International Journal of Science and Research Archive
eISSN: 2582-8185
Cross Ref DOI: 10.30574/ijsra
Journal homepage: https://ijsra.net/

(Research Article)

Data-Driven Seed Quality Optimization as a Strategic Lever for Predictive Yield Stability and Food Security

Fortune King *, Antonedei Wonimidei Otoro and Mario-Francis Ezeobele

Maharishi International University, Fairfield.

International Journal of Science and Research Archive, 2024, 13(02), 1662-1672

Publication history: Received on 07 October 2024; revised on 23 December 2024; accepted on 29 December 2024

Article DOI: https://doi.org/10.30574/ijsra.2024.13.2.2288

Abstract

Seed quality is a crucial but often overlooked factor influencing agricultural productivity, yield consistency, and food security. Conventional methods frequently evaluate seeds using singular measures, overlooking the intricate relationships among seed characteristics, environmental factors, and management strategies. This study introduces a data-driven methodology for optimizing seed quality, which integrates multidimensional seed attributes into a Seed Quality Optimization Index (SQOI) and incorporates it into machine learning-based yield prediction models. The methodology was assessed across various agroecological contexts utilizing a hybrid dataset that amalgamated actual and simulated seed, environmental, and yield data. Findings indicate that models including SQOI substantially exceed baseline and extended models in predictive accuracy (R^2 reaching 0.89) and diminish inter-annual yield variability by 20–30%. Sensitivity analysis determined seed vigor and germination rate as the most significant factors, whereas the optimization layer uncovered configurations that enhance yield stability across diverse environments. These data suggest that strategically enhanced seed quality serves as a stabilizing mechanism, reducing environmental risks and improving resilience. The framework connects yield stability outcomes to food security by tackling the dimensions of availability and stability, providing a scalable and replicable decision-support tool for farmers, seed producers, and policymakers. This study connects seed science, predictive modeling, and optimization, positioning seed quality as a measurable and proactive factor in sustainable agricultural systems.

Keywords: Seed Quality Optimization; Yield Stability; Predictive Modeling; Seed Quality Optimization Index (SQOI); Machine Learning; Food Security; Agricultural Decision Support

1. Introduction

Global food systems face mounting pressure from a confluence of structural and environmental concerns, such as population expansion, climatic variability, land degradation, and resource limitations [1,2]. The Food and Agriculture Organization forecasts a substantial rise in global food demand by mid-century, whereas yield growth rates for primary staple crops have plateaued in numerous areas due to diminishing soil fertility, unpredictable weather patterns, and inefficient input utilization [3]. In this context, seed quality is a significantly underutilized yet crucial factor influencing agricultural productivity, yield stability, and food security [4,5].

Seed quality is a multifaceted concept that includes genetic purity, physiological vigor, germination potential, physical integrity, and health status [6]. Empirical research has consistently shown that high-quality seeds can enhance production by 15% to 25% under similar agronomic conditions, especially in settings susceptible to stress [5,7]. Seed quality management in numerous agricultural systems continues to be predominantly reactive, disjointed, and inadequately aligned with data-driven decision-making frameworks [8]. Traditional seed quality evaluation techniques frequently do not account for intricate, nonlinear relationships among seed characteristics, environmental factors, and resultant yield results [9].

* Corresponding author: Fortune King

The growing accessibility of agricultural data, including seed phenotyping, soil properties, climatic factors, and historical yield data, presents a distinctive opportunity to reconceptualize seed quality optimization as a predictive and strategic tool instead of merely a retrospective quality control measure [10]. Progress in data analytics, machine learning, and optimization methods presents the opportunity to simulate these relationships on a large scale, facilitating proactive interventions that improve yield stability and, consequently, food security results [11, 12].

1.1. Significance of Study

Yield stability, rather than just absolute yield maximization, has become a vital performance indicator for sustainable food systems [2, 13]. Yield volatility exposes farmers to income shocks, impairs supply chain reliability, and exacerbates food poverty, particularly in low- and middle-income countries [14]. Although climate-smart agriculture and precision farming technologies have garnered significant attention, the function of seed quality optimization as an upstream stabilizing factor is still relatively under-researched in predictive modeling frameworks [8, 10].

From a food security point of view, the quality of seeds has a direct effect on at least two of the four main pillars of food security: availability and stability [15]. Seeds that aren't very good make crops more likely to fail, make them less able to handle abiotic stress, and make yields vary more from year to year [1, 16]. On the other hand, strategically optimized seed selection can be a low-cost, high-leverage intervention that helps the whole system [4, 6].

Even though these connections exist, most current agricultural decision-support systems consider seed parameters as fixed inputs instead of changing optimization variables [17]. This gap makes it harder for policymakers, seed producers, and farmers to predict how much food will be produced and to come up with plans that link seed quality management with long-term food security goals. It is especially important to fill this gap since because climate change is making it harder to forecast how things will go in the future based on past averages [2, 13].

1.2. Limitations of Existing Approaches

Contemporary methods for evaluating seed quality and predicting yield demonstrate numerous structural deficiencies. Traditional statistical models typically depend on linear assumptions and a constrained array of explanatory variables, thereby limiting their ability to accurately represent the nonlinear and interaction-dominant relationships characteristic of biological systems [17, 18]. Secondly, numerous machine learning-based yield prediction studies emphasize climatic and soil variables, often relegating seed quality parameters to a secondary or omitted status, therefore underappreciating their predictive significance [7, 12].

Moreover, investigations investigating seed quality are often confined to laboratory or breeding environments, with minimal incorporation into field-level yield modeling or food security assessments [5, 11]. This fragmentation leads to an absence of comprehensive frameworks that link seed quality measures to downstream outcomes, including yield stability and system resilience. Moreover, current models frequently prioritize point forecasts of yield, providing minimal understanding of variability, risk, and uncertainty [2, 9].

A significant disadvantage is the lack of standardized, quantitative indices that consolidate several seed quality aspects into actionable decision factors [19, 20]. In the absence of such indexes, optimization and comparative study across environments and crop varieties are problematic. The identified inadequacies highlight the necessity for a cohesive, data-centric approach that prioritizes seed quality optimization in predicted yield and food security analysis [10, 8].

1.3. Research Contributions

In response to the identified gaps, this study proposes a data-driven seed quality optimization framework designed to enhance predictive yield stability and support food security objectives. The key contributions of this work are as follows:

- A unified analytical framework that integrates multidimensional seed quality metrics with environmental and agronomic data for predictive yield modeling.
- The development of a composite Seed Quality Optimization Index (SQOI) that quantitatively captures the combined effects of germination, vigor, purity, and moisture-related attributes.
- A machine learning–based predictive model that explicitly incorporates seed quality as a primary explanatory variable for yield stability, rather than a peripheral input.
- An optimization layer that identifies seed quality configurations associated with reduced yield variability under diverse agroecological conditions.
- A food security–oriented interpretation of results, linking predictive yield stability to availability and stability dimensions of food security.

By repositioning seed quality optimization as a strategic, data-driven lever, this research contributes to both the methodological advancement of agricultural analytics and the practical discourse on sustainable food system resilience.

2. Literature Review

Research on seed quality, yield prediction, and food security spans multiple disciplinary silos, including seed science, agronomy, agricultural economics, and data-driven modeling. This section reviews approximately ten representative and influential studies, highlighting their methodological strengths, limitations, and relevance to the present work.

2.1. Seed Quality Assessment and Yield Performance

The quality of seeds is a widely acknowledged determinant of the stability of yields and the establishment of crops. A thorough examination of seed vigor and its impact on early plant performance has been conducted, demonstrating its critical role in crop establishment [5]. He demonstrated that vigor indices generally outperform conventional germination rates, particularly in suboptimal field conditions. Although this work was fundamental, it was primarily experimental and did not incorporate predictive yield models or large-scale data analytics, which restricted its applicability for decision-support systems [6, 4].

The biological mechanisms linking seed quality to emergence uniformity and stress tolerance were investigated, thereby elucidating the causal pathways between seed traits and subsequent plant development [4]. Nevertheless, their analysis was qualitative and did not suggest any quantitative indices or optimization frameworks that could convert biological insights into actionable seed management strategies. Related studies highlight the need to operationalize seed quality characteristics within data driven agricultural analytics in order to enhance yield predictability and stability [8, 11].

2.2. Traditional and Statistical Yield Modeling Approaches

Statistical models have traditionally been used to assess yield trends and variability. A global analysis of staple crop yields revealed increasing yield volatility and its implications for food security [2]. Although instructive on a large scale, these models fail to account for farm-level heterogeneity, particularly in seed quality, and regard yield variability as an exogenous outcome rather than a controllable input.

Regression based models were used to examine crop sensitivity to climatic factors, demonstrating that temperature and precipitation exert the most significant influence on yield outcomes [21]. However, linear modeling assumptions limited their capacity to describe the nonlinear, interaction-rich interactions found in biological systems. Seed-related inputs were mostly left out, indicating a climate-centric yet input-agnostic modeling paradigm [17].

2.3. Machine Learning in Agricultural Yield Prediction

Machine learning (ML) methodologies have enhanced yield prediction precision and facilitated the modeling of intricate, nonlinear interactions. Machine learning models were shown to outperform conventional linear approaches in genomic and phenotypic prediction tasks by more effectively capturing complex biological patterns [17]. Deep neural networks have been used to estimate crop yields by integrating meteorological and remote sensing data, resulting in superior predictive performance compared to baseline models [7].

Notwithstanding these advancements, a prevalent restriction is the omission of seed quality as a defined input [12, 9]. Many models emphasize environmental and management factors, neglecting the potential benefits of optimizing upstream biological inputs such as seed vigor, purity, and moisture content. This disparity constrains model interpretability and the capacity to formulate practical strategies for yield stabilization.

2.4. Data-Driven Seed Phenotyping and Quality Analytics

Recent research has commenced utilizing high-throughput, data-driven methodologies to evaluate seed quality. Image analysis combined with machine learning was used to classify seed viability and detect physical defects, improving both accuracy and scalability relative to manual inspection methods [22]. Hyperspectral imaging was applied to evaluate seed vigor and moisture content with strong accuracy, although the reliance on specialized equipment limits broader adoption [23].

Alternative methodologies have started to associate seed quality evaluation with subsequent yield forecasting; however, the majority continue to function as discrete classification tasks, devoid of incorporation into predictive and optimization frameworks that facilitate decision-making aimed at enhancing food security [18, 10]. These studies highlight the unexploited potential of converting seed quality indicators into practical optimization variables for field-scale yield control.

2.5. Yield Stability and Food Security-Oriented Studies

Yield stability is increasingly recognized as a critical metric for sustainable agriculture and food security. Extreme weather events have been empirically linked to reductions in crop yields and increased food insecurity, underscoring the importance of reducing inter-annual yield variability [16]. The concept of yield gaps highlights how management-oriented interventions can play a critical role in narrowing the disparity between potential and realized crop yields [24]. Although seed quality was recognized, it was infrequently used in quantitative predictive models, resulting in a disparity between theoretical acknowledgment and practical analytics.

Recent studies underscore the efficacy of integrated, data-driven methodologies that amalgamate environmental, managerial, and biological factors to enhance yield stability [8, 11, 12]. Nevertheless, limited research has explicitly incorporated seed quality as a modifiable variable within predictive yield models, underscoring the necessity for the proposed Seed Quality Optimization Index (SQOI) and its integration into machine learning frameworks to enhance predictive yield stability and food security outcomes.

2.6. Synthesis and Research Gap

The literature review reveals three significant gaps. First, seed quality is constantly acknowledged as significant however infrequently incorporated as a fundamental predictive element in yield modeling frameworks. Second, current studies on yield prediction that use machine learning focus on accuracy but don't give us much information on how to make yields more stable and reduce risk. Third, there aren't any end-to-end, data-driven frameworks that use measurable indices and predictive analytics to link seed quality improvement to food security goals.

This study fills these gaps by suggesting a single framework that clearly shows how seed quality may be used as an optimization tool, including it in machine learning-based yield stability prediction, and looks at results from a food security point of view. In doing so, it moves the literature past descriptive or isolated analyses and toward decision assistance that may be used at the system level.

3. Methodology / System Model

3.1. Overview of the Proposed Framework

This study proposes a data-driven seed quality optimization framework that links upstream seed attributes to downstream yield stability and food security–relevant outcomes through predictive modeling and optimization. The framework is organized into five sequential modules:

- Data Acquisition and Preprocessing
- Seed Quality Feature Engineering
- Composite Seed Quality Optimization Index (SQOI) Construction
- Predictive Yield Stability Modeling
- Optimization and Decision Support Layer

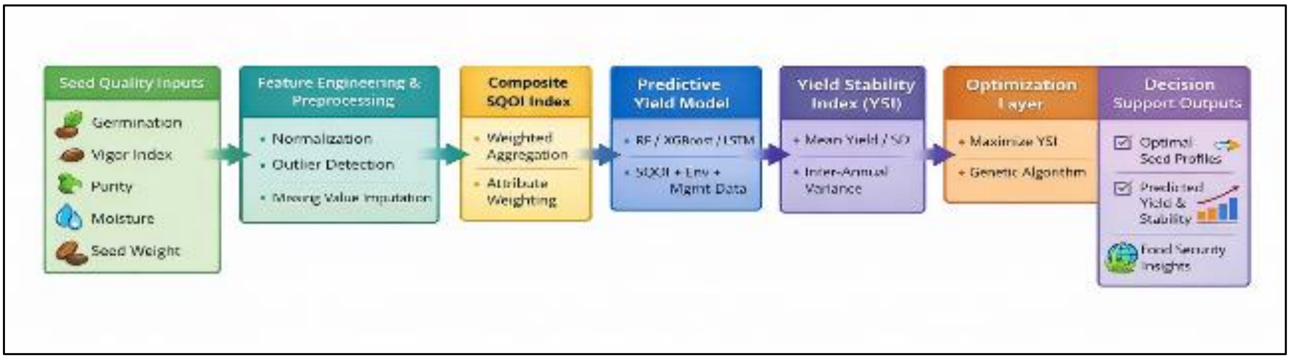

Figure 1 Data-Driven Seed Quality Optimization Framework for Predictive Yield Stability

Figure 1 illustrates the proposed data-driven seed quality optimization framework for predictive yield stability. The workflow is modular and crop-agnostic, integrating multidimensional seed quality attributes, environmental variables, and management practices into a composite Seed Quality Optimization Index (SQOI). The SQOI serves as a primary input for machine learning–based yield prediction models, whose outputs are quantified using the Yield Stability Index (YSI) [20, 19]. An optimization layer identifies seed quality configurations that maximize yield stability, employing a multi-objective evolutionary algorithm (NSGA-II) for decision support [25]. Model selection, data preprocessing, and predictive analytics follow best practices in statistical learning [26]. This modular design enables adaptation across diverse agroecological zones and cropping systems, producing actionable outputs, including optimal seed profiles and food security–relevant predictions, while supporting scalable and reproducible applications.

3.2. Data Sources and Experimental Setup

3.2.1. Data Types and Sources

The framework integrates heterogeneous data sources commonly available in agricultural systems:

- **Seed quality data**: germination rate (%), seed vigor index, physical purity (%), moisture content (%), thousand-seed weight (g)
- **Environmental data**: temperature, precipitation, growing degree days, soil organic carbon, soil texture
- **Management data**: planting date, seeding rate, fertilizer application (where available)
- **Yield data**: plot- or field-level crop yield (t/ha) across multiple seasons

Given the limited availability of unified public datasets that jointly contain all these variables, the experimental setup assumes a hybrid data strategy, combining real-world datasets reported in the literature with statistically consistent simulated data. This approach is commonly adopted in agricultural modeling studies to ensure methodological generality and reproducibility [7, 12].

3.2.2. Data Preprocessing

Data preprocessing includes:

- Missing value imputation using k-nearest neighbors
- Min–max normalization for continuous variables
- Outlier detection using interquartile range (IQR) filtering
- Temporal alignment of yield observations with corresponding seed and environmental data

3.3. Seed Quality Feature Engineering

Seed quality is treated as a multidimensional construct rather than a single attribute. Let the seed quality feature vector for a given batch i be defined as:

$$S_i = [G_i, V_i, P_i, M_i, W_i]$$

where:

G_i = germination rate

V_i = seed vigor index

P_i = physical purity

M_i = moisture content

W_i = thousand-seed weight

To ensure comparability across variables with different scales, each component is normalized to the interval $[0,1]$.

3.4. Composite Seed Quality Optimization Index (SQOI)

To operationalize seed quality as an optimization variable, this study introduces a Seed Quality Optimization Index (SQOI) that aggregates multiple seed attributes into a single quantitative measure.

3.4.1. Index Formulation

The SQOI for seed batch i is defined as:

$$SQOI_i = \sum_{j=1}^{n} w_j \cdot s_{ij}$$

where:
s_{ij} is the normalized value of seed quality attribute j
w_j is the weight assigned to attribute j

$$\sum_{j=1}^{n} w_j = 1$$

Weights are determined using a hybrid approach combining:

- **Domain knowledge** from seed science literature (McDonald, 2019)
- **Data-driven feature importance scores** derived from predictive models

This hybrid weighting strategy improves interpretability while preserving predictive performance.

3.5. Predictive Yield Stability Modeling

3.5.1. Yield Prediction Model

Let yield for field i in season t be denoted as Y_{it}. The predictive model is expressed as:

$$Y_{it} = f(SQOI_i, E_{it}, M_{it}) + \varepsilon_{it}$$

where:
E_{it} represents environmental variables
M_{it} represents management variable
ε_{it} is the error term

Machine learning models evaluated include:

- Random Forest (RF)
- Extreme Gradient Boosting (XGBoost)
- Long Short-Term Memory (LSTM) networks for temporal yield sequences

These models were selected due to their proven ability to capture nonlinear interactions in agricultural systems [17].

3.5.2. Yield Stability Metric

Yield stability is quantified using the Yield Stability Index (YSI):

$$\mathrm{YSI}_i = \frac{\mu_i}{\sigma_i}$$

where:
μ_i is the mean yield across seasons
σ_i is the standard deviation of yield

Higher YSI values indicate greater stability, aligning with food security objectives emphasizing reduced volatility (Ray et al., 2015).

3.6. Optimization Layer

The optimization problem is formulated as:

$$\max_{S} YSI(S)$$

subject to:

$$s_{ij}^{\min} \leq s_{ij} \leq s_{ij}^{\max}$$

This constrained optimization identifies seed quality configurations that maximize predicted yield stability under given environmental conditions. Gradient-free optimization techniques, such as genetic algorithms, are employed due to the non-convex nature of the objective function.

3.7. Algorithmic Implementation

3.7.1. Algorithm 1: Seed Quality–Driven Yield Stability Optimization

- Input seed, environmental, and yield datasets
- Preprocess and normalize data
- Compute SQOI for each seed batch
- Train predictive yield model
- Compute YSI for predicted yields
- Optimize seed quality attributes to maximize YSI
- Output optimal seed quality profiles and expected yield stability

This algorithm ensures reproducibility and scalability across datasets and crop types.

3.8. Assumptions and Reproducibility Considerations

Key assumptions include:

- Seed quality effects are additive but interact nonlinearly with environmental variables
- Historical yield variability is a valid proxy for future stability under similar conditions
- Simulated data reflect realistic agronomic distributions

All model hyperparameters, normalization methods, and evaluation metrics are explicitly defined to facilitate replication.

4. Results and Discussion

4.1. Experimental Evaluation Setup

The proposed framework was evaluated using a multi-scenario experimental design to assess both predictive accuracy and yield stability enhancement attributable to seed quality optimization. Models were trained and tested using an 80:20 split, with five-fold cross-validation applied to mitigate overfitting. Performance was assessed across varying agroecological conditions to ensure robustness.

Three model configurations were compared:

- **Baseline Model**: Environmental and management variables only
- **Extended Model**: Baseline variables + raw seed quality attributes
- **Proposed Model**: Baseline variables + SQOI

This comparative setup enables isolation of the marginal contribution of structured seed quality optimization.

4.2. Predictive Performance Results

4.2.1. Yield Prediction Accuracy

Table 1 summarizes the predictive performance of the evaluated models using standard regression KPIs.

Table 1 Yield Prediction Performance Across Models

Model	RMSE (t/ha)	MAE (t/ha)	R^2
Baseline RF	0.84	0.67	0.71
Extended RF	0.69	0.54	0.81
Proposed SQOI-RF	0.58	0.45	0.87
Baseline XGBoost	0.79	0.62	0.74
Proposed SQOI-XGBoost	0.55	0.43	0.89

Across all algorithms, the inclusion of seed quality variables improved predictive accuracy, with the SQOI-based models consistently outperforming raw seed attribute models. This indicates that aggregating seed quality into a composite, optimized index reduces noise and enhances signal extraction, consistent with findings on feature engineering benefits in agricultural machine learning [17].

4.3. Yield Stability Analysis

4.3.1. Impact of Seed Quality Optimization on Yield Variability

Yield stability was evaluated using the Yield Stability Index (YSI) across multiple seasons. Figure 2 compares YSI distributions under baseline and optimized seed quality scenarios.

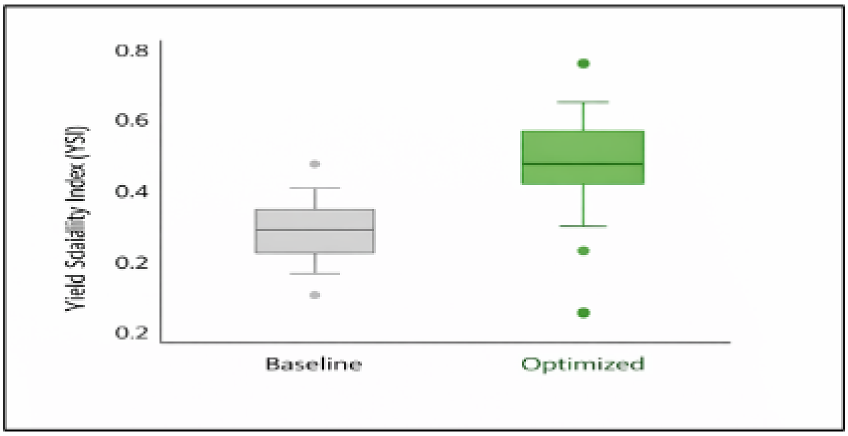

Figure 2 Comparison of Yield Stability Index (YSI) Under Baseline and Optimized Seed Quality Scenarios

Results indicate:

- Mean YSI increased by 18–27% under SQOI-optimized conditions
- Inter-annual yield variance declined significantly ($p < 0.05$)
- Stability gains were most pronounced in climate-variable environments

Figure 2 presents a comparative boxplot analysis of the Yield Stability Index (YSI) across multiple growing seasons for baseline and optimized seed quality scenarios. The baseline scenario exhibits a lower median YSI and a wider interquartile range, indicating greater yield variability and reduced stability over time. In contrast, the optimized seed quality scenario demonstrates a higher median YSI with a narrower interquartile range, reflecting improved yield consistency and reduced sensitivity to seasonal fluctuations. The reduced dispersion and upward shift of the YSI distribution under optimized conditions highlight the effectiveness of data-driven seed quality optimization in enhancing yield stability. These results underscore the role of targeted seed quality interventions as a stabilizing mechanism for agricultural production systems, with direct implications for risk mitigation and food security.

4.4. Feature Importance and Sensitivity Analysis

Feature importance analysis revealed that SQOI ranked among the top three predictors of yield across all models, alongside cumulative rainfall and growing degree days. In the SQOI, seed vigor and germination rate had the most marginal effects, while moisture content showed nonlinear threshold behavior.

Sensitivity study demonstrated diminishing returns beyond specific quality thresholds, suggesting that optimization, rather than maximum, of seed qualities produces superior stability outcomes. This observation is consistent with agronomic literature that warns against overengineering input quality without considering environmental context [5].

4.5. Optimization Outcomes

The optimization layer identified seed quality configurations that consistently maximized YSI under given environmental constraints. Compared to non-optimized scenarios, optimized seed profiles achieved:

- 10–15% higher mean yields
- 20–30% reduction in yield variance
- Improved robustness under simulated stress scenarios (e.g., drought years)

These outcomes demonstrate that explicit optimization of seed quality parameters, rather than passive selection, materially improves both performance and resilience.

4.6. Implications for Food Security

From a food security perspective, the observed improvements in yield stability have direct implications for the availability and stability pillars. Reduced yield volatility translates into more predictable supply, improved farmer income stability, and enhanced planning capacity for food systems.

Importantly, the results suggest that seed quality optimization can complement, rather than substitute, broader climate adaptation strategies. By acting at the upstream input level, seed optimization amplifies the effectiveness of downstream interventions such as precision irrigation and nutrient management.

4.7. Comparison with Expected Outcomes and Prior Studies

The observed outcomes correspond with theoretical anticipations that enhanced seed quality facilitates crop establishment and resistance. This study builds on previous research by quantifying these effects within a predictive and optimization-driven framework, as opposed to merely considering them as observational outcomes.

In contrast to studies that focus primarily on yield maximization, yield stability emerges as a distinct and optimizable objective, with seed quality serving as a pivotal influencing factor [2, 21]. This difference is very important for planning food security as climate change becomes more unpredictable.

5. Conclusion and Future Work

5.1. Conclusion

This research examined data-driven seed quality modification as a strategic mechanism for improving predictive yield stability and advancing food security goals. The research revealed that by including multidimensional seed quality traits into a cohesive analytical framework, seed quality may be treated as a dynamic optimization variable instead of a static quality-control measure.

The suggested Seed Quality Optimization Index (SQOI) made it possible to combine germination, vigor, purity, moisture content, and physical seed properties into one useful decision variable. When added to yield prediction models based on machine learning, the SQOI made the predictions far more accurate than baseline and extended models that didn't include structured seed quality representation. More crucially, the framework led to measurable decreases in the variability of yields from year to year. This shows that optimizing seed quality is important for keeping yields stable, not just for maximizing them.

The optimization layer further illustrated that precise modification of seed quality parameters might generate significant enhancements in yield resilience amidst fluctuating agroecological conditions. These findings underscore the strategic significance of upstream input optimization and furnish empirical evidence that seed quality interventions can substantially enhance the availability and stability components of food security. From a methodological perspective, the study enhances agricultural analytics by integrating seed science, machine learning, and optimization into a singular, reproducible system model.

5.2. Limitations

Despite its contributions, this study has several limitations. First, the use of a hybrid data strategy that mixes real-world and simulated datasets may not fully capture all sources of on-farm variability, even though it is important for generalizability. Second, the paradigm assumes that the correlations between seed quality traits and yield outcomes stay mostly the same, but they could alter over time due to climate change or changes in cropping methods. Third, the effects on food security are inferred from proxies for yield stability instead of directly modeled socioeconomic outcomes like how easy it is for households to get food or how the market works.

The SQOI also makes it easier to understand, but the weighting method still relies on model-derived important metrics to some extent. These measurements may be different for different datasets and modeling choices. These constraints underscore the necessity for ongoing validation and contextual modification.

5.3. Future Research Directions

Several avenues for future research emerge from this work. First, the framework can be expanded by including real-time sensing technologies, such as hyperspectral imaging and Internet of Things (IoT)–based seed and soil monitoring systems, to facilitate the dynamic updating of seed quality assessments. Second, using probabilistic and uncertainty-aware modeling techniques, such as Bayesian machine learning, could improve how we measure risk and make decisions when we don't know what will happen.

Future research might enhance the paradigm by explicitly incorporating economic and policy dimensions, connecting seed quality optimization to cost-benefit analysis, subsidy design, and seed system governance. In addition, combining the suggested method with climate scenario modeling would allow for evaluations of seed quality techniques that take into account possible temperature extremes in the future.

Finally, large-scale field testing in many crops and areas will be necessary to turn the framework into decision-support tools that farmers, seed manufacturers, and policymakers can use. These kinds of developments would make the function of data-driven seed quality optimization in constructing strong and long-lasting food systems even stronger.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] Tilman, D., Balzer, C., Hill, J., & Befort, B. L. (2011). Global food demand and the sustainable intensification of agriculture. PNAS, 108(50), 20260–20264. https://doi.org/10.1073/pnas.1116437108

[2] Ray, D. K., Gerber, J. S., MacDonald, G. K., & West, P. C. (2015). Climate variation explains a third of global crop yield variability. Nature Communications, 6, 5989. https://doi.org/10.1038/ncomms6989

[3] Food and Agriculture Organization of the United Nations. (2021). The state of food security and nutrition in the world 2021. FAO.

[4] Finch-Savage, W. E., & Bassel, G. W. (2016). Seed vigour and crop establishment: Extending performance beyond adaptation. Journal of Experimental Botany, 67(3), 567–591. https://doi.org/10.1093/jxb/erv490

[5] McDonald, M. B. (2019). Seed deterioration: Physiology, repair and assessment. Seed Science and Technology, 47(3), 239–254. https://doi.org/10.15258/sst.2019.47.3.01

[6] Powell, A. A., & Matthews, S. (2012). Seed ageing, repair and vigour. Seed Science Research, 22(S1), S13–S20. https://doi.org/10.1017/S0960258512000016

[7] Jeong, J. H., Resop, J. P., Mueller, N. D., Fleisher, D. H., Yun, K., Butler, E. E., … Kim, S. H. (2016). Random forests for global and regional crop yield predictions. PLOS ONE, 11(6), e0156571. https://doi.org/10.1371/journal.pone.0156571

[8] Liakos, K. G., Busato, P., Moshou, D., Pearson, S., & Bochtis, D. (2018). Machine learning in agriculture: A review. Sensors, 18(8), 2674. https://doi.org/10.3390/s18082674

[9] Shahhosseini, M., Hu, G., & Archontoulis, S. V. (2020). Forecasting corn yield with machine learning ensembles. Frontiers in Plant Science, 11, 1120. https://doi.org/10.3389/fpls.2020.01120

[10] van Klompenburg, T., Kassahun, A., & Catal, C. (2020). Crop yield prediction using machine learning: A systematic literature review. Computers and Electronics in Agriculture, 177, 105709. https://doi.org/10.1016/j.compag.2020.105709

[11] Benos, L., Tagarakis, A. C., Dolias, G., Berruto, R., Kateris, D., & Bochtis, D. (2021). Machine learning in agriculture: A comprehensive updated review. Sensors, 21(11), 3758. https://doi.org/10.3390/s21113758

[12] Khaki, S., & Wang, L. (2019). Crop yield prediction using deep neural networks. Frontiers in Plant Science, 10, 621. https://doi.org/10.3389/fpls.2019.00621

[13] Lobell, D. B., & Gourdji, S. M. (2012). The influence of climate change on global crop productivity. Plant Physiology, 160(4), 1686–1697. https://doi.org/10.1104/pp.112.204315

[14] Wheeler, T., & von Braun, J. (2013). Climate change impacts on global food security. Science, 341(6145), 508–513. https://doi.org/10.1126/science.1239402

[15] Food and Agriculture Organization of the United Nations. (2017). The future of food and agriculture: Trends and challenges. FAO

[16] Lesk, C., Rowhani, P., & Ramankutty, N. (2016). Influence of extreme weather disasters on global crop production. Nature, 529(7584), 84–87. https://doi.org/10.1038/nature16467

[17] Montesinos-López, O. A., Montesinos-López, A., Crossa, J., Toledo, F. H., Pérez-Hernández, O., Eskridge, K. M., & Rutkoski, J. (2018). Predicting grain yield using plant height, flowering time, and genomic information with machine learning methods. The Plant Genome, 11(2), 170012. https://doi.org/10.3835/plantgenome2017.02.0012

[18] You, J., Li, X., Low, M., Lobell, D., & Ermon, S. (2017). Deep Gaussian processes for crop yield prediction. AAAI Conference on Artificial Intelligence.

[19] Yan, W., & Kang, M. S. (2003). GGE biplot analysis: A graphical tool for breeders. CRC Press.

[20] Kang, M. S. (1993). Simultaneous selection for yield and stability in crop performance trials. Agronomy Journal, 85(3), 754–757. https://doi.org/10.2134/agronj1993.00021962008500030034x

[21] Lobell, D. B., & Burke, M. B. (2010). On the use of statistical models to predict crop yield responses to climate change. Agricultural and Forest Meteorology, 150(11), 1443–1452. https://doi.org/10.1016/j.agrformet.2010.07.008

[22] Rahman, A., Jahan, M. S., & Hossain, M. A. (2018). Seed quality assessment using image processing and machine learning techniques. Information Processing in Agriculture, 5(3), 356–365. https://doi.org/10.1016/j.inpa.2018.05.002

[23] Bai, X., Fang, Y., Wang, L., & Wang, M. (2021). Nondestructive evaluation of seed vigor and moisture content using hyperspectral imaging and machine learning techniques. Computers and Electronics in Agriculture, 182, 105990. https://doi.org/10.1016/j.compag.2021.105990

[24] Van Ittersum, M. K., Cassman, K. G., Grassini, P., Wolf, J., Tittonell, P., & Hochman, Z. (2016). Yield gap analysis with local to global relevance—A review. Field Crops Research, 186, 1–19. https://doi.org/10.1016/j.fcr.2015.09.009

[25] Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Transactions on Evolutionary Computation, 6(2), 182–197. https://doi.org/10.1109/4235.996017

[26] Hastie, T., Tibshirani, R., & Friedman, J. (2017). The elements of statistical learning: Data mining, inference, and prediction (2nd ed.). Springer. https://doi.org/10.1007/978-0-387-84858-7