an identical laboratory-scale tofu production protocol to control process variability: beans were soaked for 7 hours to achieve uniform conditions, then ground and extracted to produce soy extract, after which coagulation was induced with a defined mineral coagulant mixture ($CaCl_2$+$MgCl_2$) at controlled temperature of 86 °C. The resulting curd was pressed in a standardized mold, and the fresh tofu's weight was recorded immediately in order to calculate the raw yield. In Phase 3 (Data Integration and Statistical Modeling), the biological data from Phase 1 were paired with the technological outcomes from Phase 2, followed by data cleaning in which samples producing zero values due to technical processing failures were excluded; the resulting valid dataset was then analyzed using Spearman correlation with False Discovery Rate (FDR).

Quality Assurance

Number of repeated analyses: NIR measurements were performed as three scans per sample.

Number of experiment replications: The products (tofu) were measured in triplicate.

Calibration: NIR system utilizes the Evonik AMINONIR® global calibration network. Accuracy is ensured by a continuous "ring test" system that aggregates data from multiple parties and reference laboratories to update and validate the central calibration curves, eliminating the need for local standard-set calibration.

Scale: The Kern 572-35 precision balance was calibrated by an accredited laboratory. Calibration was performed on October 16, 2024 (Certificate No. 4261/321.21/24, provided by the Slovak Legal Metrology, with a validity period extending to October 16, 2026.

Laboratory accreditation: Experiments were performed in a commercial laboratory with no accreditation.

Data Access

Data are available to the first author of this article upon reasonable request due to privacy constraints regarding supplier trade secrets.

Statistical Analysis

Data were processed using Microsoft Excel. Descriptive statistics (mean μ, standard deviation σ) were computed for all quantitative variables in order to assess data distribution. For the construction of the predictive model, key variables were standardized using Z-score normalization: $z = (x - \mu)/\sigma$.

Bivariate relationships were evaluated using both Pearson's correlation coefficient (r) for linear relationships among physical traits and Spearman's rank correlation coefficient (ρ) for non-linear or monotonic relationships, particularly between amino acid profiles and the yield of tofu. Statistical significance of correlations was determined using the t-test approximation. To rigorously control the multiple comparisons problem inherent in screening 18 different amino acids, the False Discovery Rate (FDR) was calculated using the Benjamini-Hochberg (BH) procedure. Significance was defined as an FDR-adjusted q-value < 0.05.

Reporting and transparency statement

This study was observational, randomized, and blinding was not applicable as it involved the analysis of incoming commercial samples. All samples with complete data were included in the analysis (samples with zero values were excluded as technical errors).

RESULTS AND DISCUSSION

Analysis of 187 soybean samples revealed several key correlations. The Spearman correlation between hydration and dry matter content is $r_s \approx 0.402$, whereas the Pearson correlation is $r_p \approx 0.356$. This statistically significant positive correlation suggests that seeds with higher dry matter content tend to achieve higher hydration coefficients during soaking. The linear model describing this relationship is: hydration coefficient = (0.0283 × dry matter) – 0.519, with $R^2 \approx 0.127$. Although R^2 is suggesting that other factors also play a role (e.g. seed coat permeability, seed structure, initial damage, temperature), dry matter is identified in the analysis as the strongest marker of hydration **[12]**. The results of the experiment are shown in Figure 1.

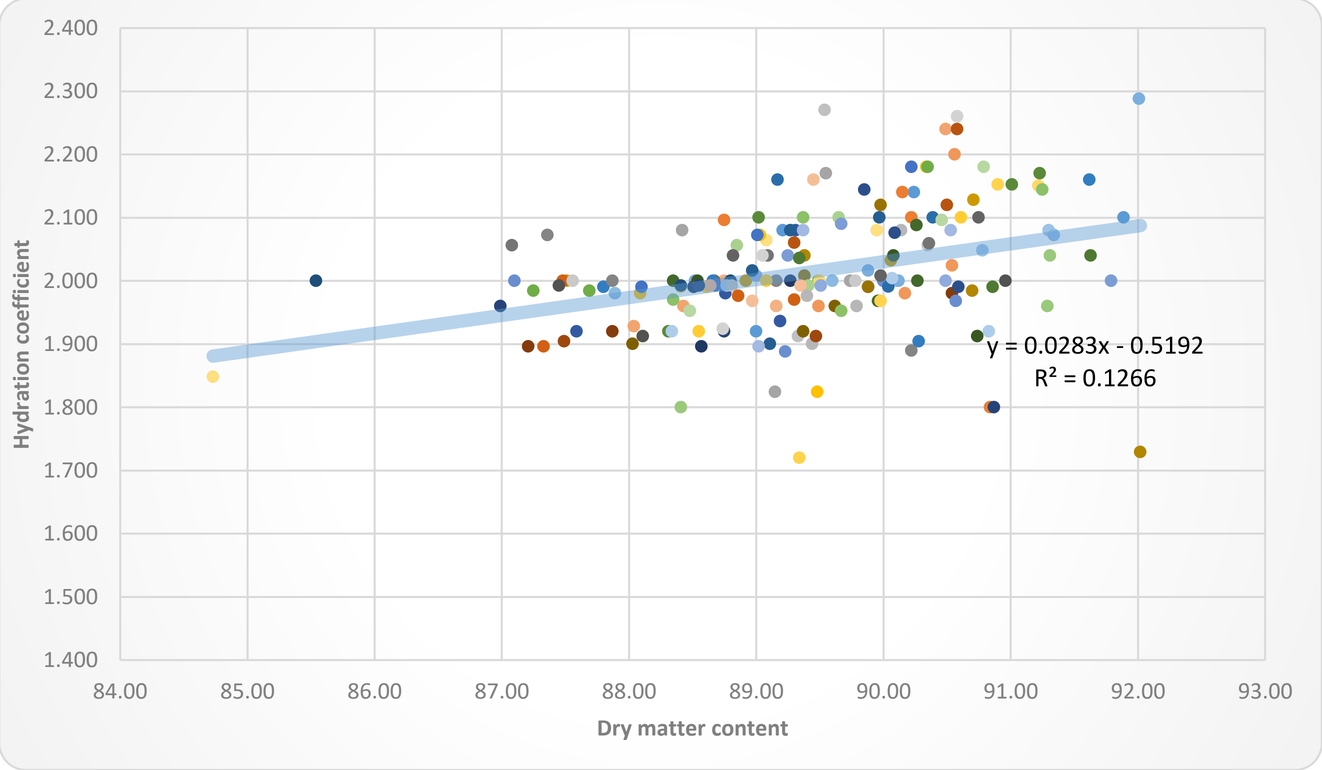

Figure 1 Results of regression analysis.

Note: Results of the regression analysis illustrating the relationship between the hydration coefficient and dry matter content. The scatter plot with the fitted linear regression curve shows a positive trend, indicating that higher dry matter content is associated with greater water absorption. This supports the role of dry matter content as a key marker of hydration intensity, reflecting the stronger imbibition driving force in seeds with lower initial moisture and highlighting the influence of intrinsic seed structure and composition on absorption behaviour.

Figure 1 shows a positive trend: higher dry matter content is associated with greater water absorption. This visually supports the claim that dry matter content is the strongest marker of hydration. Seeds with higher initial dry matter content (lower moisture) show stronger imbibition driving force and absorb more water during soaking. Researchers also linked absorption behavior to physical traits such as seed density, indicating that hydration is closely tied to intrinsic seed structure and composition **[4]**. The positive dry matter–hydration relationship observed is consistent with soybean hydration literature, showing that water absorption behavior is affected by the seed's initial moisture status (i.e., dry matter), such that initial moisture level measurably affects water uptake during soaking **[13]**. Water absorption significantly increased the extraction yields of protein, fat, and solids into soy extract, directly linking hydration behavior to extraction performance. In addition, the profile of proteins leached during soaking can shift with temperature. For example, MALDI-TOF/TOF-MS evidence showed increased release of basic 7S globulin at high temperature, indicating that hydration conditions can change which protein fractions enter the processing stream **[3]**. Seed hydration capacity is strongly variety dependent and is further modulated by moisture regime and soil type **[14]**. Differences in soybean raw material have been shown to translate into measurable differences in tofu performance and composition. For example, comparing non-GMO and GMO soybean varieties, higher tofu yield and protein content were reported for the non-GMO material, while sensory attributes and amino-acid profiles also differed between varieties **[15]**. High moisture during storage can lead to the "hard-to-cook" phenomenon or to lipid oxidation, both of which prevent water absorption **[2]**. Our dataset confirms this, showing that effective hydration is a necessary condition for yield. Okara (water-insoluble residue) accounted for approximately 29,6% of the original bean dry matter, 20,2% of the original protein, and 11,4% of the original oil, providing direct evidence that a substantial fraction of material remains outside the curd and therefore limits potential tofu yield. At the same time, this mass balance highlights okara as a nutrient-rich by-product, creating clear opportunities for its industrial valorization (e.g., as a functional ingredient) rather than treating it solely as waste **[10]**.

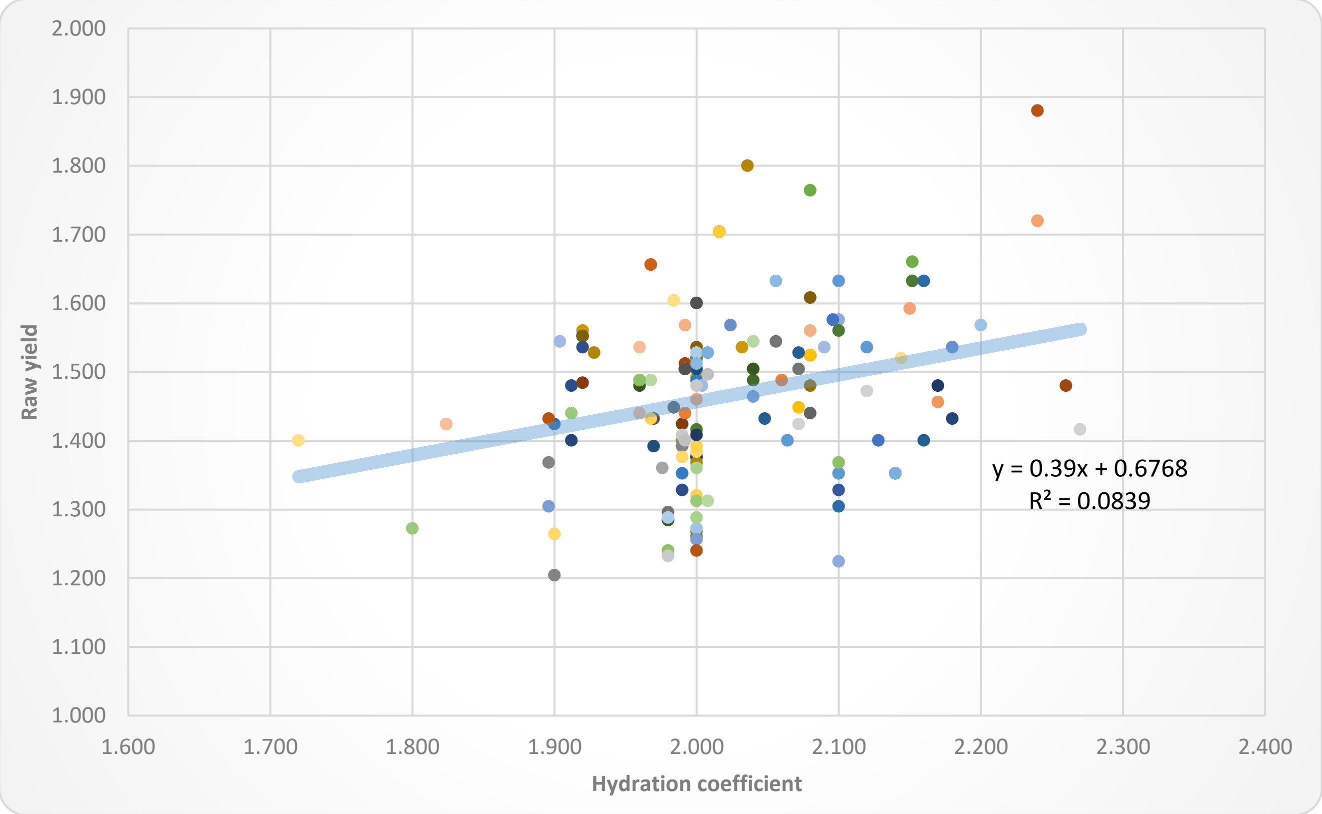

Figure 2 Results of regression analysis.

Note: The regression analysis shows a dependence of raw yield on the hydration coefficient. The solid line represents the fitted linear regression model.

Analysis of 169 samples (selected from a total of 187 measurements after excluding zero-yield measurements to maintain statistical integrity) revealed relationships between yield and the hydration coefficient. The Spearman correlation between raw yield and the hydration coefficient is $r_s \approx 0.256$, while the Pearson correlation reaches a value of $r_p \approx 0.290$. This statistically significant positive correlation suggests that seeds with a higher hydration coefficient during soaking tend to yield higher raw yields in subsequent processing. Although the $R^2 \approx 0.084$ value suggests that other factors also affect the final yield (e.g., variety genotype, extraction efficiency of components into soy milk, grinding fineness, or specific amino acid profile), the analysis confirms that the hydration coefficient is a valuable technological indicator. A higher degree of hydration directly correlates with an increase in the processed mass's volume. The results of the experiment are shown in Figure 2.

Figure 2 illustrates the regression analysis of the relationship between two key variables. The light blue regression line clearly indicates a positive correlation: as the hydration coefficient increases, the raw yield tends to rise. This visually confirms our statistical finding that efficient water absorption is a fundamental prerequisite for high yield. In the low-hydrate region (below 1,900), there are few samples with a high yield (above 1,500). This suggests that if the seed does not swell sufficiently, it is physically nearly impossible to achieve a peak yield, regardless of protein quality. Figure 2 shows that optimizing the soaking process stabilizes production, but achieving premium yields requires selecting varieties with a suitable biochemical profile, which accounts for the vertical variance between data points at identical hydration levels. These findings are consistent with several authors who show that hydration/soaking is a gateway to yield because it directly affects protein extraction into soy extract. For example, Li et al. (2019) demonstrated that higher water absorption in seeds significantly increases protein extraction yields in soy extract **[3]**. From a practical processing perspective, this is also supported by the review by Guan et al. (2021): properly optimized soaking conditions lead to higher protein extraction rates and higher protein content in soy extract, which, in turn, improve gel network formation and the water-holding capacity of tofu **[16]**. A comparative cooking study reported that high-temperature pressure cooking increased tofu yield and improved texture, attributing this to the incorporation of more water or dry matter into the gel network, alongside altered coagulation behavior **[17]**. Our positive hydration–yield relationship supports the interpretation that insufficient swelling limits the transfer of extractable protein into soy extract, thereby constraining curd mass. This mechanism is consistent with independent tofu datasets, in which incomplete hydration is captured by ‘hard beans’ (seeds that fail to absorb water), which are associated with reduced soy extract weight and lower tofu yield; i.e., hydration failure propagates downstream through reduced soy extract mass and ultimately reduced curd yield **[18]**. In bench-scale tofu manufacture, soy extract is separated from the water-insoluble residue (okara) by vacuum filtration, and factorial processing trials show that downstream parameters (thermal treatment, coagulant concentration, coagulation/draining conditions) significantly affect tofu yield and hardness, underscoring that yield is strongly modulated by processing conditions in addition to raw-material composition, which directly influences how much material is retained in okara/residue versus how much remains in the soy extract stream and is subsequently captured in the curd **[19]**. Our data show that hydration is associated with raw tofu yield. Consistent with this processing logic, independent work has shown that hydration metrics (e.g., water uptake, volume change, leaching loss) significantly correlate with soy extract yield and can predict it, supporting hydration behavior as a practical upstream predictor of yield-related performance **[20]**. Okara has also been shown to affect tofu yield directly when incorporated into the formulation. In okara-supplemented tofu, yield increased almost linearly with increasing okara ratio. Our results show that hydration is a key upstream constraint on raw tofu yield, supporting the interpretation that “raw yield” is highly sensitive to both extraction efficiency and to water-insoluble-solids retention during curd formation and dewatering **[21]**.

Among our results, we identified specific amino acids that negatively correlate with tofu yield. Spearman correlation analysis identified a set of amino acids that showed a significant negative association with raw tofu yield (Raw Yield). Specifically, aspartic acid (ASP), histidine (HIS), alanine (ALA), and glycine (GLY) exhibited consistently negative correlation coefficients. Low false discovery rate values (q; FDR), indicate that these amino acids act as negative markers of tofu yield (Table 1).

Table 1 Spearman correlations of amino acids versus raw yield of tofu.

Amino acid	Spearman r (vs. Raw Yield)	q-value (FDR)	Interpretation
Aspartic acid (ASP)	-0.271	0.0065	moderate negative marker
Histidine (HIS)	-0.253	0.0079	moderate negative marker
Alanine (ALA)	-0.240	0.0099	moderate negative marker
Glycine (GLY)	-0.233	0.0099	moderate negative marker
Proline (PRO)	0.185	0.0562	weak positive marker

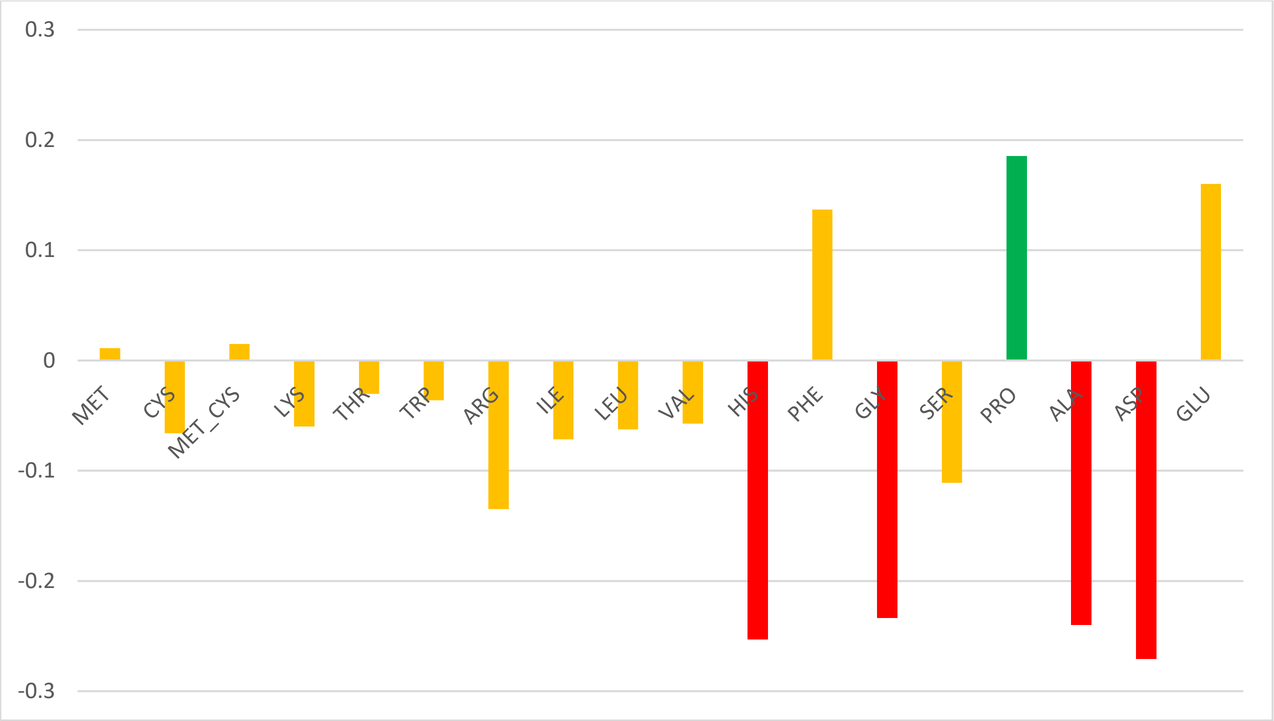

Figure 3 Correlation profile (Spearman) of amino acids and tofu raw yield.
Note: The analysis highlights negative associations of specific amino acids with raw yield, while proline shows a weak positive relationship and cysteine exhibits negligible correlation.

Figure 3 shows the visualization of the negative impact of specific amino acids on the tofu raw yield, which can be identified as markers of lower-quality protein fractions**.** The correlation profile visually confirms the negative effect of these specific amino acids, supporting the interpretation that their accumulation is associated with reduced Raw Yield. In contrast, proline (PRO) showed a weak positive association (r ≈ 0.185; q ≈ 0.057), suggesting a potential role as a weak positive marker. Cysteine (CYS), often discussed in the context of disulfide cross-linking, showed negligible direct correlations with Raw Yield in this dataset (r ≈ 0.07), implying either limited variability or masking by stronger, dominant drivers **[22]**. The finding that aspartic acid (ASP), histidine (HIS), alanine (ALA), and glycine (GLY) negatively significantly correlate with Raw Yield is best interpreted not as a direct causal effect of individual amino acids, but as a proxy signal for underlying differences in protein composition, particularly shifts in the contribution and behavior of the major soybean storage globulins (7S β-conglycinin and 11S glycinin). Glycinin (11S) and β-conglycinin (7S) constitute the dominant storage-protein fractions in soybean protein isolates, and glycinin is characterized by acidic and basic subunits linked by disulfide bonds. Amino-acid profiling of peptides derived from these globulins shows that several amino acids identified here as yield-associated markers (e.g., ASP, GLY, ALA) are prominent constituents of globulin-derived peptides, supporting their interpretation as compositional indicators related to underlying storage-protein profiles rather than direct causal drivers **[23]**. Histidine emerged as a negative yield-associated amino acid in our dataset, while independent multi-location seed-composition studies have shown that amino-acid profiles (including histidine and glutamate) can shift systematically with growing environment; thus, part of the observed amino-acid signal may reflect environment-driven compositional variation in addition to genetic differences **[24]**.

Seed-composition data support the view that amino-acid profiles can act as compositional indicators rather than direct causal drivers of processing outcomes. In a genotype series screening of 18 amino acids, several amino acids exhibited significant genotype-related shifts over release years, while proline and cysteine showed no genotype influence. Amino acids such as ASP, HIS, ALA, and GLY displayed stronger variability in the historical dataset **[25]**. Tofu yield has been reported to depend on storage-protein composition: SDS-PAGE densitometry across soybean varieties showed positive associations of yield with the 11S (glycinin) fraction and specific storage-protein polypeptides, supporting our interpretation that amino-acid signatures can act as proxies for underlying fraction shifts relevant to raw yield **[26]**. Amino-acid composition differed among soybean varieties, indicating a clear genotype effect on seed biochemical profiles **[27]**. Prior studies have shown that cultivar-dependent variation in storage-protein composition and subunit patterns affects processing functionality, such as solubility, thermal behavior, coagulation efficiency, and gel-network formation, ultimately determining tofu performance **[28]**, and **[29]**. These amino acids likely reflect compositional shifts in storage-protein fractions; however, confirmation would require direct quantification of 7S and 11S fractions.

The rising global production of tofu and soy extract has led to increased okara byproduct generation, creating a need for sustainable valorisation strategies to reduce environmental burdens **[30]**. Okara, as a by-product of soybean processing, produced in large quantities, is underused by the food industry. Its nutritional quality (e.g., when subjected to solid-state fermentation using S. cerevisiae) could improve the use of this raw material **[31]** in the food industry.

The final product of our experiments is shown in Figure 4.

Figure 4 Final tofu product obtained from soybeans.

Note: The figure shows the pressed tofu block (left) and a cross-section illustrating the internal structure (right).

Limitations

The main limitation of this study is the lack of direct quantification of protein fractions (11S/7S), which forces us to rely on the amino acid profile as a proxy. The experiment was performed using a laboratory method, which may not fully reflect the conditions of continuous industrial production involving high-pressure pressing and industrial manufacturing conditions.

CONCLUSION

This study provides a data-driven evaluation of soybean metrics that govern tofu yield and demonstrates that yield is controlled by the combined effect of seed physical properties and biochemical composition, rather than by protein content alone. The hydration coefficient is a technologically critical predictor and is strongly linked to seed dry matter content, establishing a predictable relationship that can be used to standardize raw material preparation. Dry matter content emerged as a practical marker of hydration potential, supported by a significant positive association with hydration (Spearman $r \approx 0{,}402$, Pearson $r \approx 0{,}356$, n = 187). Hydration also showed a positive relationship with raw tofu yield (Spearman $r \approx 0{,}256$, n = 169 tofu), consistent with the interpretation that insufficient swelling limits protein solubilization and transfer into soy extract, thereby increasing losses to okara or whey. In parallel, seed physical structure contributed to yield performance: 1000-seed weight was positively associated with tofu raw yield (Spearman $r \approx 0{,}256$, n = 169). An original contribution of this work is the identification of a negative amino-acid signature associated with reduced tofu yield. Aspartic acid, histidine, alanine, and glycine showed significant negative correlations with the raw yield of tofu that remained significant after FDR correction (ASP r = −0,271; HIS r = −0,253; ALA r = −0,240; GLY r = −0,233; q < 0,01; n = 187). These findings suggest that increased levels of these amino acids reflect compositional differences in storage-protein fractions, resulting in a less favorable profile for efficient coagulation and curd capture. By integrating NIR-derived composition with physical measurements and correlation–regression analyses, this paper provides an actionable framework for raw-material screening and selection in tofu manufacturing: maximize hydration efficiency (guided by dry matter) and prioritize varieties with a favorable biochemical profile to achieve premium yields and more stable processing outcomes.

REFERENCES

1. OECD-FAO Agricultural Outlook 2025-2034. In OECD-FAO Agricultural Outlook. OECD Publishing. https://doi.org/10.1787/601276cd-en

2. Wang, D., Chen, G., Yang, B., Chen, X., Song, J., Kong, X., & Kan, J. (2021). Kinetic study on soybean hydration during soaking and resulting softening kinetic during cooking. Journal of Food Science, 87(1), 266–279. https://doi.org/10.1111/1750-3841.15984

3. Li, X., Liu, X., Hua, Y., Chen, Y., Kong, X., & Zhang, C. (2019). Effects of water absorption of soybean seed on the quality of soymilk and the release of flavor compounds. RSC Advances, 9(6), 2906–2918. https://doi.org/10.1039/c8ra08029a

4. Hsu, K. H., Kim, C. J., & Wilson, L. A. (1983). Factors affecting water uptake of soybeans during soaking. Cereal Chemistry, 60(3), 208–211.

5. Gowen, A., Abu-Ghannam, N., Frias, J., & Oliveira, J. (2007). Influence of pre-blanching on the water absorption kinetics of soybeans. Journal of Food Engineering, 78(3), 965–971. https://doi.org/10.1016/j.jfoodeng.2005.12.009

6. Perera, D., Devkota, L., Shrestha, S., Bhattarai, S. P., Panozzo, J., & Dhital, S. (2026). Hard-to-cook (HTC) phenomenon associated changes in protein and lipid profile of faba and adzuki beans. Food Hydrocolloids, 171, 111803. https://doi.org/10.1016/j.foodhyd.2025.111803

7. Rekha, C. R., & Vijayalakshmi, G. (2011). Influence of processing parameters on the quality of soycurd (tofu). Journal of Food Science and Technology, 50(1), 176–180. https://doi.org/10.1007/s13197-011-0245-z

8. Bhardwaj, H. L., Bhagsari, A. S., Joshi, J. M., Rangappa, M., Sapra, V. T., & Rao, M. S. S. (1999). Yield and Quality of Soymilk and Tofu Made from Soybean Genotypes Grown at Four Locations. Crop Science, 39(2), 401–405. https://doi.org/10.2135/cropsci1999.0011183x0039000200017x

9. Saio, K., Kamiya, M., & Watanabe, T. (1969). Food Processing Characteristics of Soybean 11S and 7S Proteins. Agricultural and Biological Chemistry, 33(9), 1301–1308. https://doi.org/10.1080/00021369.1969.10859465

10. Wang, H. L., & Cavins, J. F. (1989). Yield and amino acid composition of fractions obtained during tofu production. Cereal Chemistry, 66(5), 359–361.

11. Zhu, Z., Chen, S., Wu, X., Xing, C., & Yuan, J. (2018). Determination of soybean routine quality parameters using near-infrared spectroscopy. Food Science & Nutrition, 6(4), 1109–1118. https://doi.org/10.1002/fsn3.652

12. Meyer, C. J., Steudle, E., & Peterson, C. A. (2006). Patterns and kinetics of water uptake by soybean seeds. Journal of Experimental Botany, 58(3), 717–732. https://doi.org/10.1093/jxb/erl244

13. Smith, A. K., Nash, A. M., & Wilson, L. I. (1961). Water absorption of soybeans. Journal of the American Oil Chemists' Society, 38(3), 120–123. https://doi.org/10.1007/bf02641230

14. Gourkhede, P. H., Baraptre, D. D., & Pendke, M. S. (2024). Studies on effect of seed hydration value of various soybean varieties with respect to soil types and moisture regimes. International Journal of Advanced Biochemistry Research, 8(4S), 479–491. https://doi.org/10.33545/26174693.2024.v8.i4sf.1035

15. Hendrawati, T. Y., Audini, K., Ismiyati, Ramadhan, A. I., & Gustia, H. (2021). Effects and characterization of different soybean varieties in yield and organoleptic properties of tofu. Results in Engineering, 11, 100238. https://doi.org/10.1016/j.rineng.2021.100238

16. Guan, X., Zhong, X., Lu, Y., Du, X., Jia, R., Li, H., & Zhang, M. (2021). Changes of Soybean Protein during Tofu Processing. Foods, 10(7), 1594. https://doi.org/10.3390/foods10071594

17. Zuo, F., Chen, Z., Shi, X., Wang, R., & Guo, S. (2016). Yield and textural properties of tofu as affected by soymilk coagulation prepared by a high-temperature pressure cooking process. Food Chemistry, 213, 561–566. https://doi.org/10.1016/j.foodchem.2016.07.008

18. Döttinger, C. A., Steige, K. A., Hahn, V., Bachteler, K., Leiser, W. L., Zhu, X., & Würschum, T. (2025). Unravelling the genetic architecture of soybean tofu quality traits. Molecular Breeding, 45(1). https://doi.org/10.1007/s11032-024-01529-x

19. Benassi, V. de T., Yamashita, F., & Prudencio, S. H. (2011). A statistical approach to define some tofu processing conditions. Food Science and Technology (Campinas), 31(4), 897–904. https://doi.org/10.1590/s0101-20612011000400011
20. Alpaslan, M., & Hayta, M. (2002). Hydration properties, soymilk and okara yield of soybean affected by agronomic factors. Nahrung/Food, 46(3), 141. https://doi.org/10.1002/1521-3803(20020501)46:3<141::aid-food141>3.0.co;2-8
21. Joo, K. H., Kerr, W. L., & Cavender, G. A. (2023). The Effects of Okara Ratio and Particle Size on the Physical Properties and Consumer Acceptance of Tofu. Foods, 12(16), 3004. https://doi.org/10.3390/foods12163004
22. Bainy, E. M., Corredig, M., Poysa, V., Woodrow, L., & Tosh, S. (2010). Assessment of the effects of soy protein isolates with different protein compositions on gluten thermosetting gelation. Food Research International, 43(6), 1684–1691. https://doi.org/10.1016/j.foodres.2010.05.010
23. Vasconcellos, F. C. S., Woiciechowski, A. L., Soccol, V. T., Mantovani, D., & Soccol, C. R. (2014). Antimicrobial and antioxidant properties of β-conglycinin and glycinin from soy protein isolate. *International Journal of Current Microbiology and Applied Sciences, 3*(8), 144–157.
24. McClure, T., Cocuron, J.-C., Osmark, V., McHale, L. K., & Alonso, A. P. (2017). Impact of Environment on the Biomass Composition of Soybean (Glycine max) seeds. Journal of Agricultural and Food Chemistry, 65(32), 6753–6761. https://doi.org/10.1021/acs.jafc.7b01457
25. de Borja Reis, A. F., Tamagno, S., Moro Rosso, L. H., Ortez, O. A., Naeve, S., & Ciampitti, I. A. (2020). Historical trend on seed amino acid concentration does not follow protein changes in soybeans. Scientific Reports, 10(1). https://doi.org/10.1038/s41598-020-74734-1
26. Mujoo, R., Trinh, D. T., & Ng, P. K. W. (2003). Characterization of storage proteins in different soybean varieties and their relationship to tofu yield and texture. Food Chemistry, 82(2), 265–273. https://doi.org/10.1016/s0308-8146(02)00547-2
27. Asrat, U., & Gudisa, A. (2025). Qualitative variabilities among different soybean (Glycine max (L.) Merrill) varieties. Cogent Food & Agriculture, 11(1). https://doi.org/10.1080/23311932.2025.2562166
28. Wu, C., Hua, Y., Chen, Y., Kong, X., & Zhang, C. (2016). Effect of 7S/11S ratio on the network structure of heat-induced soy protein gels: a study of probe release. RSC Advances, 6(104), 101981–101987. https://doi.org/10.1039/c6ra22388e
29. Dang, Y., Ren, J., Guo, Y., Yang, Q., Liang, J., Li, R., Zhang, R., Yang, P., Gao, X., & Du, S. (2023). Structural, functional properties of protein and characteristics of tofu from small-seeded soybeans grown in the Loess Plateau of China. Food Chemistry: X, 18, 100689. https://doi.org/10.1016/j.fochx.2023.100689
30. Rohrer, K., Whitfield, F., Aussanasuwannakul, A., Ningrum, A., Hugi, C., & Breitenmoser, L. (2025). Comparative Screening Life Cycle Assessments of Okara Valorisation Scenarios. *Environments*, *12*(3), 93. https://doi.org/10.3390/environments12030093
31. Santos, V. A. Q, Nascimento, C.G., Schmidt, C.A.P., Mantovani,D., Dekker,R.F.H., Cunha, M.A.A. (2018). Solid-state fermentation of soybean okara: Isoflavones biotransformation, antioxidant activity and enhancement of nutritional quality, LWT, 92, 509-515, https://doi.org/10.1016/j.lwt.2018.02.067

Funds:

This work was supported by the VEGA grant no. 1/0454/25.

Acknowledgments:

-

Competing Interests:

No potential conflict of interest was reported by the author(s).

Ethical Statement:

This article does not contain any studies that would require an ethical statement

AI Statement:

This article was written by the authors, with AI tools used solely for grammar and language correction.

Contact Address:

Daniela Lukáčová
Affiliation: The Slovak University of Agriculture in Nitra, Faculty of Biotechnology and Food Sciences, Department of Food Science, Tr. A. Hlinku 2, 949 76 Nitra, Slovakia,
Tel.: +421 37 641 5825
E-mail: dalukac@gmail.com
ORCID: https://orcid.org/0009-0003-0294-3694
Author contribution: methodology, writing – original draft, writing – review & editing.

Martina Fikselová
Affiliation: The Slovak University of Agriculture in Nitra, Faculty of Biotechnology and Food Sciences, Department of Food Science, Tr. A. Hlinku 2, 949 76 Nitra, Slovakia,
Tel.: +421 37 641 5825
E-mail: martina.fikselova@uniag.sk
ORCID: https://orcid.org/0000-0002-0962-3437
Author contribution: conceptualisation, writing – original draft, writing, conceptualisation

Radoslav Židek
Affiliation: The Slovak University of Agriculture in Nitra, Faculty of Biotechnology and Food Sciences, Department of Food Science, Tr. A. Hlinku 2, 949 76 Nitra, Slovakia,
Tel.: +421 37 641 5825
E-mail: radoslav.zidek@uniag.sk
ORCID: https://orcid.org/0000-0003-4751-1257
Author contribution: conceptualisation, writing – original draft.

Jozef Čapla
Affiliation: The Slovak University of Agriculture in Nitra, Faculty of Biotechnology and Food Sciences, Department of Food Science, Tr. A. Hlinku 2, 949 76 Nitra, Slovakia,
Tel.: +421 37 641 4371
E-mail: jozef.capla@uniag.sk
ORCID: https://orcid.org/0000-0001-9475-6359
Author contribution: conceptualisation, writing – review & editing, funding acquisition.

Peter Zajác
Affiliation: The Slovak University of Agriculture in Nitra, Faculty of Biotechnology and Food Sciences, Department of Food Science, Tr. A. Hlinku 2, 949 76 Nitra, Slovakia,
Tel.: +421 37 641 4371
E-mail: peter.zajac@uniag.sk
ORCID: https://orcid.org/0000-0001-9475-6359
Author contribution: conceptualisation, writing – original draft, writing – review & editing.

Corresponding author: **Martina Fikselová**

Copyright notice:

© 2026 The Author(s). Published by HACCP Consulting, Slovakia, European Union, on *Scifood* (https://scifood.eu), the official website of the journal. *Scifood* is owned and operated by HACCP Consulting s.r.o. (https://www.haccp.sk). This article is published as **Open Access** under the **Creative Commons Attribution–NonCommercial–NoDerivatives 4.0 International License** https://creativecommons.org/licenses/by-nc-nd/4.0/ (**CC BY-NC-ND 4.0**). This license permits non-commercial use, sharing, distribution, and reproduction in any medium or format, provided that appropriate credit is given to the author(s) and the original publication, a link to the license is provided, and the work is not modified, adapted, or transformed.