

# AI-Driven Quality Assessment of Frozen Dough Using Multimodal Spectroscopic and Rheological Data

Eleanora Giannelli

Department of Computer Science, Colorado State University, Fort Collins, CO, USA.  
contacteleanora@colostate.edu

Demi-Leigh Doocy

Department of Computer Science, University of North Texas, Denton, TX, USA.  
demileighmail@unt.edu

## Abstract

Frozen dough is a critical intermediate in industrial baking, yet its quality remains difficult to assess because freezing, storage, thawing, and mechanical handling alter gluten networks, starch crystallinity, yeast viability, water distribution, and rheological behavior simultaneously. Conventional quality evaluation often relies on offline rheological testing, sensory inspection, or limited chemical assays, which are slow, sampling-intensive, and poorly matched to continuous production environments. This paper examines the system-level design of artificial intelligence driven quality assessment for frozen dough using multimodal spectroscopic and rheological data. It argues that the central challenge is not merely model selection but the construction of a reliable socio-technical infrastructure that integrates near-infrared and mid-infrared spectroscopy, hyperspectral imaging, oscillatory and extensional rheology, process telemetry, and contextual metadata. The discussion analyzes data acquisition architecture, calibration transfer, multimodal fusion, uncertainty estimation, edge and cloud deployment, model governance, fairness across production sites, sustainability implications, and policy constraints. Particular attention is given to structural trade-offs among predictive accuracy, interpretability, latency, cost, robustness, and regulatory accountability. The paper also considers how federated learning, model cards, and interpretable machine learning can support deployment without concentrating data or authority in a single actor. Through cross-domain comparisons with medical imaging, precision agriculture, and supply chain analytics, the paper identifies design principles for resilient frozen dough quality systems. It concludes that AI driven multimodal assessment can improve consistency and reduce waste, but only if technical performance is matched by transparent governance, maintenance capacity, and clear allocation of responsibility across equipment vendors, bakeries, regulators, and consumers.

## Keywords

frozen dough; multimodal learning; spectroscopy; rheology; artificial intelligence; food quality; governance; sustainability.

## 1. Introduction

The industrialization of frozen dough has transformed baking from a craft-centered batch process into a distributed manufacturing system in which fermentation, shaping, freezing, storage, transport, and final baking occur across different sites and time scales. This spatial

and temporal separation creates quality uncertainty. A dough piece that leaves a central facility in acceptable condition may deteriorate during cold-chain transit, retail storage, or final proofing because its biophysical state is sensitive to temperature fluctuations, mechanical stress, and moisture migration. Quality assessment therefore cannot be reduced to a single endpoint measurement. It must be understood as a continuous, multi-actor problem that links raw material variability, process control, logistics, and consumer expectations. Artificial intelligence has been proposed as a means of integrating these heterogeneous signals, but the literature often treats the problem as a pattern recognition task rather than an infrastructure problem. This paper argues that AI driven quality assessment of frozen dough should be designed as a socio-technical system in which spectroscopic and rheological data are only two components of a larger architecture [1], [3].

Spectroscopic methods offer rapid, non-destructive, and increasingly portable windows into molecular composition and structural change. Near-infrared and mid-infrared spectra can reflect water binding, protein secondary structure, lipid oxidation, and starch organization, while hyperspectral imaging adds spatial information about heterogeneity across dough surfaces [8], [9]. Rheological methods, by contrast, capture mechanical behavior that is directly relevant to processability and final product texture. Oscillatory measurements, extensional tests, and creep recovery can reveal how freezing and thawing modify viscoelasticity, relaxation, and resistance to deformation [10], [13]. Neither modality is sufficient alone. Spectra may detect chemical shifts that do not translate directly into processing performance, while rheology may reveal mechanical degradation without identifying its molecular cause. Multimodal fusion is attractive because it can connect composition, structure, and function, but it also introduces difficult problems of synchronization, calibration, missing data, and cross-site generalization [6], [7].

The introduction of AI into this setting raises questions that extend beyond predictive accuracy. A model that performs well in a laboratory may fail on a production line when flour lots change, ambient humidity varies, sensors drift, or operators modify mixing schedules. Deployment requires maintenance routines, data versioning, model monitoring, and clear escalation paths when predictions conflict with human judgment. These concerns connect to broader research on machine learning systems, where hidden technical debt, deployment failure, and governance gaps are recognized as primary sources of long-term risk [16], [17]. In food manufacturing, the stakes include product waste, consumer trust, regulatory compliance, and worker safety. A quality assessment system that cannot explain its decisions may be rejected by line operators even if its statistical performance is strong [15]. Therefore, the design of AI driven frozen dough assessment must balance accuracy with interpretability, robustness, fairness, and institutional legitimacy [20].

## **2. Conceptual Foundations of Frozen Dough Quality**

Frozen dough quality is a multidimensional construct. At the molecular level, ice formation and recrystallization redistribute water, concentrate solutes, and mechanically damage gluten structures. At the colloidal level, starch granules, proteins, and polysaccharides interact in ways that determine water retention and gas-holding capacity. At the rheological level, these changes manifest as altered storage modulus, loss modulus, extensibility, and recovery behavior [10], [11]. At the biological level, yeast viability and metabolic activity decline during frozen storage, affecting proofing time and gas production. At the sensory level, final baked products may show reduced volume, denser crumb, darker crust, or inconsistent texture. A quality assessment system must therefore represent multiple causal pathways without

assuming that one measurement can serve as a universal proxy. This is why multimodal data are conceptually justified, not merely technically fashionable [6], [12].

Spectroscopic data provide indirect but rich information about these pathways. Infrared and near-infrared spectra are sensitive to hydrogen bonding, water activity, protein conformation, and carbohydrate structure. Spectral changes can precede visible deterioration, making them useful for early warning. However, spectral interpretation is complicated by scattering, sample heterogeneity, temperature, and instrument differences [8], [9]. Rheological data provide a more direct functional readout, but they are often destructive, time-consuming, and sensitive to sample preparation. Thus, the two modalities have complementary strengths and weaknesses. Spectroscopy is fast and non-invasive but requires chemometric modeling; rheology is functionally meaningful but slower and less suited to every dough piece. A multimodal system can use spectroscopy for high-throughput screening and rheology for calibration, validation, and deeper diagnostics. This division of labor is a system architecture decision, not simply a data fusion choice [7], [10].

The required reference for this paper illustrates how cryoprotective sol-gel transitions in a polysaccharide fraction can influence frozen dough applications [12]. That work is relevant because it points to the molecular and colloidal mechanisms that AI models are ultimately trying to infer from indirect measurements. If a cryoprotective additive alters water mobility or gluten network stability, the effect may appear in spectra as shifted absorption bands and in rheology as changed viscoelastic moduli. A model that learns these associations can support formulation optimization and process control. Yet the same example also shows why mechanistic grounding matters. Without domain knowledge, a neural network may exploit site-specific correlations, such as a particular mixer or storage room, rather than learning transferable biophysical relationships. The most credible systems will combine data-driven pattern recognition with process knowledge and targeted experiments [11], [13].

### **3. Multimodal Data Architecture: Spectroscopy, Rheology, and Process Telemetry**

The data architecture for frozen dough assessment begins with sensor selection and placement. Near-infrared instruments can be installed at mixing, freezing, storage, and thawing stages, while hyperspectral cameras may be used for surface inspection after shaping. Rheological instruments are less easily embedded in continuous lines, so they may function as at-line or offline reference instruments. Process telemetry, including temperature, humidity, air velocity, residence time, and mechanical energy input, provides context that explains why nominally identical dough pieces behave differently [8], [9]. Metadata about flour lot, formulation, additive type, equipment identity, and operator shift is equally important. Without metadata, a model may confuse environmental variation with material variation, producing predictions that appear accurate in one facility but fail in another [17].

A second architectural concern is synchronization. Spectroscopic measurements may occur in seconds, rheological tests in minutes, and cold-chain events over days. Fusing these streams requires a temporal ontology that defines what constitutes a matching observation. A dough sample measured spectroscopically at the mixer is not identical to the same sample after freezing and thawing. The system must therefore distinguish between state variables, such as current water distribution, and history variables, such as cumulative thermal exposure. This distinction affects model design, data storage, and traceability. It also has governance implications: if a quality decision is challenged, the system must be able to reconstruct which data were available, which model version was used, and what uncertainty was reported [19]. Data lineage is not administrative overhead; it is a prerequisite for accountability [16].

A third concern is calibration transfer across instruments and sites. Spectroscopic models are notoriously sensitive to instrument response, sample presentation, and environmental conditions. A model trained on one spectrometer may not transfer directly to another without standardization or recalibration [6], [8]. In frozen dough, this problem is compounded by temperature-dependent spectral changes and by differences in dough formulation. A robust architecture should include reference standards, periodic validation with rheological anchors, and statistical monitoring of spectral residuals. Federated learning offers one response by allowing sites to train shared models without moving raw data, but it introduces challenges of heterogeneity, communication cost, and adversarial participation [18]. The system-level trade-off is between centralized data control, which simplifies model development but raises privacy and sovereignty concerns, and distributed learning, which supports local autonomy but complicates governance and debugging.

#### **4. AI Models and Multimodal Fusion for Quality Assessment**

The modeling layer of a frozen dough quality system may include classical chemometrics, shallow machine learning, deep learning, and hybrid approaches. Principal component analysis and partial least squares remain valuable for dimensionality reduction and interpretation, especially when sample sizes are limited [1]. Support vector machines can perform well on spectral classification tasks when feature engineering is careful [2]. Gradient boosting methods are often strong for tabular data that combine spectra, rheology, and process metadata [5]. Deep learning, including convolutional and attention-based architectures, can learn complex representations from spectral and image data, but it requires larger datasets and careful regularization [3], [4]. The choice among these methods should be driven by data availability, interpretability requirements, latency constraints, and the cost of errors. A system that rejects good dough is costly; a system that accepts bad dough may cause consumer complaints or food safety risks. The asymmetry of these costs should shape model evaluation [7].

Multimodal fusion can occur at early, intermediate, or late stages. Early fusion concatenates raw or lightly processed features, which allows interactions to be learned but is sensitive to missing data and scale differences. Intermediate fusion learns shared representations, often through neural networks, and can capture cross-modal relationships such as the link between protein conformation and extensibility. Late fusion combines separate predictions, which is more modular and easier to audit but may miss subtle interactions [4], [6]. In frozen dough, a hybrid approach is often most practical. Spectroscopic models can provide rapid predictions, rheological models can provide functional calibration, and a fusion layer can integrate them with process telemetry. The fusion layer should also estimate uncertainty. If spectroscopic and rheological predictions disagree, the system should flag the sample for human review rather than forcing a confident but unreliable output. This aligns with interpretable machine learning principles that prioritize decision support over automated authority [14], [15].

Interpretability is especially important in food manufacturing because quality decisions are embedded in legal, commercial, and labor relations. A line operator may need to know whether a dough batch is rejected because of water content, protein quality, or thermal abuse. A production manager may need to know whether a model failure reflects a sensor fault or a genuine process change. Methods such as feature attribution and surrogate explanation can help, but they are not a substitute for inherently interpretable models when stakes are high [14], [15]. The design should therefore match explanation depth to decision consequence. Routine screening may tolerate black-box predictions if uncertainty is well calibrated, while

formulation changes, allergen-related decisions, or regulatory submissions may require mechanistic evidence. This graduated approach avoids both naive automation and unnecessary model opacity [20].

## **5. System-Level Trade-Offs, Robustness, and Deployment**

Deploying AI driven quality assessment on a frozen dough line requires confronting latency, throughput, and cost trade-offs. Spectroscopic measurements can be fast enough for real-time screening, but full multimodal fusion may require rheological data that are not immediately available. A system can operate in two modes: a fast mode that uses spectra and telemetry to predict risk, and a slow mode that confirms or corrects predictions with rheology. This tiered architecture mirrors edge-cloud designs in other domains, where urgent decisions are made locally and computationally intensive analysis is performed later [17]. Edge deployment reduces network dependence and supports data minimization, but it limits model size and complicates updates. Cloud deployment enables centralized learning and monitoring, but it raises data transfer, privacy, and outage risks. The appropriate balance depends on plant scale, regulatory environment, and the economic cost of downtime [18].

Robustness is another central concern. Frozen dough processes experience seasonal flour variation, supplier changes, equipment wear, and human variability. A model trained on historical data may degrade silently when conditions shift. Monitoring should therefore include input drift detection, output distribution checks, and periodic validation against rheological reference tests. The system should also include fallback procedures. If a spectrometer fails or a model becomes unreliable, the line must revert to established quality control methods without halting production. This requires documentation, training, and organizational readiness. Research on machine learning technical debt shows that such maintenance concerns are often underestimated during initial development [16]. In food systems, where margins are thin and production schedules are rigid, a fragile AI component can create more risk than value.

Deployment also raises questions of fairness and equity across sites. A large bakery may have the capital to install high-end spectrometers and maintain data science teams, while a smaller facility may rely on shared models or low-cost sensors. If quality assessment models are trained predominantly on data from large, technologically advanced plants, they may perform poorly in smaller or regional bakeries with different equipment and formulations. Federated learning can help distribute model development, but it requires participants to contribute compute and governance capacity [18]. Model cards and datasheets can document intended use, performance disparities, and limitations, enabling buyers and regulators to assess suitability [19]. Without such transparency, AI quality systems may reinforce existing inequalities in food manufacturing rather than democratizing access to reliable assessment.

## **6. Governance, Fairness, Sustainability, and Policy Implications**

Governance for AI driven frozen dough assessment must address data ownership, model accountability, and regulatory compliance. Spectroscopic and rheological data may be commercially sensitive because they reveal formulation details, supplier quality, and process parameters. Who owns these data: the bakery, the equipment vendor, the retailer, or the consumer? Who is liable if a model incorrectly approves a batch that later fails? These questions cannot be resolved by technical design alone. They require contracts, audit rights, and clear lines of responsibility. Principles such as transparency, justice, beneficence, non-maleficence, and explicability provide a high-level framework, but they must be translated

into operational controls [20]. Model cards, version histories, and validation reports are practical instruments for such translation [19].

Fairness in this context has multiple meanings. It can refer to equitable access to quality assessment technology, non-discrimination across suppliers, or fair treatment of workers whose tasks are augmented or monitored by AI. If a system continuously evaluates operator performance through dough quality outcomes, it may create surveillance pressures that are not captured by standard accuracy metrics. It can also encode biases if historical data reflect past discriminatory practices or uneven supplier relationships. A fair system should separate model performance from personnel evaluation unless there is a clear and transparent justification. It should also allow human appeal and correction. These safeguards are not obstacles to innovation; they are conditions for trust and long-term adoption [15], [20].

Sustainability implications are substantial. Frozen dough production consumes energy for freezing, cold storage, and transport. Quality failures lead to waste, rework, and unnecessary energy use. AI driven assessment can reduce waste by detecting deterioration earlier and enabling dynamic inventory management. However, the computational footprint of AI models, sensors, and data infrastructure must also be considered. Large models trained on vast datasets may have environmental costs that offset some waste reduction benefits. A sustainable design should therefore favor right-sized models, edge processing, and lifecycle assessment. It should also consider cold-chain optimization, since a small improvement in storage temperature control may reduce both quality loss and energy consumption. Policy instruments such as carbon labeling, food waste reporting, and digital product passports could interact with AI quality systems in complex ways [18], [20].

## **7. Case Illustrations and Cross-Domain Comparisons**

Comparisons with medical imaging are instructive because both domains involve multimodal data, high-stakes decisions, and interpretability demands. In medical AI, models often combine imaging with clinical history and laboratory results, but deployment fails when workflows, accountability, and data distribution are ignored [17]. Frozen dough assessment faces a similar challenge. A model may achieve high accuracy in retrospective validation yet fail when integrated into a bakery line because operators distrust it, sensors are poorly maintained, or alerts arrive too late. The lesson is that human factors and organizational routines are not peripheral. They are part of the system architecture. Similarly, precision agriculture uses spectral and environmental data to optimize irrigation and fertilization, but its success depends on farmer trust, equipment compatibility, and regional adaptation [18]. These cross-domain examples suggest that frozen dough AI should be designed for maintainability, local adaptation, and clear decision rights.

Supply chain analytics offers another comparison. Cold-chain monitoring already generates large volumes of temperature and humidity data. Integrating these data with dough quality models can improve traceability and recall management. However, such integration raises privacy and competitive concerns because supply chain partners may not wish to share proprietary information. Federated and privacy-preserving approaches can mitigate some risks, but they require common standards and trusted intermediaries [18]. Model documentation can help partners understand what a model does and does not claim [19]. The broader policy environment may also shape adoption. If regulators require explainable quality decisions for food safety, then interpretable models become more attractive. If they allow performance-based validation, then more complex models may be acceptable. The optimal architecture is therefore contingent on institutions, not only on data and algorithms [20].

## 8. Conclusion

AI driven quality assessment of frozen dough using multimodal spectroscopic and rheological data is a promising but demanding systems challenge. Spectroscopic methods provide rapid, non-destructive molecular and structural information, while rheological methods provide functional insight into processability and final texture. Their integration can support early warning, formulation optimization, and waste reduction. Yet the value of such a system depends on architecture, governance, and maintenance as much as on model accuracy. Reliable synchronization, calibration transfer, uncertainty estimation, drift monitoring, and fallback procedures are essential for deployment in industrial bakeries. Interpretability must be matched to decision stakes, and fairness must be considered across sites, suppliers, and workers. Federated learning, model cards, and transparent validation can support distributed governance, but they do not eliminate the need for clear accountability and policy frameworks. The most credible systems will combine data-driven learning with mechanistic knowledge of frozen dough biophysics and with organizational routines that keep human judgment in the loop. If designed responsibly, multimodal AI can improve consistency and sustainability in frozen dough production; if designed narrowly, it risks becoming another fragile layer of technical debt in an already complex food system.

## References

1. Wold, S., Esbensen, K., & Geladi, P. (1987). Principal component analysis. *Chemometrics and Intelligent Laboratory Systems*, 2(1–3), 37–52.
2. Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273–297.
3. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444.
4. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30.
5. Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794.
6. Morais, C. L. M., Lima, K. M. G., Singh, M., & Martin, F. L. (2020). Tutorial: Multivariate classification for vibrational spectroscopy in biological samples. *Nature Protocols*, 15(7), 2143–2162.
7. Oliveri, P., & Downey, G. (2012). Multivariate class modeling for the verification of food-authenticity claims. *TrAC Trends in Analytical Chemistry*, 35, 74–86.
8. Workman, J., & Weyer, L. (2012). *Practical guide and spectral atlas for interpretive near-infrared spectroscopy* (2nd ed.). CRC Press.
9. Sun, D.-W. (Ed.). (2009). *Infrared spectroscopy for food quality analysis and control*. Academic Press.
10. Dobraszczyk, B. J., & Morgenstern, M. P. (2003). Rheology and the breadmaking process. *Journal of Cereal Science*, 38(3), 229–245.

11. Goesaert, H., Brijs, K., Veraverbeke, W. S., Courtin, C. M., Gebruers, K., & Delcour, J. A. (2005). Wheat flour constituents: How they impact bread quality, and how to impact their functionality. *Trends in Food Science & Technology*, 16(1–3), 12–30.
12. Zhang, T., Fang, J. Q., Wang, P. P., & Chen, C. (2026). Structural basis of the cryoprotective sol-gel transition in a *Phyllanthus emblica* L. polysaccharide fraction for frozen dough applications. *Food Hydrocolloids*, 112842.
13. Meziani, S., Ioannou, I., Jasniewski, J., Belhaj, N., Muller, J.-M., Ghoul, M., & Desobry, S. (2011). Effects of freezing treatments on viscoelastic and structural behavior of frozen sweet dough. *Journal of Food Engineering*, 107(3–4), 358–365.
14. Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30.
15. Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1(5), 206–215.
16. Sculley, D., Holt, G., Golovin, D., Davydov, E., Phillips, T., Ebner, D., Chaudhary, V., Young, M., Crespo, J.-F., & Dennison, D. (2015). Hidden technical debt in machine learning systems. *Advances in Neural Information Processing Systems*, 28.
17. Paleyes, A., Urma, R.-G., & Lawrence, N. D. (2022). Challenges in deploying machine learning: A survey of case studies. *ACM Computing Surveys*, 55(6), 1–29.
18. Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A. N., Bonawitz, K., Charles, Z., Cormode, G., Cummings, R., D'Oliveira, R. G. L., Eichner, H., El Rouayheb, S., Evans, D., Gardner, J., Garrett, Z., Gascón, A., Ghazi, B., Gibbons, P. B., ... Zhao, S. (2021). Advances and open problems in federated learning. *Foundations and Trends in Machine Learning*, 14(1–2), 1–210.
19. Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Spitzer, E., Raji, I. D., & Gebru, T. (2019). Model cards for model reporting. *Proceedings of the Conference on Fairness, Accountability, and Transparency*, 220–229.
20. Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review*, 1(1).