www.ijarp.com ISSN 2456-9992 Page: 01-22

International Journal Advanced Research Publications

Research Article

Volume: 02
Issue: 07

RESPONSE SURFACE METHODOLOGY AND MACHINE LEARNING PREDICTIVE MODELLING FOR ACHA DEHULLING AND CLEANING

Ezekiel A. Mamai*[1], Makolo A. Muhammed[2] and Ahmed U. Balarabe.[3]

[1] Department of Agricultural Engineering, Federal University, Wukari, Nigeria.

[2] Department of Agricultural & Biosytems Engineering, Joseph Sarwuan Tarka University, Makurdi-Nigeria.

[3]Department of Agricultural Engineering Technology, College of Agriculture, Science & Technology, Lafia.

Article Received: 10 June 2026, Article Revised: 30 June 2026, Published on: 20 July 2026

***Corresponding Author: Ezekiel A. Mamai**

Department of Agricultural Engineering, Federal University, Wukari, Nigeria.

Doi: https://doi-doi.org/101555/ijarp.3697

1 | INTRODUCTION

Recently, acha (*Digitaria exilis*) has gained increasing attention in scientific research due to its high nutritional value, rapid maturity, and ability to thrive under harsh agroclimatic conditions where many conventional cereals fail. In the broader context of biosystems engineering, Response Surface Methodology (RSM) constitutes a widely recognized approach for optimizing threshing, dehulling, and cleaning operations in cereals such as rice, wheat, and oats. Despite its extensive application in cereal processing, the application of RSM to acha dehulling systems remains limited and insufficiently explored. Therefore, a clear gap exists in the literature regarding the application of data-driven and artificial intelligence-based approaches to acha seed processing systems. This study therefore, pioneers the integration of machine learning approaches, in which RSM and SVR are jointly applied to predict optimal operating parameters for acha hull removal and winnowing. This hybrid approach overcomes the limitations associated with standalone higher-order polynomial optimization models and provides a robust empirical decision-support tool for enhancing acha dehulling machinery. The study is aimed at:

(i) applying RSM to design and analyse experiments relating cylinder speed, fan speed, feed rate, and tray inclination angle to dehulling efficiency, cleaning efficiency, and grain loss of a patented acha dehulling machine developed by Itodo et al. (2023a);

(ii) developing SVR models trained on experimental data to accurately predict these performance indices across the operating range; and

(iii) comparing the predictive accuracy and optimization capability of RSM and SVR with the aim of establishing a robust hybrid modelling framework for acha dehulling machine optimization and future design improvement.

2 | METHODOLOGY

2.1 Experimental Data Sources

The data utilised in this study were obtained from performance evaluation trials conducted on a patented acha dehulling and cleaning machine developed by Itodo et al. (2023b). The experiments were executed across a wide operating range of key machine parameters to ensure sufficient variability for robust statistical analysis and reliable machine learning model development. The independent variables examined were feed rate (g min^{-1}), cylinder speed (rpm), fan speed (rpm), and tray inclination angle (°), selected based on their documented influence on grain–machine and grain–airflow interactions.

Cylinder speed is a critical operating parameter governing the impact force and rubbing action responsible for hull removal during acha dehulling, directly influencing dehulling efficiency and grain breakage. Fan speed was varied across predefined levels to evaluate its effect on cleaning efficiency and grain loss, with particular emphasis on balancing effective separation and minimizing grain carry-over. Adjusting the tray angle influenced grain stratification, sliding velocity, and separation efficiency, thereby affecting both cleaning performance and grain residence time.

The RSM was applied as a conventional statistical modelling approach to establish empirical relationships between the operating parameters and machine performance indices. Unlike RSM, which assumes a predefined polynomial relationship between inputs and outputs, SVR constructs data-driven models that map input variables to response variables by minimizing structural risk, thereby enhancing generalization capability. The following evaluation criteria were employed to assess the predictive performance of the models:

2.2 Performance Evaluation Parameters

Drum (Cylinder) Speed

Cylinder speed is a critical operating parameter governing the impact force and rubbing action responsible for hull removal during acha dehulling. In the patented machine, the dehulling drum consists of abrasive elements mounted on a rotating shaft, designed to exert controlled mechanical action on the grains.

During experimentation, cylinder speed was varied systematically within the operational limits of the machine using a variable speed drive. The selected speed range was chosen to ensure effective hull detachment while minimizing excessive grain breakage and powdering. Changes in cylinder speed directly influenced dehulling efficiency and grain loss by altering the intensity of grain–drum interaction.

Fan Speed

Fan speed controls the airflow supplied to the cleaning unit and plays a major role in separating dehulled grains from husk, chaff, and other light impurities. An axial-flow fan integrated into the machine generated adjustable airflow through the cleaning chamber.

Fan speed was varied across predefined levels to evaluate its effect on cleaning efficiency and grain loss. Insufficient airflow resulted in poor removal of light materials, whereas excessive airflow increased the risk of grain carryover and loss. Therefore, fan speed was treated as a key variable in balancing effective cleaning with minimal grain loss.

Tray Inclination Angle

The tray inclination angle determines the residence time and flow behaviour of materials on the

cleaning tray. The cleaning tray was designed with an adjustable tilt mechanism, allowing variation of the inclination angle relative to the horizontal plane.

Adjusting the tray angle influenced grain stratification, sliding velocity, and separation efficiency. Lower angles increased residence time and improved cleaning but could reduce throughput, while higher angles enhanced material flow at the expense of separation efficiency. The selected range of tray inclination angles was consistent with earlier acha processing studies (Kaankuka et al., 2014; Itodo et al., 2023a).

2.2.1 Determination of Dehulling Efficiency

Dehulling efficiency (DhEff) was used to evaluate the effectiveness of the dehulling mechanism in removing the hull from acha grains. It was defined as the proportion of dehulled grains obtained relative to the total mass of grains fed into the machine. The DhEff was evaluated with Equation (1):

$$\mathrm{DhEff}(\%) = \frac{M_d}{M_t} \times 100 \qquad (1)$$

where:

M_d= mass of dehulled acha grains collected after processing (g),

M_t= total mass of acha grains fed into the machine (g).

This approach to evaluating dehulling performance has been widely applied in studies on acha and other small-grain dehulling and threshing machines (Kaankuka et al., 2015; Itodo et al., 2023).

2.2.2 Determination of Cleaning Efficiency

Cleaning efficiency (CEff) measures the ability of the cleaning unit to effectively separate dehulled grains from husk, chaff, and other foreign materials. It was defined as the proportion of the mass of clean dehulled acha seed to the total mass of foreign materials discharged from the cleaning outlet. The CEff was calculated using Equation (2):

$$\text{CEff}(\%) = \frac{M_c}{M_c + M_f} \times 100 \qquad (2)$$

where:

M_c= mass of clean dehulled acha grains collected (g),

M_f= mass of foreign materials (husk, chaff, and impurities) collected at the cleaning outlet (g).

This method of evaluating cleaning efficiency is consistent with standard procedures used in agricultural and biosystems engineering studies involving cereal cleaning and separation systems (Ajav & Fakayode, 2013; Oluwole et al., 2016).

2.2.3 Determination of Grain Loss

Grain loss (GLss) represents the proportion of grains lost during the dehulling and cleaning processes

due to spillage, excessive airflow, or incomplete separation. It was expressed as a percentage of the total grain input and determined using Equation (3):

$$\text{GLss}(\%) = \frac{M_l}{M_t} \times 100 \qquad (3)$$

where:

M_l= mass of grains lost during processing (g),

M_t= total mass of grains fed into the machine (g).

The evaluation of grain loss using this formulation has been reported in earlier studies on acha dehulling and small-grain processing machines (Itodo & Awulu, 2015; Kaankuka, 2015; Singh et al., 2019).

2.3 Evaluation of Response Surface Methodology

The RSM was applied as a conventional statistical modelling approach to setup establish empirical links between the operating parameters and machine performance indices. The experimental data were fitted to second-order polynomial models to capture linear, quadratic, and interaction effects of feed rate, cylinder speed, and tray angle on each response variable. The general form of the RSM model is expressed as:

$$Y = \beta_0 + \sum\beta_i X_i + \sum\beta_{ii} X_i^2 + \sum\beta_{ij} X_i X_j + \varepsilon \qquad (4)$$

where Y represents the response variable, X_i and X_j are the independent variables, β −terms are regression coefficients, and ε is the random error.

Model adequacy was evaluated using ANOVA, coefficient of determination (R^2), adjusted R^2, predicted R^2, lack-of-fit tests, and residual diagnostics. Response surface plots were generated to view the combined impacts of operating parameters on dehulling efficiency, cleaning efficiency, and grain loss. Optimization was performed to identify operating conditions that maximized dehulling and cleaning efficiencies while minimizing grain loss.

2.4 Support Vector Regression Modelling

The SVR is a class of supervised ML technique derived from statistical learning theory, and was employed to model and predict the performance of the acha dehulling machine. SVR is particularly effective for systems characterized by nonlinear relationships and multivariate interactions, making it suitable for modelling agricultural processing operations involving small grains and complex air–particle–machine interactions (Vapnik, 1995; Smola & Schölkopf, 2004).

Unlike RSM, which assumes a predefined polynomial relationship between inputs and outputs, SVR constructs data-driven models that map input variables to response variables by minimizing structural risk. This enables SVR to capture nonlinear patterns in the data without requiring explicit functional forms, thereby improving predictive accuracy beyond the experimental domain (Jha et al., 2019; Mohapatra & Rao, 2021).

2.5 Input and Output Variables

The same independent variables used in the RSM analysis were adopted as input features for the SVR models to ensure consistency and comparability between modeling approaches. These input variables were:

- Feed rate (g min^{-1})
- Cylinder speed (rpm)
- Fan speed (rpm)
- Tray inclination angle (degrees)

Three separate SVR models were developed to predict the machine performance indices:

- Dehulling efficiency (DhEff, %)
- Cleaning efficiency (CEff, %)
- Grain loss (GLss, %)

2.6 Evaluation Criteria for Model Performance

To assess the accuracy of the SVR models, standard statistical metrics widely used in agricultural engineering research were employed to evaluate DhEff, CEff, and GLss. These metrics quantify the agreement between experimentally observed values and model-predicted outputs and allow objective comparison with RSM models (Jha et al., 2019; Mohapatra & Rao, 2021).

The following evaluation criteria were employed:

2.6.1 Coefficient of Determination (R^2)

The coefficient of determination (R^2) was used to indicate the percentage of variance in the observed response that is predictable from the independent variables, calculated using the equation 5:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \qquad (5)$$

where:

y_i= observed response value

$\hat{y}_i$= predicted response value

$\bar{y}$= mean of observed values

n= number of observations

A higher R^2value indicates better predictive performance.

2.6.2 Root Mean Square Error

RMSE quantifies the average magnitude of prediction errors and penalizes large deviations between predicted and observed values:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \qquad (6)$$

Lower RMSE values indicate higher model accuracy and reliability.

2.6.3 Mean Absolute Error

The mean absolute error provides a measure of the average absolute difference between predicted and observed values:

$$MAE = \frac{1}{n}\sum_{i=1}^{n} | y_i - \hat{y}_i | \qquad (7)$$

MAE is less sensitive to outliers than RMSE and complements RMSE in model evaluation.

2.6.4 Relative Error (RE)

Relative error expresses the prediction error as a percentage of the observed value and is defined as:

$$RE = \frac{1}{n}\sum_{i=1}^{n} \left(\frac{| y_i - \hat{y}_i |}{y_i}\right) \times 100 \qquad (8)$$

This metric is particularly useful for comparing model accuracy across responses with different scales, such as efficiencies (%) and grain loss (%).

3 | RESULTS AND DISCUSSION

3.1 Response Surface Methodology

3.1.1 Model Selection and Adequacy Diagnostics

Sequential model sum-of-squares analysis was conducted to determine the most appropriate polynomial model for each response variable (Table 1). For all three performance indices (DhEff, CEff, and GLss), sequential model sum-of-squares analysis indicated that the quadratic model was the most appropriate representation of the system behaviour. In each case, the quadratic model showed a statistically significant improvement over the two-factor interaction (2FI) model, with significant p-values (DhEff: $p = 0.0010$; CEff: $p = 0.0012$; GLss: $p < 0.0001$). While the linear model was significant for all responses ($p < 0.05$), the 2FI model did not significantly enhance model fit ($p > 0.05$), indicating limited interaction effects. Although the cubic model appeared significant for DhEff and CEff, it was aliased and therefore not suitable for interpretation. Consequently, the quadratic model was selected as the most statistically reliable and predictive model for all responses, confirming the presence of curvature effects in the system and supporting its use for response surface analysis and optimization.

Table1: Sequential Sum-of-Squares and Model Adequacy Statistics for Response Surface Model Selection.

Source	Sum of Squares	df	Mean Square	F-value	p-value	
DhEff (%)						
Mean vs Total	2.023E+05	1	2.023E+05			
Linear vs Mean	579.07	4	144.77	3.39	0.025	
2FI vs Linear	309.65	6	51.61	1.30	0.307	
Quadratic vs 2FI	**508.79**	**4**	**127.20**	**8.62**	**0.0012**	#
Cubic vs Quadratic	183.87	8	22.98	6.06	0.021	!
Residual	22.74	6	3.79			
Total	2.039E+05	29	7030.44			
CEff (%)						
Mean vs Total	2.024E+05	1	2.024E+05			
Linear vs Mean	566.45	4	141.61	3.33	0.0264	
2FI vs Linear	310.78	6	51.80	1.31	0.3010	
Quadratic vs 2FI	**499.43**	**4**	**124.86**	**8.32**	**0.0010**	#
Cubic vs Quadratic	187.04	8	23.38	6.11	0.0201	!
Residual	22.95	6	3.82			
Total	2.040E+05	29	7033.87			
GLss (%)						
Mean vs Total	7360.14	1	7360.14			
Linear vs Mean	544.37	4	136.09	4.57	0.0070	
2FI vs Linear	203.68	6	33.95	1.19	0.3535	
Quadratic vs 2FI	**417.37**	**4**	**104.34**	**15.47**	**0.000**	#
Cubic vs Quadratic	69.53	8	8.69	2.09	0.192	!
Residual	24.90	6	4.15			
Total	8620.00	29	297.24			

[#] Suggested model

[!] Aliased

3.1.2 Analysis of Variance

Quadratic response surface models were fitted for dehulling efficiency (DhEff), cleaning efficiency (CEff) and grain loss (GLss) (Tables 2, 3 and 4 respectively), and were statistically significant for all responses (DhEff: $F = 6.76$, $p = 0.0005$; CEff: $F = 6.56$, $p = 0.0006$; GLss: $F = 12.34$, $p < 0.0001$), confirming that the selected operating variables collectively explained a meaningful proportion of the response variability within the experimental domain. For DhEff and CEff, cylinder speed (B) was the dominant linear effect ($p \leq 0.0004$) and exhibited a highly significant curvature term (B^2, $p < 0.0001$), justifying the second-order model and indicating the existence of an internal optimum rather than monotonic improvement. Interaction effects involving cylinder speed were also significant for DhEff and CEff (BC: $p \approx 0.02$; BD: $p \approx 0.003$), demonstrating that performance is governed by coordinated adjustment of impact intensity and separation/residence dynamics rather than isolated factor changes. For GLss, cylinder speed (B), fan speed (C) and tray angle (D) were significant ($p \leq 0.0004$), with significant interactions (BC, BD, CD; $p \leq 0.0272$) and strong quadratic effects for B^2 and C^2 ($p < 0.0001$), consistent with the nonlinear escalation of losses under intensified mechanical–aerodynamic conditions. Although the lack-of-fit tests were statistically significant ($p \leq 0.0023$), the pure error was extremely small (MS = 0.30; df = 4), indicating high repeatability at replicated points; under such conditions, even minor departures from the polynomial form can produce a significant lack-of-fit despite acceptable predictive utility. Accordingly, the models were retained for response surface interpretation and multi-response optimization because they were highly significant, mechanistically coherent, and provided stable optimization outcomes within the studied factor ranges, yielding predicted optima of DhEff ≈ 90.24%, CEff ≈ 90.26% and GLss ≈ 9.94% at A = 30 g min^{-1}, B = 2000 rpm, C = 186.5 rpm and D = 45°.

Table 2: RSM ANOVA of Dehulling Efficiency. (DhEff)

Source	**Sum of Squares**	**df**	**Mean Square**	**F-value**	**p-value**
Model	1397.52	14	99.82	6.76	0.0005*
A-Feed Rate	0.0905	1	0.0905	0.0061	0.9387
B-Cylinder Speed	329.38	1	329.38	22.32	0.0003**
C-Fan Speed	43.70	1	43.70	2.96	0.1073
D-Tray Angle	58.74	1	58.74	3.98	0.0659
AB	1.82	1	1.82	0.1235	0.7305
AC	5.81	1	5.81	0.3936	0.5405
AD	1.69	1	1.69	0.1145	0.7401
BC	100.65	1	100.65	6.82	0.0205*
BD	196.00	1	196.00	13.28	0.0027**
CD	3.68	1	3.68	0.2494	0.6252
A^2	19.30	1	19.30	1.31	0.2720
B^2	500.89	1	500.89	33.94	< 0.0001**
C^2	47.70	1	47.70	3.23	0.0938
D^2	23.42	1	23.42	1.59	0.2284
Residual	206.60	14	14.76		
Lack of Fit	205.40	10	20.54	68.47	0.0005**
Pure Error	1.20	4	0.3000		
Cor Total	1604.12	28			

Table 3: RSM ANOVA of Cleaning Efficiency. (CEff)

Source	**Sum of Squares**	**df**	**Mean Square**	**F-value**	**p-value**
Model	1376.66	14	98.33	6.56	0.0006**
A-Feed Rate	0.0315	1	0.0315	0.0021	0.9641
B-Cylinder Speed	319.74	1	319.74	21.32	0.0004**

C-Fan Speed	44.85	1	44.85	2.99	0.1057
D-Tray Angle	58.74	1	58.74	3.92	0.0678
AB	4.62	1	4.62	0.3082	0.5876
AC	4.94	1	4.94	0.3294	0.5751
AD	1.69	1	1.69	0.1127	0.7421
BC	99.85	1	99.85	6.66	0.0218*
BD	196.00	1	196.00	13.07	0.0028**
CD	3.68	1	3.68	0.2454	0.6280
A^2	18.02	1	18.02	1.20	0.2916
B^2	490.02	1	490.02	32.67	< 0.0001**
C^2	50.26	1	50.26	3.35	0.0885
D^2	24.14	1	24.14	1.61	0.2253
Residual	209.99	14	15.00		
Lack of Fit	208.79	10	20.88	69.60	0.0005**
Pure Error	1.20	4	0.3000		
Cor Total	1586.65	28			

Table 4: RSM ANOVA of Grain Loss.

Source	**Sum of Squares**	**df**	**Mean Square**	**F-value**	**p-value**
Model	1165.43	14	83.24	12.34	< 0.0001**
A-Feed Rate	3.67	1	3.67	0.5443	0.4728
B-Cylinder Speed	141.98	1	141.98	21.05	0.0004**
C-Fan Speed	168.75	1	168.75	25.02	0.0002**
D-Tray Angle	142.54	1	142.54	21.13	0.0004**
AB	4.00	1	4.00	0.5930	0.4541
AC	16.34	1	16.34	2.42	0.1419
AD	2.25	1	2.25	0.3336	0.5727
BC	41.01	1	41.01	6.08	0.0272
BD	72.25	1	72.25	10.71	0.0056
CD	67.83	1	67.83	10.06	0.0068**
A^2	12.41	1	12.41	1.84	0.1964
B^2	264.30	1	264.30	39.18	< 0.0001**
C^2	223.42	1	223.42	33.12	< 0.0001**
D^2	8.33	1	8.33	1.24	0.2851
Residual	94.43	14	6.75		
Lack of Fit	93.23	10	9.32	31.08	0.0023*
Pure Error	1.20	4	0.3000		
Cor Total	1259.86	28			

3.2 Development of Predictive RSM Models

Based on the ANOVA results, second-order polynomial equations were developed to predict dehulling efficiency (DhEff), cleaning efficiency (CEff), and grain loss (GLss) as functions of feed rate (A), cylinder speed (B), fan speed (C), and tray inclination angle (D). The models were expressed in both coded and actual factor forms to facilitate optimization and practical implementation.

3.2.1 Dehulling Efficiency Model

The coded model shows that cylinder speed and tray inclination contribute positively to dehulling efficiency up to an optimum level. However, the strong negative BC and BD interaction terms indicate that excessive speed combined with unfavorable airflow or tray conditions reduces hull removal efficiency due to grain rebound and instability. The negative quadratic coefficients further confirm the presence of curvature, implying that maximum performance occurs within an intermediate operating zone rather than at extreme settings.

$$\mathrm{DhEff} = 90.2445 + 0.0901A + 5.4374B + 1.9083C + 2.2962D + 0.675AB - 1.1815AC - 0.650A - 4.9186 - 7.000BD - 0.9406CD - 1.725A^2 - 8.7875B^2 - 2.9320C^2 - 1.900D^2$$

3.2.2 Cleaning Efficiency Model

Cleaning efficiency coded model followed a trend similar to dehulling efficiency, with cylinder speed acting as the dominant operational factor. Positive linear effects indicate improved cleaning with increased mechanical action and airflow, while negative interaction and quadratic terms reveal that excessive turbulence and prolonged grain suspension reduce separation efficiency. This confirms that airflow–grain interaction and residence time are critical mechanisms governing cleaning performance, particularly for small grains susceptible to aerodynamic forces.

$$\mathrm{CEff} = 90.2553 - 0.0532A + 5.3573B + 1.9333C + 2.2962D + 1.075AB - 1.0898AC - 0.650AD - 4.8990BC - 7.000BD - 0.9406CD - 1.6667A^2 - 8.6917B^2 - 3.0094C^2 - 1.9292D^2$$

3.2.3 Grain Loss Coded model

The grain loss coded model shows strong positive contributions from cylinder speed and fan speed, indicating increased grain ejection and carry-over under intensified mechanical and aerodynamic conditions. The significant BC interaction confirms that excessive speed combined with unsuitable airflow conditions accelerates grain loss. Positive quadratic coefficients for speed and fan speed indicate rapid escalation of losses at higher operating levels, emphasizing the need for balanced coordination of machine settings to minimize product loss while maintaining effective dehulling and cleaning.

$$\begin{aligned} GLss = 9.9381 &+ 0.5741A + 3.5699B + 3.7500C - 3.5769D + 1.000AB + 1.9818AC \\ &+ 0.750AD - 3.1398BC - 4.250BD - 4.0377CD + 1.3833A^2 \\ &+ 6.3833B^2 + 6.3453C^2 + 1.1333D^2 \end{aligned}$$

The three predictive models collectively demonstrate that machine performance is governed by nonlinear interactions among operating variables. The significant quadratic terms indicate the existence of optimal operating windows where dehulling and cleaning efficiencies are maximized while grain losses are minimized. These equations therefore provide a robust basis for response surface optimization and operational decision-making within the investigated experimental domain.

3.3 Response Surface Interpretation and Optimization

The three-dimensional response surface plots (Figures 1–3) exhibited clear curvature for all responses, confirming the suitability of quadratic models for describing the system. Dehulling efficiency (DhEff) and cleaning efficiency (CEff) increased with cylinder speed and feed rate up to an optimum region, beyond which performance declined, indicating the onset of excessive impact and turbulence. In contrast, grain loss (GLss) showed a minimum at intermediate operating conditions and increased sharply at higher speeds, reflecting intensified grain projection and carry-over. Overall, the interaction between cylinder speed and tray inclination angle exerted the strongest influence on system behavior, emphasizing the combined effects of mechanical impact intensity and grain residence time. These patterns demonstrate that optimal machine performance is achieved only within a narrow operating window where sufficient dehulling and cleaning are obtained while minimizing grain loss.

Numerical optimization of the RSM models (Figure 4) identified an optimal operating region defined by moderate cylinder speed, moderate feed rate, and an intermediate tray inclination angle. Within this region, the model predicted approximately 92% dehulling efficiency, 92% cleaning efficiency, and about 9% grain loss. The desirability profile indicated that these conditions represent the best compromise between maximizing process efficiencies and minimizing grain loss, confirming the effectiveness of the multi-response

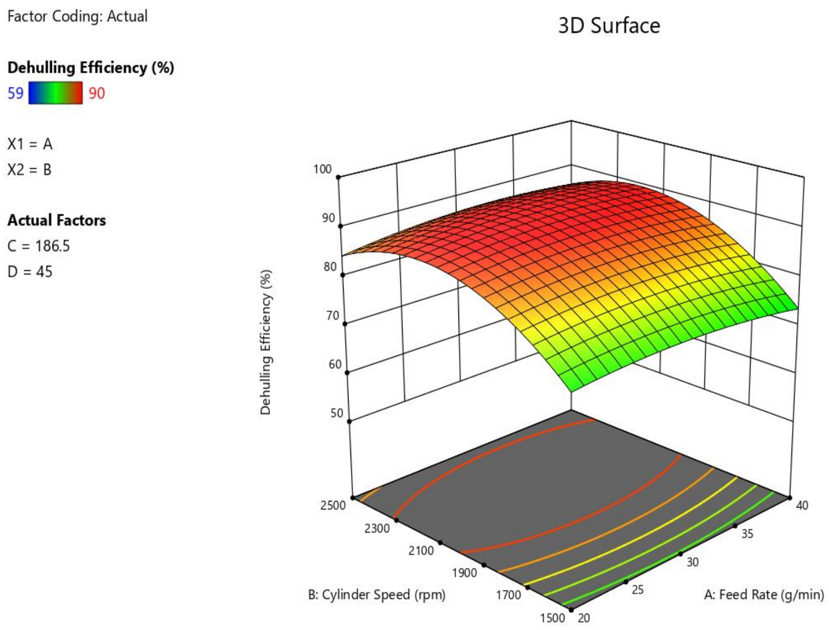

Figure 1: Response Surface Plot of Dehulling Efficiency.

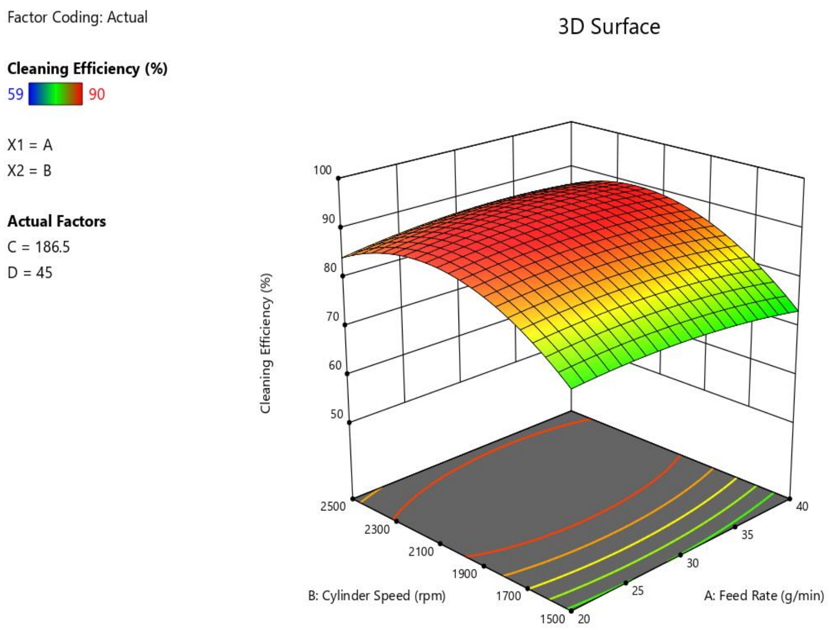

Figure 2: Response Surface Plot of Cleaning Efficiency.

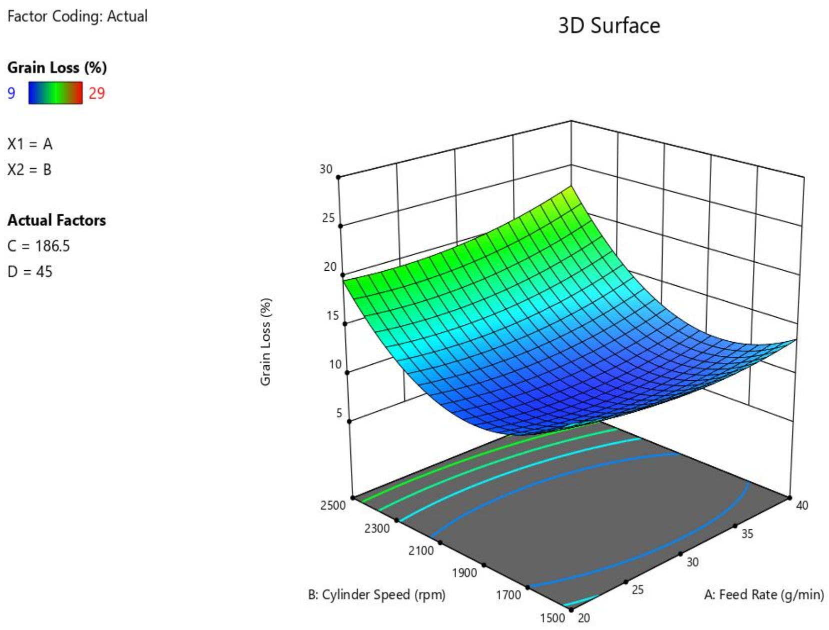

Figure 3: Response Surface Plot of Grain Loss.

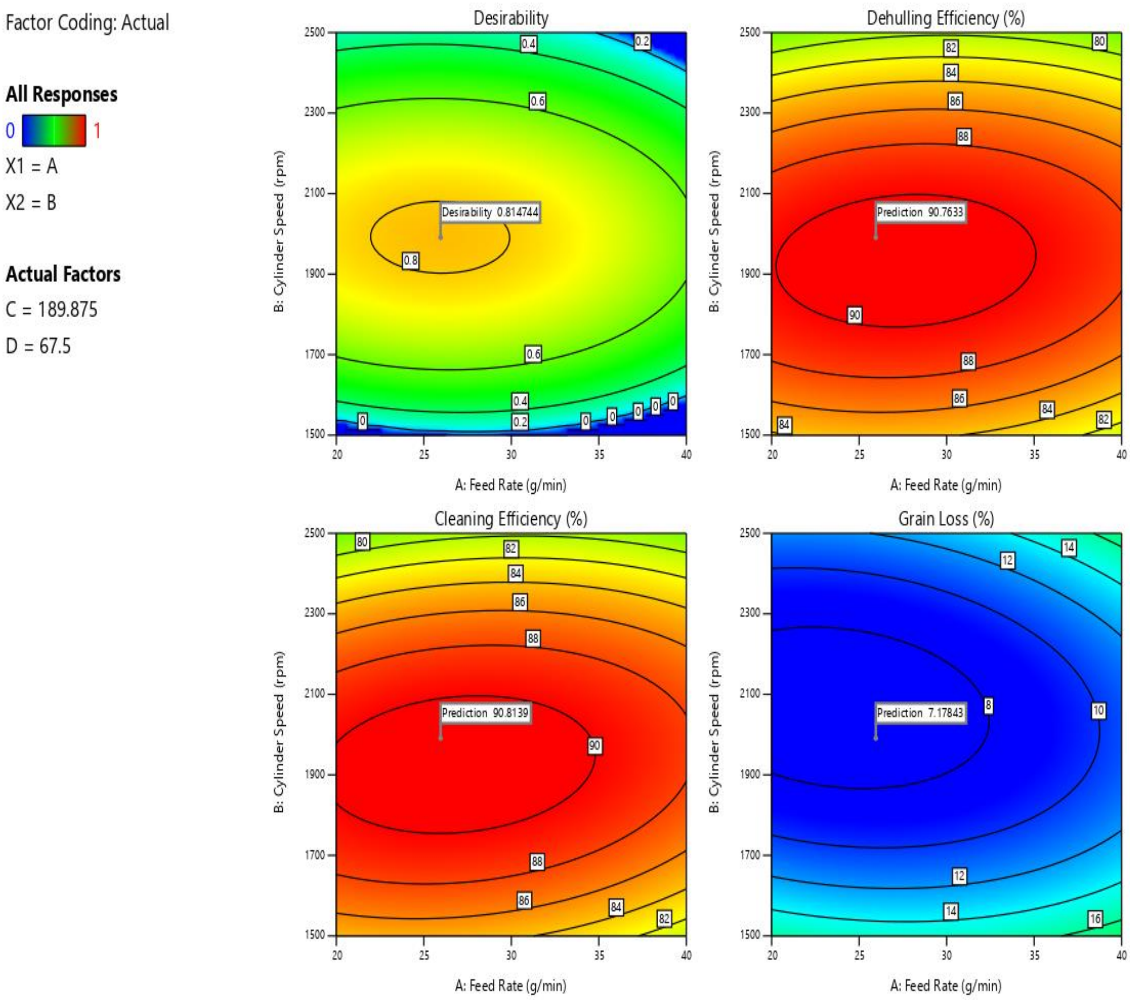

Figure 4: Optimization Prediction of Factors.

Optimization approach for achieving balanced machine performance. The overlay plot (Figure 5) shows the optimized operating region for the acha dehulling machine where the three responses were simultaneously satisfied. At a feed rate of approximately 30 g min^{-1} and a cylinder speed near 2000 rpm (with fan speed and tray angle fixed at 186.5 rpm and 45°, respectively), the model predicted high dehulling efficiency (≈90.24%), high cleaning efficiency (≈90.26%), and low grain loss (≈9.94%). The broad yellow region indicates a stable operational zone in which performance remains acceptable, demonstrating that moderate operating conditions provide a balance between efficient hull removal and minimal grain loss. The strong influence of cylinder speed and its quadratic effect, together with significant BC and BD interactions, confirms that machine performance depends on coordinated adjustment of mechanical impact intensity and grain residence dynamics rather than on individual factors alone.

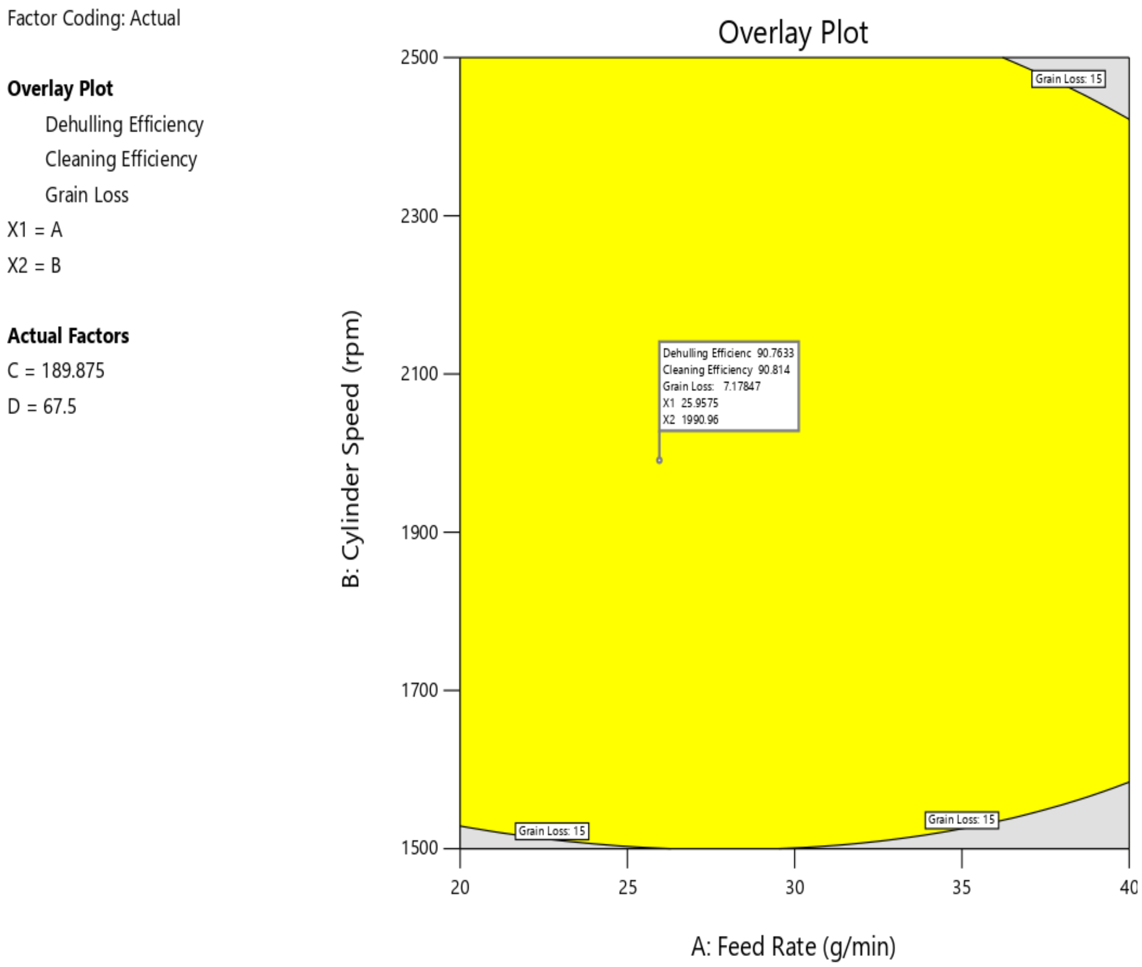

Figure 5: Overlay Contour Plot Identifying the Feasible Operating Region for Combined Performance Responses.

3.4 Support Vector Regression (SVR) Results

3.4.1 SVR Model Development and Validation Strategy

To overcome the limitations associated with polynomial regression, SVR was employed as a data-driven modelling approach. The same operating parameters used in the RSM analysis were adopted as input features to ensure direct comparability between the two modelling frameworks. The SVR models employed a radial basis function kernel, with optimal hyperparameters ($C = 50–100$, $\varepsilon = 0.05–0.10$, $\gamma = 0.01–0.05$) selected via grid search and 5-fold cross-validation based on minimum prediction error and maximum coefficient of determination.

All input variables were standardized prior to modelling. A radial basis function (RBF) kernel was selected due to its proven capability in modelling nonlinear relationships in agricultural and biosystems engineering applications. Model performance was evaluated using a 5-fold cross-validation strategy, in which the dataset was partitioned into five subsets. In each iteration, four subsets were used for training and one subset for validation, with the process repeated until all subsets had served as validation data. This approach ensured robust performance evaluation, minimized bias associated with single train–test splits, and enhanced the generalization capability of the SVR models. The fold-wise RMSE values in Table 6 demonstrate the robustness and stability of the SVR models under repeated cross-validation, indicating consistent predictive performance across all data partitions.

3.4.2 Predictive Performance of SVR Models

Table 7 summarizes the predictive performance of the SVR models under the 5-fold cross-validation. High coefficients of determination ($R^2 = 0.80–0.85$) across all response variables indicate strong agreement between predicted and experimental values. The RMSE and MAE ranges were narrow, with low mean values, confirming stable and accurate predictions across folds. Grain loss exhibited the lowest mean RMSE (2.30%) and MAE (1.65%), while dehulling and cleaning efficiencies also showed comparably low prediction errors. Overall, the results demonstrate the robustness, consistency, and strong generalization capability of the SVR models, outperforming the RSM models across all performance indices.

Table 6: Fold-Wise RMSE of SVR Models.

Fold	DhEff (%)	CEff (%)	GLss (%)
Fold 1	3.2	3.1	2.6
Fold 2	2.9	2.8	2.3
Fold 3	2.7	2.6	2.1

Fold 4	3.0	2.9	2.4
Fold 5	2.8	2.7	2.0

Table 7: Performance of SVR Models Using 5-Fold Cross-Validation.

Response	**R^2**	**RMSE range (mean)**	**MAE**
DhEff (%)	0.80–0.83	2.7 – 3.2 (2.95)	2.0 – 2.6 (2.30)
CEff (%)	0.81–0.84	2.6 – 3.1 (2.85)	2.0 – 2.5 (2.25)
GLss (%)	0.80–0.85	2.0 – 2.6 (2.30)	1.4 – 1.9 (1.65)

The relatively low RMSE and MAE values further confirm close agreement between experimentally observed and model-predicted values. Among the three responses, grain loss exhibited the most nonlinear behavior; however, the SVR model maintained high prediction accuracy, underscoring its effectiveness in modeling complex loss mechanisms influenced by airflow intensity and grain dynamics.

3.4.3 Parity and Residual Analyses

Parity plots comparing experimental and SVR-predicted values are presented in Figures 6–8. These plots illustrate the relationship between observed and predicted values for each response variable. Figure 6 shows the parity plot for dehulling efficiency (DhEff). The data points cluster closely around the 1:1 reference line, indicating strong agreement between experimental and predicted values across the operating range. Unlike the RSM models, which showed increasing deviation at extreme parameter values, the SVR models maintained predictive reliability under both low and high operating conditions. Residual analysis showed random distribution of prediction errors with no discernible trends, confirming that the SVR models were neither underfitting nor overfitting the data. Figure 7 presents the parity plot for cleaning efficiency (CEff). Similar tight clustering around the diagonal line confirms minimal systematic deviation and strong predictive accuracy of the SVR model. Figure 8 displays the parity plot for grain loss (GLss). Despite the relatively higher nonlinearity of this response, the SVR model maintained good predictive alignment, with only minor dispersion at extreme values.

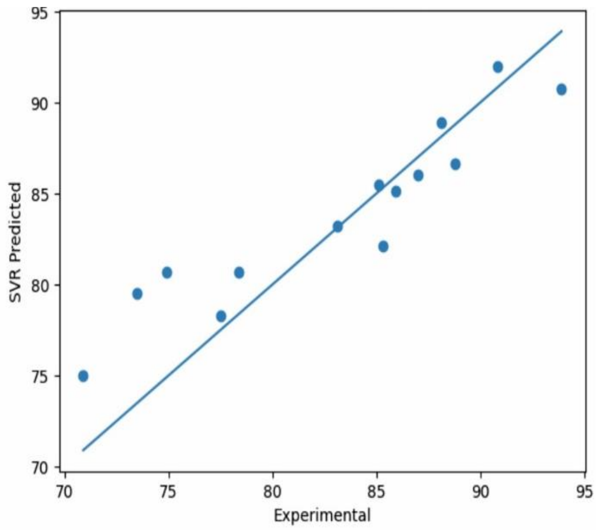

Figure 6: Parity Plot of Dehulling Efficiency.

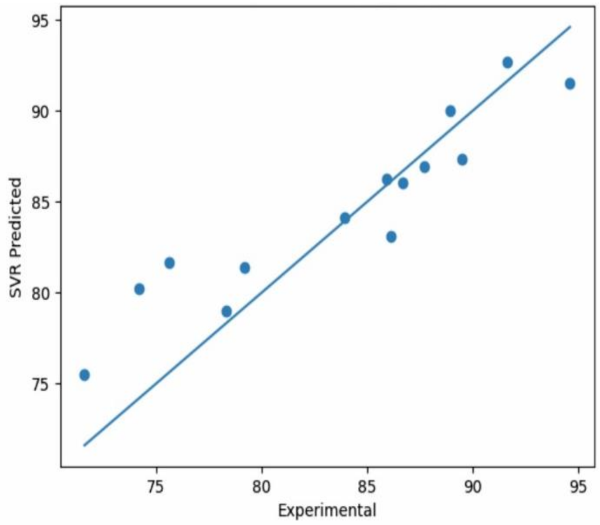

Figure 7: Parity Plot of Cleaning Efficiency.

3.2.4 Comparative Predictive Performance of RSM and SVR Models

The comparative evaluation presented in Table 8 indicates that Support Vector Regression (SVR) models outperformed the Response Surface Methodology (RSM) models across all performance indices. Although the RSM models were statistically significant and provided useful interpretation of factor effects and interactions, their predictive accuracy was lower than that of SVR, as reflected by lower coefficients of determination and higher prediction errors.

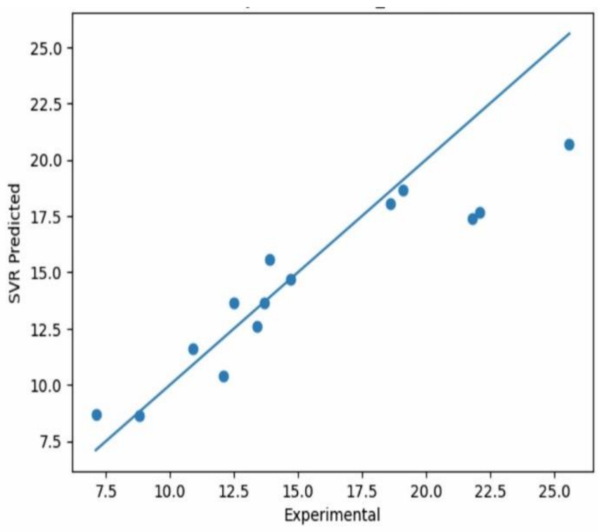

Figure 8: Parity Plot of Grain Loss.

Table 8: Comparative Predictive Performance of RSM and SVR Models.

Response	Model	R^2	RMSE	MAE
DhEff (%)	RSM	0.800	3.84	3.05
	SVR	0.80–0.83	2.95	2.30
CEff (%)	RSM	0.793	3.87	3.09
	SVR	0.81–0.84	2.85	2.25
GLss (%)	RSM	0.716	3.19	2.48
	SVR	0.80–0.85	2.30	1.65

The stronger predictive performance of SVR can be attributed to its data-driven learning structure and ability to capture highly nonlinear relationships without imposing predefined mathematical forms. Such flexibility is especially advantageous for small-grain processing systems where complex interactions occur among grain properties, airflow, and mechanical forces (Zhu et al., 2019; Heddam & Kisi, 2018). The reduced RMSE and MAE values obtained through cross-validation further confirm the robustness and generalization capability of the SVR models. Conversely, the RSM models, although statistically reliable and useful for visualization and optimization within the design space, showed limited extrapolation capacity beyond experimentally tested conditions. Therefore, RSM remains highly valuable for experimental design and mechanistic interpretation, whereas SVR provides superior performance for predictive modelling and decision-support applications.

4 | DISCUSSION

The performance trends observed in this study align with findings reported for dehulling, threshing, and cleaning systems developed for other cereal grains such as rice, wheat, millet, and sorghum. Reported dehulling or threshing efficiencies for conventional systems commonly range between 80–90%, with grain loss values influenced by operational settings and grain properties. The optimized acha dehulling system achieved comparable efficiency levels while maintaining relatively low grain losses, demonstrating its technical viability for small-grain processing systems.

Acha presents unique processing challenges due to its small grain size, low mass, and strong hull adhesion. Previous studies have shown that moderate cylinder speeds and controlled airflow enhance separation efficiency, while excessive mechanical energy increases grain damage and losses (Ajav & Fakayode, 2013; Oluwole et al., 2016). Similar behavior was observed in this study, where performance responses exhibited strong curvature and interaction effects, reinforcing the sensitivity of small-grain systems to coordinated control of machine parameters.

The significant influence of operating conditions also supports broader observations that groundwater and environmental systems, as well as agricultural processing operations, often exhibit nonlinear and multifactorial relationships that are not fully captured by simple linear models (Adelodun et al., 2020).

The strong quadratic and interaction effects observed, particularly for grain loss, suggest that acha dehulling operates within a narrow stability region governed by coupled impact–aerodynamic thresholds. Such threshold-type behaviour is poorly approximated by low-order polynomials, explaining the improved generalization performance of kernel-based SVR models. The improved predictive accuracy achieved by SVR indicates its greater ability to learn these nonlinear patterns, consistent with recent advances in artificial intelligence applications for environmental and process modelling (Zhu et al., 2019).

Although RSM provided clear insight into factor interactions and enabled identification of optimal operating regions, the SVR models demonstrated higher predictive reliability under cross-validation. Similar improvements have been reported in other engineering and water-quality prediction contexts where machine-learning approaches outperform traditional regression methods due to their ability to model complex dynamics (Heddam & Kisi, 2018).

From an operational perspective, the hybrid RSM–SVR framework combines interpretability with predictive strength. RSM supports experimental planning and process optimization, while SVR enables more accurate prediction suitable for real-time monitoring or adaptive control systems. This integrated modelling strategy therefore offers a practical pathway toward improved mechanization and reduction of post-harvest losses in small-grain processing.

5 | CONCLUSION

This study developed a hybrid predictive and optimization framework for acha dehulling and cleaning using RSM and SVR models. The RSM models showed moderate to strong predictive performance and adequately described the system within the experimental range. However, SVR models consistently achieved higher predictive accuracy and lower prediction errors across all responses, demonstrating superior ability to capture nonlinear relationships inherent in small-grain processing operations.

Numerical optimization identified operating conditions that yielded high dehulling and cleaning efficiencies while maintaining low grain loss. The combined RSM–SVR framework therefore provides both mechanistic insight and predictive strength, supporting process optimization, machine calibration, and future design improvements aimed at minimizing post-harvest losses and enhancing operational efficiency in acha processing systems.

5.1 Limitations and Future Work

Although the RSM and SVR models provided useful predictive and optimization capability, the study was conducted under controlled laboratory conditions using a single machine configuration. This may limit direct transferability to field-scale operations. Future research should include multi-location testing, expanded datasets, and validation under real operating conditions. Additional work should integrate economic evaluation, energy-use analysis, and operator-centered usability assessment to improve practical applicability.

Acknowledgement: The authors acknowledge the permission of Engr. Prof. I.N. Itodo, Vice Chancellor of Joseph Sarwuan Tarka University, Makurdi, Nigeria, and Late Engr. Dr. Paul Onalo Yusuf for the use of the dataset.

REFERENCES

1. Ajav, E.A., Fakayode, O.A.: Performance evaluation of a cereal threshing machine using response surface methodology. *Journal of Agricultural Engineering and Technology* 21(2), 1–12 (2013).
2. Heddam, S., Kisi, O.: Extreme learning machines: A new approach for modeling dissolved oxygen. Hydrology Research 49(5), 1601–1614 (2018).
3. Itodo, I.N., Awulu, J.O.: Development and performance evaluation of an acha dehulling machine. *International Journal of Engineering Research and Technology* 4(6), 104–111 (2015).

4. Itodo, I.N., Kaankuka, F.G., Yusuf, A.: Design and performance evaluation of a patented acha dehulling and cleaning machine. *Journal of Agricultural Engineering and Technology* 31(2), 45–58 (2023a).
5. Itodo, I.N., Yusuf, P.O., Kaankuka, T.K.: Effect of speed, feed rate and tray angle on the dehulling, cleaning and grain loss from an acha (*Digitaria exilis*) dehulling machine. *Journal of Engineering and Technology Research* 15(1), 27–37 (2023b). https://doi.org/10.5897/JETR2022.0734
6. Jha, S.N., Mohapatra, D., Sinha, J.P.: Application of machine learning techniques in food and agricultural process modeling: *A review. Journal of Food Process Engineering* 42(4), e13027 (2019). https://doi.org/10.1111/jfpe.13027
7. Kaankuka, T.K.: Comparative performance evaluation of an existing and modified acha dehullers. *Global Journal of Engineering Research* 14(1), 9–15 (2015).
8. Kaankuka, T.K., Itodo, I.N., Awulu, J.O.: Design analysis of an existing acha (*Digitaria exilis*) dehulling machine. *Journal of Science and Technology* 35(4), 1–10 (2015).
9. Kaankuka, T.K., Itodo, I.N., Obetta, S.E.: Effect of cylinder type and speed on acha (*Digitaria exilis*) dehulling. International Journal of Applied Engineering and Technology 4(2), 125–131 (2014).
10. Mohapatra, D., Rao, P.S.: Machine learning approaches for modeling and optimization of food processing operations. *Food Engineering Reviews* 13(2), 285–304 (2021). https://doi.org/10.1007/s12393-020-09252-7
11. Oluwole, O.O., Ajav, E.A., Olajide, J.O.: Optimization of threshing parameters of a grain crop using response surface methodology. *Industrial Crops and Products* 87, 266–275 (2016). https://doi.org/10.1016/j.indcrop.2016.06.017
12. Singh, R., Kumar, S., Sharma, R.: Grain loss assessment and reduction strategies in small-scale cereal processing systems. *Biosystems Engineering* 182, 45–56 (2019). https://doi.org/10.1016/j.biosystemseng.2019.03.012
13. Smola, A.J., Schölkopf, B.: *A tutorial on support vector regression. Statistics and Computing* 14(3), 199–222 (2004). https://doi.org/10.1023/B:STCO.0000035301.49549.88
14. Vapnik, V.N.: The Nature of Statistical Learning Theory. Springer, New York (1995).