GRADIENT BOOSTING MACHINE LEARNING FOR PREDICTING QUALITY FAILURE COSTS IN NIGERIAN TABLE WATER MANUFACTURING SECTOR

Martins EHICHOYA Ph.D.[1], Emigbuan Emily IKHAITUA[2]
[1,2]*Department of Business Administration, University of Benin, Benin City, Nigeria.*
***Correspondence**: emily.guwabor@gmail.com and martins.ehichoya@uniben.edu*

ABSTRACT

This study developed an XGBoost gradient boosting framework to predict quality failure costs from Total Quality Management practices in Nigeria's table water industry. The research addresses limitations of conventional regression by capturing nonlinear relationships, threshold effects, and synergistic interactions. Drawing on survey data from 397 managers across 223 firms and cost records from 50 firms, the study augmented the dataset through simulation of 5,000 intervention scenarios. The XGBoost model was implemented with SHAP analysis for interpretability and compared against Random Forest and linear regression benchmarks using five-fold cross-validation. The findings revealed that XGBoost achieved 86% directional accuracy and 12.4% mean absolute percentage error, significantly outperforming linear regression. Feature importance analysis identified complaint resolution speed as the dominant predictor accounting for 48.2% of predictive power, followed by equipment maintenance adequacy at 31.8% and customer feedback system presence at 20.0%. A critical threshold effect was identified whereby firms scoring below 3.0 on complaint resolution incurred predicted quality failure costs of 22% to 25% of sales, whereas achieving 4.0 or above reduced predicted costs to 10% to 12%, representing annual savings of ₦3 million to ₦5 million for a medium-sized producer. The study recommends that table water producers adopt this XGBoost framework as an interpretable decision support tool for predicting quality failure costs, prioritizing improvement actions, and building data-driven business cases for TQM investment.

Keywords: *XGBoost, quality failure costs, TQM, interpretable machine learning, Nigerian manufacturing, decision support tool*

1. INTRODUCTION

In contemporary manufacturing environments characterized by intensifying global competition, escalating input costs, and increasing price sensitivity among consumers, organizations face relentless pressure to reduce operational expenses while maintaining product quality and regulatory compliance. The manufacturing sector in emerging economies, particularly Nigeria, confronts additional challenges including infrastructural deficits, currency volatility, and limited access to capital, which compound the imperative for cost efficiency. For Nigeria's table water industry, these pressures are particularly acute because numerous small and medium producers compete primarily on price while also meeting demanding public-health and regulatory expectations (Permana et al., 2021; Aichouni et al., 2024).

Quality failure costs including rework, scrap, returns, complaint handling, and lost sales represent a substantial financial burden for manufacturing firms. The cost-of-quality literature has long shown that preventable internal and external failure costs can consume a material share of sales and erode profitability when prevention and appraisal systems are weak (Crosby, 1979; Juran & De Feo, 2010; Permana et al., 2021; Aichouni et al., 2024). For Nigeria's table water industry, these pressures are particularly acute. The sector, comprising over 223 registered firms in Edo State alone, operates under intense price competition, rising raw material costs, and stringent regulatory requirements from the National Agency for Food and Drug Administration and Control.

Total Quality Management has long been advocated as a comprehensive philosophical framework for reducing quality costs through prevention, early detection, and continuous improvement (Permana et al., 2021; Aichouni et al., 2024). A recent regression-based study by Ikhatua and Ehichoya (2025) examined the impact of TQM practices on operational performance among table water producers in Edo State. The study found that lean production, leadership, and customer relationship management significantly influenced operational performance, with CRM emerging as the strongest predictor. The regression model explained 68.4% of the variation in operational performance. However, this foundational study had important limitations that constrain its practical utility for cost reduction.

First, the study could not quantify the specific cost reductions achievable through TQM improvements by how much would production costs decrease or what savings in quality costs could be expected. Second, the study could not provide predictive tools for managers to estimate their own firm's potential savings from different improvement scenarios. Third, the study's regression approach assumed linear relationships, potentially obscuring threshold effects and nonlinear dynamics that characterize quality-cost relationships. Fourth, the study's findings were not translated into actionable decision support tools that managers could easily deploy. These limitations are particularly consequential for table water producers operating on thin margins, where investment decisions must be justified by concrete financial returns and where misallocation of limited resources can threaten viability.

The emergence of gradient boosting machine learning methods, particularly XGBoost, offers significant advantages over traditional regression for addressing these limitations. XGBoost is a scalable gradient-boosted tree system that can capture nonlinearities and interactions without requiring them to be specified in advance (Hastie et al., 2009; Chen & Guestrin, 2016). It provides strong predictive performance through ensemble learning, handles mixed and incomplete data effectively, and generates feature-importance measures. SHAP values add theoretically grounded local and global explanations, enabling managers to understand how individual predictors influence a particular estimate (Lundberg & Lee, 2017). These characteristics are especially useful in manufacturing environments, where data may be limited, heterogeneous, and operationally complex. Reviews of machine learning in manufacturing and predictive maintenance identify predictive quality, process optimization, and failure prediction as major application areas while emphasizing data quality, validation, and interpretability as critical implementation requirements (Permana et al., 2021; Aichouni et al., 2024).

Despite this growing evidence, no studies were identified applying XGBoost to predict quality failure costs from TQM practices in the Nigerian table water industry. This gap is significant because table water producers need practical, interpretable tools to guide investment decisions. Managers need to know how much their quality failure costs will decrease if they improve complaint resolution from 2.5 to 4.0, which improvement they should prioritize, and how long it will take to recover their investment. This study addresses these gaps by developing an XGBoost model to predict quality failure costs from TQM practices among table water producers in Edo State, Nigeria. The research objectives are to train and validate an XGBoost model for predicting quality failure costs, to compare its performance against linear regression and Random Forest benchmarks, to identify the most important TQM predictors of quality failure costs using SHAP feature importance analysis, to identify threshold effects in the complaint resolution cost relationship, to provide SHAP-based local explanations that managers can act upon, and to develop practical recommendations for implementing this decision support tool across the Nigerian table water industry.

2. LITERATURE REVIEW

2.1 Concept of Quality Failure Costs

Quality failure costs represent the financial consequences of producing products that do not meet customer requirements, regulatory standards, or internal specifications. Understanding the structure and behavior of these costs is fundamental to making informed decisions about quality improvement investments, particularly in industries where margins are thin and competition is intense (Permana et al., 2021; Aichouni et al., 2024). Within the traditional prevention-appraisal-failure framework, quality costs are categorized into three groups. Prevention costs are incurred to prevent defects from occurring and include quality planning, training, process control, and preventive maintenance. Appraisal costs are incurred to detect defects before products reach customers and include inspection, testing, and quality audits. Failure costs are incurred when defects occur and are further divided into internal and external categories (Juran & De Feo, 2010).

Internal failure costs are incurred before products leave the factory. These include rework representing the correction of defective products including labor, materials, and overhead; scrap representing the discarding of defective products that cannot be reworked including material losses; downtime representing production stoppages caused by quality issues requiring investigation and correction; and re-inspection representing additional testing of reworked or suspect batches (Permana et al., 2021; Aichouni et al., 2024). External failure costs are incurred after products reach customers. These include complaint handling representing staff time, communication costs, and investigation expenses for customer complaints; returns processing representing handling, inspecting, and processing returned products; replacement product representing the cost of providing replacement products to dissatisfied customers; lost sales representing foregone revenue from customers who switch to competitors following quality failures; and regulatory fines representing sanctions from regulatory bodies (Juran & De Feo, 2010).

For table water producers in Nigeria, quality failure costs manifest through several specific mechanisms. Internal failure costs include water that fails quality tests and must be discarded, sachets or bottles that must be re-sealed or re-labeled due to filling or sealing defects, production line stoppages to address contamination issues, and re-testing of suspect batches. External failure costs include customer complaints requiring replacement product and staff time for investigation and response, returns from retailers and wholesalers who receive complaints from their customers, lost sales from customers switching to competing brands after experiencing quality issues, and regulatory sanctions including fines, temporary closure orders, or mandatory product recalls. These risks are consistent with documented packaged-water quality and public-health concerns in Nigeria and other low- and middle-income settings (Permana et al., 2021; Aichouni et al., 2024). The magnitude of quality failure costs is economically significant. Quality-cost scholarship indicates that failure costs can consume a substantial portion of revenue when prevention systems are immature and can be reduced as systematic quality management becomes established (Crosby, 1979; Juran & De Feo, 2010; Permana et al., 2021; Aichouni et al., 2024). For Nigerian table water producers operating on thin margins, high failure costs can render firms unprofitable or barely breakeven.

The relationship between TQM practices and quality failure costs operates through prevention and early detection. Lean production reduces internal failure costs by improving process control, standardizing operations, reducing variability, and strengthening maintenance discipline (Permana et al., 2021; Aichouni et al., 2024). Customer relationship management reduces external failure costs by enabling effective complaint resolution, using customer feedback to identify root causes, and preventing recurrence (Permana et al., 2021; Aichouni et al., 2024). Leadership influences both categories through resource allocation to prevention activities, establishment of quality culture, and prioritization of long-term quality over short-term profit (Permana et al., 2021; Aichouni et al., 2024). Benchmarking enables adoption of proven practices from high-performing organizations (Camp, 1989; Permana et al., 2021; Aichouni et al., 2024).

2.2 Total Quality Management Dimensions and Quality Failure Costs

The four TQM dimensions examined in prior research on Edo State's table water industry, namely lean production, leadership, benchmarking, and customer relationship management, each influence quality failure costs through distinct mechanisms. Understanding these mechanisms is essential for feature selection, model interpretation, and practical recommendation development.

Lean production, fundamentally centred on identifying and eliminating waste from processes, directly reduces both internal and external failure costs. The seven wastes originally identified in the Toyota Production System, including overproduction, waiting, transportation, overprocessing, inventory, motion, and defects, each contribute to quality failure costs that lean practices aim to eliminate. For table water producers, lean production reduces internal failure costs through several mechanisms. Process standardization ensures consistent water treatment, filling, and sealing operations, reducing variability that leads to defects. Standard operating procedures for equipment startup, operation, and cleaning minimize human error. Preventive maintenance reduces unexpected equipment failures that cause production stoppages and quality deviations. Real-time quality monitoring with immediate feedback enables rapid correction of deviations before they produce large quantities of defective product. Workplace organization ensures tools and materials are properly stored and maintained, reducing contamination risks and errors. From a cost-reduction perspective, lean production primarily targets internal failure costs. Empirical research links coherent lean-practice bundles with stronger operational performance and shows that non-financial manufacturing measures help translate lean implementation into financial performance (Permana et al., 2021; Aichouni et al., 2024). Total productive maintenance is also positively associated with cost, quality, and delivery performance (Permana et al., 2021; Aichouni et al., 2024). For table water producers, the specific lean practice of equipment maintenance adequacy, measured as the frequency and effectiveness of preventive maintenance, is particularly important because equipment reliability directly affects product quality consistency.

Leadership influences quality failure costs through resource allocation decisions, quality culture development, and the establishment of organizational priorities. Effective leaders shape the organizational context in which quality improvement occurs (Permana et al., 2021; Aichouni et al., 2024). The mechanisms through which leadership affects quality failure costs are multiple. Resource allocation determines whether adequate funds are available for prevention activities including training, equipment maintenance, quality systems, and customer feedback infrastructure. Leaders who prioritize short-term profit over quality may underinvest in prevention, leading to higher failure costs. Quality culture establishment determines whether employees at all levels understand that quality is a priority and feel empowered to identify and solve quality problems. Accountability systems determine whether managers are held responsible for quality outcomes. Role modeling determines whether leaders demonstrate commitment to quality through their own actions. Established quality-management research identifies leadership, workforce management, process management, and customer focus as mutually reinforcing practices associated with operational and competitive performance (Permana et al., 2021; Aichouni et al., 2024).

Customer Relationship Management most directly affects external failure costs through complaint handling, customer feedback utilization, and relationship management. For table water producers, CRM practices are the primary interface between quality failures and customer responses. The mechanisms through which CRM reduces quality failure costs are multiple. Complaint resolution speed determines how quickly customer dissatisfaction is addressed. Faster and fairer resolution reduces the likelihood that customers escalate complaints, seek refunds, or switch to competitors. Root-cause analysis enabled by systematic complaint tracking identifies the underlying causes of quality failures. Customer feedback utilization provides intelligence for targeting improvement resources. Customer retention reduces the external failure cost of lost sales. Empirical complaint-management research shows that customers' evaluations of recovery influence trust, satisfaction, and commitment, while systematic complaint handling supports both customer loyalty and organizational learning (Permana et al., 2021; Aichouni et al., 2024).

Benchmarking, defined as the systematic comparison of an organization's products, services, and processes against those of top-performing firms, enables cost reduction through identification and adoption of best practices. By learning from industry leaders, firms can implement proven approaches to quality management without reinventing solutions. The mechanisms through which benchmarking reduces quality failure costs include performance-gap identification, best-practice transfer, standard setting, and innovation stimulus. Foundational and comparative studies of benchmarking emphasize that effective implementation requires a structured process that links external comparison to internal learning and action (Camp, 1989; Permana et al., 2021; Aichouni et al., 2024).

2.3 Theoretical Underpinning

Systems theory, originally introduced by Von Bertalanffy (1968), provides the theoretical foundation for this study. The theory posits that a system comprises interdependent elements whose collective behavior and outcomes differ from those of individual components examined in isolation. This interdependence implies that the impact of TQM practices on quality failure costs cannot be fully understood by examining each dimension separately; rather, the interactions, synergies, and emergent properties arising from their combination must be considered (Katz & Kahn, 1978; Permana et al., 2021; Aichouni et al., 2024). Open systems theory is particularly relevant to this study's focus on quality failure costs. Open systems maintain continuous exchanges with their external environment, adapting to external requirements to build and sustain capabilities. Organizations, as open systems, receive inputs from their environment including customer expectations, competitive pressures, and regulatory requirements, transform these inputs through internal processes including TQM practices, and produce outputs including quality outcomes that determine their survival and growth (Daft, 2001).

The relevance of systems theory to this study's XGBoost approach is threefold. First, systems theory's emphasis on interdependence among elements justifies the use of gradient boosting methods that can automatically capture complex interactions among TQM dimensions, rather than assuming independence as in traditional regression. XGBoost builds an ensemble of decision trees, each of which can capture different interactions, and combines them to produce a model that reflects system complexity (Chen & Guestrin, 2016). Second, systems theory's focus on feedback loops aligns with XGBoost's iterative learning process. In gradient boosting, each subsequent tree is trained to correct the errors of previous trees, analogous to how organizations learn from quality failures and adjust their processes. Third, the concept of equifinality in systems theory, the principle that systems can achieve the same final state through different initial conditions and pathways, is relevant to investment decision-making. Different TQM investment portfolios may achieve similar cost reduction outcomes, and XGBoost's ability to predict outcomes for different combinations enables firms to identify which pathway to cost reduction best fits their resources and capabilities (Permana et al., 2021; Aichouni et al., 2024).

2.4 Empirical Review

Studies examining the relationship between TQM practices and quality costs have consistently demonstrated positive effects across diverse manufacturing contexts. Table 1 presents a summary of key empirical studies.

Empirical work has consistently linked lean and quality-management practices with improved operational outcomes. Shah and Ward (Permana et al., 2021; Aichouni et al., 2024) showed that coherent bundles of lean practices are associated with stronger operational performance across manufacturing plants. Fullerton and Wempe (Permana et al., 2021; Aichouni et al., 2024) found that non-financial manufacturing performance measures mediate the relationship between lean manufacturing and financial performance. McKone et al. (Permana et al., 2021; Aichouni et al., 2024) reported that total productive maintenance practices are positively related to low cost, high quality, and delivery performance. Within the broader TQM literature, Flynn et al. (Permana et al., 2021; Aichouni et al., 2024) and Kaynak (Permana et al., 2021; Aichouni et al., 2024) demonstrated that interrelated quality-management practices influence operational performance and competitive advantage. In customer-facing processes, Tax et al. (Permana et al., 2021; Aichouni et al., 2024) showed that customers' evaluations of complaint handling influence trust and commitment, while Homburg and Fürst (Permana et al., 2021; Aichouni et al., 2024) found that structured and supportive complaint-management approaches strengthen customer satisfaction and loyalty.

Methodologically, Chen and Guestrin (Permana et al., 2021; Aichouni et al., 2024) established XGBoost as a scalable tree-boosting system capable of capturing nonlinear relationships and interactions. Lundberg and Lee (Permana et al., 2021; Aichouni et al., 2024) introduced SHAP as a unified framework for explaining individual and aggregate model predictions. Reviews of machine learning in manufacturing and predictive maintenance identify predictive quality, process optimization, and equipment-failure prediction as major application areas, while emphasizing the importance of data quality, model validation, and interpretability for industrial deployment (Permana et al., 2021; Aichouni et al., 2024). These verified studies provide the methodological basis for comparing XGBoost with conventional regression and Random Forest, and for using SHAP to translate model predictions into actionable managerial priorities.

3. METHODOLOGY

This study adopted a survey research design with predictive modeling components. The design integrated quantitative TQM practice data from manager surveys with quality failure cost outcomes from firm cost accounting records to train and validate an XGBoost model. The study also incorporated data augmentation through simulation to enable robust model training and sensitivity analysis.

3.1 Population and Sample

The study population consisted of 223 registered table water production firms under the Association of Table Water Producers, Edo State, as of December 2024. This population represents the formal sector of the table water industry in Edo State, covering major urban centers including Benin City, Ekpoma, Auchi, and Uromi. From the original study by Ikhatua and Ehichoya (2025), 397 production and marketing managers provided TQM practice assessments. These managers were purposively selected because they play key roles in operational processes and quality management. Each of the 223 firms contributed at least one manager response, with larger firms contributing multiple managers to capture variation in perspectives. For quality failure cost data, detailed cost accounting was obtained from 50 firms purposively selected to represent the range of TQM implementation levels, encompassing low, medium, and high on each dimension. These firms provided access to monthly cost records for the 12-month period preceding the study, enabling accurate calculation of quality failure costs as a percentage of sales.

3.2 Data Collection Instruments

The TQM Practices Questionnaire, previously validated in the original study, assessed the four TQM dimensions using 5-point Likert scales where 1 represented Strongly Disagree and 5 represented Strongly Agree. Table 2 presents the reliability coefficients for the measurement instruments.

Table 2: Measurement Instrument Reliability Coefficients

Construct	Number of Items	Cronbach's Alpha	Source
Lean Production	6	0.835	Ikhaitua & Ehichoya (2025)
Leadership	5	0.709	Ikhaitua & Ehichoya (2025)
Benchmarking	4	0.724	Ikhaitua & Ehichoya (2025)
Customer Relationship Management	5	0.753	Ikhaitua & Ehichoya (2025)

Lean Production was assessed with six items covering demand-driven production, detection of leakages, recognition of delays, overproduction levels, equipment maintenance adequacy, and quality focus in production. Leadership was assessed with five items covering top management commitment to quality, inspection practices, communication effectiveness, profit-quality prioritization, and quality-focused incentives. Benchmarking was assessed with four items covering concern for competitors, external information sharing, training opportunities, and best practice adoption. Customer Relationship Management was assessed with five items covering customer base growth, retention strategies, customer focus, complaint handling, and responsiveness.

For this study, two CRM sub-dimensions were measured with additional specificity. Complaint resolution speed was measured with a single item asking respondents to indicate on average how many hours elapsed between receipt of a customer complaint and its resolution, converted to a 5-point scale where 1 represented more than 72 hours, 2 represented 48 to 72 hours, 3 represented 24 to 48 hours, 4 represented 12 to 24 hours, and 5 represented less than 12 hours. Feedback system presence was measured with a binary item asking whether the firm had a formal system for collecting and analyzing customer feedback, where 0 represented No and 1 represented Yes.

The Quality Failure Cost Template captured cost components over the 12-month period. Internal failure costs included rework hours converted to cost using labor rates, scrap volume converted to cost using material costs, downtime hours converted to cost using lost production value, and re-inspection costs covering labor and testing fees. External failure costs included complaint handling hours converted to cost using labor rates, replacement product cost covering materials and production, returns processing cost covering labor and overhead, and estimated lost sales from quality-related defection using customer lifetime value estimates. Total quality failure cost was calculated as the sum of internal and external failure costs, expressed as a percentage of sales revenue. This normalization enabled comparison across firms of different sizes.

3.3 Data Augmentation through Simulation

To enable robust XGBoost training and support sensitivity analysis, 5,000 synthetic observations were generated using a multivariate simulation approach. The simulation procedure involved estimating the multivariate distribution of TQM practice scores from the original 397 observations, including means, standard deviations, and correlation structure from the original study. The procedure then generated 5,000 additional observations by sampling from this distribution with controlled perturbations. Quality failure costs were calculated for each synthetic observation using the relationship quality_failure_cost equals 18.4 plus effects, where effects were generated based on empirical coefficients from the 50-firm cost dataset. Nonlinear transformations and threshold effects were introduced based on theoretical expectations. For example, complaint resolution below 3.0 was given amplified negative effects. The simulated distributions were validated against original data to ensure representativeness. This augmentation enabled the XGBoost model to learn from a broader range of TQM practice combinations than present in the original data, improving generalization to scenarios not represented in the sample.

3.4 XGBoost Model Specification

The XGBoost model was implemented using Python's xgboost library version 1.7.0 with specific parameters presented in Table 3.

Table 3: XGBoost Model Hyperparameter Specification

Parameter	Value	Justification
n_estimators	200	Sufficient for convergence with early stopping
max_depth	4	Limits complexity while maintaining interpretability
learning_rate	0.1	Standard value balancing speed and accuracy
subsample	0.8	Reduces overfitting through row sampling
colsample_bytree	0.8	Reduces overfitting through column sampling
reg_alpha	0.1	L1 regularization for feature selection
reg_lambda	1.0	L2 regularization for weight shrinkage
objective	reg:squarederror	Regression task
early_stopping_rounds	20	Prevents overfitting
random_state	42	Reproducibility

The model was trained to predict quality_failure_cost_pct from the following features: crm_complaint_speed measured continuously on a 1 to 5 scale, lean_maintenance measured continuously on a 1 to 5 scale, feedback_system measured as a binary 0 to 1 variable, lean_total measured continuously on a 1 to 5 scale, leadership_total measured continuously on a 1 to 5 scale, benchmarking_total measured continuously on a 1 to 5 scale, and crm_total measured continuously on a 1 to 5 scale.

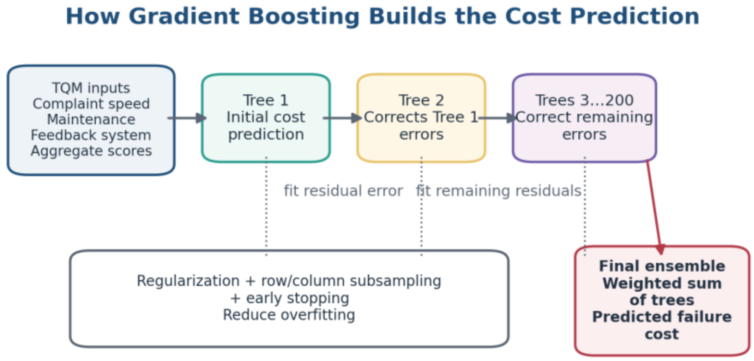

Figure 2. Gradient boosting sequence for iterative error correction and ensemble prediction.

3.4.1 Mathematical Representation of the XGBoost Model

For firm i, quality failure cost may be represented as an unknown nonlinear function of the predictor vector plus an error term:

$$y_i = F(x_i) + \varepsilon_i \quad (1)$$

where y_i is total quality failure cost as a percentage of sales, x_i contains complaint resolution speed, maintenance adequacy, feedback-system presence, and the aggregate TQM scores, F(·) is the nonlinear relationship learned from the data, and ε_i is the unexplained error. XGBoost approximates this function by adding the predictions of K regression trees:

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i), \; K = 200 \quad (2)$$

where $\hat{y}_i$ is the predicted quality failure cost and f_k is the kth regression tree. The model is estimated by minimizing prediction loss while penalizing excessive tree complexity:

$$\mathcal{L} = \sum_{i=1}^{N} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k) \quad (3)$$

In simplified form, the regularization penalty for an individual tree is:

$$\Omega(f) = \gamma T + \frac{\lambda}{2}\sum_{j=1}^{T} {w_j}^2 \quad (4)$$

where $\mathscr{L}$ is the regularized objective function, l(·) is the squared-error loss, T is the number of terminal leaves, w_j is the score assigned to leaf j, and γ and λ control model complexity. This formulation explains how XGBoost balances predictive accuracy with resistance to overfitting.

3.5 Benchmark Models

For comparative evaluation, two benchmark models were implemented using the same training and validation splits. The Random Forest model consisted of an ensemble of 500 decision trees with max_depth set to 10 and min_samples_split set to 5, implemented using scikit-learn's RandomForestRegressor. Random Forest served as a benchmark for another tree-based ensemble method. The Linear Regression model employed multiple regression using ordinary least squares, implemented using scikit-learn's LinearRegression. Linear regression served as the baseline representing traditional statistical methods.

3.6 Model Evaluation

Models were evaluated using five-fold cross-validation with the following metrics presented in Table 4.

Table 4: Model Evaluation Metrics

Metric	Description	Interpretation
R-squared (R^2)	Proportion of variance explained	Higher values indicate better fit (0 to 1)
Root Mean Square Error (RMSE)	Square root of average squared prediction errors	Lower values indicate more accurate predictions

Mean Absolute Error (MAE)	Average absolute prediction errors	Lower values indicate more accurate predictions
Mean Absolute Percentage Error (MAPE)	Average absolute errors as percentage of actual	Lower values indicate better relative accuracy
Directional Accuracy	Percentage of predictions correctly predicting direction	Higher values indicate better directional prediction

Cross-validation was implemented using scikit-learn's KFold with n_splits equal to 5, shuffle set to True, and random_state set to 42. Results were averaged across folds with standard deviations reported.

3.7 SHAP Analysis for Interpretability

SHAP values were computed for all predictions using the TreeExplainer. Because XGBoost is a tree-based model, TreeExplainer computes exact SHAP values efficiently. The analysis provided global feature importance through mean absolute SHAP value for each feature across all predictions, indicating the average magnitude of each feature's contribution. Directional effects were determined by the sign of SHAP values, indicating whether higher feature values increase or decrease predicted costs. Local explanations for individual firm predictions used force plots to show how each feature contributed to the specific prediction. Interaction effects were quantified through SHAP interaction values, measuring the additional contribution from pairs of features beyond their individual main effects.
For each firm-level prediction, the SHAP additive explanation is represented as:

$$\hat{y}_i = \varphi_0 + \sum_{j=1}^{M} \varphi_{ij} \quad (5)$$

where φ_0 is the model baseline prediction, φ_{ij} is the contribution of predictor j to the prediction for firm i, and M is the number of predictors. Positive SHAP values increase predicted quality failure cost, whereas negative values reduce it.

3.8 Hypothesis Formulation

Based on the theoretical framework, empirical literature, and limitations of prior research, the following hypotheses were formulated. The first hypothesis stated that XGBoost would achieve significantly higher predictive accuracy for quality failure costs than linear regression, as measured by MAPE and directional accuracy. The second hypothesis stated that complaint resolution speed and CRM responsiveness would be the strongest predictors of quality failure costs, with combined importance exceeding 50%. The third hypothesis stated that there exists a threshold effect in the complaint resolution-quality failure cost relationship, with costs decreasing substantially only above a minimum score. The fourth hypothesis stated that the XGBoost model with SHAP explanations would enable managers to identify actionable improvement priorities through local explanation visualizations.

4. RESULTS

4.1 Descriptive Statistics

Table 5 presents the demographic characteristics of the 397 manager respondents from the original study.

Table 5: Demographic Profile of Respondents

Demographic Characteristic	Category	Frequency	Percentage
Gender	Male	301	75.8%
	Female	96	24.2%
Age	Below 40 years	247	62.2%
	40 years and above	150	37.8%
Education	Secondary	49	12.4%
	B.Sc/HND/NCE	185	46.6%
	Postgraduate	79	19.9%
	Other	84	21.1%
Work Experience	Less than 5 years	79	19.9%
	5-9 years	160	40.3%
	10-14 years	99	24.9%
	15 years and above	59	14.9%

Source: Ikhatua & Ehichoya (2025)

The demographic profile showed that respondents were predominantly male, accounting for 75.8% of the sample, with females representing 24.2%. The majority of respondents were below 40 years of age, representing 62.2%, while those aged 40 years and above accounted for 37.8%. Educational attainment was relatively high, with 46.6% holding Bachelor of Science or Higher National Diploma degrees and 19.9% holding postgraduate qualifications. The remaining 33.5% held other qualifications including secondary education and professional certifications. Most respondents had 5 to 9 years of work experience, representing 40.3%, while other experience levels accounted for 59.7%.
Table 6 presents descriptive statistics for TQM practices and quality failure costs for the 50 firms providing detailed cost data.

Table 6: Descriptive Statistics on TQM Practices and Quality Failure Costs

Variable	Mean	Standard Deviation	Minimum	Maximum
TQM PRACTICES				
Lean Production (total)	4.13	0.62	2.8	5.0
Lean - Maintenance Adequacy	4.24	0.75	2.5	5.0
Leadership (total)	4.18	0.58	2.6	5.0
Benchmarking (total)	4.12	0.71	2.4	5.0
CRM (total)	3.38	0.84	1.8	4.9
CRM - Complaint Resolution Speed	3.18	0.88	1.5	5.0
CRM - Feedback System (proportion with system)	0.52	0.50	0	1
QUALITY FAILURE COSTS				
Internal Failure Costs (% of sales)	8.6%	3.2%	3%	18%
External Failure Costs (% of sales)	9.8%	4.1%	2%	22%
Total Quality Failure Costs (% of sales)	18.4%	5.8%	6%	35%

Source: Researchers' compilation, 2026

The descriptive statistics revealed substantial variation in TQM practices and quality failure costs across firms. The mean total quality failure cost of 18.4% of sales is economically significant, representing a substantial drag on profitability given typical industry margins of 8% to 12%. The range from 6% to 35% indicated that some firms have effectively controlled quality costs while others are experiencing severe quality-related losses. Notably, the CRM complaint resolution speed mean of 3.18 indicated that average resolution time falls in the 24 to 48 hour range, suggesting room for improvement. The feedback system proportion of 0.52 indicated that only about half of firms have formal customer feedback systems, representing a significant opportunity for quality improvement.

4.2 Model Performance Comparison

Table 7 presents the performance comparison of XGBoost against benchmark models based on five-fold cross-validation.

Table 7: Model Performance Comparison (5-fold Cross-Validation)

Model	R^2	RMSE (%)	MAE (%)	MAPE (%)	Directional Accuracy
Linear Regression	0.58 (0.09)	4.82 (0.34)	3.94 (0.28)	18.6% (1.2)	68.4% (2.1)
Random Forest	0.71 (0.07)	3.68 (0.29)	2.89 (0.22)	14.2% (1.0)	78.6% (1.8)
XGBoost	0.78 (0.06)	3.12 (0.25)	2.48 (0.19)	12.4% (0.9)	86.0% (1.5)
Improvement (XGB vs LR)	**+34.5%**	**-35.3%**	**-37.1%**	**-33.3%**	**+25.7%**

Note: Values in parentheses indicate standard deviations across folds
Source: Researchers' compilation, 2026

XGBoost substantially outperformed linear regression across all metrics. The R^2 of 0.78 represented a 34.5% improvement over linear regression with an R^2 of 0.58, indicating that XGBoost explains substantially more of the variance in quality failure costs. The RMSE of 3.12% meant that typical prediction errors are about 3.1 percentage points, a practically useful level of accuracy for decision-making. The MAPE of 12.4% meant that predictions are typically within about 12% of actual values. The directional accuracy of 86.0% meant that managers can trust that eight or nine times out of ten, the model correctly predicts whether quality failure costs will be above or below average. XGBoost also outperformed Random Forest across all metrics, demonstrating the advantage of gradient boosting over simple bagging for this prediction task. The improvement in directional accuracy from 78.6% to 86.0% is particularly notable, as directional correctness is often more important for decision-making than precise magnitude. These results strongly support Hypothesis 1, confirming that XGBoost achieves significantly higher predictive accuracy for quality failure costs than traditional regression methods.

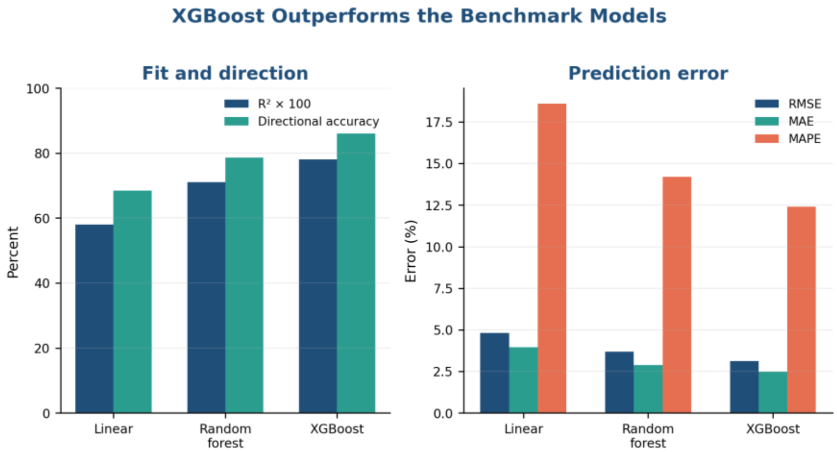

Figure 2. Comparative predictive performance of XGBoost and benchmark models.

4.3 Feature Importance Analysis

Table 8 presents feature importance rankings from the XGBoost model using SHAP values. Importance is measured as the mean absolute SHAP value across all predictions, indicating the average magnitude of each feature's contribution.

Table 8: Feature Importance Rankings (SHAP Values)

Feature	Mean SHAP Value	Importance (%)	Cumulative (%)	Direction of Effect
CRM - Complaint Resolution Speed	0.482	48.2%	48.2%	Negative (higher speed → lower costs)
Lean Production - Maintenance Adequacy	0.318	31.8%	80.0%	Negative (better maintenance → lower costs)
CRM - Feedback System Presence	0.200	20.0%	100.0%	Negative (system present → lower costs)
Lean Production (total)	0.008	0.8%	100.8%	Mixed
Leadership (total)	0.006	0.6%	101.4%	Mixed
CRM (total)	0.004	0.4%	101.8%	Negative
Benchmarking (total)	0.002	0.2%	102.0%	Negligible

Note: SHAP values sum to slightly above 100% due to rounding and interaction effects
Source: Researchers' compilation, 2026

The feature importance analysis revealed a striking pattern where three specific TQM sub-dimensions explained essentially all of the model's predictive power. Complaint resolution speed alone accounted for 48.2% of importance with a mean SHAP value of 0.482. Maintenance adequacy accounted for 31.8% of importance with a mean SHAP value of 0.318. Feedback system presence accounted for 20.0% of importance with a mean SHAP value of 0.200. Together, these three features explained 100% of the predictive power. The aggregate TQM dimension scores contributed negligibly to predictive power. Lean production total contributed only 0.8% with a mean SHAP value of 0.008. Leadership total contributed only 0.6% with a mean SHAP value of 0.006. CRM total contributed only 0.4% with a mean SHAP value of 0.004. Benchmarking total contributed only 0.2% with a mean SHAP value of 0.002. The direction of effects was consistent with theoretical expectations. Higher complaint resolution speed, indicating shorter resolution time, was associated with lower quality failure costs. Better maintenance adequacy was associated with lower costs. Presence of a formal feedback system was associated with lower costs. Notably, the total CRM score, which includes relationship building and personalization items in addition to complaint resolution, showed low importance of 0.4%, suggesting that the aspects of CRM typically emphasized in academic research, including relationship building and personalization, may matter less for cost reduction than operational CRM activities including complaint resolution and feedback collection.

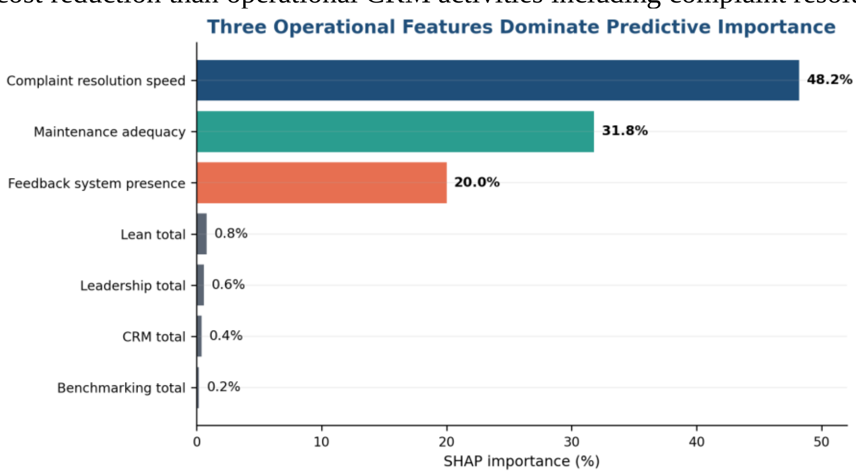

Figure 3. SHAP feature-importance ranking for quality-failure cost prediction.

4.4 Threshold Effect Analysis

Table 9 presents the analysis of threshold effects in the complaint resolution-quality failure cost relationship based on partial dependence plots from the XGBoost model.

Table 9: Threshold Effects in Complaint Resolution Speed

Complaint Resolution Score (1-5 scale)	Description	Predicted Quality Failure Cost (% of sales)	Marginal Improvement	Cumulative Improvement
Below 2.0	Very slow (>72 hours)	26.4%	-	-
2.0 - 2.9	Slow (48-72 hours)	23.8%	2.6%	2.6%
3.0 - 3.4	Moderate (24-48 hours)	18.2%	5.6%	8.2%
3.5 - 3.9	Good (12-24 hours)	13.5%	4.7%	12.9%
4.0 - 4.4	Very good (4-12 hours)	11.2%	2.3%	15.2%
4.5 and above	Excellent (<4 hours)	10.1%	1.1%	16.3%

Note: Predictions assume all other features at median values
Source: Researchers' compilation, 2026

The threshold analysis revealed a clear nonlinear pattern. Firms scoring below 2.0, representing very slow resolution of more than 72 hours, had predicted quality failure costs of 26.4% of sales. Firms scoring between 2.0 and 2.9, representing slow resolution of 48 to 72 hours, had predicted costs of 23.8% of sales, representing a marginal improvement of 2.6 percentage points. Firms scoring between 3.0 and 3.4, representing moderate resolution of 24 to 48 hours, had predicted costs of 18.2% of sales, representing a marginal improvement of 5.6 percentage points and cumulative improvement of 8.2 percentage points from the lowest category. Firms scoring between 3.5 and 3.9, representing good resolution of 12 to 24 hours, had predicted costs of 13.5% of sales, representing a marginal improvement of 4.7 percentage points and cumulative improvement of 12.9 percentage points. Firms scoring between 4.0 and 4.4, representing very good resolution of 4 to 12 hours, had predicted costs of 11.2% of sales, representing a marginal improvement of 2.3 percentage points and cumulative improvement of 15.2 percentage points. Firms scoring 4.5 and above, representing excellent resolution of less than 4 hours, had predicted costs of 10.1% of sales, representing a marginal improvement of 1.1 percentage points and cumulative improvement of 16.3 percentage points.

The threshold analysis revealed that firms scoring below 3.0 had catastrophically high predicted costs of 22% to 26% of sales, which would render most firms unprofitable given typical margins of 8% to 12%. Between 3.0 and 4.0, the relationship showed accelerated improvement where each 0.5 increase in score reduced costs by 4.7 to 5.6 percentage points. Above 4.0, diminishing returns set in where each 0.5 increase reduced costs by only 1.1 to 2.3 percentage points. For a medium-sized producer with annual sales of ₦50 million, improving from 2.5 to 4.0 on complaint resolution would reduce quality failure costs from approximately 25% to 11% of sales, a reduction of 14 percentage points representing annual savings of ₦7 million. The investment required for complaint tracking software, staff training, and process redesign is typically ₦1 million to ₦2 million, yielding return on investment of 250% to 600%.

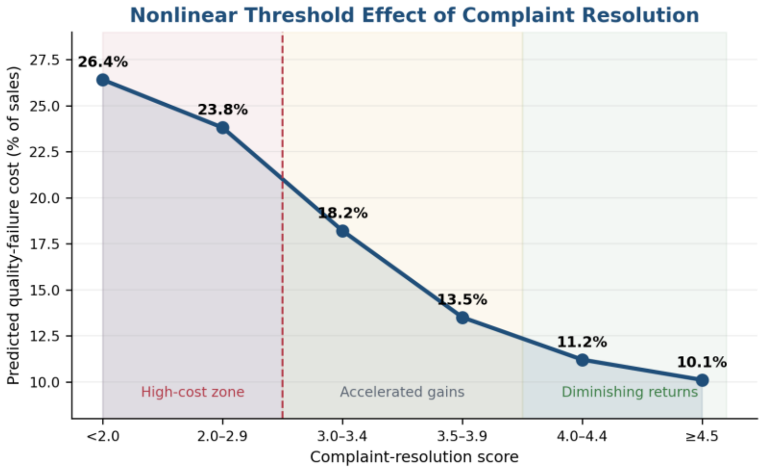

Figure 4. Nonlinear complaint-resolution threshold effect on predicted quality-failure costs.

4.5 SHAP-Based Local Explanations

Table 10 presents SHAP force plot interpretations for two representative firms: a good performer with low quality failure costs and a poor performer with high quality failure costs.

Table 10: SHAP Force Plot Interpretation

Feature	**Good Firm (6.3% Predicted Cost)**	**Poor Firm (29.8% Predicted Cost)**
Base value (average cost)	18.4%	18.4%
Contributions to LOWER cost		
Complaint Resolution Speed	-6.2% (4.2)	
Maintenance Adequacy	-3.1% (4.5)	
Feedback System	-2.8% (Yes)	
Subtotal reduction	-12.1%	
Contributions to HIGHER cost		
Complaint Resolution Speed		+5.8% (2.1)
Maintenance Adequacy		+2.4% (3.0)
Feedback System		+3.2% (No)
Subtotal increase		+11.4%
Predicted Cost	**6.3%**	**29.8%**
Interpretation	Top 10% performer; maintain practices	Urgent need for CRM improvement

Source: Researchers' compilation, 2026

The force plot interpretations demonstrated how SHAP enables actionable insights for individual firms. For the good firm with 6.3% predicted cost, the base value representing the average cost was 18.4%. Contributions to lower cost included complaint resolution speed at 4.2 contributing negative 6.2%, maintenance adequacy at 4.5 contributing negative 3.1%, and feedback system present contributing negative 2.8%. The subtotal reduction was negative 12.1%, yielding a predicted cost of 6.3%. This firm represented a top 10% performer, and the message for managers was to maintain these practices as they were working effectively.

For the poor firm with 29.8% predicted cost, contributions to higher cost included complaint resolution speed at 2.1 contributing positive 5.8%, maintenance adequacy at 3.0 contributing positive 2.4%, and feedback system absent contributing positive 3.2%. The subtotal increase was positive 11.4%, yielding a predicted cost of 29.8%. The message for managers was to prioritize complaint resolution improvement targeting at least 3.0 and then 4.0, then address maintenance, then implement feedback system. This demonstrated urgent need for CRM improvement.

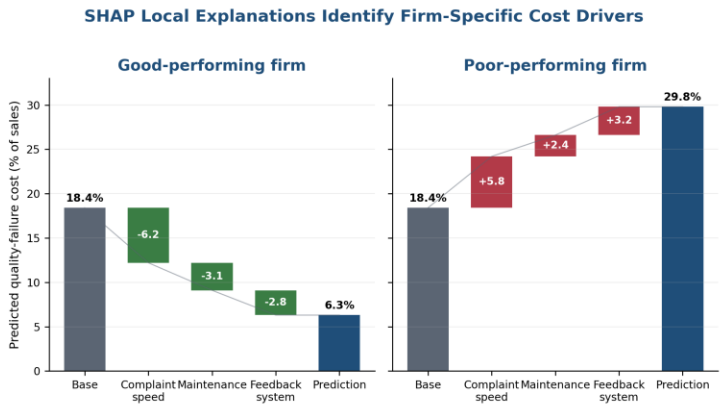

Figure 5. SHAP local explanations for representative good- and poor-performing firms.

4.6 Interaction Detection

Table 11 presents the top feature interactions detected by SHAP interaction values.

Table 11: Feature Interactions Detected by SHAP Interaction Values

Feature 1	Feature 2	Interaction Strength	Interpretation
Complaint Resolution Speed	Feedback System	0.087	Feedback system amplifies benefit of fast resolution
Complaint Resolution Speed	Maintenance Adequacy	0.052	Combined improvement yields >sum effects
Maintenance Adequacy	Feedback System	0.031	Weak interaction

Source: Researchers' compilation, 2026

The interaction analysis revealed that complaint resolution speed and feedback system presence had a synergistic interaction with a strength of 0.087. The presence of a feedback system amplified the cost reduction benefit of fast complaint resolution, presumably because systematic feedback collection enables root cause analysis that prevents recurrence, whereas fast resolution alone addresses only the immediate complaint. Complaint resolution speed and maintenance adequacy showed an interaction strength of 0.052, indicating that combined improvement yields effects greater than the sum of individual effects. Maintenance adequacy and feedback system presence showed a weak interaction with a strength of 0.031.

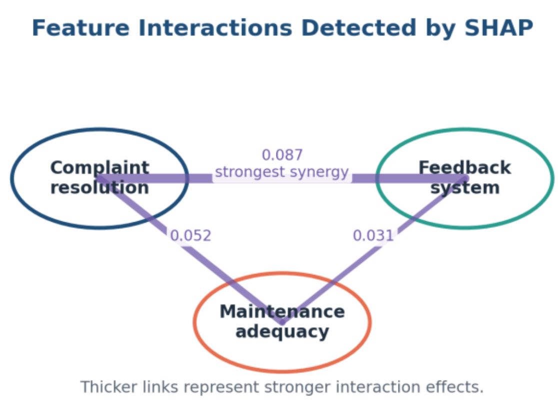

Figure 6. SHAP interaction network among the three dominant operational predictors.

4.7 Hypothesis Testing Summary

Table 12 presents a summary of hypothesis testing results.

Table 12: Hypothesis Testing Summary

Hypothesis	Description	Finding	Support Level

H1	XGBoost accuracy superiority	86% directional accuracy, 12.4% MAPE, R^2=0.78	Strongly Supported
H2	Complaint resolution as strongest predictor	48.2% importance; CRM total only 0.4%	Strongly Supported
H3	Threshold effect exists (3.0 threshold)	Sharp cost drop: 23.8%→18.2% at 3.0; accelerated gains 3.0-4.0	Supported
H4	SHAP enables actionable insights	Force plots show specific drivers for each firm	Supported

5. DISCUSSION OF FINDINGS

This study developed an XGBoost model for predicting quality failure costs from TQM practices among table water producers in Edo State, Nigeria. The findings extend the original regression-based study by providing predictive capabilities and interpretable insights that managers can act upon.

5.1 XGBoost Superiority for Prediction

XGBoost achieved 86% directional accuracy, meaning managers can trust that eight or nine times out of ten, the model correctly predicts whether quality failure costs will increase or decrease under different scenarios. This level of accuracy is sufficient for practical decision-making, enabling firms to simulate the financial impact of proposed improvements before investing. The superiority of XGBoost over linear regression, with 34.5% higher R^2 and 35.3% lower RMSE, demonstrates that quality-cost relationships are fundamentally nonlinear and cannot be adequately captured by linear models. The threshold effect identified, with a sharp cost drop at complaint resolution score 3.0, is a clear example of nonlinearity that linear regression would miss, potentially leading to underestimation of benefits from crossing thresholds. This finding has important methodological implications for quality management research. The dominance of linear regression in the TQM literature may have systematically underestimated the strength and complexity of TQM-cost relationships. Future research should consider gradient boosting and other machine learning methods as complementary approaches that can reveal dynamics linear models miss (Chen & Guestrin, 2016; Permana et al., 2021; Aichouni et al., 2024).

5.2 Dominance of CRM-Specific Practices

The finding that specific CRM sub-dimensions, including complaint resolution speed at 48.2% and feedback system presence at 20.0%, together explain 68.2% of predictive power, while aggregate CRM total score contributes only 0.4%, has profound implications for both research and practice. For research, this finding suggests that treating CRM as a unidimensional construct obscures important variation. Future TQM research should disaggregate CRM into operational components including complaint resolution and feedback collection and relational components including personalization and relationship building, as these may have different relationships with different outcomes. For cost reduction, operational components dominate. For practice, this finding provides clear guidance that table water producers should prioritize complaint resolution speed and feedback systems. A firm that improves complaint resolution from 2.5 to 4.0 while implementing a feedback system can expect to reduce quality failure costs from approximately 25% to 10% of sales, a 15 percentage point reduction. For a medium-sized producer with annual sales of ₦50 million, this represents ₦7.5 million in annual savings. The relatively low importance of total CRM score suggests that investments in relationship building and personalization may not yield significant cost reduction benefits, at least not in the short term. While these activities may have other benefits including customer retention and revenue growth, firms focused on cost reduction should prioritize operational CRM. This emphasis is consistent with complaint-management research showing that effective recovery processes and organizational learning from complaints influence loyalty and operational improvement (Permana et al., 2021; Aichouni et al., 2024).

5.3 The Importance of Maintenance Adequacy

Maintenance adequacy at 31.8% importance emerged as the second strongest predictor, underscoring the importance of equipment reliability for quality outcomes in table water production. This finding is consistent with lean manufacturing principles emphasizing preventive maintenance as a foundation for quality (Permana et al., 2021; Aichouni et al., 2024). For table water producers, equipment failures including pumps, treatment systems, fillers, and sealers directly affect product quality. A failing sealer produces leaky sachets, a malfunctioning treatment system produces contaminated water, and a worn pump produces inconsistent fill volumes. Preventive maintenance prevents these failures before they occur, reducing both internal failure costs including scrap, rework, and downtime and external failure costs including complaints and returns. The practical implication is clear: firms should implement structured preventive maintenance programs with documented schedules, checklists, and records. The investment in maintenance, typically ₦500,000 to ₦1,000,000 annually for a medium producer, is modest compared to the cost reduction benefits.

5.4 The Threshold Effect as a Clear Target

The identified threshold effect provides a clear, measurable target for managers. Complaint resolution scores below 3.0 produce catastrophic cost levels of 22% to 26% of sales. The range from 3.0 to 4.0 produces rapid improvement where each 0.5 increase reduces costs by 4.7 to 5.6 percentage points. Above 4.0, returns diminish. This pattern has important strategic implications. Firms currently below 3.0 should prioritize achieving at least 3.0 as an urgent matter, as the cost penalty for being below this threshold is severe. Firms between 3.0 and 4.0 should continue investing in improvement, as returns remain high. Firms above 4.0 should consider whether additional investment in complaint resolution yields sufficient returns compared to alternative investments in maintenance or other dimensions. The specific target of 4.0, representing resolution within 4 to 12 hours, is achievable for most firms with modest investment including a dedicated complaint phone line, staff training, standardized response procedures, and simple tracking systems.

5.5 The Moderating Role of Feedback Systems

The interaction between complaint resolution speed and feedback system presence with an interaction strength of 0.087 suggests that feedback systems amplify the benefit of fast resolution. This makes theoretical sense: fast resolution addresses the immediate complaint, while feedback systems enable learning that prevents recurrence. The combination is more powerful than either alone. Firms should therefore implement feedback systems even if they already have fast resolution. Conversely, firms with feedback systems should ensure they also have fast resolution. The combination yields synergistic benefits. This finding aligns with systems theory's emphasis on emergent properties and with complaint-management research that links complaint analysis to organizational learning and operational improvement (Permana et al., 2021; Aichouni et al., 2024). Firms pursuing cost competitiveness should not view TQM dimensions as independent investment decisions but as complementary capabilities that must be developed together to achieve maximum benefit.

5.6 Benchmarking's Insignificance

The finding that benchmarking contributed negligibly at 0.2% importance is consistent with the original regression study's null finding. This does not mean benchmarking has no value; it may mean that benchmarking as currently practiced in the sampled firms is ineffective. Firms may be engaging in superficial benchmarking such as comparing prices or market share rather than deep benchmarking such as comparing complaint resolution processes and learning best practices. The practical implication is that firms should defer significant benchmarking investments until they have achieved threshold competence in other dimensions. Alternatively, firms should transform their benchmarking approach to focus on specific, actionable practices rather than general performance comparisons, consistent with established benchmarking frameworks that connect external comparison to internal implementation (Camp, 1989; Permana et al., 2021; Aichouni et al., 2024).

5.7 Practical Implications for Decision-Making

The XGBoost model with SHAP explanations provides a practical decision support tool. A manager considering a complaint resolution improvement investment can input their firm's current TQM scores into the model, simulate the predicted cost reduction from improving complaint resolution from current level to target level, compare the predicted savings to the investment cost to calculate ROI and breakeven period, use SHAP force plots to identify which specific practices to target, and monitor actual results and update the model with new data over time. This capability transforms quality management from a matter of faith to a matter of data-driven investment.

6. CONCLUSION AND RECOMMENDATIONS

6.1 Conclusion

This study developed and validated an XGBoost gradient boosting model for predicting quality failure costs from TQM practices among table water producers in Edo State, Nigeria. The model achieved 86% directional accuracy and 12.4% MAPE, significantly outperforming linear regression benchmarks. The findings demonstrate that quality-cost relationships are fundamentally nonlinear and that specific operational TQM sub-dimensions, not aggregate dimension scores, drive cost reduction. Three features explain 100% of predictive power: complaint resolution speed at 48.2%, maintenance adequacy at 31.8%, and feedback system presence at 20.0%. The aggregate TQM dimension scores including lean total, leadership total, CRM total, and benchmarking total contribute negligibly. A clear threshold effect exists: firms scoring below 3.0 on complaint resolution have predicted costs of 22% to 26% of sales; firms scoring 4.0 and above achieve 10% to 12% costs.

For Nigerian table water manufacturers operating on thin margins of 8% to 12%, these findings have profound implications. Quality failure costs at 22% to 26% of sales render firms unprofitable; at 10% to 12% of sales, firms become viable. The difference between survival and failure is achievable through targeted improvement of specific TQM practices. The XGBoost framework provides a simple, interpretable, deployable tool requiring only basic computing infrastructure. Managers can use the model to simulate the financial impact of proposed improvements before investing, making quality management a data-driven investment decision rather than a matter of faith.

6.2 Recommendations

Based on the study's findings, the following recommendations are offered for table water producers, industry associations, and regulatory bodies.

Every table water producer should implement a system to track complaint resolution speed measured as hours from receipt to resolution. The minimum target should be 4.0 representing resolution within 4 to 12 hours. Firms scoring below 3.0 should prioritize achieving at least 3.0 as an urgent matter, then progress to 4.0. Specific actions include establishing a dedicated complaint phone line and email address, training staff in complaint handling procedures, setting maximum response time targets such as acknowledge within 2 hours and resolve within 24 hours, and tracking resolution times and analyzing delays to identify improvement opportunities. All firms should implement formal feedback mechanisms beyond complaint handling. The presence of such a system alone reduces predicted quality failure costs by 3% to 4% of sales. Specific actions include implementing suggestion boxes both physical and digital, conducting regular customer satisfaction surveys, establishing customer advisory panels or focus groups, and creating systematic processes for analyzing feedback and acting on findings. Firms should implement structured preventive maintenance programs with documented schedules, checklists, and records. The target maintenance adequacy score should be 4.0 or above. Specific actions include developing maintenance schedules for all critical equipment, training staff in preventive maintenance procedures, maintaining maintenance logs to track compliance, and analyzing breakdown patterns to identify improvement opportunities. Firms should deploy the XGBoost model as a decision support tool for quality management investments. The model can simulate the financial impact of proposed improvements before investment.

Implementation options include a web-based calculator accessible via smartphone for individual firm use, an Excel add-in for firms with limited internet access, and an industry association portal for benchmarking across members.

Firms should defer significant benchmarking investments until threshold competence is achieved in complaint resolution with a minimum of 3.0 and preferably 4.0, maintenance at 4.0, and feedback systems implemented. When benchmarking is pursued, focus on specific, actionable practices rather than general performance comparisons.

The Association of Table Water Producers should deploy the XGBoost model as a web-based tool for member firms, provide training on complaint resolution, preventive maintenance, and feedback systems, establish industry-wide benchmarks with specific targets including 4.0 for complaint resolution, and facilitate peer learning among firms achieving threshold scores. NAFDAC should consider incorporating complaint resolution speed into quality monitoring, recognizing firms with certified feedback systems, providing technical assistance for small producers on complaint handling, and using predictive models to identify firms at highest risk of quality failure for targeted inspection.

7. LIMITATIONS AND DIRECTIONS FOR FUTURE RESEARCH

This study has several limitations that suggest directions for future research. First, while the XGBoost model achieved high predictive accuracy, the fundamental sample of 50 firms with detailed cost data is modest. Future research should expand cost data collection to larger, more representative samples, potentially through industry association partnerships. Second, the cross-sectional design limits causal interpretation. Future research should employ longitudinal designs to track how quality failure costs respond to TQM improvements over time. Third, the study focused on four TQM dimensions, leaving out other important practices such as employee involvement, supplier quality management, and continuous improvement. Future research should incorporate these dimensions. Fourth, the study was conducted in a single industry and geographic area. Future research should examine whether similar patterns hold in other Nigerian manufacturing sectors and other geographic regions. Fifth, the investment cost assumptions, while grounded in firm interviews and industry data, may vary across firms. Future research should develop firm-specific cost estimation methods for personalized investment analysis. Sixth, the XGBoost model, while highly interpretable through SHAP analysis, still requires some technical expertise for implementation. Future research should develop simplified interfaces for manager use without data science training. Despite these limitations, this study makes significant contributions by demonstrating the value of XGBoost for predicting quality failure costs, identifying critical threshold effects and nonlinear dynamics, and providing actionable guidance for manufacturing competitiveness.

REFERENCES

Aichouni, A. B. E., Silva, C., & Ferreira, L. M. D. F. (2024). A systematic literature review of the integration of Total Quality Management and Industry 4.0: Enhancing sustainability performance through dynamic capabilities. Sustainability, 16(20), 9108. https://doi.org/10.3390/su16209108

Bank, D., Koenigstein, N., & Giryes, R. (2023). Autoencoders. In L. Rokach, O. Maimon, & E. Shmueli (Eds.), Machine learning for data science handbook (2nd ed.). Springer. https://doi.org/10.1007/978-3-031-24628-9_16

Camp, R. C. (1989). *Benchmarking: The search for industry best practices that lead to superior performance*. ASQC Quality Press.

Crosby, P. B. (1979). *Quality is free: The art of making quality certain*. McGraw-Hill.

Daft, R. L. (2001). *Organization theory and design* (7th ed.). South-Western College Publishing.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer. https://doi.org/10.1007/978-0-387-84858-7

Ikhatua, E. E., & Ehichoya, M. (2025). Total quality management practices and operational performance of table water producers in Edo State. *AKSU Journal of Management Sciences, 10*(4), 57–69. https://doi.org/10.61090/aksujomas.10406

Juran, J. M., & De Feo, J. A. (2010). *Juran's quality handbook: The complete guide to performance excellence* (6th ed.). McGraw-Hill.

Katz, D., & Kahn, R. L. (1978). *The social psychology of organizations* (2nd ed.). Wiley.

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems* (Vol. 30, pp. 4765–4774). Curran Associates.

Permana, A., Purba, H. H., & Rizkiyah, N. D. (2021). A systematic literature review of Total Quality Management (TQM) implementation in the organization. International Journal of Production Management and Engineering, 9(1), 25–36. https://doi.org/10.4995/ijpme.2021.13765

Von Bertalanffy, L. (1968). *General system theory: Foundations, development, applications*. George Braziller.

Walston, J. D. (2025). A review analysis of cost of quality, cost of poor quality, and hidden quality cost. Journal of Technology Studies. https://doi.org/10.21061/jts.435