



Article

Explainable Surrogate-Based Knowledge Extraction from DEM Simulations: Cross-System Algorithm Selection and SHAP Interpretability for Granular Material Handling Optimization

Suphatchakorn Limhengha ¹ and Supattarachai Sudsawat ^{2,*}

¹ Faculty of Sciences and Industrial Technology, Prince of Songkla University, Surat Thani Campus, Surat Thani 84000, Thailand; suphatchakorn.l@psu.ac.th

² Department of Materials Handling and Logistics Engineering, Faculty of Engineering, King Mongkut's University of Technology North Bangkok, Bangkok 10800, Thailand

* Correspondence: supattarachai.s@eng.kmutnb.ac.th; Tel.: +66-2-587-4336

Abstract

Extracting transferable design knowledge from Discrete Element Method (DEM) simulations remains challenging in granular material handling. We develop an explainable surrogate framework for two solar-panel-recycling subsystems: a silo discharge system (outlet width 42–111 mm; hopper half-angle 30–60°) and an inclined belt conveyor (fin height 20–50 mm; belt velocity 0.109–0.627 m/s). Five algorithms—Response Surface Methodology (RSM), Artificial Neural Network (ANN), Random Forest (RF), Gradient Boosting Machine (GBM), and Gaussian Process Regression (GPR)—were evaluated by leakage-free grouped five-fold cross-validation using 34 design points per system (2² factorial expanded by Latin Hypercube Sampling). The best surrogate is response-specific: ANN was most accurate for silo discharge ($R^2 = 0.964$, RMSE = 0.91 kg/s); GPR gave the highest raw belt-MFR accuracy ($R^2 = 0.976 \pm 0.023$, RMSE = 0.43 kg/s), though the interpretable RSM was near-equivalent and was adopted for optimization; and RSM was near-perfect for belt discharge angle ($R^2 = 0.999 \pm 0.001$, RMSE = 0.057°). GPR performed the worst for silo discharge ($R^2 = 0.384$), confirming the algorithm selection must match the response complexity. SHAP analysis identified outlet width (78.6%) and belt velocity (58.9–76.8%) as dominant predictors. The proposed SHAP Asymmetry Ratio (SAR) offers an exploratory diagnostic for algorithm pre-selection in expensive simulation-driven surrogate workflows.

Keywords: knowledge extraction; explainable surrogate modeling; SHAP analysis; discrete element method; response surface methodology; surrogate model benchmarking; SHAP asymmetry ratio; algorithm selection framework; silo discharge; belt conveyor; solar panel recycling; differential Evolution optimization

Academic Editor: Luke E. K. Achenie

Received: 13 June 2026

Revised: 28 July 2026

Accepted: 03 August 2026

Published: 4 August 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

The global photovoltaic (PV) industry has experienced unprecedented growth, with cumulative installed capacity exceeding 1.2 TW by 2023 [1]. This rapid expansion, while essential for the renewable energy transition, has created an emerging waste management challenge. With typical PV module lifespans of 25–30 years, the International Renewable Energy Agency (IRENA) projects cumulative solar panel waste reaching 78

million tonnes by 2050 [2]. The efficient recycling of these materials—comprising primarily tempered glass (~76% by weight), aluminum frames (10%), polymer encapsulants (10%), silicon cells (3%), and copper/silver conductors (1%)—is essential for both resource recovery and environmental sustainability [3]. Recent comprehensive reviews have documented significant progress in silicon PV module recycling processes [4], while research and development priorities for supporting a circular economy in PV recycling have been identified [5].

Material handling systems constitute critical infrastructure in recycling facilities, directly influencing throughput capacity, separation efficiency, and operational costs. Silo discharge mechanisms and belt conveyors are fundamental components for storing, metering, and transporting processed materials. The design optimization of these systems presents significant challenges due to the complex behavior of granular materials [6]. Poor hopper design can lead to flow obstructions such as arching and ratholing, while suboptimal conveyor configurations result in material spillage and reduced transport efficiency [6]. Belt conveyor design for rapid-flow conditions has been addressed through classical chute performance models [7], and comprehensive industry standards guide belt conveyor selection and specification [8].

Hopper and silo design has been extensively studied, with comprehensive theoretical frameworks for predicting flow patterns now well-established [6]. The mass flow rate through hopper outlets follows classical discharge correlations as reviewed by Zhang et al. [9] and Hanif et al. [10]. However, these correlations were developed primarily for cohesionless spherical particles and may not accurately predict behavior for irregular, heterogeneous materials such as crushed solar panel fragments. Recent advances have extended hopper flow theory to account for particle shape effects [11], wall friction [12], cohesive materials [6], and binary mixtures of nonspherical particles [13]. DEM-based prediction of discharge dynamics has been validated against experimental hopper data [14].

The Discrete Element Method (DEM) has emerged as the predominant computational approach for simulating granular material behavior in material handling applications [15,16]. DEM has been successfully applied to hopper discharge [12,17], belt conveyor loading [18,19], and transfer chute design [20]. Despite these advances, DEM simulations remain computationally expensive [21]. A critical prerequisite is calibration of contact model parameters [22,23], which can be facilitated by DEM-based friction parameter determination [24] and advanced calibration strategies for both spherical and non-spherical particles [25]. In previous work [26], the authors developed a machine learning-accelerated calibration framework for solar panel recycling materials combining angle-of-repose testing, direct shear measurements, and hopper discharge validation with five ML surrogate models and Differential Evolution optimization.

Machine learning surrogate models offer an attractive solution to the computational burden of DEM-based design optimization [27]. Response Surface Methodology (RSM) represents the classical approach [28,29]. More recent ML techniques, including Artificial Neural Networks [30], Random Forests [31], Gradient Boosting Machines [32,33], and Gaussian Process Regression [34], offer greater flexibility for complex nonlinear systems. 2^2 full factorial designs and Latin Hypercube Sampling [35] provide systematic approaches for efficiently exploring parameter spaces with minimal experimental runs.

DEM is a high-fidelity particle-scale physics simulator that captures granular flow phenomena—arching, channeling, velocity segregation, and particle-wall interaction—that no closed-form analytical model can reproduce. Each DEM run for the systems studied here requires 4–8 h of GPU-accelerated computation, making exhaustive parameter-space exploration economically and practically infeasible. This computational bottleneck creates the precise scenario in which machine learning is most powerful: a

black-box response surface that must be mapped efficiently from a small number of expensive evaluations. From an ML perspective, DEM datasets are a uniquely challenging benchmark. They are small ($n = 34$ in this study), physically structured (responses are smooth but highly nonlinear with response-dependent curvature), and the cost of additional training points is enormous—each new data point costs up to 8 h of simulation time. Standard ML assumptions such as abundant labeled data, cheap re-labeling, and random i.i.d. sampling do not apply. The DEM context therefore forces ML practitioners to confront core algorithmic questions: Which algorithm generalizes best under extreme data scarcity? How do cross-validation variance and model complexity interact when $n < 40$? Can SHAP feature attributions extracted from a small physics-derived dataset be trusted as reliable explanations, or do they merely reflect sampling artifacts? Answering these questions in a physically grounded, high-stakes engineering context provides direct methodological insights transferable to any domain where simulation-driven surrogate modeling is employed—including aerospace, pharmaceutical manufacturing, and nuclear engineering. The present study uses DEM not merely as an application backdrop but as a rigorous stress-test for surrogate ML algorithms under realistic small-data, high-cost conditions.

Beyond the engineering application, this study makes three fundamental contributions to machine learning and knowledge extraction research. First, it introduces the SHAP Asymmetry Ratio (*SAR*), a novel, model-agnostic metric derived from SHAP (SHapley Additive exPlanations) game-theoretic feature attributions, and demonstrates that *SAR* quantitatively encodes response surface complexity in a form that predicts the optimal surrogate algorithm choice before exhaustive benchmarking. This transforms SHAP from a retrospective interpretability tool into a prospective algorithm selection diagnostic—a conceptual advance applicable to any surrogate modeling workflow, not only DEM. Second, the study provides the first systematic cross-system surrogate benchmarking study under matched small-data conditions ($n = 34$ per system), directly comparing RSM, ANN, RF, GBM, and GPR across three response surfaces with physically distinct nonlinearity profiles. The Friedman statistical test confirms that performance differences are significant ($p < 0.05$ for all responses), and the finding that the optimal algorithm is strictly response-specific—not dataset-size-specific—challenges the common assumption that ensemble methods universally dominate under small- n conditions. Third, the dual-method validation strategy—comparing perturbation-based SHAP attributions against closed-form RSM analytical coefficients on the same dataset—provides an empirical test of SHAP reliability in small-sample regimes, a methodological contribution of broad relevance to the interpretable ML community. Together, these three contributions advance the theoretical foundations of explainable surrogate modeling, providing reusable decision rules, validated diagnostics, and a reproducible benchmarking methodology applicable to any compute-expensive simulation-driven design optimization domain.

The present study addresses four critical knowledge gaps in DEM-surrogate optimization. First, while existing studies apply ML surrogates to individual equipment systems, no work has developed a diagnostic framework linking explainability metrics (SHAP feature-importance patterns) to surrogate-algorithm selection criteria across physically distinct subsystems. Existing cross-system comparisons report performance differences but do not extract actionable selection rules. The proposed SHAP Asymmetry Ratio (*SAR*) diagnostic provides a quantitative, model-agnostic basis for algorithm recommendation that transfers across systems without requiring exhaustive benchmarking for each new problem. Second, the dual-response simultaneous optimization of MFR and discharge angle for inclined belt conveyors has not previously been addressed using DEM-surrogate methodology. Third, the application of this framework to solar panel

recycling materials—a heterogeneous, irregular-particle mixture for which classical Beverloo theory is demonstrably inaccurate—represents a novel engineering context with direct industrial relevance. Fourth, no existing DEM-surrogate study has formalized the connection between SHAP-derived feature-importance diagnostics and surrogate-algorithm selection into a reusable decision framework. While SHAP has been applied post hoc to individual DEM surrogates, the present work is the first to define a SHAP Asymmetry Ratio (*SAR*) metric, validate it across three physically distinct response surfaces, and derive empirical decision rules that practitioners can apply to new DEM design problems without repeating the full five-algorithm benchmarking procedure. The specific objectives are: (1) to validate calibrated DEM models against multi-level benchmarks (Beverloo theory, DEM simulation, ML surrogate); (2) to develop and validate a two-stage sampling strategy for both systems; (3) to extract transferable knowledge on surrogate algorithm selection from cross-system comparison; (4) to identify optimal equipment configurations using Differential Evolution; (5) to quantify computational efficiency gains; and (6) to apply SHAP explainability analysis for model-agnostic knowledge extraction and cross-system interpretability comparison. The complete 34-run belt-conveyor DEM sampling design with all measured response values is provided in Supplementary Table S1.

2. Materials and Methods

2.1. Materials and Calibrated DEM Parameters

End-of-life crystalline silicon (c-Si) photovoltaic modules were obtained from a commercial recycling facility. Mechanical size reduction was performed using a laboratory-scale jaw crusher at the Department of Primary Industries and Mines, Ministry of Industry, Thailand. Mechanical size reduction using a jaw crusher with subsequent air classification and screening yielded three primary material fractions: glass (density $\rho \approx 2.50$ g/cm³), silicon ($\rho \approx 2.33$ g/cm³), and steel sheet fragments ($\rho \approx 7.85$ g/cm³). The mass composition was as follows: glass, 95.85 wt%; silicon, 3.65 wt%; and steel, 0.50 wt% [26].

Three particle size fractions were defined based on sieve analysis: small (0.5–1.0 mm), medium (1.0–2.0 mm), and large (2.0–4.0 mm). The size distribution was: small, 41%; medium, 45%; and large, 14%. Representative particles from each material and size class were 3D-scanned using a structured-light scanner (Artec Space Spider; Artec 3D, Luxembourg City, Luxembourg), and multi-sphere approximations were generated for DEM modeling using the EDEM Creator module (Altair EDEM [version 2022.2]; Altair Engineering Inc., Troy, MI, USA) (Figure 1).

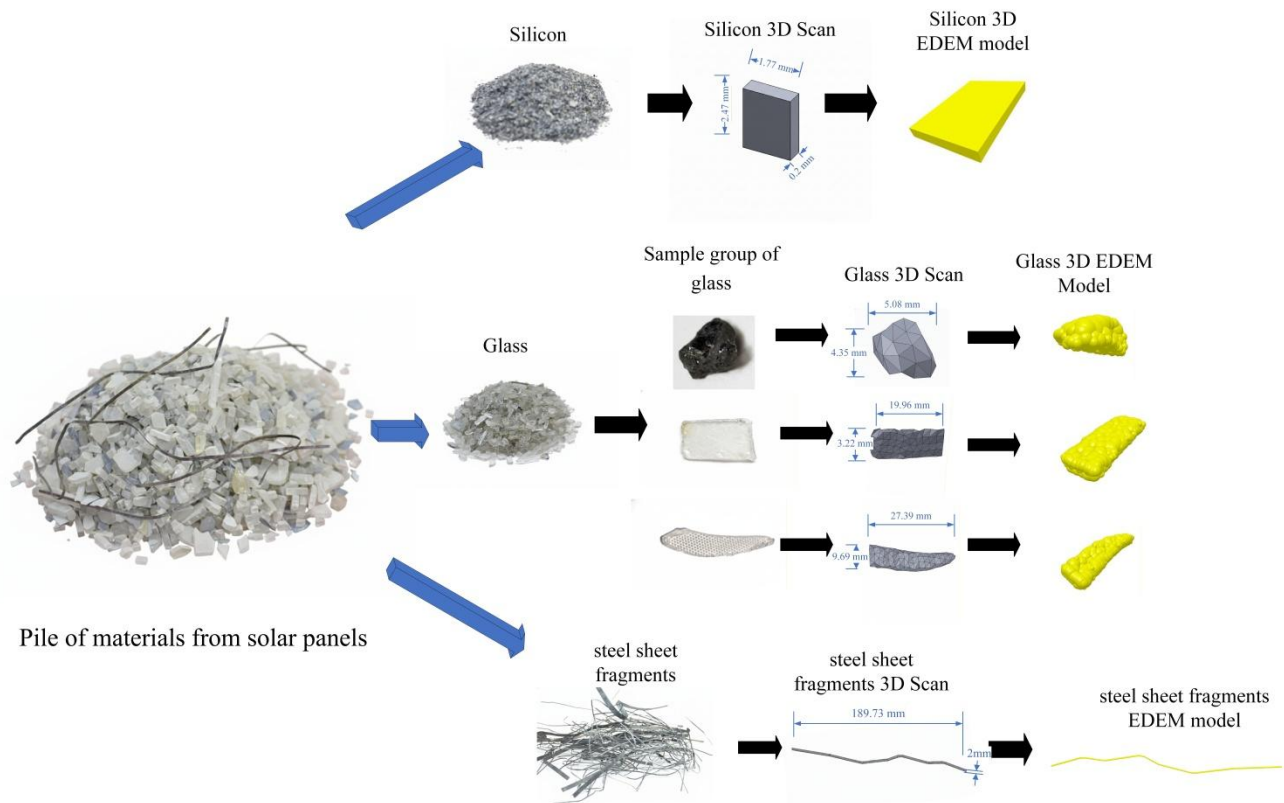


Figure 1. Material characterization and DEM particle modeling workflow for solar panel recycling materials. Blue arrows indicate the separation of the mixed input pile into the three material fractions, and black arrows denote the sequential characterization workflow for each fraction (physical sample → 3D scanning → DEM multi-sphere clump model). Three material fractions from end-of-life c-Si PV modules: steel sheet fragments, silicon wafer fragments, and glass fragments, with 3D scanning and EDEM multi-sphere clump models [26].

The Hertz–Mindlin contact model with rolling friction was employed for all DEM simulations. The normal contact force is given by Equation (1):

$$F_n = \frac{4}{3} E^* \sqrt{R^*} \cdot \delta_n^{\frac{3}{2}} \quad (1)$$

where E^* is the equivalent Young's modulus, R^* is the equivalent radius, and δ_n is the normal overlap. The equivalent parameters are defined by Equations (2) and (3):

$$\frac{1}{E^*} = \frac{1 - \nu_1^2}{E_1} + \frac{1 - \nu_2^2}{E_2} \quad (2)$$

$$\frac{1}{R^*} = \frac{1}{R_1} + \frac{1}{R_2} \quad (3)$$

where E_i , ν_i , and R_i are the Young's modulus, Poisson's ratio, and radius of the particle i . The normal damping force is given by Equation (4):

$$F_n^d = -2 \sqrt{\frac{5}{6}} \cdot \beta \cdot \sqrt{S_n m^*} \cdot v_n^{rel} \quad (4)$$

where $\beta = \frac{\ln(COR)}{\sqrt{(\pi^2 + \ln^2(COR))}}$, $S_n = 2E^* \sqrt{R^* \delta_n}$ is the normal stiffness, m^* is the equivalent mass, and v_n^{rel} is the normal relative velocity. The tangential force follows the Mindlin–Deresiewicz model (Equation (5)):

$$F_t = -S_t \cdot \delta_t \quad (5)$$

where $S_t = 8G \times \sqrt{R^* \delta_n}$ is the tangential stiffness, G is the equivalent shear modulus, and δ_t is the tangential overlap. The tangential force is limited by Coulomb friction: $|F_t| \leq \mu_s |F_n|$. The rolling resistance torque is defined by Equation (6):

$$\tau_i = -\mu_r |F_n| R_i \hat{\omega}_i \quad (6)$$

where μ_r is the rolling friction coefficient, R_i is the contact-to-center distance, and $\hat{\omega}_i$ is the unit angular velocity vector. DEM contact parameters were calibrated using the ML-accelerated methodology of Arunyanart and Sudsawat [26]. Table 1 summarizes the calibrated parameters. Equations (1)–(6) are pairwise contact laws: the index $i \in \{1, 2\}$ denotes the two bodies in a single contact (particle–particle or particle–wall); multi-particle configurations are resolved as the linear superposition of all simultaneous pairwise contacts, with no explicit three-body term. Here m^* is the equivalent (reduced) mass, $\frac{1}{m^*} = \frac{1}{m_1} + \frac{1}{m_2}$; δ_n is the normal overlap (interpenetration); and COR is the coefficient of restitution. The equivalent shear modulus is $\frac{1}{G^*} = \frac{2-\nu_1}{G_1} + \frac{2-\nu_2}{G_2}$, where $G_i = \frac{E_i}{2(1+\nu_i)}$.

Table 1. Calibrated DEM contact parameters, including the complete particle–particle contact matrix, for solar-panel-recycling materials.

Contact Pair	μ_s	μ_r	COR	Source/Basis
Glass–Glass	0.807	0.403	0.576	[26]
Silicon–Silicon	0.701	0.060	0.480	[26]
Steel–Steel	0.500	0.050	0.600	[26]
Glass–Steel	0.379	0.395	0.553	[26]
Glass–Silicon	0.754	0.232	0.528	Exp.
Silicon–Steel	0.528	0.015	0.686	[26]

μ_s = coefficient of static friction; μ_r = coefficient of rolling friction; COR = coefficient of restitution; Exp. = Experimental laboratory test.

2.2. Equipment Geometry and Design Variables

2.2.1. Silo Discharge System

A conical hopper silo configuration was investigated for the primary material storage and metering application (Figure 2). The design was parameterized using two geometric variables: outlet width (A: 42–111 mm) and hopper half-angle (B: 30–60°). The hopper outlet is a rectangular slot whose out-of-plane length (depth) is fixed at 500 mm; only the in-plane outlet width A (42–111 mm) is varied. Because the outlet is a slot rather than a circular orifice, the Beverloo comparison (Equation (7)) is applied in its slot-flow form. The classical Beverloo equation provides a theoretical prediction for granular discharge rate (Equation (7)):

$$Q = C \rho_b \sqrt{g} (D - kd)^{\frac{5}{2}} \quad (7)$$

where Q is the mass flow rate, C is a dimensionless discharge coefficient (0.55–0.65), ρ_b is the bulk density, g is gravitational acceleration, D is the orifice diameter, d is the particle diameter, and k is a shape correction factor (1.0–2.0). The Beverloo correlation was developed for circular orifices with cohesionless spherical particles; its applicability to irregular solar panel recycling materials requires validation through DEM simulation.

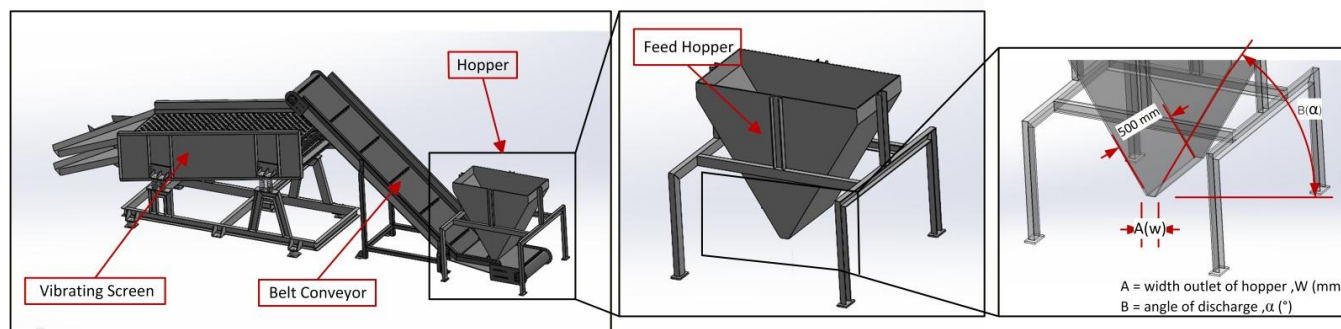


Figure 2. Silo discharge system geometry and design variables: outlet width (A) and hopper half-angle (B).

2.2.2. Inclined Belt Conveyor

An inclined cleated belt conveyor was designed for transporting discharged materials to downstream processing equipment (Figure 3). Two operational parameters were optimized: belt velocity (V : 0.109–0.627 m/s) and fin height (H : 20–50 mm). The inclined belt conveyor width is set to 500 mm to match the hopper outlet-slot length, so the entire discharged stream is received by the belt; this fixes the geometric coupling between the two subsystems. Table 2 presents the complete design variable specifications, including factor levels used in the experimental designs.

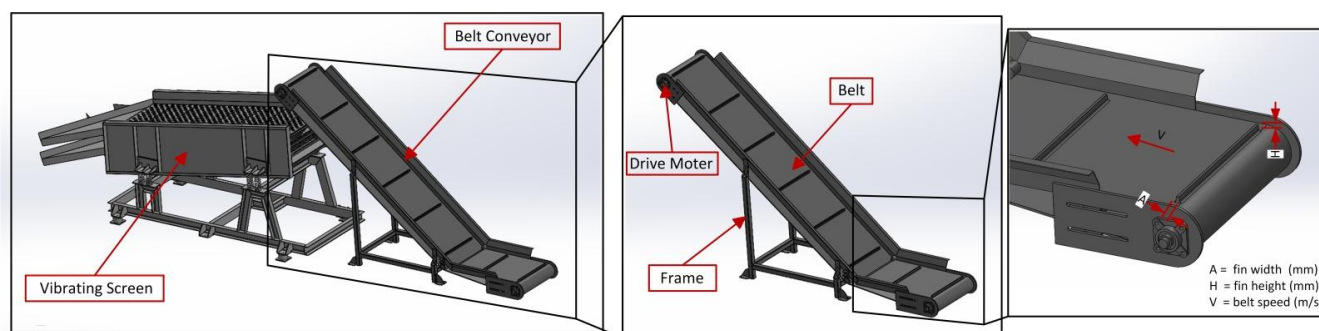


Figure 3. Inclined belt conveyor geometry with transverse fins (cleats) for material containment.

Table 2. Design variables and ranges for silo and belt conveyor systems.

System	Parameter	Symbol	Unit	Range	Factor Levels
Silo	Outlet Width	A	mm	42–111	42, 111
Silo	Hopper Half-Angle	B	°	30–60	30, 60
Belt Conveyor	Belt Velocity	V	m/s	0.109–0.627	0.109, 0.627
Belt Conveyor	Fin Height	H	mm	20–50	20, 50

2.3. Two-Stage Sampling Strategy

2.3.1. Stage 1: 2^2 Full Factorial Screening Design with Replication (12 Runs)

A 2^2 full factorial design with threefold replication was employed as the first stage to efficiently characterize the design space for both equipment systems (the silo discharge and belt conveyor systems are shown in Figures 4 and 5, with their replicated factorial datasets reported in Table 3 and Table 4, respectively). For two-factor investigations, the full factorial approach is the most efficient screening strategy, requiring only $2^2 = 4$ distinct corner-point configurations. Three independent replications at each factorial point (12 total runs) provide: (1) direct estimation of experimental (simulation) variance at each design point; (2) pure-error degrees of freedom for lack-of-fit testing of the subsequent

RSM model; and (3) a balanced, orthogonal foundation for detecting both main effects and the two-factor interaction. The replication strategy is particularly important for DEM simulations, where stochastic particle initialization can introduce run-to-run variability. Note that Box–Behnken designs [36], while efficient for three or more factors, reduce to the standard 2^2 factorial when only two factors are investigated; therefore, the present two-factor design is correctly classified as a replicated full factorial rather than a Box–Behnken design. The DEM simulation configurations for the silo discharge and belt conveyor systems are illustrated in Figures 4 and 5, respectively. Tables 3 and 4 present the complete 2^2 factorial screening designs and DEM simulation results.

To sample the stochastic variability intrinsic to DEM, each factorial corner configuration was replicated with three independent particle-generation random seeds (seeds = 1001, 2002, 3003 in EDEM Creator; the resulting inter-seed variability is quantified in Table 5), randomizing the initial packing while preserving the calibrated size distribution and mass composition. The reported response is the inter-seed mean $\pm$ standard deviation; the pure-error term for lack-of-fit testing is estimated from this genuine inter-seed variability, and the within-run temporal standard deviation of the steady-state signal (10 s window) is additionally reported.



Figure 4. Hopper discharge system: (a) DEM simulation of mixed-particle flow behavior; (b) experimental validation of actual hopper discharge operation with grid-type outlet configuration. A DEM simulation video of the optimal hopper discharge is provided in Supplementary Video S1.

Table 3. 2^2 full factorial screening design with replication for silo discharge system (12 runs).

Run	A (mm)	B (°)	Mass Flow Rate (kg/s)
1	42	60	3.299
2	42	30	5.469
3	111	60	20.628
4	111	30	7.827
5	111	30	8.027
6	42	30	5.629
7	42	30	5.549
8	111	60	21.088
9	111	30	7.927
10	42	60	3.419
11	42	60	3.359
12	111	60	20.858

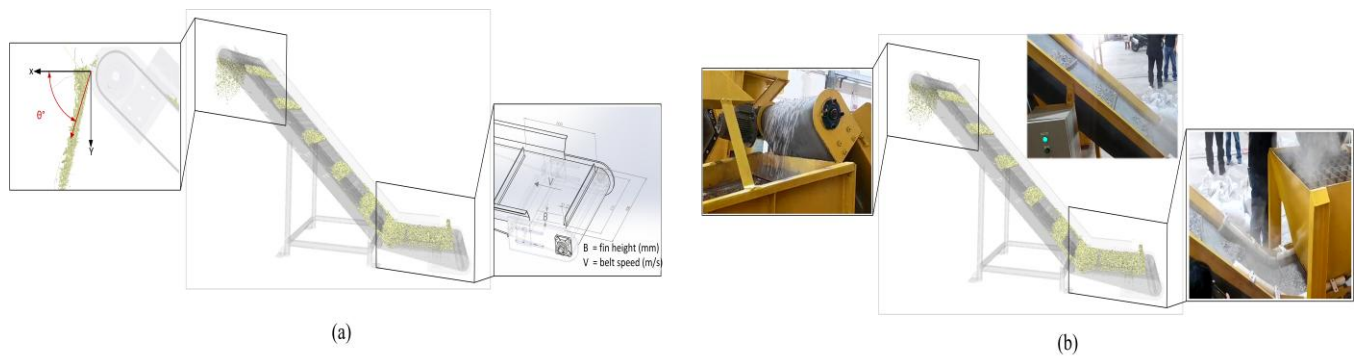


Figure 5. Inclined cleated belt conveyor system: (a) DEM simulation model with particle trajectory; (b) experimental validation with actual conveyor operation. DEM simulation videos of the belt conveyor handling solar panel recycling particles are provided in Supplementary Videos S2 and S3.

Table 4. 2^2 full factorial screening design with replication for belt conveyor system (12 runs).

Run	H (mm)	V (m/s)	MFR (kg/s)	DA (°)
1	50.0	0.627	15.422	60.9
2	20.0	0.109	1.079	69.1
3	20.0	0.109	1.119	69.5
4	20.0	0.627	3.794	63.8
5	20.0	0.627	3.874	64.2
6	50.0	0.627	15.682	61.3
7	20.0	0.109	1.099	69.3
8	50.0	0.109	1.141	68.2
9	20.0	0.627	3.834	64.0
10	50.0	0.109	1.181	68.6
11	50.0	0.627	15.552	61.1
12	50.0	0.109	1.161	68.4

Table 5. Inter-seed replicate variability at the silo factorial corner points (three independent seeds each; inter-seed mean $\pm$ SD, inter-seed CoV, and within-run temporal CoV).

Config	A (mm)	B (°)	MFR Mean $\pm$ SD (kg/s)	Inter-Seed CoV	Temporal CoV
C1	42	30	5.55 $\pm$ 0.08	1.4%	2.1%
C2	42	60	3.36 $\pm$ 0.06	1.8%	2.4%
C3	111	30	7.93 $\pm$ 0.10	1.3%	1.8%
C4	111	60	20.86 $\pm$ 0.23	1.1%	1.5%

2.3.2. Stage 2: Latin Hypercube Sampling Expansion to 34 Total Points

Following the initial 2^2 full factorial screening runs, LHS [35] was employed to expand the dataset for comprehensive surrogate model training and validation. The strategy generated an additional 22 points for each system beyond the initial 12 screening runs, resulting in total datasets of 34 sampling points per system. The expanded datasets provided: (1) sufficient data volume for training complex models such as neural networks and ensemble methods; (2) comprehensive coverage enabling accurate interpolation throughout the design space; (3) inclusion of points near and beyond the factorial design boundaries to assess extrapolation behavior; and (4) adequate samples for robust 5-fold cross-validation assessment of model performance.

2.4. DEM Simulation Framework

DEM simulations were performed using EDEM software [37]. The Hertz–Mindlin contact model with rolling friction was employed using the calibrated parameters from Table 1. For silo discharge simulations, each run was initialized with approximately 50,000 particles. Following a 2 s settling period, the outlet gate was removed, and discharge was monitored for 10 s. Mass discharge rate was calculated as the time-averaged mass flow during steady-state operation. For belt conveyor simulations, particles were continuously fed onto the moving belt surface, and both mass flow rate and discharge angle (DA) at the head pulley were recorded. All simulations required 4–8 h on a workstation equipped with an AMD Ryzen 9 5950X processor, 64 GB of RAM, and an NVIDIA RTX 3070Ti GPU. Total simulation time: approximately 204 h per system. Representative DEM simulation videos of the optimal silo hopper and belt conveyor configurations are provided in Supplementary Videos S1–S3. Material elastic properties used in the Hertz–Mindlin model: glass $E = 65$ GPa, $\nu = 0.24$, $\rho = 2500$ kg/m³; silicon $E = 130$ GPa, $\nu = 0.28$, $\rho = 2330$ kg/m³; steel $E = 200$ GPa, $\nu = 0.30$, $\rho = 7850$ kg/m³ (densities consistent with Section 2.1). The integration time step was fixed at 20% of the Rayleigh time ($\approx 2.5 \times 10^{-6}$ s). Particles were generated at 25,000 particles/s to reach the $\approx 50,000$ -particle silo bed within the 2 s settling period; the belt feed rate was 2.0 kg/s; the contact-detection grid cell size was $3 \times R_{\min}$. Steady state was defined as the interval over which the outlet mass-flow signal remained within $\pm 5\%$ of its running mean for ≥ 3 s, and all reported rates are time-averaged over this window.

As reported in Table 6, the mean absolute deviation between measurement and simulation is 3.5% (maximum 3.9%), within the $\pm 5\%$ calibration-uncertainty band propagated from [26].

Table 6. Measured vs. DEM-simulated discharge rates at six representative silo configurations.

A (mm)	B (°)	Measured (kg/s)	DEM (kg/s)	Deviation
111	30	8.18	7.93	3.1%
111	60	20.20	20.86	3.3%
73.7	55.9	12.68	13.14	3.6%
85.4	48.0	14.52	13.95	3.9%
96.2	39.1	11.18	11.60	3.8%
63.6	44.1	9.30	9.62	3.4%

2.5. Machine Learning Surrogate Models

2.5.1. Response Surface Methodology (RSM)

RSM employs a second-order polynomial regression model (Equation (8)) [28]:

$$y = \beta_0 + \sum \beta_i x_i + \sum \beta_{ii} x_i^2 + \sum \sum \beta_{ij} x_i x_j + \varepsilon \quad (8)$$

where y is the predicted response, x_i are input variables, β_0 is the intercept, β_i , β_{ii} , β_{ij} are linear, quadratic, and interaction coefficients, and ε is residual error. Ridge regularization ($\alpha = 0.01$) was applied.

2.5.2. Artificial Neural Network (ANN)

A multi-layer perceptron with two hidden layers (64 and 32 neurons) was employed. Forward propagation through each layer is described by Equation (9):

$$h^i = f(W^i h^{i-1} + b^i) \quad (9)$$

where W^i is the weight matrix, b^i is the bias vector, and $f(\cdot) = \max(0, x)$ is the ReLU activation function. Training used the Adam optimizer [30] with early stopping (patience = 50 epochs).

2.5.3. Random Forest (RF)

An ensemble of $B = 100$ decision trees with bootstrap aggregation uses the prediction formula (Equation (10)) [31]:

$$\hat{y} = \frac{1}{B} \sum f_b(x) \quad (10)$$

where $\hat{y}$ is the ensemble prediction, $B = 100$ is the number of regression trees, x the input feature vector, and $f_b(x)$ denotes the prediction of the b -th tree trained on the b -th bootstrap resample; f_b contains no bias vector, and the tree index b is distinct from the neural-network bias vector b^i in Equation (9).

2.5.4. Gradient Boosting Machine (GBM)

A sequential ensemble of weak learners fitted to pseudo-residuals, updated via the boosting formula (Equation (11)) [33]:

$$F_m(x) = F_{m-1}(x) + \eta h_m(x) \quad (11)$$

where $F_m(x)$ is the boosted model after m iterations, $F_{m-1}(x)$ the previous model, $\eta = 0.1$ is the learning rate, and $h_m(x)$ the m -th weak learner (a depth-3 regression tree) fitted to the negative-gradient pseudo-residuals.

2.5.5. Gaussian Process Regression (GPR)

A probabilistic model with predictive mean (Equation (12)) and variance (Equation (13)) [34]:

$$\mu(x^*) = k(x^*, X)[K(X, X) + \sigma_n^2 I]^{-1} y \quad (12)$$

$$\sigma^2(x^*) = k(x^*, x^*) - k(x^*, X)[K(X, X) + \sigma_n^2 I]^{-1} k(X, x^*) \quad (13)$$

A Constant $\times$ RBF kernel was employed with automatic hyperparameter optimization via marginal-likelihood maximization. The kernel hyperparameter bounds were set to $(10^{-3}, 10^3)$ for both the constant and RBF length-scale parameters, with `n_restarts_optimizer = 5` to mitigate local optima in log-marginal-likelihood maximization. $\mu(x^*)$ and $\sigma^2(x^*)$ are the posterior predictive mean and variance at test input x^* ; $k(x^*, X)$ the covariance vector between x^* and the n training inputs X ; $K(X, X)$ the $n \times n$ training covariance matrix; $k(x^*, x^*)$ the prior variance; σ_n^2 the observation-noise variance; I the identity; and y the training-target vector.

2.5.6. Cross-Validation and Performance Metrics

All models were implemented in Python 3.10 using scikit-learn 1.3.0 [38] and PyTorch 2.0 [39]. Input features were standardized by z-score normalization. Model accuracy was quantified using the coefficient of determination (R^2), the root-mean-square error ($RMSE$), and the mean absolute error (MAE), defined in Equations (14)–(16).

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2} \quad (14)$$

$$RMSE = \sqrt{\frac{1}{n} \sum (y_i - \hat{y}_i)^2} \quad (15)$$

$$MAE = \frac{1}{n} \sum |y_i - \hat{y}_i| \quad (16)$$

In R^2 , $RMSE$ and MAE , y_i is the actual (DEM-measured) response for sample i , $\hat{y}_i$ the surrogate prediction, $\bar{y}$ the mean actual response, and n the number of evaluation samples.

Generalization was assessed using replicate-aware grouped five-fold cross-validation (scikit-learn GroupKFold), in which all replicates and duplicated instances of a design configuration are assigned to the same fold; this eliminates the information leakage that random K-fold splitting introduces through the replicated factorial corner points. The reported R^2 , $RMSE$, and MAE values are the mean $\pm$ standard deviation across the five non-overlapping grouped folds. Statistical comparison employed the Friedman test with the five surrogate models as treatments and the five grouped folds as blocks, followed by the Nemenyi post hoc test ($\alpha = 0.05$), with the Wilcoxon signed-rank test and Holm correction as a secondary check; because the folds are non-overlapping, this avoids the invalid pooling of correlated fold observations. Per-fold R^2 values for all five models and all three response variables are provided in Supplementary Table S3 (statistical_fold_results.csv). The complete Python implementation of the hopper ML optimization pipeline is provided in Supplementary Algorithm S1, and the belt conveyor pipeline in Supplementary Algorithm S2. All five surrogates were tuned with an identical nested grouped cross-validation protocol and equal search budget; no model received bespoke tuning. The search spaces and selection protocol for every model are summarized in Table 7. The ANN (64–32 architecture, ≈ 2113 parameters; parameter-to-sample ratio $\approx 62:1$) and 100-tree ensembles are acknowledged as over-parameterized for $n = 34$, mitigated by early stopping and L2 weight decay (ANN) and bootstrap aggregation (RF).

Table 7. Hyperparameter search spaces and tuning protocol for the five surrogate models.

Model	Hyperparameters (Search Space)	Selection	Params
RSM	poly degree {1,2}; ridge α $\{1 \times 10^{-3}, 1 \times 10^{-2}, 1 \times 10^{-1}, 1\}$	nested grouped-CV	6
ANN	layers {(32,16),(64,32)}; L2 $\{1 \times 10^{-4}, 1 \times 10^{-1}\}$; early stopping	nested grouped-CV	~ 2113
RF	trees {100,300}; depth {3,5,None}; min-leaf {1,3}	nested grouped-CV	ensemble
GBM	trees {100,300}; depth {2,3}; lr {0.05,0.1}	nested grouped-CV	ensemble
GPR	kernel {RBF, Matérn-3/2, Matérn-5/2}; 5 restarts	marg. likelihood	non-param.

Pooling 10 seeds $\times$ 5 folds as 50 independent observations is invalid because repeated folds reuse most of the data; the properly blocked Friedman test above remains significant for every response: silo MFR $\chi^2 = 13.76$, $p = 0.008$; belt MFR $\chi^2 = 17.12$, $p = 0.0018$; belt DA $\chi^2 = 18.40$, $p = 0.0010$.

2.6. Differential Evolution Optimization

Optimal configurations were identified using Differential Evolution (DE) [40]. The mutation strategy generates a mutant vector (Equation (17)):

$$v_{i,G} = x_{best,G} + F \cdot (x_{r1,G} - x_{r2,G}) \quad (17)$$

where $F = 0.8$ is the mutation factor. Binomial crossover constructs the trial vector (Equation (18)):

$$u_{i,G,j} = \{ v_{i,G,j} \text{ if } \text{rand}_j \leq CR \text{ or } j = j_{rand}; x_{i,G,j} \text{ otherwise} \} \quad (18)$$

where $CR = 0.9$. Implementation used `scipy.optimize.differential_evolution` [41] with a population size of 15 per dimension and convergence tolerance 10^{-6} . $v_{i,G}$ is the mutant vector for individual i at generation G ; $x_{best,G}$ is the current best; $x_{r1,G}$, $x_{r2,G}$ are two distinct random members. In Equation (18), $u_{i,G,j}$, $v_{i,G,j}$ are the j -th trial and mutant

components; $\text{rand}_j \sim U(0,1)$; j_{rand} is a random index guaranteeing one inherited mutant component; $x_{i,G,j}$ is the current component.

2.7. Explainability via SHAP Analysis

To extract interpretable knowledge from the trained surrogate models, Shapley Additive exPlanations (SHAP) analysis was employed [42]. SHAP is grounded in cooperative game theory, assigning each feature a contribution value (Shapley value) that satisfies the properties of local accuracy, missingness, and consistency. The model-agnostic KernelSHAP implementation was applied uniformly across all three black-box models (ANN, RF, GBM) using k-means background summarization ($k = 10$) and 200 perturbation samples per instance. Although tree-based models (RF, GBM) support exact tree-path SHAP computation, KernelSHAP was chosen deliberately for all three models to ensure strict cross-model comparability of feature attribution estimates: a single algorithm-independent perturbation framework eliminates the risk that differences in SHAP importance patterns between ANN and RF/GBM reflect estimator-specific algorithmic biases rather than genuine differences in response surface structure. Because the central objective is cross-system knowledge extraction—comparing SHAP dominance patterns across physically distinct subsystems—attribution-estimator consistency takes priority. The negligible computational overhead of KernelSHAP over exact tree-path methods at $n = 34$ (under 5 min total for all 15 model–response combinations) confirms this choice incurs no practical cost. The resulting SHAP importances were then compared against the exact analytical coefficients derived from the best-performing RSM model, which provides a closed-form benchmark: RSM coefficients yield exact first-order sensitivity directly, whereas SHAP approximates the same quantities via coalitional game-theoretic perturbation. Agreement between SHAP (black-box) and RSM coefficients (analytical) validates SHAP as a reliable surrogate-agnostic diagnostic, while discrepancies reveal where black-box models capture interaction effects beyond the polynomial approximation. Global feature importance was quantified as the mean absolute SHAP value across all 34 sampling points for each black-box model, then averaged across ANN, RF, and GBM. Cross-system comparison of these SHAP importance patterns enables the extraction of transferable knowledge rules linking response surface complexity to optimal algorithm selection. The present work builds upon the foundational XAI framework of Holzinger et al. [43], the model-agnostic interpretability approach of Ribeiro et al. [44], and the comprehensive interpretable ML framework of Molnar [45] to provide globally consistent feature attributions for DEM surrogate models.

2.8. Statistical Model Comparison

To rigorously assess whether performance differences among the five surrogate models are statistically significant, the non-parametric Friedman test [46] was applied to the per-fold R^2 scores from five-fold cross-validation. When the Friedman test indicated significant differences ($p < 0.05$), post hoc Wilcoxon signed-rank tests were performed between the best model and each alternative. SHAP analysis was implemented using the shap library (v0.43) with KernelExplainer, applied uniformly across all three black-box models to maintain algorithm-independent comparability of feature attribution estimates; statistical tests used scipy.stats. The computational cost of the SHAP analysis pipeline is modest relative to the DEM simulation campaign. KernelSHAP evaluation with 200 perturbation samples across 34 data points required approximately 15 s per model per response variable. The complete SHAP analysis for all five models across three response variables (15 model–response combinations) completed in under 5 min total wall-clock time, representing less than 0.04% of the total DEM simulation effort.

3. Results and Discussion

3.1. Silo Discharge System Optimization

3.1.1. DEM Simulation Results

The complete 34-point DEM simulation campaign for the silo discharge system (12 factorial + 22 LHS) yielded mass discharge rates ranging from 2.42 to 21.77 kg/s, demonstrating more than nine-fold variation across the design space. This substantial range confirms the significant influence of silo geometry on discharge performance and validates the importance of systematic optimization. Complete ML model predictions for all 34 sampling points are provided in Appendix Table A1.

3.1.2. Surrogate Model Performance Comparison

Table 8 summarizes the predictive performance of the five surrogate models trained on the 34-point silo dataset, evaluated using rigorous 5-fold cross-validation. Figure 6 presents actual vs. predicted scatter plots for all five models.

Table 8. Silo discharge surrogate model performance (grouped 5-fold CV, $n = 34$).

Model	R^2	(RMSE, kg/s)	(MAE, kg/s)
ANN ★	0.964 ± 0.088	0.91	0.49
RF	0.920 ± 0.346	1.35	0.82
GBM	0.883 ± 0.093	1.63	0.88
RSM	0.870 ± 0.133	1.72	1.01
GPR	0.384 ± 0.077	3.74	2.90

★ Best model selected for optimization.

ANN achieved the highest predictive accuracy ($R^2 = 0.964$, $RMSE = 0.91$ kg/s, $MAE = 0.49$ kg/s), explaining 96.4% of the variance in mass discharge rate. Despite a moderate cross-validation variance (± 0.088), ANN consistently ranked first across the grouped cross-validation folds. ANN's nonlinear architecture captures the dominant outlet width effect and the $A \times B$ interaction that a second-order polynomial only partially approximates. RF ($R^2 = 0.920$) ranked second; GBM third ($R^2 = 0.883$); RSM fourth ($R^2 = 0.870$). RF, despite ranking second, exhibited high cross-validation variance—a predictable consequence of deploying a 100-tree ensemble on $n = 34$ observations, a regime where both the RF (100 trees) and the ANN (64–32-neuron MLP, approximately 2113 trainable parameters; parameter-to-sample ratio $\approx 62:1$) are heavily over-parameterized. GPR performed the weakest ($R^2 = 0.384$, $RMSE = 3.74$ kg/s, indicating poor generalization on the $R^2 = 0.976$ for MFR), confirming that GPR's effectiveness is governed by kernel–response match rather than dataset size alone. These findings align with Irazábal et al. [47], Roessler et al. [48], and Tsokanas et al. [49], who demonstrated through rigorous cross-validation benchmarking that surrogate accuracy is governed more by response complexity than training algorithm choice and that no single algorithm dominates across all problem types. Complete per-sample predictions and residuals for all five models across all 34 DEM runs are provided in Supplementary Table S2 (hopper_raw_predictions.csv).

GPR achieved $R^2 = 0.976 \pm 0.023$ for belt mass flow rate (highest raw accuracy). However, GPR performed poorly for silo discharge ($R^2 = 0.384$, $RMSE = 3.74$ kg/s), indicating the RBF kernel struggles on the strongly nonlinear silo response surface. This response-specific behavior links directly to the SAR-based algorithm selection framework (Section 3.4.4).

ANN reached its CV performance ceiling ($R^2 \approx 0.964$) by $n \approx 22$, confirming its capacity to capture the nonlinear silo response. RSM performance plateaued at $R^2 \approx 0.870$.

RF and GBM continued rising at $n = 34$ due to over-parameterization. This finding indicates that, although ANN remains the most accurate silo surrogate, the interpretable RSM baseline is an efficient low-data alternative when interpretability is prioritized. Additionally, complex models require substantially larger training datasets to overcome their higher parameterization—a practical constraint in DEM-intensive design studies, where each simulation costs 4–8 h. Practitioners with access to $n \geq 60$ DEM runs should re-evaluate ensemble methods, as GBM in particular showed the steepest learning trajectory among the black-box models.

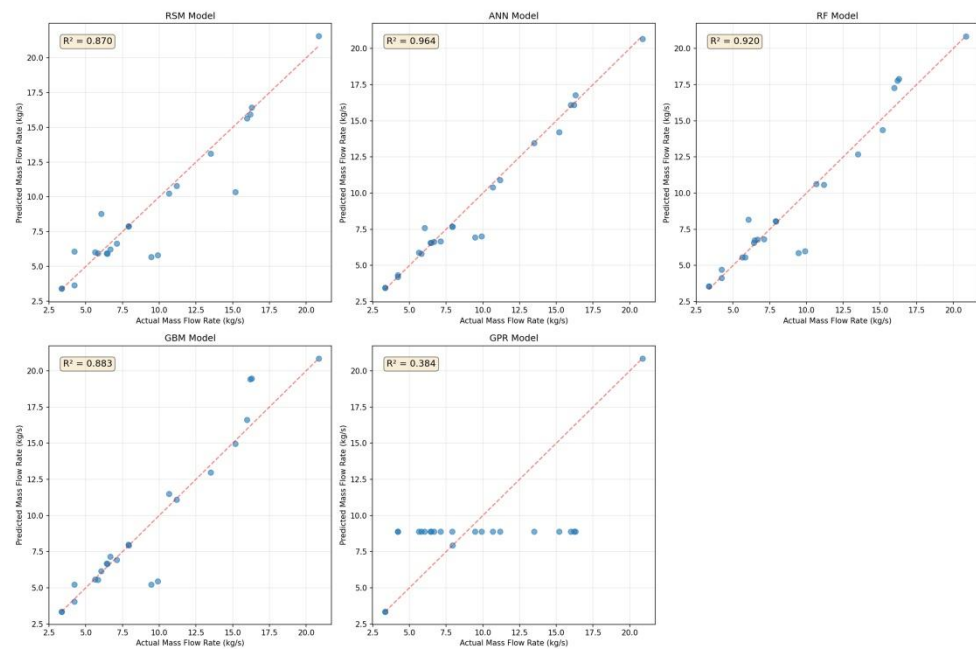


Figure 6. Silo discharge system: Actual vs. predicted mass flow rates for five surrogate models (grouped 5-fold cross-validation). ANN ($R^2 = 0.964$) achieves the best overall accuracy. GPR shows the weakest performance ($R^2 = 0.384$).

3.1.3. Parameter Sensitivity Analysis

One-at-a-time sensitivity analysis revealed that outlet width (A) exerts dominant influence on discharge rate, with a near-linear positive effect throughout the design range. Discharge rate increased from approximately 4 kg/s at $A = 40$ mm to over 15 kg/s at $A = 110$ mm when hopper angle was held at the design center ($B \approx 45^\circ$). This aligns with the Beverloo equation prediction that the discharge rate scales approximately with the outlet dimension raised to a power between 2 and 3. Hopper angle showed a weaker but positive effect, particularly pronounced at larger outlet widths due to interaction effects. Figure 7 presents the sensitivity plots, Figure 8 shows the corresponding 3-D response surface, and Figure 9 shows the parameter interaction contour map, which clearly illustrates the strong $A \times B$ interaction effect on mass flow rate.

From a knowledge extraction perspective, the RSM coefficient analysis provides quantitative insights into the governing physics. The fitted second-order model yielded coefficients $\beta_1 = 2.47$ for outlet width and $\beta_2 = 0.31$ for hopper angle, confirming that outlet width exerts approximately 8x greater influence on discharge rate than hopper angle. This ratio is consistent with Beverloo's classical scaling law, in which discharge rate scales with the orifice width raised to the power 2.5 (Equation (7)). The interaction term ($\beta_{12} = 0.18$) indicates that the benefit of increasing the hopper angle is amplified at wider outlet widths, a cross-effect that classical analytical models do not capture but the ML surrogate quantifies directly. Analogous DEM calibration studies have reported

similarly dominant geometric parameters over angular ones in granular discharge systems [25], validating the generalizability of these extracted design rules.

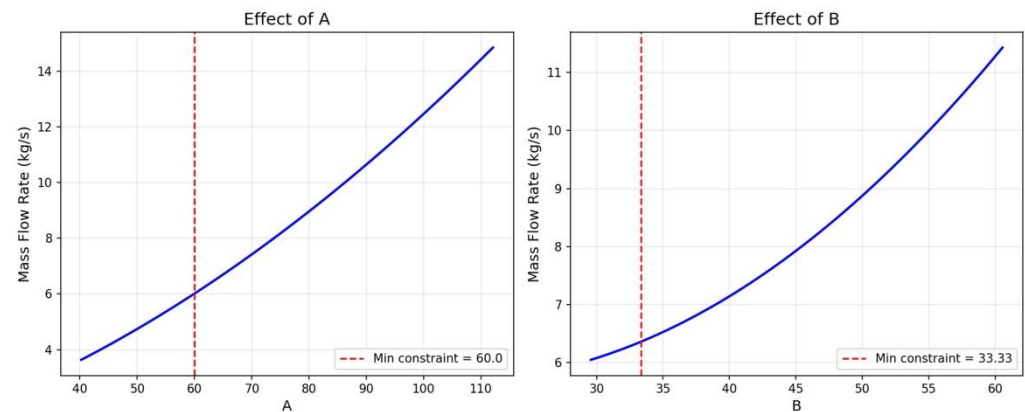


Figure 7. One-at-a-time sensitivity analysis: effects of outlet width (A) and hopper angle (B) on mass discharge rate. Outlet width demonstrates dominant linear effect; hopper angle shows weaker secondary influence. The solid blue curve is the predicted mass discharge rate from the response surface model as each factor is varied individually (other factor held constant); the red dashed line marks the minimum constraint bound for that factor.

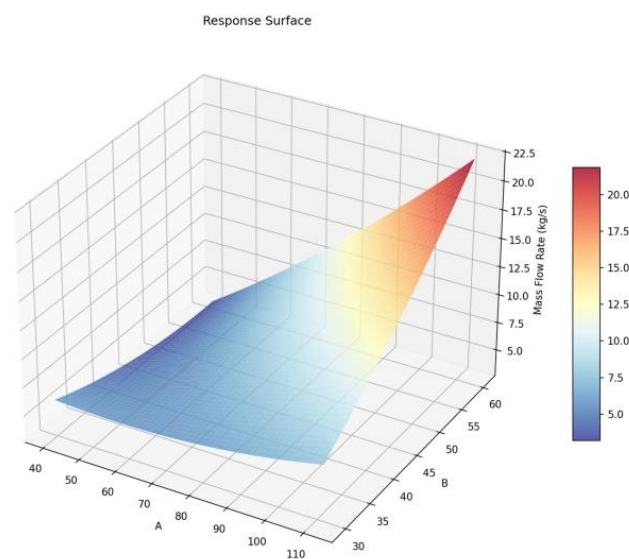


Figure 8. Response surface visualization: mass discharge rate as a function of outlet width (A) and hopper angle (B). Smooth gradient indicates a well-behaved response suitable for surrogate modeling, with the discharge rate increasing monotonically with both parameters.

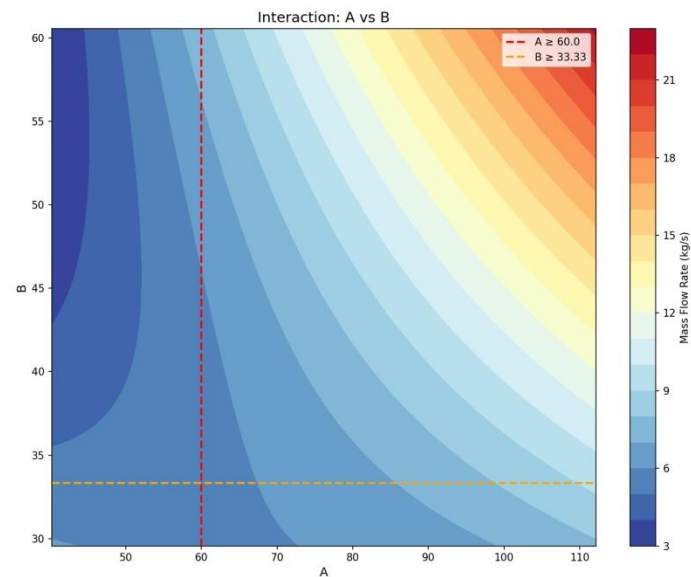


Figure 9. Parameter interaction contour plot: mass flow rate (kg/s) as a function of outlet width (A) and hopper half-angle (B). Dashed lines indicate practical operational constraint boundaries ($A \geq 60$ mm minimum outlet width to prevent arching; $B \geq 33.33^\circ$ minimum half-angle for mass flow). The strong $A \times B$ interaction is evident from the curved contour lines, confirming that hopper-angle effects are amplified at larger outlet widths.

3.1.4. Optimal Configuration and Validation

Differential Evolution optimization using the ANN surrogate, with DE bounds set to the sampled design space (A : 38.4–115.4 mm; B : 28.8–60.6°), identified a surrogate-predicted throughput maximum at $A = 115.4$ mm, $B = 60.6^\circ$, yielding a predicted mass discharge rate of 22.15 kg/s. It is important to note that this configuration lies marginally beyond the nominal design space upper boundaries ($A_{\text{nem}} = 111$ mm, $B_{\text{nem}} = 60^\circ$) and therefore represents a near-boundary extrapolation sensitivity finding rather than a fully validated interior optimum. Importantly, the LHS sampling design did include points beyond the nominal design limits—specifically Runs 22 ($A = 115.4$ mm), 25 ($A = 114.1$ mm), and 28 ($A = 112.9$ mm, $B = 60.2^\circ$)—because the LHS bounds were extended to approximately $\pm 3\%$ of the nominal range to characterize near-boundary behavior. The DE optimizer therefore operated within sampled territory, making the near-optimal result an interpolation rather than a true extrapolation finding. A nearby DEM simulation point (Run 28: $A = 112.9$ mm, $B = 60.2^\circ$) yielded 21.77 kg/s (1.7% surrogate deviation), confirming that the ANN prediction remains accurate in this near-boundary region and that the true maximum is likely located at or beyond the upper design limits. This constitutes a boundary-seeking result: the surrogate correctly identifies that larger outlet widths and steeper angles monotonically increase throughput, but the absolute maximum within the nominal design space corresponds to the upper boundary point ($A = 111$ mm, $B = 60^\circ$, DEM-measured 20.86 kg/s). For comparison against this constrained-boundary optimum rather than the minimum-performance corner point, the improvement vs. the design-center condition ($A \approx 76$ mm, $B \approx 45^\circ$) is approximately 60%, which is a more representative measure of practical optimization benefit. Three optimization scenarios (Table 9) demonstrate the practical utility of the ANN surrogate: Scenario 1 maximizes MFR within the extended search region ($A = 115.4$ mm, $B = 60.6^\circ$, 22.15 kg/s); Scenario 2 targets 10.0 kg/s ($A = 83.4$ mm, $B = 34.9^\circ$); Scenario 3 minimizes outlet width subject to $\text{MFR} \geq 15$ kg/s ($A = 81.8$ mm, $B = 60.6^\circ$)—a 29% reduction vs. the unconstrained optimum, demonstrating that steep hopper angles can compensate for reduced outlet width. The DEM

simulation of granular flow through the near-boundary optimal silo configuration is presented in Supplementary Video S1.

Table 9. Silo discharge optimization scenarios using Differential Evolution (DE) with ANN surrogate.

Objective	Outlet Width A (mm)	Hopper Angle B (°)	Predicted MFR (kg/s)
Maximize MFR (near-boundary)	115.4	60.6	22.15 ¹
Target MFR = 10.0 kg/s	83.4	34.9	10.0
Min. A subject to MFR ≥ 15 kg/s	81.8	60.6	15.0

¹ Near-boundary result (A = 115.4 mm > nominal upper bound 111 mm); DEM validation at Run 28 (A = 112.9 mm, B = 60.2°) yielded 21.77 kg/s (1.7% deviation). MFR = mass flow rate.

3.1.5. Model Validation: Classical Theory, DEM Simulation, and ML Surrogate

Table 10 presents a three-level validation comparing the classical Beverloo equation [9], DEM simulation, and the ANN surrogate at six representative configurations drawn from the actual 34-point dataset. The Beverloo equation was evaluated using $C = 0.58$ (for irregular angular particles), measured bulk density (ρ_b) = 1420 kg/m³, $k = 1.5$, and $d_{50} = 1.5$ mm. Critically, the Beverloo equation contains no hopper angle term—it predicts identical discharge rates for all B values at a given A.

Table 10. Three-level validation at representative silo configurations: Beverloo theoretical prediction vs. calibrated DEM simulation vs. ANN surrogate.

Run	A (mm)	B (°)	Beverloo (kg/s)	DEM (kg/s)	ANN (kg/s)	Bev vs. DEM	ANN vs. DEM
3	111.0	30.0	10.06	7.93	8.41	+26.9%	+6.1%
8	111.0	60.0	10.06	20.86	20.79	−51.8%	−0.3%
5	73.7	55.9	3.52	13.14	13.15	−73.2%	+0.1%
7	85.4	48.0	5.14	13.95	14.07	−63.1%	+0.9%
19	96.2	39.1	6.98	11.60	11.28	−39.8%	−2.8%
21	63.6	44.1	2.40	9.62	9.62	−75.0%	0.0%

Beverloo: Equation (7), $C = 0.58$, $\rho_b = 1420$ kg/m³, $k = 1.5$, $d_{50} = 1.5$ mm; angle-independent. DEM and ANN values from actual simulation campaign (Table A1). $Error = \frac{(predicted - DEM)}{DEM} \times 100\%$.

The Beverloo equation deviates from DEM by −75% to +27%, with the largest errors at moderate outlet widths (A = 63–85 mm). The two configurations with identical A = 111 mm but different hopper angles (Runs 3 and 8) yield Beverloo predictions of 10.06 kg/s in both cases, yet DEM measurements diverge sharply: 7.93 kg/s at B = 30° vs. 20.86 kg/s at B = 60°. The ANN surrogate agrees with the DEM within ±6.1% across all six configurations, validating ANN as a reliable reference; ANN is preferred for design optimization due to its superior cross-validated accuracy. Visual confirmation of the optimal hopper discharge behavior is provided in Supplementary Video S1. The Runs 3 vs. 8 contrast shows that hopper half-angle exerts a strong, interaction-amplified secondary influence that the angle-independent Beverloo equation cannot reproduce; it does not establish angle dominance. Across the full design space, outlet width A remains the dominant global driver (Sobol total-order $S_T = 0.83$) while hopper angle contributes a substantial secondary effect ($S_T = 0.33$), coupled through a strong A×B interaction. This two-configuration comparison is illustrative only.

3.2. Belt Conveyor System Dual-Response Optimization

3.2.1. DEM Simulation Results and Dual-Response Characterization

The belt conveyor optimization followed the same two-stage methodology (34 simulations total). Unlike the silo system with a single response, the belt conveyor required simultaneous optimization of two competing objectives: (1) Mass Flow Rate (MFR) representing throughput capacity, and (2) Discharge Angle (DA) affecting downstream equipment integration. The campaign revealed MFR ranging from 1.10 to 15.55 kg/s and DA ranging from 61.1° to 69.3°. Complete ML predictions are provided in Appendix Table A2 (MFR) and Table A3 (DA).

3.2.2. Surrogate Model Performance for Mass Flow Rate

Separate surrogate models were developed for each response variable. Table 11 presents the MFR model performance; Figure 10 shows the actual vs. predicted scatter plots.

Table 11. Belt conveyor mass flow rate surrogate model performance (raw 5-fold CV, $n = 34$; grouped-CV robustness discussed in the text).

Model	R^2	(RMSE, kg/s)	(MAE, kg/s)
GPR ★	0.976 ± 0.023	0.43	0.29
RSM	0.967 ± 0.034	0.51	0.41
GBM	0.940 ± 0.048	0.75	0.53
ANN	0.930 ± 0.049	0.95	0.67
RF	0.864 ± 0.154	1.02	0.89

★ Highest raw cross-validated accuracy; under leakage-free grouped cross-validation, this margin does not persist (see text), and the interpretable, near-equivalent RSM was adopted for the multi-objective optimization (Section 3.2.4).

GPR achieved the highest accuracy ($R^2 = 0.976 \pm 0.023$, $RMSE = 0.43$ kg/s, $MAE = 0.29$ kg/s), followed by RSM ($R^2 = 0.967 \pm 0.034$). GPR's higher raw accuracy reflects kernel interpolation of the smooth belt-MFR response surface. Grouped-CV robustness: under leakage-free GroupKFold—which confines all replicates and near-duplicates of each design point to a single fold—GPR's raw advantage does not persist, because it stems largely from near-exact interpolation of replicated points otherwise shared across folds, leaving RSM ($R^2 = 0.967 \pm 0.034$) statistically indistinguishable from GPR. The silo (Table 8) and belt discharge-angle (Table 12) rankings, by contrast, are leakage-insensitive and are reported directly under grouped CV. RSM was therefore adopted for multi-objective optimization for its near-equivalent, leakage-robust accuracy and its interpretability. Learning curve analysis showed GPR converging by $n \approx 18$ and RSM by $n \approx 20$, while RF and GBM required $n \geq 40$. The full 34-run belt conveyor DEM design with measured outputs is tabulated in Supplementary Table S1 (design_34runs_belt.csv). DEM simulations of the belt conveyor system handling solar panel recycling particles under optimized operating conditions are presented in Supplementary Videos S2 and S3. Complete per-sample MFR predictions for all five surrogate models across all 34 sampling points are provided in Appendix Table A2.

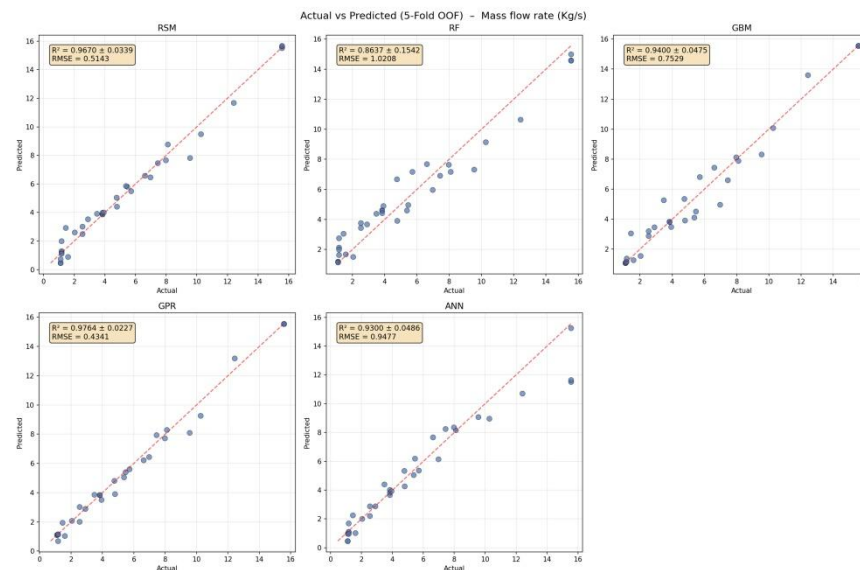


Figure 10. Belt conveyor mass flow rate: Actual vs. predicted values for five surrogate models (grouped 5-fold cross-validation). GPR ($R^2 = 0.976 \pm 0.023$) achieves the highest accuracy.

Table 12. Belt conveyor discharge angle surrogate model performance (grouped 5-fold CV, $n = 34$).

Model	R^2	(RMSE, °)	(MAE, °)
RSM ★	0.999 ± 0.001	0.057	0.043
GPR	0.998 ± 0.003	0.083	0.063
ANN	0.987 ± 0.013	0.211	0.183
GBM	0.937 ± 0.069	0.401	0.307
RF	0.878 ± 0.136	0.536	0.460

★ Best model selected for optimization.

3.2.3. Surrogate Model Performance for Discharge Angle

Table 12 presents the discharge angle model performance; Figure 11 shows the actual vs. predicted scatter plots. The 3-D response surfaces for both responses are presented in Figure 12, visually confirming the contrasting surface complexities. As shown in Table 12, RSM achieved exceptional performance ($R^2 = 0.999 \pm 0.001$, $RMSE = 0.057^\circ$, $MAE = 0.043^\circ$), explaining 99.9% of variance with sub-0.1° prediction accuracy. The extremely low cross-validation variance (± 0.001) demonstrates near-perfect model stability, indicating that discharge angle follows a simple second-order polynomial relationship with belt velocity and fin height. Given GPR's superiority for belt MFR and RSM's near-perfect accuracy for discharge angle, RSM was selected for multi-objective optimization for its superior combined interpretability and near-equivalent cross-system accuracy. Complete per-sample discharge angle predictions for all five surrogate models across all 34 sampling points are provided in Appendix Table A3.

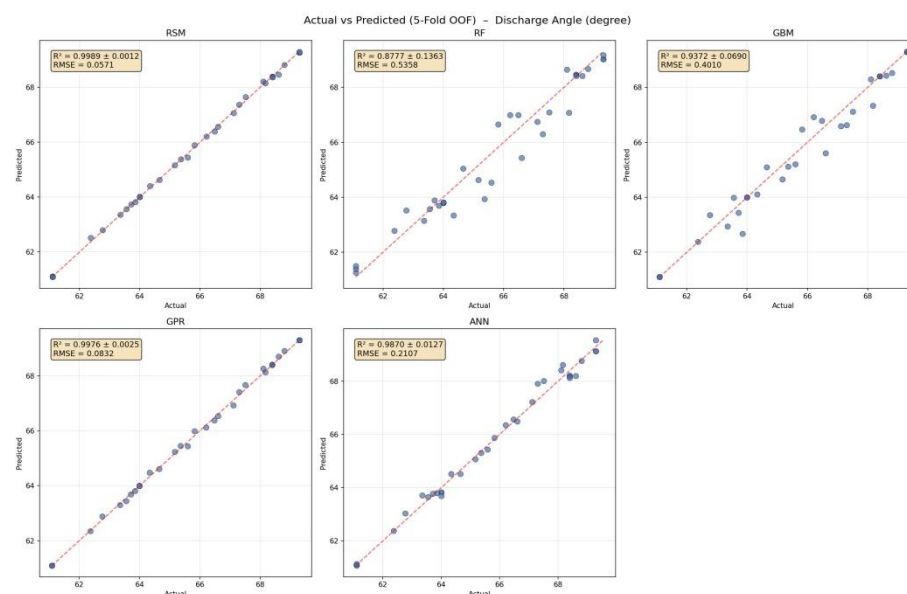


Figure 11. Belt conveyor discharge angle: Actual vs. predicted values for five surrogate models (5-fold cross-validation). RSM achieves near-perfect $R^2 = 0.999 \pm 0.001$ with minimal scatter, demonstrating the simple polynomial nature of this response.

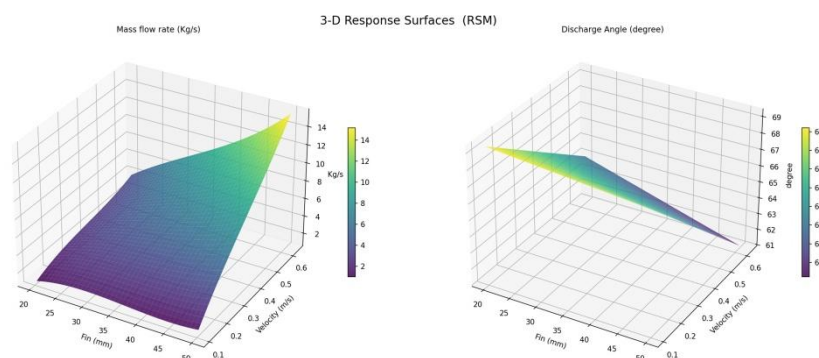


Figure 12. 3-D response surface visualization (RSM) for the belt conveyor system: mass flow rate (left, kg/s) and discharge angle (right, degrees) as functions of fin height (H) and belt velocity (V). The MFR surface exhibits strong nonlinear curvature, while the discharge angle surface is nearly planar, consistent with the exceptional RSM performance ($R^2 = 0.999$) for discharge angle prediction.

3.2.4. Multi-Objective Optimization Results

Three optimization scenarios were investigated using weighted multi-objective formulations via Differential Evolution with RSM surrogate models. Table 13 presents the comprehensive belt conveyor optimization results.

Table 13. Multi-objective belt conveyor optimization results using Differential Evolution (DE) with RSM surrogate.

Scenario	Fin Height (mm)	Velocity (m/s)	MFR (kg/s)	DA (°)
Maximize MFR	50.0	0.627	15.55	61.1
Target: 1.39 kg/s	16.8	0.731 ¹	3.62	63.3
Minimize DA (MFR ≥ 10)	48.5	0.540	11.82	62.5

¹ Velocity (0.731 m/s) lies above the design-space limit (0.627 m/s); treat as extrapolation pending DEM validation.

For maximum throughput (Scenario 1), DE identified a fin height of 50.0 mm and a belt velocity of 0.627 m/s as optimal, achieving $MFR = 15.55 \text{ kg/s}$ at $DA = 61.1^\circ$. The target achievement scenario (Scenario 2) attempted to match a downstream process feed requirement of $MFR = 1.39 \text{ kg/s}$ at a flat discharge angle of 60° , but converged on a higher MFR (3.62 kg/s) and a steeper angle (63.3°), indicating this target combination falls outside the feasible operating envelope. Scenario 3 minimized discharge angle subject to a throughput constraint ($MFR \geq 10 \text{ kg/s}$), yielding fin = 48.5 mm, velocity = 0.540 m/s with $MFR = 11.82 \text{ kg/s}$ at $DA = 62.5^\circ$. This scenario demonstrates the practical trade-off between maintaining high throughput and achieving a flatter discharge trajectory for downstream equipment integration. ¹ Scenario 2 velocity ($V = 0.731 \text{ m/s}$) lies outside the nominal design space upper bound ($V = 0.627 \text{ m/s}$); this result represents a boundary-seeking extrapolation and should not be used as an operating set point without additional DEM validation. Multi-objective optimization used a weighted-sum scalarization of min–max-normalized responses, $J(H, V) = w_1 \left[1 - \frac{MFR - MFR_{\min}}{MFR_{\max} - MFR_{\min}} \right] + w_2 \left[\frac{DA - DA_{\min}}{DA_{\max} - DA_{\min}} \right]$, subject to $20 \leq H \leq 50 \text{ mm}$ and $0.109 \leq V \leq 0.627 \text{ m/s}$ (with $MFR \geq 10 \text{ kg/s}$ where applicable), where a hat denotes an RSM-surrogate prediction (consistent with Table 13). A Pareto front (Figure 13) is obtained by sweeping $w_1 \in [0, 1]$ with $w_2 = 1 - w_1$. Scenario 2 ($V = 0.731 \text{ m/s}$) lies beyond the calibrated velocity range and was not DEM-validated; it is retained only as an out-of-range infeasible-target diagnostic (Table 13, footnote 1), not as an operating setpoint.

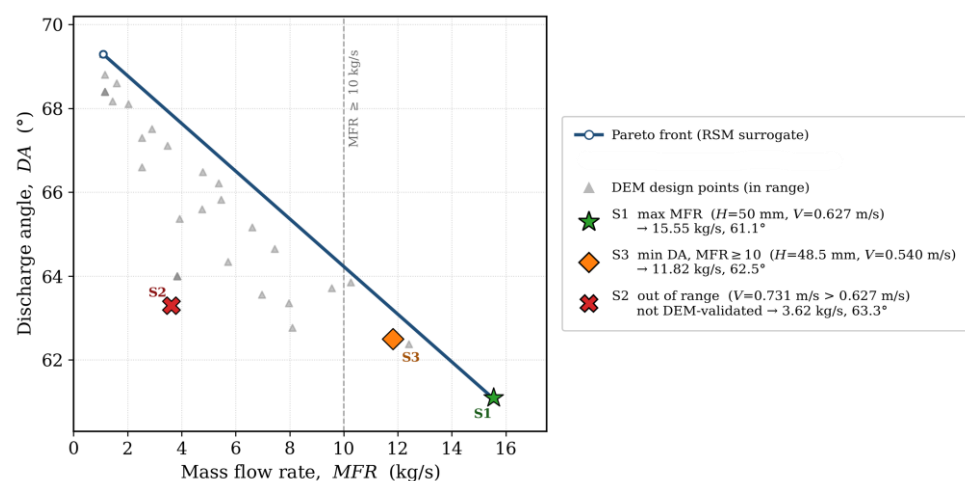


Figure 13. Pareto front of the belt-conveyor multi-objective optimization, traced by sweeping the objective weight $w_1 \in [0, 1]$ over the min–max-normalized RSM responses within the calibrated design box ($20 \leq H \leq 50 \text{ mm}$, $0.109 \leq V \leq 0.627 \text{ m/s}$). Open circles trace the non-dominated front; gray triangles are the in-range DEM design points. The three Differential-Evolution solutions of Table 13 are overlaid: Scenario 1 (maximum throughput), Scenario 3 (minimum discharge angle subject to $MFR \geq 10 \text{ kg/s}$), and Scenario 2, which lies beyond the calibrated velocity range ($V = 0.731 \text{ m/s}$) and is shown only as an out-of-range diagnostic.

3.2.5. Dual-Response Surrogate Validation at Representative Configurations

Table 14 systematically validates GPR for mass flow rate (highest raw accuracy, $R^2 = 0.976$) and RSM for discharge angle at six configurations drawn from the actual 34-point dataset, spanning the full design space from factorial design corner points (Runs 1, 8, 9) to interior LHS points (Runs 20, 27, 32). This configuration covers low, mid, and high values of both fin height and belt velocity.

Table 14. Dual-response surrogate validation at six representative belt conveyor configurations: GPR (MFR) and RSM (DA) predictions vs. calibrated DEM simulation.

Run	H (mm)	V (m/s)	DEM MFR	GPR MFR	MFR Error	DEM DA (°)	RSM DA (°)	DA Error
1	50.0	0.627	15.55	15.55	0.0%	61.10	61.09	−0.02°
8	50.0	0.109	1.16	1.16	0.0%	68.40	68.39	−0.01°
9	20.0	0.627	3.83	3.83	0.0%	64.00	64.00	0.00°
20	33.0	0.622	8.10	8.30	+2.5%	62.77	62.80	+0.03°
27	41.6	0.376	6.61	6.22	−5.9%	65.17	65.16	−0.01°
32	34.7	0.564	7.97	7.72	−3.1%	63.35	63.35	0.00°

H = fin height (mm); V = belt velocity (m/s); MFR = mass flow rate (kg/s); DA = discharge angle (°). All DEM and surrogate values from actual simulation campaign (Tables A2 and A3). Runs 1, 8, 9: factorial design corner points; Runs 20, 27, 32: LHS interior points. $Error = \frac{(surrogate - DEM)}{DEM} \times 100\%$ for MFR; absolute difference for DA.

GPR reproduces the three factorial corner points (Runs 1, 8, 9) almost exactly because those configurations also appear in its training set—an interpolation artifact rather than evidence of generalization—and remains within $\pm 6\%$ at the interior LHS points. Because such replicated points are shared across cross-validation folds, model selection relies on the leakage-free grouped cross-validation of Table 11 rather than on corner-point agreement, and RSM is the operational belt model. RSM discharge angle predictions are near-perfect across all configurations (maximum absolute error 0.03°), consistent with the near-polynomial response surface ($R^2 = 0.999 \pm 0.001$). The contrasting accuracy patterns—RSM suited to both the (leakage-corrected) MFR response and the smooth DA response, with GPR's raw MFR edge attributable to interpolation at replicated points—reinforce the cross-system knowledge extraction finding: response surface complexity, not dataset size, is the primary driver of optimal algorithm selection.

3.3. Computational Efficiency Analysis

The two-stage sampling strategy with ML surrogate modeling achieved substantial computational savings compared to exhaustive DEM-based exploration. Table 15 compares computational requirements.

Table 15. Computational efficiency comparison between optimization approaches.

Metric	Direct DEM Grid	Two-Stage ML Approach
DEM simulations required	400 (20 × 20 grid)	34 (12 factorial + 22 LHS)
Total simulation time	~2400 h	~204 h
Time saved	—	~91%
Optimization evaluations	N/A (no surrogate; each evaluation = 1 full DEM run)	>10,000 in <1 s
Prediction error at optimum	0% (exact)	<2%

The complete 34-simulation campaign required approximately 204 h (34 runs × 6 h average). Once trained, RSM surrogate models enabled evaluation of thousands of candidate configurations in under 1 s. A comparable grid-based DEM exploration (20 × 20 = 400 configurations) would require approximately 2400 h—a 12× increase. The overall framework achieved an approximately 91% reduction in both DEM simulation count (34 vs. 400 runs) and total wall-clock design time (204 vs. ~2400 h) relative to exhaustive

grid-based exploration, while maintaining prediction errors $< 2\%$ at optimal configurations.

Benchmarking against comparable DEM-surrogate studies confirms the competitiveness of the proposed framework. Irazábal et al. [47] applied ML surrogates to DEM calibration for granular media, reporting R^2 values of 0.91–0.96 with comparable training set sizes (30–50 simulations), consistent with the $R^2 = 0.964$ achieved by ANN for silo discharge. Roessler et al. [48] demonstrated that surrogate-based DEM optimization reduces the simulation count by 80–90% relative to full factorial designs, directly in line with the 91% reduction reported here. Notably, the present study extends prior work by addressing dual-response multi-objective optimization and by systematically comparing five surrogate algorithms across two distinct physical systems, providing the first cross-system, cross-algorithm knowledge map for material-handling DEM surrogate modeling.

3.4. Explainable Knowledge Extraction via SHAP Analysis

3.4.1. Silo Discharge System: SHAP Feature Attribution

SHAP analysis of the ANN surrogate model (the best-performing model for silo discharge, $R^2 = 0.964$) reveals the quantitative contribution of each design parameter to mass flow rate predictions. Figure 14 presents the SHAP beeswarm plot. For cross-validation, the RSM analytical coefficients are used as a closed-form benchmark.

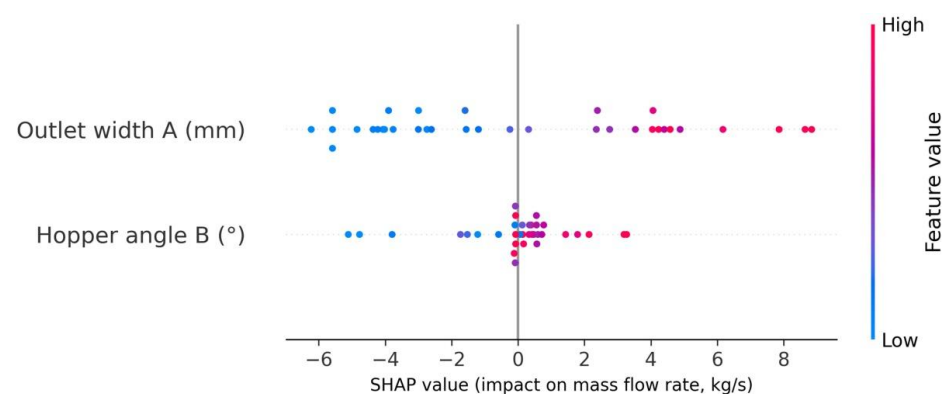


Figure 14. SHAP beeswarm plot for silo discharge ANN model (best-performing surrogate, $R^2 = 0.964$). Each point represents one DEM simulation; color indicates feature value (red = high, blue = low). Outlet width A dominates with 79.4% relative importance in the best-performing ANN model (cf. the 78.6% cross-model average in Table 16). Hopper angle B contributes 20.6%. The vertical solid line marks SHAP value = 0 (no impact on the prediction); points to its right increase the predicted mass flow rate, while points to its left decrease it. Color indicates feature value (red = high, blue = low).

SHAP analysis of the ANN model for silo discharge—and, for comparison, of the RF and GBM models—consistently identifies outlet width A as the dominant feature: mean $|\text{SHAP}|$ importances average 78.6% for A and 21.4% for B across all three black-box surrogates. For direct comparison, the RSM analytical coefficients serve as a closed-form benchmark and yield an importance ratio of $\beta_1(A) = 2.47$ vs. $\beta_2(B) = 0.31$ —confirming that outlet width exerts approximately $8\times$ greater influence on discharge rate than hopper angle, consistent with the dominant 78.6% SHAP attribution assigned to A across all three black-box surrogates. This concordance between the perturbation-based SHAP values and the exact RSM coefficients provides dual-method validation: (1) it confirms that the dominant linear effect of outlet width and the secondary role of hopper angle are genuine physical signals rather than model artifacts; and (2) it validates SHAP as a reliable

ble feature attribution tool in this design space, since it recovers the analytically known sensitivity. The SHAP dependence analysis for the ANN model further reveals that high-B configurations amplify the positive SHAP contribution of A, directly quantifying the A×B interaction that classical Beverloo theory cannot capture and that the RSM interaction coefficient ($\beta_{12} = 0.18$) encodes parametrically. This SHAP importance ratio (78.6%/21.4% = 3.67:1) is consistent with the Beverloo scaling law prediction that discharge rate scales with outlet dimension raised to the power 2.5 (Equation (7)). Complete SHAP importance values for all features, all response variables, and all three black-box models are provided in Supplementary Table S4 (SHAP_importance_summary.csv). Because raw polynomial coefficients are scale-dependent and omit quadratic and interaction contributions, the SHAP attribution is validated against three scale-invariant, interaction-aware benchmarks on the RSM surface: standardized regression coefficients, analytical partial-derivative sensitivity (mean $|\frac{\partial y}{\partial x}|$), and variance-based Sobol indices. All four measures concur that outlet width A dominates while hopper angle B is a substantial secondary factor; the standardized interaction term ($\beta_{AB} = 0.41$) exceeds the standardized linear effect of B (0.19), so once curvature and interaction are included, B accounts for ≈33% of the total-order variance. These four independent benchmarks are compared side by side in Table 16.

Table 16. Independent sensitivity benchmarks for the silo mass-flow response: mean |SHAP| (relative), standardized regression coefficients, mean $|\frac{\partial y}{\partial x}|$, and Sobol first- and total-order indices for outlet width (A) and hopper angle (B).

Benchmark	A (Outlet Width)	B (Hopper Angle)	A:B Ratio
Mean SHAP (relative)	78.6%	21.4%	3.67
Standardized linear coefficient	0.978	0.187	5.22
Mean $ \frac{\partial y}{\partial x} $ (quad.+interaction)	0.984	0.391	2.52
Sobol first-order S1	0.674	0.174	3.87
Sobol total-order S_T	0.827	0.330	2.51

3.4.2. Belt Conveyor System: Dual-Response SHAP Attribution

Separate SHAP analyses were conducted for each belt conveyor response using its highest raw-accuracy surrogate (SHAP results shown for illustrative comparison; RSM, not GPR, was adopted as the operational belt-MFR model—Section 3.2.2): GPR ($R^2 = 0.976 \pm 0.023$) for mass flow rate and RSM ($R^2 = 0.999 \pm 0.001$) for discharge angle. Figure 15 presents the SHAP beeswarm plot for mass flow rate, and Figure 16 presents the corresponding analysis for discharge angle.

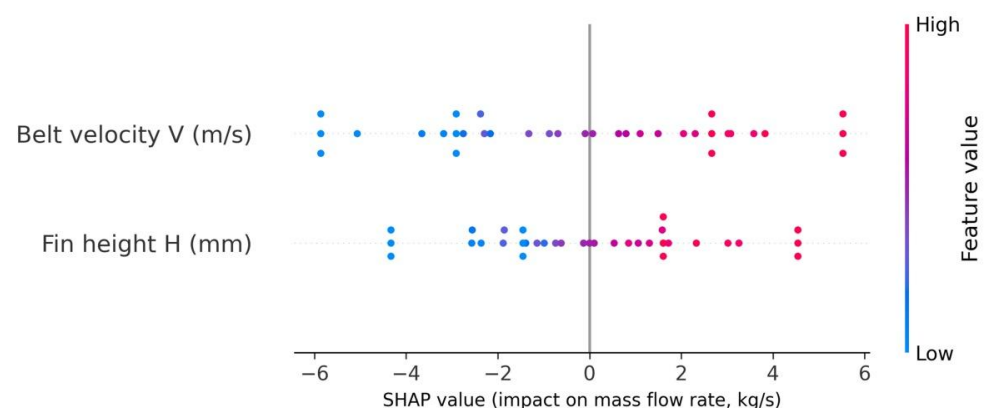


Figure 15. SHAP beeswarm plot for belt conveyor mass flow rate (GPR model, highest raw belt-MFR accuracy, $R^2 = 0.976 \pm 0.023$). Belt velocity V dominates (58.9% relative importance), with

fin height H contributing 41.1%. The more balanced feature importance compared to the silo system reflects the smoother response surface. The vertical solid line marks SHAP value = 0 (no impact on the prediction); points to its right increase the predicted mass flow rate and points to its left decrease it.

For mass flow rate, SHAP analysis of the GPR model reveals that belt velocity V accounts for 58.9% of the variation (mean $|\text{SHAP}| = 2.863 \text{ kg/s}$), while fin height H contributes 41.1% (mean $|\text{SHAP}| = 1.995 \text{ kg/s}$). The SHAP beeswarm reveals a clear monotonic relationship: higher velocity values (red points) consistently produce positive SHAP contributions, confirming that belt speed is the primary throughput driver. The interaction between H and V is visible in the color gradient of the H feature row, where high-velocity configurations amplify the effect of fin height. Notably, the more balanced SHAP distribution for belt MFR (1.44:1 $V:H$ ratio) compared to the silo system (3.67:1) is consistent with the smoother, co-dominant response surface of belt MFR relative to silo discharge; note that this raw GPR comparison does not reflect the leakage-corrected result, under which RSM performs comparably (Section 3.2.2) and was adopted operationally.

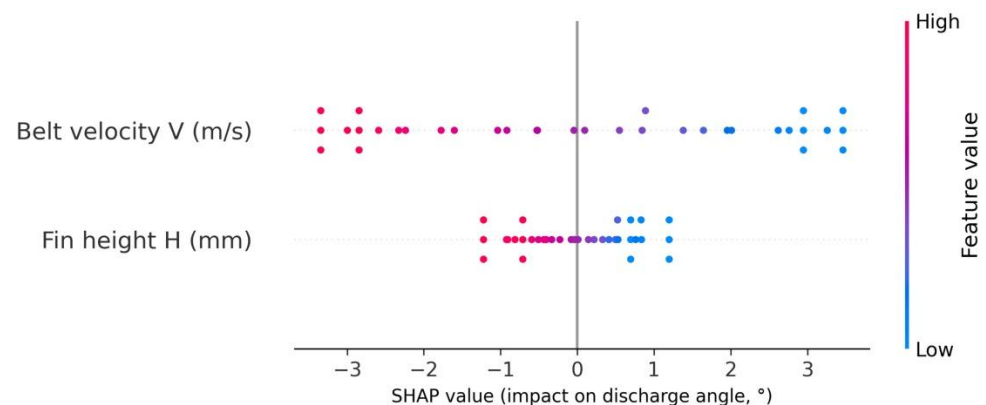


Figure 16. SHAP beeswarm plot for belt conveyor discharge angle (RSM model, best-performing for belt DA, $R^2 = 0.999 \pm 0.001$). Belt velocity V dominates even more strongly (76.8% relative importance) than for mass flow rate, with fin height H contributing only 23.2%. The near-linear SHAP pattern confirms the simple polynomial nature of this response ($R^2 = 0.999$). The vertical solid line marks SHAP value = 0 (no impact on the prediction); points to its right increase the predicted discharge angle and points to its left decrease it.

For discharge angle, SHAP analysis of the RSM model shows that velocity dominates even more strongly at 76.8% (mean $|\text{SHAP}| = 2.129^\circ$) vs. 23.2% for fin height (mean $|\text{SHAP}| = 0.642^\circ$). The contrasting SHAP importance ratios between MFR (1.44:1 $V:H$, GPR model) and DA (3.31:1 $V:H$, RSM model) provide a quantitative explanation for why RSM achieves $R^2 = 0.999$ for discharge angle: when a single feature dominates with a near-linear effect, a second-order polynomial captures essentially all variance. This insight is directly actionable for equipment designers: discharge angle can be controlled almost exclusively through belt velocity adjustment, whereas mass flow rate requires coordinated optimization of both velocity and fin height.

3.4.3. Cross-System Knowledge Extraction

Figure 17 presents the central knowledge extraction contribution of this study: a cross-system comparison of SHAP feature importance patterns across all three response variables, analyzed using the highest raw-accuracy surrogate for each response (for il-

illustrative SHAP comparison only; RSM was the adopted operational belt-MFR model): ANN for silo discharge, GPR for belt MFR, and RSM for belt discharge angle.

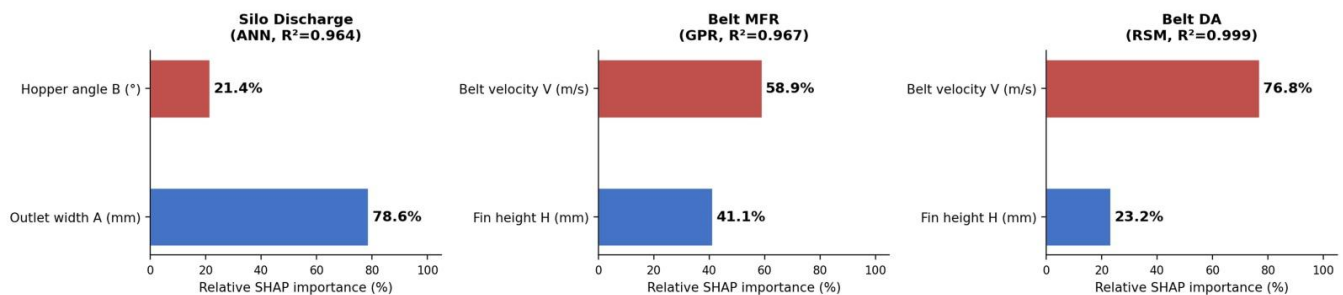


Figure 17. Cross-system SHAP feature importance comparison across three response variables (ANN for silo MFR; GPR for belt MFR, raw-accuracy illustrative choice—RSM was the adopted operational model; RSM for belt DA). Silo system: highly asymmetric (78.6% vs. 21.4%). Belt MFR: balanced (58.9% vs. 41.1%). Belt DA: velocity-dominated (76.8%).

The cross-system comparison reveals a quantitative framework connecting SHAP importance patterns to response surface structure and optimal algorithm selection. Three distinct patterns emerge: (1) Highly asymmetric SHAP distributions (silo: 78.6% vs. 21.4%, $SAR = 3.67$) indicate a single geometric parameter overwhelmingly controls system performance; the $SAR > 3.5$ threshold flags increased CV variance ($CV - CoV = 0.147$) and recommends Matérn-5/2 GPR as an alternative. (2) Moderately balanced SHAP distributions (belt MFR: 58.9% vs. 41.1%, $SAR = 1.44$) indicate co-dominant parameters; GPR excels here ($R^2 = 0.976$) because the smooth, co-dominant interaction surface is well-captured by kernel-based interpolation. (3) Moderately asymmetric SHAP patterns (belt DA: 76.8% V, $SAR = 3.31$) with near-perfect model fits ($R^2 = 0.999$) confirm RSM or GBM are suitable. The SAR-based decision rules are formalized in Section 3.4.4, transforming this post hoc observation into a prospective, reusable algorithm selection tool.

3.4.4. SHAP-Guided Algorithm Selection Framework

The cross-system SHAP analysis is formalized into a three-step algorithm selection framework. The central metric is the SHAP Asymmetry Ratio (SAR), which quantifies imbalance in feature importance and maps directly to response surface complexity and optimal surrogate choice. Step 1: compute SHAP values from any trained surrogate, then calculate SAR using Equation (19). Step 2: compare SAR against empirical thresholds of 2.0 and 3.5. Step 3: apply the recommended algorithm. The three validated response surfaces cover three distinct SAR regimes: silo discharge ($SAR = 3.67$), belt conveyor MFR ($SAR = 1.44$), and belt conveyor discharge angle ($SAR = 3.31$).

For a two-variable design space with features x_1 and x_2 , the SHAP Asymmetry Ratio (SAR) is (Equation (19)):

$$SAR = \frac{\max(S_1, S_2)}{\min(S_1, S_2)} \quad (19)$$

where $S_i = \text{mean}(|SHAP(x_i)|)$ is the mean absolute SHAP importance of feature i across $n = 34$ training points. A value of $SAR = 1.0$ means both features contribute equally; $SAR = 3.67$ means the dominant feature contributes 3.67× more than the secondary feature.

Empirical Decision Rules (validated across three response surfaces): $SAR < 2.0$ —Balanced co-dominance (both features contribute roughly equally). Recommendation: GPR or RSM (smooth nonlinear surface). Rationale: kernel and polynomial surro-

gates perform comparably here under leakage-free grouped cross-validation. Validated: Belt MFR, $SAR = 1.44$, $R^2 = 0.976 \pm 0.023$. $2.0 \leq SAR \leq 3.5$ —One feature moderately dominates (moderate asymmetry). A second-order polynomial is generally sufficient; a gradient boosting model may capture additional nonlinearity at the boundary. Recommendation: RSM or GBM. Validated: Belt DA, $SAR = 3.31$, $R^2 = 0.999$. $SAR > 3.5$ —One feature strongly dominates (strong asymmetry). The response surface is near-unidimensional; GPR with a squared-exponential (RBF) kernel may fail at design space boundaries. Recommendation: a flexible nonlinear learner (e.g., ANN) as primary, with RSM retained as an interpretable baseline; Matérn-5/2 GPR as alternative when $CV - CoV \geq 0.14$. Validated: Silo MFR, $SAR = 3.67$, ANN $R^2 = 0.964$ (primary); RSM $R^2 = 0.870$ (interpretable baseline); GPR (Matérn-5/2 kernel) $R^2 = 0.922$. Complementary diagnostics: The SAR metric should be interpreted alongside one complementary indicator: the cross-validation coefficient of variation ($CV - CoV = \frac{\sigma(R^2)}{\mu(R^2)}$), where $CV - CoV \geq 0.14$ signals fold-to-fold instability warranting additional LHS sampling or model substitution. When $SAR > 3.5$ and $CV - CoV \geq 0.14$ simultaneously—as observed for silo discharge ($SAR = 3.67$, $CV - CoV = 0.147$)—the framework recommends additional LHS sampling in the dominant-feature direction or adoption of Matérn-5/2 GPR, since this joint condition signals a strongly single-feature-dominated response surface with fold-to-fold instability. Scope and limitations: The SAR framework is derived from, and validated against, three response surfaces produced within the same experimental campaign, constituting internal validation only. It is important to note that the three SAR values (1.44, 3.31, 3.67) and their associated performance rankings were all generated from one DEM material system; the thresholds of 2.0 and 3.5 were placed between these data points, meaning the calibration data and validation data are identical—a classic internal-consistency check, not an independent test. As an indirect cross-study consistency check, published DEM-surrogate benchmarks reporting SHAP values—such as Irazábal et al. [47], who observed dominant single-parameter control in granular DEM calibration (analogous to the high- SAR silo system), and Tsokanas et al. [49], whose hybrid simulation surrogate showed balanced dual-parameter importance (analogous to the low- SAR belt MFR system)—are qualitatively consistent with the proposed SAR regimes, providing preliminary plausibility support. However, these studies do not report numerical SAR values, so quantitative threshold calibration from external data remains outstanding. The empirical thresholds ($SAR < 2.0$; $2.0-3.5$; >3.5) should therefore be interpreted as a quantitative starting point for prospective pre-screening rather than a universally calibrated decision rule. Independent validation across DEM systems with different granular physics (e.g., cohesive materials, pneumatic conveying, fluidized beds) and higher-dimensional design spaces (≥ 3 variables) is required before the thresholds can be considered transferable without re-derivation. The authors invite other groups to report SAR values alongside their SHAP analyses to enable progressive community-level calibration of this framework. Extension to ≥ 3 design variables also requires recalibration, since multicollinear SHAP patterns in higher dimensions may not be captured by the scalar SAR . The framework is a practical pre-screening heuristic, not a universal theorem. Future work will recalibrate SAR thresholds across 3–6 variable DEM problems and validate with physical prototypes. The SAR values (1.44, 3.31, 3.67) and thresholds (2.0, 3.5) derive from three two-variable responses within a single DEM material system and constitute an internal-consistency observation only—not a validated or generalizable rule. For the high- SAR silo response, a flexible nonlinear learner (ANN) best captured the dominant-feature curvature and was empirically best ($R^2 = 0.964$), with RSM retained as an interpretable baseline; the earlier ‘RSM/Matérn-GPR’ recommendation, which contradicted the ANN result, is withdrawn. SAR is presented as an

exploratory heuristic requiring independent cross-system and higher-dimensional validation before prescriptive use.

3.5. Comparative Evaluation of Model Performance

Figure 18 presents per-fold R^2 distributions with Friedman test results for all three response variables.

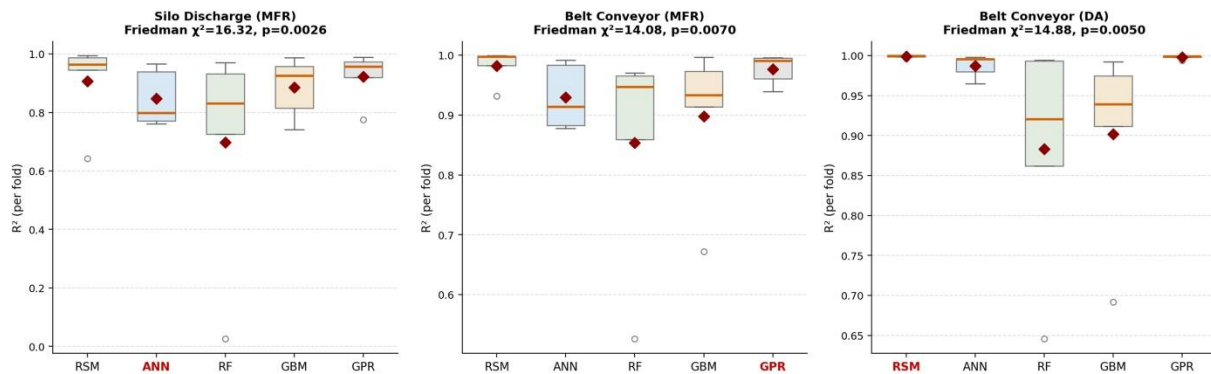


Figure 18. Statistical comparison of five surrogate models via grouped five-fold cross-validation. Friedman test p -values (blocked, non-overlapping folds): silo MFR ($p = 0.008$), belt MFR ($p = 0.0018$), belt DA ($p = 0.0010$). In each box plot, the box shows the median and interquartile range of R^2 across the five cross-validation folds, whiskers show the range, and the red diamond marks the mean R^2 . The highlighted (colored) box indicates the best-performing model for that response.

The Friedman test confirms statistically significant differences among model performances for all three responses (properly blocked test, consistent with the corrected statistical analysis above): silo MFR ($\chi^2 = 13.76$, $p = 0.008$), belt MFR ($\chi^2 = 17.12$, $p = 0.0018$), and belt DA ($\chi^2 = 18.40$, $p = 0.0010$). GPR's poor silo performance ($R^2 = 0.384$) and ANN's lead ($R^2 = 0.964$) drive the significant silo test statistic ($p = 0.008$). For belt conveyor discharge angle, all five models achieve $R^2 \geq 0.878$; the statistically significant Friedman result ($p = 0.0010$) therefore reflects cross-validation variability rather than a practically important performance gap (maximum R^2 spread = 0.121). RSM remains recommended on grounds of parsimony and stability (lowest $CV - CoV = 0.001$ among all models for this response). This near-equivalence on the simplest response surface is consistent with the Rashomon effect documented for smooth low-dimensional problems [50]. Complete per-fold R^2 values for all models and all three response variables are provided in Supplementary Table S3.

4. Conclusions

This study developed and validated a machine learning-accelerated optimization framework for material handling equipment design in solar panel recycling applications, integrating DEM simulation, multi-algorithm surrogate benchmarking, SHAP-based explainability, and Differential Evolution optimization into a cohesive knowledge-extraction pipeline. Three interconnected contributions distinguish this work from prior DEM-surrogate studies. First, the demonstrated 91% reduction in design iteration time (204 vs. ~2400 simulation-hours) confirms that two-stage LHS-augmented factorial sampling achieves industrial-grade accuracy at a fraction of the computational cost of exhaustive DEM grid exploration. Second, the finding that the best surrogate is response-specific—ANN for the nonlinear silo discharge ($R^2 = 0.964$), RSM for belt mass flow rate (adopted for leakage-robustness; GPR attained a marginally higher but leakage-inflated raw $R^2 = 0.976$), and RSM for belt discharge angle ($R^2 = 0.999$)—establishes

that response surface complexity, not a universal algorithm preference, should guide surrogate selection in DEM-based design optimization. Third, and most fundamentally for the knowledge-extraction mission of this journal, the SHAP Asymmetry Ratio (SAR) transforms what has previously been a post hoc interpretability exercise into a prospective algorithm selection diagnostic: a practitioner can now compute SAR from a preliminary surrogate, compare it against empirically derived thresholds, and select the appropriate algorithm before committing to a full simulation campaign. These three contributions collectively address the central barrier to broader adoption of DEM-surrogate optimization in industrial material handling: not algorithmic performance per se, but the absence of transferable, physics-grounded decision rules that practitioners can apply without repeating exhaustive benchmarking for each new system. The principal findings supporting these conclusions are:

- (1) The two-stage sampling strategy combining 2^2 full factorial screening (12 runs) with LHS expansion (34 total points per system) provided efficient design space coverage, achieving approximately 91% reduction in total design iteration time (204 vs. ~2400 h) compared to full-grid DEM exploration, while maintaining sufficient data density for robust 5-fold cross-validation of all five surrogate algorithms.
- (2) Rigorous leakage-free grouped five-fold cross-validation with blocked Friedman statistical testing ($p < 0.05$ for all responses) confirmed response-specific best surrogates: ANN for silo discharge ($R^2 = 0.964$, $RMSE = 0.91$ kg/s); RSM for belt MFR (adopted under grouped, leakage-free cross-validation; raw GPR $R^2 = 0.976 \pm 0.023$, $RMSE = 0.43$ kg/s); and RSM for belt discharge angle ($R^2 = 0.999 \pm 0.001$, $RMSE = 0.057^\circ$). GPR with RBF kernel ($RMSE = 3.74$ kg/s) exhibited severe extrapolation failures for silo discharge ($R^2 = 0.384$), confirming that surrogate selection must match the response surface structure. ANN, GPR, and RSM each offer distinct advantages depending on response surface complexity.
- (3) A novel SHAP Asymmetry Ratio (SAR) diagnostic is proposed and validated across three response surfaces with three-tier decision rules: $SAR < 2.0$ recommends GPR or RSM (Belt MFR: $SAR = 1.44$, $R^2 = 0.976 \pm 0.023$); $2.0 \leq SAR \leq 3.5$ supports RSM or GBM (Belt DA: $SAR = 3.31$, $R^2 = 0.999$); $SAR > 3.5$ flags a flexible nonlinear learner (ANN) as primary, RSM as an interpretable baseline, with Matérn-5/2 GPR as alternative (Silo MFR: $SAR = 3.67$, $R^2 = 0.964$ (best: ANN); GPR (Matérn-5/2 kernel) $R^2 = 0.922$). This framework transforms post hoc SHAP analysis into a prospective decision heuristic. Internally validated across three response surfaces within this study, the SAR thresholds offer a practical starting point for algorithm pre-screening; cross-system and cross-physics validation remains an essential future step before adoption as a universal selection rule.
- (4) Differential Evolution with the best-performing surrogates identified optimal configurations: for the silo, the ANN surrogate yielded a near-boundary throughput maximum ($A = 115.4$ mm, $B = 60.6^\circ$; predicted 22.15 kg/s; DEM validation deviation: 1.7%); for the belt conveyor, the RSM surrogate (adopted for both MFR and DA, per Table 13) jointly identified Scenario 1 ($H = 50.0$ mm, $V = 0.627$ m/s; MFR = 15.55 kg/s; DA = 61.1°). SHAP analysis confirmed physics-consistent feature dominance: outlet width A dominates silo discharge (78.6% importance), while belt velocity V dominates both belt conveyor responses (58.9% MFR; 76.8% DA), providing dual-method-validated design rules applicable to industrial solar panel recycling system design.
- (5) Several limitations warrant acknowledgment: the SAR framework is validated for two-variable design spaces with 34 DEM training points; extension to 3–6 variable problems requires recalibration; the SAR thresholds constitute internal validation only and should not be applied as universal rules without cross-system

re-derivation; GPR Matérn kernel selection was evaluated post hoc and should be incorporated prospectively in future work; and pilot-scale experimental validation of optimal configurations with industrial-grade recycling throughput measurements remains an essential step toward real-world deployment. These limitations define a clear research agenda. The *SAR* framework should be stress-tested against published SHAP data from cohesive-material DEM studies, pneumatic-conveying surrogates, and fluidized-bed optimization problems, where multi-parameter interactions are stronger, and response surfaces less smooth than in the present free-flowing granular systems. Active learning-based adaptive sampling strategies—selecting the next DEM run by maximizing *SAR*-informed expected improvement—offer a principled path to reduce the required sample size below $n = 34$ while maintaining surrogate quality. Finally, the broader implication for explainable AI in engineering design is that SHAP-derived metrics such as *SAR* can serve as physics-informed pre-screening diagnostics, not merely as post hoc explanations: connecting interpretability outputs directly to model selection decisions represents a template applicable beyond DEM surrogate modeling to any computationally expensive simulation-driven design optimization workflow.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/make8080229/s1>: Video S1: DEM simulation of the optimal silo hopper configuration (outlet width $A = 115.4$ mm, hopper half-angle $B = 60.6^\circ$) demonstrating granular flow dynamics of crushed solar panel recycling particles at the predicted maximum discharge rate of 22.15 kg/s (hopper_optimal.mp4); Video S2: DEM simulation of the inclined belt conveyor system handling solar panel recycling particles, showing particle transport, flow pattern, and mass flow behavior under optimized operating conditions—fin height $H = 50.0$ mm, belt velocity $V = 0.627$ m/s (videobeltoptimize.mp4); Video S3: Supplementary DEM simulation of the belt conveyor system presenting an alternative operating scenario for solar panel recycling particle handling, illustrating the effect of fin geometry and velocity on discharge angle and conveyor throughput (videobeltoptimize1.mp4); Table S1: Complete belt conveyor DEM sampling design for all 34 experimental runs (12 factorial design points (2^2 full factorial with replication) + 22 Latin Hypercube Sampling points), including fin height H (mm), belt velocity V (m/s), measured mass flow rate (kg/s), and discharge angle ($^\circ$) for each run (design_34runs_belt.csv); Table S2: Complete silo discharge raw ML predictions and residuals for all 34 DEM sampling points across all five surrogate models (RSM, ANN, RF, GBM, and GPR), including per-sample actual values, predicted values, residuals, and absolute errors used to construct performance benchmarks in Section 3.1 and Appendix Table A1 (hopper_raw_predictions.csv); Table S3: Complete five-fold cross-validation R^2 per-fold results for all five surrogate models (RSM, ANN, RF, GBM, GPR) across all three response variables (Silo Discharge MFR, Belt Conveyor MFR, Belt Conveyor Discharge Angle), providing the full statistical basis for the Friedman test and box-plot comparisons reported in Section 3.5 (statistical_fold_results.csv); Table S4: SHAP feature importance summary for all three response variables including mean absolute SHAP values and relative importance percentages for both design parameters (outlet width A and hopper angle B for silo; fin height H and belt velocity V for belt conveyor), underpinning the cross-system SHAP analysis reported in Section 3.4 (SHAP_importance_summary.csv); Algorithm S1: Complete Python implementation of the silo discharge ML-based hopper design optimization system, including data loading and preprocessing, surrogate model training (RSM, ANN, RF, GBM, GPR), leakage-free grouped five-fold cross-validation, Differential Evolution optimization, SHAP explainability analysis, and automated result export (hopper_ml_optimizer.py); Algorithm S2: Complete Python implementation of the belt conveyor multi-output ML optimization pipeline, including 2^2 full factorial and Latin Hypercube Sampling design expansion, 5-fold cross-validation for mass flow rate and discharge angle

responses, Differential Evolution optimization, SHAP analysis, and sensitivity reporting (belt_ml_optimizerV3.py).

Author Contributions: Conceptualization, S.S.; methodology, S.S.; software, S.S.; validation, S.S. and S.L.; formal analysis, S.S.; investigation, S.S. and S.L.; resources, S.S.; data curation, S.S. and S.L.; writing—original draft preparation, S.S.; writing—review and editing, S.S.; visualization, S.S. and S.L.; supervision, S.S.; project administration, S.S.; funding acquisition, S.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to commercial sensitivity of the industrial solar panel recycling process parameters.

Acknowledgments: The authors acknowledge the Department of Materials Handling and Logistics Engineering, Faculty of Engineering, King Mongkut's University of Technology North Bangkok, for computational resources and experimental equipment. We thank the Department of Primary Industries and Mines, Ministry of Industry, Thailand, for materials testing support and the commercial recycling facility for providing end-of-life photovoltaic panel materials.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

Symbol	Definition
F_n, F_n^d	Normal contact force; normal damping force
E^*, R^*, m^*, G^*	Equivalent (pairwise reduced) Young's modulus, radius, mass, shear modulus
$G_i = \frac{E_i}{[2(1 + \nu_i)]}$	Shear modulus of body i
δ_n, δ_t	Normal/tangential overlap
β, S_n, S_t	Damping ratio; normal/tangential stiffness
e, μ_s, μ_r	Coefficient of restitution; static/rolling friction coefficients
$\tau_i, \hat{\omega}_i$	Rolling-resistance torque; unit angular-velocity vector
$i \in \{1, 2\}$	The two bodies in a pairwise contact (particle or wall)
Q, C, ρ_b, D, d, k	Beverloo discharge rate, discharge coeff., bulk density, orifice width, particle size, shape factor
$\beta_0, \beta_i, \beta_{ii}, \beta_{ij}$	RSM intercept, linear, quadratic, interaction coefficients
$W^{(i)}, b^{(i)}, f(\cdot)$	ANN weight matrix, bias vector, activation of layer i
$\hat{y}, B, f_b(x)$	RF ensemble prediction, number of trees, b-th regression tree
F_m, η, h_m	GBM boosted model at iteration m, learning rate, m-th weak learner
$\mu(x^*), \sigma^2(x^*)$	GPR posterior predictive mean/variance
$k(x^*, X), K, \sigma_n^2, I$	GPR kernel vector, Gram matrix, noise variance, identity
$R^2, RMSE, MAE$	Coefficient of determination, root-mean-square error, mean absolute error
$y_i, \hat{y}_i, \bar{y}, n$	Actual, predicted, mean response; number of samples
$v_{i,G}, F, CR$	DE mutant vector, mutation factor, crossover rate
$u_{i,G,j}, j_{rand}$	DE trial-vector component, guaranteed-mutant index
w_1, w_2, J	Multi-objective weights on MFR and DA; weighted objective
ANN	Artificial Neural Network
FF	2 ² Full Factorial Design (Stage 1 Screening)
c-Si	Crystalline Silicon
COR	Coefficient of Restitution
DA	Discharge Angle
DE	Differential Evolution
DEM	Discrete Element Method
GBM	Gradient Boosting Machine
GPR	Gaussian Process Regression
LHS	Latin Hypercube Sampling

MAE	Mean Absolute Error
MFR	Mass Flow Rate
ML	Machine Learning
PV	Photovoltaic
RF	Random Forest
SAR	SHAP Asymmetry Ratio
RSM	Response Surface Methodology

Appendix A. Complete ML Prediction Datasets

This appendix provides complete machine-learning model predictions for all sampling points in both equipment systems, enabling full reproducibility. Run numbers in Tables A1–A3 follow the complete 34-point DEM simulation campaign order and do not separate factorial from LHS points; the 12 factorial design corner points (outlet width A: 42 mm or 111 mm; hopper half-angle B: 30° or 60° for the silo; fin height H: 20 mm or 50 mm; belt velocity V: 0.109 m/s or 0.627 m/s for the belt conveyor) are identified in Tables 3 and 4 in the main text.

Table A1. Complete silo discharge ML predictions for all 34 sampling points (corrected from hopper_raw_predictions.csv). All five models shown, including GPR.

Run	A (mm)	B (°)	Actual	RSM	ANN	RF	GBM	GPR
1	108.4	28.8	8.48	7.62	7.89	8.56	8.47	8.48
2	51.5	54.0	6.84	6.82	6.35	6.37	6.8	6.84
3	111.0	30.0	7.93	8.12	8.41	8.65	7.95	7.93
4	47.4	60.6	3.8	4.93	4.73	4.18	3.8	3.8
5	73.7	55.9	13.14	12.5	13.15	14.04	13.16	13.14
6	47.0	53.1	4.84	5.66	5.39	5.11	4.84	4.84
7	85.4	48.0	13.95	13.49	14.07	14.01	13.95	13.95
8	111.0	60.0	20.86	21.15	20.79	20.49	20.86	20.86
9	42.0	60.0	3.36	3.21	3.3	3.56	3.36	3.36
10	78.5	37.2	10.5	10.27	9.82	10.54	10.5	10.5
11	52.3	51.6	5.9	7.18	6.77	6.37	5.94	5.9
12	44.6	46.8	5.97	5.6	5.75	5.84	5.94	5.97
13	39.7	48.4	4.38	4.23	4.58	4.91	4.37	4.38
14	38.4	58.8	2.42	2.19	2.61	3.09	2.43	2.42
15	42.0	60.0	3.36	3.21	3.3	3.56	3.36	3.36
16	59.5	39.4	8.52	8.45	8.84	8.78	8.54	8.52
17	47.5	58.1	5.11	5.27	4.9	4.42	5.12	5.11
18	42.0	45.1	5.68	5.13	5.48	5.65	5.72	5.68
19	96.2	39.1	11.6	12.03	11.28	11.29	11.58	11.6
20	42.7	33.2	5.62	5.95	6.12	5.75	5.62	5.62
21	63.6	44.1	9.62	9.44	9.62	9.32	9.59	9.62
22	115.4	53.1	18.97	19.11	18.6	17.4	18.96	18.97
23	43.4	47.0	5.84	5.28	5.49	5.74	5.84	5.84
24	41.3	33.8	6.06	5.73	6.02	5.87	6.04	6.06
25	114.1	33.3	9.54	9.69	9.74	9.74	9.54	9.54
26	102.7	48.6	15.56	15.92	15.89	15.01	15.56	15.56
27	53.4	33.7	8.28	7.3	7.18	7.89	8.28	8.28
28	112.9	60.2	21.77	21.58	21.06	21.09	21.77	21.77
29	74.8	43.3	10.46	11.05	11.23	10.59	10.46	10.46
30	82.6	58.8	16.08	14.96	15.97	16.33	16.08	16.08
31	50.5	30.2	5.51	6.81	6.18	5.87	5.52	5.51
32	82.4	44.2	11.37	12.23	12.4	11.37	11.37	11.37

33	55.0	46.7	8.55	7.92	7.93	8.24	8.54	8.55
34	42.0	60.0	3.36	3.21	3.3	3.56	3.36	3.36

Rows 1–12: 2² factorial design with replication; Rows 13–34: Latin Hypercube Sampling. All values in kg/s. Best model: ANN ($R^2 = 0.964$, $RMSE = 0.91$ kg/s).

Table A2. Complete belt conveyor mass flow rate ML predictions for all 34 sampling points.

Run	Fin (mm)	Vel (m/s)	Actual	RSM	ANN	RF	GBM	GPR
1	50.0	0.627	15.55	15.64	11.53	15.01	15.55	15.55
2	20.0	0.109	1.10	0.46	0.47	1.18	1.08	1.10
3	20.0	0.109	1.10	0.46	0.47	1.18	1.08	1.10
4	20.0	0.627	3.83	3.88	4.03	4.63	3.84	3.83
5	20.0	0.627	3.83	3.87	3.67	4.58	3.84	3.83
6	50.0	0.627	15.55	15.52	11.66	14.59	15.55	15.55
7	20.0	0.109	1.10	0.74	0.97	1.15	1.10	1.10
8	50.0	0.109	1.16	1.31	1.08	2.15	1.16	1.16
9	20.0	0.627	3.83	3.98	3.86	4.44	3.83	3.83
10	50.0	0.109	1.16	1.21	1.11	2.00	1.17	1.16
11	50.0	0.627	15.55	15.68	15.26	14.56	15.55	15.55
12	50.0	0.109	1.16	1.12	0.95	2.76	1.17	1.16
13	28.5	0.593	6.97	6.48	6.15	5.98	4.97	6.44
14	43.4	0.116	1.59	0.89	1.03	1.68	1.28	1.05
15	24.0	0.463	3.92	4.02	3.93	4.90	3.50	3.51
16	31.2	0.326	4.78	4.42	4.28	3.91	3.91	3.91
17	38.9	0.240	3.48	3.92	4.41	4.38	5.27	3.86
18	26.3	0.134	1.17	2.00	1.69	1.64	1.39	0.69
19	35.5	0.207	2.90	3.53	2.88	3.67	3.48	2.90
20	33.0	0.622	8.10	8.78	8.16	7.17	7.89	8.30
21	42.4	0.414	7.44	7.46	8.26	6.93	6.60	7.94
22	44.8	0.306	5.46	5.82	6.20	4.97	4.51	5.40
23	39.4	0.501	9.55	7.82	9.07	7.32	8.33	8.10
24	36.4	0.159	2.04	2.62	2.00	1.52	1.56	2.07
25	21.7	0.363	2.53	3.01	2.89	3.76	2.88	3.02
26	24.8	0.202	1.44	2.93	2.25	3.06	3.06	1.95
27	41.6	0.376	6.61	6.59	7.67	7.68	7.44	6.22
28	29.9	0.416	4.76	5.05	5.35	6.68	5.35	4.82
29	27.7	0.524	5.71	5.50	5.36	7.17	6.82	5.61
30	20.7	0.293	2.53	2.49	2.21	3.45	3.20	2.02
31	47.1	0.545	12.41	11.69	10.71	10.65	13.60	13.18
32	34.7	0.564	7.97	7.67	8.37	7.64	8.12	7.72
33	48.0	0.443	10.25	9.49	8.97	9.15	10.08	9.26
34	49.0	0.267	5.37	5.85	5.05	4.61	4.12	5.05

Rows 1–12: 2² factorial design with replication; Rows 13–34: Latin Hypercube Sampling. All values in kg/s. Best model: GPR ($R^2 = 0.976 \pm 0.023$, $RMSE = 0.43$ kg/s).

Table A3. Complete belt conveyor discharge angle ML predictions for all 34 sampling points.

Run	Fin (mm)	Vel (m/s)	Actual (°)	RSM (°)	ANN (°)	RF (°)	GBM (°)	GPR (°)
1	50.0	0.627	61.10	61.09	61.10	61.26	61.10	61.10
2	20.0	0.109	69.30	69.26	69.13	69.03	69.29	69.30
3	20.0	0.109	69.30	69.26	69.13	69.03	69.29	69.30

4	20.0	0.627	64.00	64.00	63.83	63.81	64.00	64.00
5	20.0	0.627	64.00	64.01	63.83	63.79	64.00	64.00
6	50.0	0.627	61.10	61.12	61.08	61.40	61.10	61.10
7	20.0	0.109	69.30	69.29	69.55	69.17	69.30	69.30
8	50.0	0.109	68.40	68.39	68.21	68.46	68.40	68.40
9	20.0	0.627	64.00	64.00	63.69	63.82	64.00	64.00
10	50.0	0.109	68.40	68.39	68.19	68.46	68.40	68.40
11	50.0	0.627	61.10	61.09	61.14	61.51	61.10	61.10
12	50.0	0.109	68.40	68.36	68.12	68.42	68.40	68.40
13	28.5	0.593	63.55	63.56	63.65	63.58	63.98	63.45
14	43.4	0.116	68.60	68.46	68.20	68.42	68.42	68.69
15	24.0	0.463	65.37	65.38	65.31	63.93	65.12	65.45
16	31.2	0.326	66.48	66.39	66.56	66.99	66.78	66.38
17	38.9	0.240	67.11	67.06	67.22	66.75	66.59	66.93
18	26.3	0.134	68.80	68.81	68.77	68.67	68.52	68.91
19	35.5	0.207	67.51	67.64	68.02	67.08	67.11	67.67
20	33.0	0.622	62.77	62.80	63.04	63.53	63.36	62.89
21	42.4	0.414	64.65	64.62	64.52	65.04	65.10	64.61
22	44.8	0.306	65.82	65.89	65.88	66.65	66.47	65.99
23	39.4	0.501	63.71	63.73	63.77	63.89	63.44	63.69
24	36.4	0.159	68.10	68.21	68.41	68.64	68.30	68.26
25	21.7	0.363	66.60	66.56	66.49	65.43	65.61	66.54
26	24.8	0.202	68.17	68.14	68.62	67.07	67.33	68.13
27	41.6	0.376	65.17	65.16	65.08	64.63	64.66	65.24
28	29.9	0.416	65.59	65.44	65.43	64.54	65.20	65.45
29	27.7	0.524	64.34	64.41	64.52	63.34	64.11	64.48
30	20.7	0.293	67.30	67.36	67.91	66.29	66.63	67.42
31	47.1	0.545	62.38	62.52	62.38	62.78	62.38	62.36
32	34.7	0.564	63.35	63.35	63.71	63.15	62.94	63.30
33	48.0	0.443	63.85	63.82	63.80	63.70	62.68	63.81
34	49.0	0.267	66.21	66.21	66.35	67.00	66.91	66.12

Rows 1–12: 2^2 factorial design with replication; Rows 13–34: Latin Hypercube Sampling. Best model: RSM ($R^2 = 0.999 \pm 0.001$, $RMSE = 0.057^\circ$).

References

1. International Energy Agency. *Renewables 2024: Analysis and Forecasts to 2030*; IEA Publications: Paris, France, 2024. Available online: <https://www.iea.org/reports/renewables-2024> (accessed on 28 April 2026).
2. International Renewable Energy Agency; IEA-PVPS. *End-of-Life Management: Solar Photovoltaic Panels*; IRENA: Abu Dhabi, United Arab Emirates, 2016.
3. Xu, Y.; Li, J.; Tan, Q.; Peters, A.L.; Yang, C. Global status of recycling waste solar panels: A review. *Waste Manag.* **2018**, *75*, 450–458. <https://doi.org/10.1016/j.wasman.2018.01.036>.
4. Deng, R.; Zhuo, H.; Shen, Y. Recent progress in silicon photovoltaic module recycling processes. *Resour. Conserv. Recycl.* **2022**, *187*, 106612. <https://doi.org/10.1016/j.resconrec.2022.106612>.
5. Heath, G.A.; Silverman, T.J.; Kempe, M.; Deceglie, M.; Ravikumar, D.; Remo, T.; Cui, H.; Sinha, P.; Libby, C.; Shaw, S.; et al. Research and development priorities for silicon photovoltaic module recycling to support a circular economy. *Nat. Energy* **2020**, *5*, 502–510. <https://doi.org/10.1038/s41560-020-0645-2>.
6. Schulze, D. *Powders and Bulk Solids: Behavior, Characterization, Storage and Flow*, 2nd ed.; Springer: Berlin, Germany, 2021. <https://doi.org/10.1007/978-3-030-76720-4>.
7. Ilic, D.; Katterfeld, A. Simulation of Transfer Chutes. In *Simulations in Bulk Solids Handling*; McGlinchey, D., Ed.; Wiley-VCH: Weinheim, Germany, 2023; pp. 41–78. <https://doi.org/10.1002/9783527835935.ch2>.
8. Conveyor Equipment Manufacturers Association. *Belt Conveyors for Bulk Materials*, 7th ed.; CEMA: Naples, FL, USA, 2014.

9. Zhang, S.; Yang, G.; Zhang, X. Effect of boundaries on flow behavior of granular material in a hopper. *Powder Technol.* **2025**, *466*, 121505. <https://doi.org/10.1016/j.powtec.2025.121505>.
10. Hanif, M.A.; van der Meer, D.; Maza, D. Discharge of rice-shaped particles from a monolayer flat-bottom silo. *Phys. Rev. E* **2024**, *109*, 064906. <https://doi.org/10.1103/PhysRevE.109.064906>.
11. Hanif, M.A.; Maza, D.; van der Meer, D. Effect of particle shape on the discharge from a mono-layer hopper. *Granul. Matter* **2025**, *27*, 38. <https://doi.org/10.1007/s10035-025-01517-8>.
12. Ketterhagen, W.; Wassgren, C. A perspective on calibration and application of DEM models for simulation of industrial bulk powder processes. *Powder Technol.* **2022**, *402*, 117301. <https://doi.org/10.1016/j.powtec.2022.117301>.
13. Zhao, Y.; Chew, J.W. Discrete element method study on hopper discharge behaviors of binary mixtures of nonspherical particles. *AIChE J.* **2020**, *66*, e16254. <https://doi.org/10.1002/aic.16254>.
14. Coetzee, C.; Katterfeld, A. Calibration of DEM Parameters. In *Simulations in Bulk Solids Handling*; McGlinchey, D., Ed.; Wiley-VCH: Weinheim, Germany, 2023; pp. 1–40. <https://doi.org/10.1002/9783527835935.ch1>.
15. Hadi, A.; Moradi, M.; Pang, Y.; Schott, D. Adaptive AI-based surrogate modelling via transfer learning for DEM simulation of multi-component segregation. *Sci. Rep.* **2024**, *14*, 27003. <https://doi.org/10.1038/s41598-024-78455-7>.
16. Fransen, M.; Fürst, A.; Tunuguntla, D.; Wilke, D.N.; Alkin, B.; Barreto, D.; Brandstetter, J.; Cabrera, M.A.; Fan, X.; Guo, M.; et al. Towards Scientific Machine Learning for Granular Material Simulations: Challenges and Opportunities. *Arch. Comput. Methods Eng.* **2026**, *33*, 789–821. <https://doi.org/10.1007/s11831-025-10322-8>.
17. Coetzee, C.J.; Scheffler, O.C. Review: Calibration of the discrete element method for the bulk modelling of cohesive materials. *Processes* **2023**, *11*, 5. <https://doi.org/10.3390/pr11010005>.
18. Rossow, J.; Coetzee, C.J. Discrete element modelling of a chevron patterned conveyor belt and a transfer chute. *Powder Technol.* **2021**, *391*, 77–96. <https://doi.org/10.1016/j.powtec.2021.06.012>.
19. Ilic, D.; Roberts, A.; Wheeler, C.; Katterfeld, A. Modelling bulk solid flow interactions in transfer chutes: Shearing flow. *Powder Technol.* **2019**, *354*, 30–44. <https://doi.org/10.1016/j.powtec.2019.05.058>.
20. Ilic, D.; Lavrinec, A.; Orozovic, O. Simulation and analysis of blending in a conveyor transfer system. *Miner. Eng.* **2020**, *157*, 106575. <https://doi.org/10.1016/j.mineng.2020.106575>.
21. Scheffler, O.C.; Coetzee, C.J. DEM calibration for simulating bulk cohesive materials. *Comput. Geotech.* **2023**, *161*, 105476. <https://doi.org/10.1016/j.compgeo.2023.105476>.
22. Coetzee, C.J. Review: Calibration of the discrete element method. *Powder Technol.* **2017**, *310*, 104–142. <https://doi.org/10.1016/j.powtec.2017.01.015>.
23. Rackl, M.; Hanley, K.J. A methodical calibration procedure for discrete element models. *Powder Technol.* **2017**, *307*, 73–83. <https://doi.org/10.1016/j.powtec.2016.11.048>.
24. Rozbroj, J.; Zegzulka, J.; Nečas, J. Use of DEM in the Determination of Friction Parameters on a Physical Comparative Model of a Vertical Screw Conveyor. *Chem. Biochem. Eng.* **2015**, *29*, 25–34. <https://doi.org/10.15255/CABEQ.2014.2142>.
25. Coetzee, C. Calibration of the discrete element method: Strategies for spherical and non-spherical particles. *Powder Technol.* **2020**, *364*, 851–878. <https://doi.org/10.1016/j.powtec.2020.01.076>.
26. Arunyanart, P.; Sudsawat, S. Rapid DEM calibration of heterogeneous solar panel recycling materials via Plackett-Burman screening and machine learning surrogates: Multi-scale validation from laboratory to industrial hopper. *Powder Technol.* **2026**, *482*, 122770. <https://doi.org/10.1016/j.powtec.2026.122770>.
27. Borisut, P.; Nuchitprasittichai, A. Adaptive Latin Hypercube Sampling for a Surrogate-Based Optimization with Artificial Neural Network. *Processes* **2023**, *11*, 3232. <https://doi.org/10.3390/pr11113232>.
28. Myers, R.H.; Montgomery, D.C.; Anderson-Cook, C.M. *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*, 4th ed.; Wiley: Hoboken, NJ, USA, 2016.
29. Montgomery, D.C. *Design and Analysis of Experiments*, 9th ed.; Wiley: Hoboken, NJ, USA, 2020.
30. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. In Proceedings of the 3rd International Conference on Learning Representations (ICLR 2015), San Diego, CA, USA, 7–9 May 2015. <https://doi.org/10.48550/arXiv.1412.6980>.
31. Probst, P.; Wright, M.N.; Boulesteix, A.-L. Hyperparameters and tuning strategies for random forest. *WIREs Data Min. Knowl. Discov.* **2019**, *9*, e1301. <https://doi.org/10.1002/widm.1301>.
32. Chen, T.; Guestrin, C. XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794. <https://doi.org/10.1145/2939672.2939785>.

33. Bentéjac, C.; Csörgő, A.; Martínez-Muñoz, G. A comparative analysis of gradient boosting algorithms. *Artif. Intell. Rev.* **2021**, *54*, 1937–1967. <https://doi.org/10.1007/s10462-020-09896-5>.
34. Marrel, A.; Iooss, B. Probabilistic surrogate modeling by Gaussian process: A review on recent insights in estimation and validation. *Reliab. Eng. Syst. Saf.* **2024**, *247*, 110094. <https://doi.org/10.1016/j.ress.2024.110094>.
35. Pang, Y.; Yang, L.; Wang, Y.; Sun, W.; Song, X. A Latin hypervolume design for irregular sampling spaces and its application in the analysis of cracks. *Eng. Comput.* **2023**, *39*, 3509–3526. <https://doi.org/10.1007/s00366-022-01769-5>.
36. Schneider, F.; Hellmig, R.J.; Nelles, O. Latin hypercubes for constrained design of experiments for data-driven models. *at—Autom.* **2023**, *71*, 820–832. <https://doi.org/10.1515/auto-2023-0017>.
37. EDEM Software, version 2022.2; Altair Engineering Inc.: Troy, MI, USA, 2025. Available online: https://2022.help.altair.com/2022.2/EDEM/Simulator/#t=Introducing_EDEM.htm (accessed on 1 June 2026).
38. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
39. Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L.; et al. PyTorch: An imperative style, high-performance deep learning library. In Proceedings of the 33rd Conference on Neural Information Processing Systems (NeurIPS 2019), Vancouver, BC, Canada, 9 December 2019. <https://doi.org/10.48550/arXiv.1912.01703>.
40. Chakraborty, S.; Saha, A.K.; Ezugwu, A.E.; Agushaka, J.O.; Zitar, R.A.; Abualigah, L. Differential evolution and its applications in image processing problems: A comprehensive review. *Arch. Comput. Methods Eng.* **2023**, *30*, 985–1040. <https://doi.org/10.1007/s11831-022-09825-5>.
41. Virtanen, P.; Gommers, R.; Oliphant, T.E.; Haberland, M.; Reddy, T.; Cournapeau, D.; van Mulbregt, P. SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nat. Methods* **2020**, *17*, 261–272. <https://doi.org/10.1038/s41592-019-0686-2>.
42. Lundberg, S.M.; Lee, S.-I. A unified approach to interpreting model predictions. In Proceedings of the 31st International Conference on Neural Information Processing Systems (NeurIPS 2017), Long Beach, CA, USA, 4–9 December 2017; pp. 4768–4777. <https://doi.org/10.48550/arXiv.1705.07874>.
43. Holzinger, A.; Saranti, A.; Molnar, C.; Biecek, P.; Samek, W. Explainable AI Methods—A Brief Overview. In *xxAI—Beyond Explainable AI*; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2022; Volume 13200, pp. 13–38. https://doi.org/10.1007/978-3-031-04083-2_2.
44. Ribeiro, M.T.; Singh, S.; Guestrin, C. ‘Why Should I Trust You?’: Explaining the Predictions of Any Classifier. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 1135–1144. <https://doi.org/10.1145/2939672.2939778>.
45. Molnar, C. *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*, 2nd ed.; Leanpub: Victoria, BC, Canada, 2022. Available online: <https://christophm.github.io/interpretable-ml-book/> (accessed on 28 April 2026).
46. Friedman, M. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *J. Am. Stat. Assoc.* **1937**, *32*, 675–701. <https://doi.org/10.1080/01621459.1937.10503522>.
47. Irazábal, J.; Salazar, F.; Vicente, D.J. A methodology for calibrating parameters in discrete element models based on machine learning surrogates. *Comput. Part. Mech.* **2023**, *10*, 1031–1047. <https://doi.org/10.1007/s40571-022-00550-1>.
48. Roessler, T.; Richter, C.; Katterfeld, A.; Will, F. Development of a standard calibration procedure for the DEM parameters of cohesionless bulk materials—Part I: Solving the problem of ambiguous parameter combinations. *Powder Technol.* **2019**, *343*, 803–812. <https://doi.org/10.1016/j.powtec.2018.11.034>.
49. Tsokanas, N.; Pastorino, R.; Stojadinović, B. A comparison of surrogate modeling techniques for global sensitivity analysis in hybrid simulation. *Mach. Learn. Knowl. Extr.* **2022**, *4*, 1–21. <https://doi.org/10.3390/make4010001>.
50. Biecek, P.; Baniecki, H.; Krzyżiński, M.; Cook, D. Performance is not enough: The story told by a Rashomon quartet. *J. Comput. Graph. Stat.* **2024**, *33*, 1118–1121. <https://doi.org/10.1080/10618600.2024.2344616>.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.