PHYSICS-INFORMED MACHINE LEARNING FOR MULTI-OBJECTIVE OPTIMIZATION IN ADDITIVE MANUFACTURING: A DATA-EFFICIENT APPROACH

THANH-CONG TRUONG[1,*], HUYNH NGOC THANH TRUNG[1], THANH-THAO MAI[1], THANH-TAM MAI[1], VINH TRUONG HOANG[2], QUOC-PHU MA[3]

[1]Faculty of Data Science, University of Finance - Marketing, No. 778 Nguyen Kiem, Ho Chi Minh, 70000, Viet Nam

[2]Faculty of Information Technology, Ho Chi Minh City Open University, No. 97 Vo Van Tan Street, Ho Chi Minh, 70000, Viet Nam

[3]Department of Machining, Assembly and Engineering Metrology, Faculty of Mechanical Engineering, VSB-Technical University of Ostrava, 70833 Ostrava, Czech Republic

DOI: 10.17973/MMSJ.2025_10_2025094

e-mail to corresponding author: ttcong@ufm.edu.vn

Abstract

Additive manufacturing (AM) quality control relies on empirical approaches due to complex process-property relationships. While machine learning (ML) offers promising solutions, most approaches treat parameters independently without leveraging thermomechanical principles governing the properties of the printed materials. This is essential for understanding the behaviour of fused deposition modelling (FDM) printing. This study investigates whether integrating elementary thermomechanical knowledge into feature engineering improves mechanical property prediction for polylactic acid components under data-constrained conditions. Using 50 experimental samples from controlled printing conditions, three feature engineering strategies were systematically compared: raw process parameters, physics-informed features based on heat transfer and material flow principles, and polynomial interactions across five ML algorithms. Physics-informed features consistently outperformed baseline approaches, with Huber Regressor achieving coefficient of determination equal to 0.817 (51.3% improvement over raw parameters). Feature importance analysis using SHapley Additive exPlanations identified layer height and nozzle temperature as primary predictors, with engineered thermal diffusion and density features contributing significantly to model performance. This study demonstrates the potential of physics-informed feature engineering for improving prediction accuracy in data-constrained AM scenarios, providing methodological insights for thermomechanical integration and actionable guidance for industrial artificial intelligence (AI) implementation.

KEYWORDS

Additive manufacturing, Fused deposition modelling, Artificial intelligence, Machine Learning, Feature Engineering

1 INTRODUCTION

The integration of artificial intelligence (AI) applications in additive manufacturing has undergone substantial evolution, progressing from basic process monitoring to sophisticated predictive modeling, with recent developments in generative AI marking a significant advancement in the field [Truong 2025]. In particular, machine learning emerging as a transformative technology for addressing manufacturing challenges [Ng 2024]. The integration of ML techniques across design, process optimization, and quality control has demonstrated significant potential for advancing AM capabilities within Industry 4.0 frameworks [Trovato 2025]. Nevertheless, contemporary ML approaches for mechanical property prediction in fused deposition modeling (FDM) predominantly rely on raw process parameters - layer height, nozzle temperature, print speed - without leveraging the underlying thermomechanical relationships that fundamentally govern part quality and performance.

Current ML frameworks for AM suffer from a critical limitation: they treat process parameters as independent statistical variables rather than recognizing their interdependent effects governed by heat transfer, material flow, and crystallization physics [Faegh 2025]. This physics-agnostic approach constrains both model interpretability and generalization capabilities, particularly problematic in manufacturing environments where process understanding drives optimization strategies.

In view of this, physics-informed ML represents an emerging paradigm that addresses these limitations by integrating domain knowledge directly into model architectures and feature engineering strategies [Karniadakis 2021]. Recent advances in process-property modeling demonstrate substantial improvements when physical principles constrain feature spaces, enabling more accurate predictions with reduced data requirements [Faegh 2025]. Advanced techniques including reinforcement learning for process optimization and Bayesian learning for thermal prediction have shown promise in specific applications [Zhu 2024]. However, the systematic application of thermomechanical principles to feature engineering for FDM property prediction remains underexplored, particularly for small dataset scenarios common in industrial practice.

This study investigates whether integrating elementary thermomechanical principles into ML feature spaces can improve prediction accuracy for mechanical properties of FDM-printed polylactic acid (PLA) components under data-constrained conditions. Using a controlled experimental dataset of 50 specimens, three feature engineering strategies are systematically compared: conventional raw (standalone) process parameters, physics-informed features derived from heat transfer and material flow analysis, and polynomial interaction terms across five ML algorithms. The research addresses the practical challenge of deploying ML systems in manufacturing environments where training data is limited, providing both methodological insights for physics-informed feature engineering and actionable guidance for industrial implementation of AI-driven process optimization in AM.

2 METHODOLOGY

2.1 Dataset

The dataset employed in this study was collected from experimental samples manufactured using an Ultimaker S5 3D printer at the Additive Manufacturing Laboratory, TR/Selçuk University, as initially described by Okudan et al. [Okudan 2018]. The dataset was curated to investigate the influence of 3D printing process parameters on the resulting mechanical properties of printed PLA and ABS samples, making it wellaligned with the objectives of this study in evaluating the effectiveness of feature engineering for explainable models.

A total of 50 specimens, each with standardized dimensions of 170 mm × 20 mm × 4 mm, were printed using PLA and ABS filaments and tested under ISO 527-2:2012 guidelines. Each sample is described by nine input features: *(x_1) layer height (mm), (x_2) wall thickness (mm), (x_3) infill density (%), (x_4) infill pattern (grid, triangles, cubic), (x_5) nozzle temperature (°C), (x_6) bed temperature (°C), (x_7) print speed (mm/s), (x_8) material type (PLA/ABS), and (x_9) fan speed (%).*

The target outputs include: *(y_1) surface roughness (Ra, μm)* – measured using a MITUTOYO SJ-210 profilometer, *(y_2) tensile strength (MPa), and (y_3) elongation at break (%)* – obtained using a universal tensile testing machine.

A detailed breakdown of all experimental parameters, including their symbols, ranges, and associated physical significance in the printing process, is provided in Tab. 1. These annotations not only support feature selection but also reflect underlying thermal, geometric, and material dynamics essential for downstream explainability.

Table 1. Experimental Parameters and Their Physical Significance in AM

Parameter	Symbol	Unit	Physical Significance
Process Parameters			
Layer height	h_{layer}	mm	Controls interlayer bonding and thermal mass
Wall thickness	t_{wall}	-	Affects structural rigidity and wall strength
Infill density	ρ_{infill}	%	Determines material volume and internal strength
Infill pattern	-	-	Alters stress distribution and structural damping
Nozzle temperature	T_{nozzle}	°C	Governs viscosity and extrusion quality
Bed temperature	T_{bed}	°C	Affects adhesion and warping resistance
Print speed	v_{print}	mm/s	Controls thermal and geometric effects
Material type	-	-	Defines base mechanical and thermal behavior
Fan speed	v_{fan}	%	Controls material cooling
Target Properties			
Surface roughness	R_a	μm	Surface quality
Tensile strength	σ_{UTS}	MPa	Ultimate tensile strength
Elongation at break	ϵ_{break}	%	Ductility

Given the limited sample size (n=50), a stratified random sampling approach was employed for the train-test split (80%-20%), ensuring representative distribution of categorical variables (material type and infill pattern) across both subsets. This resulted in 40 samples for training and 10 samples for testing. The small dataset size necessitated careful consideration of model complexity to avoid overfitting, which influenced the choice of cross-validation strategy and tuning of the model hyperparameters.

All samples were sliced using customized G-code configurations, varying the input parameters across a defined range to ensure design space diversity (e.g., infill density: 10–100%, layer height: 0.02–0.2 mm). Preprocessing included min–max normalization for numerical inputs and one-hot encoding for categorical features.

The dataset's controlled experimental conditions and comprehensive parameter coverage make it particularly valuable for examining the impact of feature engineering on prediction accuracy in AM applications. However, the limited sample size represents a constraint that requires robust validation techniques to ensure reliable conclusions.

2.2 Feature Engineering Strategy

2.2.1 Thermomechanical Foundations

Feature engineering represents a critical component in AM prediction models, as raw process parameters may not fully capture the underlying thermomechanical phenomena governing mechanical properties. The FDM process involves complex relationships between thermal management, material flow that require systematic encoding through physics-informed transformations. Seven engineered features are systematically derived from established physical relationships to capture these underlying mechanisms that raw process parameters cannot represent independently.

Temperature Ratio - Dimensionless thermal driving force:

$$f_{temp\_ratio} = \frac{T_{nozzle}}{T_{bed}} \quad (1)$$

This dimensionless parameter represents the thermal driving force controlling interlayer bonding efficiency. Higher ratios indicate greater thermal energy available for polymer chain interdiffusion across layer boundaries, directly influencing mechanical properties through enhanced adhesion mechanisms [Jaganathan 2024].

Temperature Difference - Absolute thermal gradient:

$$f_{temp\_difference} = T_{nozzle} - T_{bed} \quad (2)$$

The absolute temperature differential serves as the driving force for heat transfer according to Newton's law of cooling: $q = h \cdot A \cdot \Delta T$. This gradient determines cooling rates, thermal stress development, and crystallization kinetics in semi-crystalline polymers like PLA [Siddiqui 2024]].

Cooling Index - Normalized cooling effectiveness:

$$f_{cooling\_index} = \frac{v_{fan}}{T_{nozzle}} \quad (3)$$

This feature quantifies cooling effectiveness relative to initial thermal energy. Higher cooling rates promote rapid solidification, affecting crystal structure formation, stress relaxation patterns, and dimensional accuracy of printed components [Ranjbar 2021]].

Speed-Layer Ratio - Shear rate approximation:

$$f_{speed\_layer\_ratio} = \frac{v_{print}}{h_{layer}} \quad (4)$$

This parameter approximates the apparent shear rate in the nozzle based on $\dot{\gamma} = \frac{du}{dy}$ for flow in narrow channels. Shear rate directly influences polymer melt viscosity through non-Newtonian behavior $\eta = \eta(\dot{\gamma})$ affecting polymer chain orientation, flow instabilities, and resulting mechanical anisotropy [Ranjbar 2021]].

Layer-Wall Ratio - Geometric aspect ratio:

$$f_{layer\_wall\_ratio} = \frac{h_{layer}}{t_{wall}} \quad (5)$$

This geometric ratio captures the structural aspect relationship affecting stress concentration and load distribution mechanisms. The ratio influences the trade-off between printing resolution and structural integrity, with implications for failure modes and mechanical performance [Beșliu-Băncescu 2023].

Density Volume - Three-dimensional material distribution:

$$f_{density\_volume} = \rho_{infill} \times h_{layer} \times t_{wall} \quad (6)$$

This feature represents the effective material content per unit length of the printed path. It captures three-dimensional packing effects where voids between deposited filament layers significantly impact tensile strength through their interaction with geometric parameters. The feature directly correlates with load-bearing capacity and mechanical properties [Beșliu-Băncescu 2023].

Efficiency Index - Manufacturing productivity measure:

$$f_{efficiency\_index} = \frac{v_{print} \times \rho_{infill}}{h_{layer}} \quad (7)$$

This productivity index quantifies material deposition rate normalized by resolution, representing the manufacturing efficiency trade-off. Higher values indicate faster material deposition but may compromise surface quality, while lower values suggest higher resolution at reduced throughput. This feature captures rheological property effects on printability and overall part quality in extrusion-based systems [Gillispie 2023].

2.2.2 Feature Engineering Approaches

Building upon the thermomechanical foundations established in Section 2.2.1, three distinct feature engineering strategies were systematically implemented to evaluate the impact of physics-informed transformations on prediction accuracy. Each approach represents a different philosophy for encoding process-structure-property relationships in ML models for AM.

Baseline Features

The baseline feature set comprises the nine original process parameters extracted from the dataset without modification, serving as the control condition for evaluating feature engineering effectiveness. This approach represents current standard practice in AM modelling studies and provides the reference point for measuring improvement.

Domain-Specific Engineered Features

The physics-informed strategy transforms the nine raw parameters into seven engineered features using the thermomechanical relationships defined in Equations (1)-(7). This approach systematically encodes domain knowledge by creating features that directly represent physical phenomena governing FDM part quality [Karniadakis 2021]. Each engineered feature captures specific thermomechanical mechanisms:

- Thermal features (Equations 1-3) encode heat transfer and crystallization effects through temperature ratio, thermal gradient, and cooling effectiveness. Temperature-based parameters can significantly influence crystallinity and mechanical properties with 20-30% improvement potential [Jaganathan 2024].
- Flow features (Equations 4-5) represent polymer rheology and geometric constraints via shear rate approximation and aspect ratio relationships. Flow-related features such as speed-layer ratio represent shear rate proxies critical for non-Newtonian flow behavior of polymer materials, where shear rates in narrow printer nozzles are typically very high and directly affect melt viscosity and deformation behavior [Ranjbar 2021].
- Material features (Equations 6-7) capture three-dimensional packing effects and manufacturing efficiency through density-volume interactions. Structural features including density volume account for three-dimensional packing effects, as voids between deposited filament layers significantly impact tensile strength and are determined by the interaction of geometric parameters [Beșliu-Băncescu 2023]

Polynomial Interaction Features

To systematically capture pairwise parameter interactions without assuming specific functional forms, polynomial features were generated using degree-2 interaction terms (excluding squared terms). This automated approach complements domain-specific engineering by exploring potential synergistic effects between process parameters that may not be evident from physical intuition alone. The initial expansion from 9 base features produced 36 interaction terms, which were reduced to 25 features through variance thresholding (variance > 0.01) to maintain a reasonable feature-to-sample ratio (1:2) and prevent overfitting given the limited dataset size (n=50). This approach enables discovery of non-linear parameter coupling effects while maintaining computational tractability and model interpretability.

2.3 Model Selection and Configuration

The model selection strategy evaluates predictive performance across different algorithmic approaches. Given the dataset size (n=50) and complex thermomechanical relationships in AM, models were selected to represent linear, robust, and ensemble approaches while maintaining conservative complexity to prevent overfitting.

Linear Methods: Ordinary Least Squares (OLS) provides an interpretable baseline with direct coefficient interpretation for understanding process parameter effects in manufacturing contexts. Ridge Regression (α = 1.0) addresses potential multicollinearity between process parameters through L2 regularization.

Robust Regression: Huber Regressor (ε = 1.35) handles outliers commonly present in manufacturing data due to measurement errors or process variations. This approach maintains efficiency for normal data while providing robustness against extreme values.

Ensemble Methods: Random Forest (n_estimators=50, max_depth=5) captures non-linear interactions and threshold effects characteristic of thermomechanical behavior in 3D printing. XGBoost (n_estimators=50, max_depth=3) provides gradient boosting with built-in regularization capabilities.

Conservative hyperparameters were chosen to prevent overfitting given the small dataset size.

It should be noted that deep learning approaches were excluded due to insufficient training data. The MultiOutputRegressor wrapper enables simultaneous prediction of three correlated mechanical properties, reflecting their interdependent nature in AM.

2.4 Metric

The evaluation strategy herein addresses the challenges of limited data (n=50) through robust validation and comprehensive metric selection. Mean Absolute Error (MAE) is used as the primary metric for its direct engineering interpretability and outlier robustness, while Mean Squared Error (MSE) provides complementary assessment by penalizing larger prediction errors more heavily. This is critical for avoiding significant outliers in mechanical property predictions. The Coefficient of Determination (R^2) averaged across three mechanical properties offers normalized overall performance assessment independent of measurement units.

The validation approach combines stratified holdout testing (80/20 split) with 5-fold cross-validation using custom multi-output scoring. Stratification ensures representative distribution of categorical variables, while cross-validation provides robust performance estimates despite the small dataset. The statistical significance was assessed through bootstrap confidence intervals and paired testing across multiple random seeds to ensure reliable comparative conclusions between feature engineering approaches.

2.5 Integrated Methodology Framework

Fig. 1 presents the methodology framework integrating all experimental components from data input through final analysis. The flowchart illustrates the systematic evaluation of three feature engineering strategies across five ML algorithms, with consistent validation procedures and interpretability analysis through SHAP.

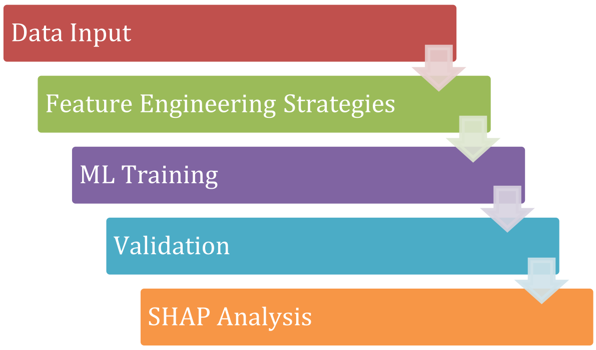

Figure 1. Methodology flowchart

This integrated approach enables direct comparison of physics-informed feature engineering against baseline and automated approaches, providing empirical validation of the thermomechanical principles encoded in Equations (1)-(7) while establishing practical guidelines for industrial AM applications.

3 RESULTS

3.1 Baseline Performance with Original Features

The baseline evaluation using original nine process parameters established reference performance levels across five ML algorithms. XGBoost demonstrated the strongest baseline performance (R^2 = 0.650, MAE = 8.41), effectively capturing non-linear relationships in the raw parameter space. Linear methods showed moderate effectiveness with Linear Regression (R^2 = 0.582) substantially outperforming Ridge Regression (R^2 = 0.463), while Huber Regressor exhibited balanced performance (R^2 = 0.540, MAE = 10.81) with robust outlier handling.

3.2 Feature Engineering Impact Assessment

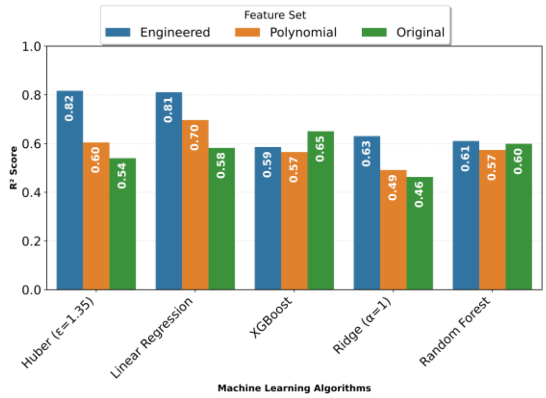

Figure 2. Coefficient of determination (R^2) performance comparison across five ML algorithms and three feature engineering strategies.

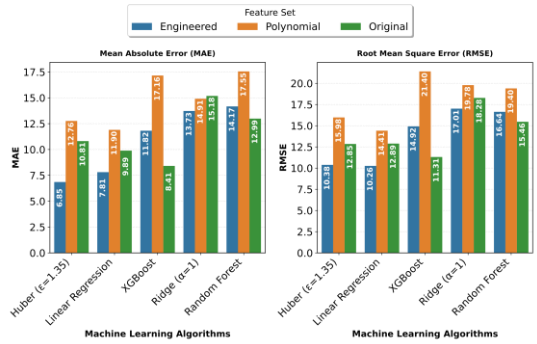

Figure 3. Error metric analysis showing MAE and Root and Mean Square Error RMSE across different algorithms and feature strategies

The implementation of thermomechanically-informed engineered features yielded substantial improvements across all algorithms. Huber Regressor achieved optimal overall performance with R^2 = 0.817, representing a 51.3% improvement over its baseline performance. This exceptional result (MAE = 6.85, RMSE = 10.38) establishes the Huber-Engineered combination as the best-performing configuration in this study.

Linear Regression demonstrated remarkable improvement with engineered features, achieving R^2 = 0.811 (39.3% increase from baseline), indicating that thermomechanical transformations effectively linearized the underlying process-property relationships. Even Ridge Regression achieved substantial improvement (R^2 = 0.631, 36.3% increase), suggesting that engineered features mitigated multicollinearity issues present in the expanded feature space.

The comprehensive performance analysis presented in Fig. 2 and Fig. 3 demonstrates the systematic superiority of physics-informed feature engineering across multiple algorithmic approaches. Fig. 2 shows that engineered features consistently achieve higher R^2 values, with Huber Regressor reaching optimal performance ($R^2 = 0.817$, representing 51.3% improvement over baseline). The corresponding error analysis in Fig. 3 confirms these findings, showing substantial reductions in both MAE and RMSE metrics for physics-informed approaches. Notably, the performance patterns reveal distinct algorithm-feature synergies. Linear methods (Linear and Ridge Regression) show exceptional improvement with engineered features, achieving R^2 values of 0.811 and 0.631 respectively, indicating successful linearization of thermomechanical relationships. This validates the hypothesis that physics-informed transformations create feature spaces more suitable for linear algorithms to capture complex process-property relationships that would otherwise require sophisticated non-linear methods. The consistent improvements across four of five algorithms provide strong empirical validation of the thermomechanical principles encoded in Equations (1)-(7), demonstrating both the robustness and practical applicability of this approach for industrial AM applications.

3.3 Comparative Performance Analysis

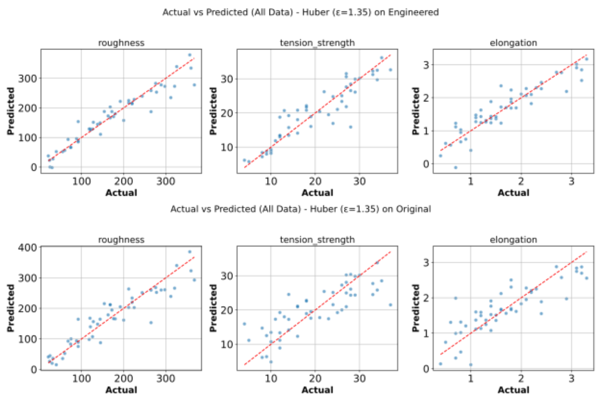

Figure 4. Feature engineering impact on Huber Regressor prediction accuracy across three mechanical properties: (top) engineered features, (bottom) original features.

Fig. 4 demonstrates the substantial improvement in Huber Regressor performance achieved through engineered features, comparing prediction accuracy across the three mechanical properties (surface roughness, tensile strength, and elongation at break). Engineered features achieved superior performance across all evaluation metrics, with the top three combinations all utilizing domain-specific features: Huber-Engineered ($R^2 = 0.817$), Linear-Engineered ($R^2 = 0.811$), and Ridge-Engineered ($R^2 = 0.631$).

The visual comparison in Fig. 4 clearly illustrates how engineered features enable the same algorithm (Huber Regressor) to achieve markedly better prediction accuracy across all three mechanical properties, with improved alignment along the ideal prediction line and reduced prediction scatter. A critical finding is the consistent underperformance of polynomial features despite their higher dimensionality (27 vs. 16 features).

3.4 Feature Importance Analysis

SHapley Additive exPlanations (SHAP) analysis was conducted on the optimal Huber-Engineered model to quantify individual feature contributions to prediction accuracy (Fig. 5).

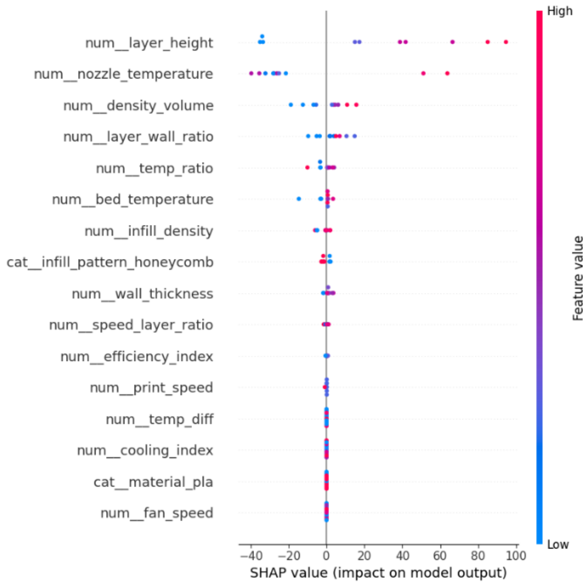

Figure 5. SHAP Feature Importance Analysis

Table 2 presents the top ten most influential features from the 16-feature engineered set, with layer height emerging as the most influential feature with a SHAP value of 46.18, followed by nozzle temperature at 35.10. The engineered density volume feature ranked third in importance (8.75), substantially outperforming its constituent original parameters when considered individually.

Table 2. Top 10 Feature Importance Ranking

Rank	Feature	SHAP Value	Type
1	layer_height	46.18	Original
2	nozzle_temperature	35.10	Original
3	density_volume	8.75	Engineered
4	layer_wall_ratio	6.32	Engineered
5	temperature_ratio	3.05	Engineered
6	bed_temperature	2.78	Original
7	infill_density	2.18	Original
8	infill_pattern_honeycomb	1.86	Original
9	wall_thickness	1.47	Original
10	speed_layer_ratio	0.69	Engineered

Among the top-ranking engineered features shown in Tab. 2, layer-wall ratio demonstrated notable importance (6.32), while thermal-based engineered features showed moderate contributions with temperature ratio achieving 3.05 and bed temperature 2.78. Original process parameters showed lower individual importance scores, with infill density (2.18), infill pattern honeycomb (1.86), and wall thickness (1.47) ranking in the middle tier.

The top 10 ranking in Tab. 2 reveals that speed-layer ratio (0.69) represents the threshold for meaningful feature contribution, with the remaining 6 engineered features showing progressively lower importance values. The SHAP value distribution demonstrates a clear hierarchy in feature relevance, with the top five features accounting for the majority of predictive influence in the optimal model configuration.

The dominance of layer height and nozzle temperature aligns with established thermomechanical principles in FDM. Layer height directly controls interlayer contact area and thermal mass, fundamentally governing mechanical properties. Nozzle temperature determines viscosity and flow characteristics of the polymer material, which are critical for layer adhesion. The remaining 18.81% contribution from engineered features represents meaningful enhancement by capturing interaction effects that individual parameters cannot express. This hierarchical importance validates both the physical understanding and engineered feature design.

3.5 Cross-Validation Performance Analysis

Tab. 3 presents the comprehensive performance comparison across all feature engineering strategies based on 5-fold cross-validation results averaged across three mechanical properties (surface roughness, tensile strength, and elongation at break):

Table 3. Statistical Validation of Feature Engineering Performance

Algorithm	Original R^2	Engineered R^2	Improvement
Huber	0.540	**0.817**	+51.3%
Linear Regression	0.582	0.811	+39.3%
Random Forest	0.599	0.610	+ 1.8%
Ridge	0.463	0.631	+36.3%
XGBoost	0.650	0.586	-9.8%

Note: Performance based on 5-fold cross-validation averaged across three target properties.

Key Performance Findings

Huber Regressor Excellence

- Achieved the highest predictive performance with engineered features (R^2 = 0.817)
- Outperformed the baseline by +51.3%, highlighting the advantage of robust regression
- Validates the strong synergy between physics-informed feature engineering and robust modeling techniques

Linear Method Effectiveness

- Linear Regression yielded a significant performance gain (R^2 = 0.811**,** +39.3% improvement)
- Confirm that the thermomechanical relationships were successfully linearized through feature transformation
- Ridge Regression also benefited considerably **(**+36.3% improvement**)**, despite the impact of regularization

Ensemble Method Analysis

- Ensemble models such as XGBoost and Random Forest showed marginal or negative improvements
- Indicates that complex models may not capitalize on explicitly engineered, physics-informed features
- Suggests that simpler or more interpretable models are better aligned with domain-informed inputs

4 DISCUSSION

4.1 Feature Engineering Effectiveness

The superior performance of thermomechanically-informed engineered features represents a significant advancement in AM prediction models. The substantial improvement achieved by the Huber-Engineered combination demonstrates that incorporating domain-specific physical principles enhances predictive capability beyond what raw process parameters can provide.

The SHAP analysis provides crucial validation of both the engineered feature design and the underlying physical assumptions. While the two most influential features remain original parameters - layer height and nozzle temperature - four of the top-ranking features are engineered transformations that capture critical physical interactions. This pattern demonstrates that engineered features effectively complement rather than replace fundamental process parameters, enhancing the model's ability to capture complex thermomechanical relationships.

The density_volume feature exemplifies successful feature engineering, achieving substantially higher importance than its constituent parameters when considered individually. This validates the hypothesis that three-dimensional geometric interactions provide superior predictive value compared to isolated parameter effects. Similarly, the layer_wall_ratio demonstrates how dimensionless geometric relationships can capture structural effects that individual measurements cannot represent.

The consistent failure of polynomial features across multiple algorithms highlights the critical importance of physics-informed feature design over automated expansion. Despite generating more features, the polynomial approach failed to create meaningful physical relationships and introduced overfitting complications in the small-dataset scenario.

4.2 Algorithm Performance Insights

Huber Regressor's exceptional performance with engineered features reveals important insights about robust regression's suitability for AM applications. The SHAP analysis reveals that the algorithm effectively exploits both high-impact individual features and moderate-impact engineered interactions, suggesting optimal alignment between the robust regression approach and the engineered feature representations.

The dominance of layer height and nozzle temperature in the feature importance ranking aligns with established thermomechanical principles, where layer thickness directly controls interlayer bonding area and thermal mass effects, while nozzle temperature governs polymer viscosity and flow characteristics. The algorithm's ability to properly weight these fundamental parameters while simultaneously leveraging engineered interaction terms demonstrates sophisticated feature utilization.

The substantial 81.28% predictive contribution from these two primary features reflects the fundamental thermal-geometric coupling in polymer processing, where layer thickness and processing temperature directly control the primary bonding mechanisms in FDM. This concentration is consistent with process-structure-property relationships observed in polymer manufacturing, where 2-3 dominant variables typically account for 70-85% of mechanical property variation. The remaining 18.72% from engineered features, particularly density_volume (8.75%) and layer_wall_ratio (6.32%), provides meaningful enhancement by capturing synergistic interaction effects that individual parameters cannot express independently. This hierarchical importance validates the physics-informed approach, demonstrating that engineered features effectively complement rather than compete with fundamental process drivers.

The strong performance of Linear Regression with engineered features becomes more interpretable through the SHAP analysis, which reveals a clear feature hierarchy that linear models can effectively exploit. The relatively balanced importance distribution among the top features suggests that the engineered feature set successfully linearized complex interactions without creating problematic multicollinearity.

Ensemble methods' failure to consistently outperform linear approaches may relate to their inability to properly weight the engineered features' physical significance. The feature importance analysis suggests that the carefully designed feature hierarchy requires algorithms capable of recognizing and preserving these physically meaningful relationships rather than treating all features as equivalent inputs.

4.3 Physical Validation and Feature Design Insights

The SHAP analysis provides compelling validation of the thermomechanical principles underlying the engineered feature design. The moderate importance of thermal-based features confirms their role in capturing thermal interaction effects while avoiding redundancy with the dominant nozzle temperature parameter. This pattern demonstrates successful feature engineering that enhances rather than duplicates existing information.

The relatively lower importance of some engineered features suggests opportunities for feature set optimization. However, their inclusion remains justified as they may capture critical effects for specific parameter combinations or contribute to prediction stability across the full experimental range.

The clear separation between high-impact features, moderate-impact engineered features, and lower-impact parameters provides practical guidance for process control prioritization in manufacturing environments where not all parameters can be optimally controlled simultaneously.

4.4 Manufacturing Applications and Practical Significance

The achieved prediction accuracy levels approach industrial requirements for AM quality control systems. The SHAP analysis provides additional practical value by identifying which parameters require the most precise control and which engineered relationships should be monitored for optimal quality outcomes.

The feature importance hierarchy suggests that manufacturing process control systems should prioritize monitoring and feedback for layer height and nozzle temperature while implementing secondary control loops for the engineered feature combinations. This layered control approach could optimize both computational efficiency and control effectiveness in real-time manufacturing environments.

The substantial performance improvement from feature engineering combined with the interpretability provided by SHAP analysis demonstrates clear return on investment for implementing physics-informed prediction models in manufacturing environments. The ability to identify and quantify the most critical parameter interactions enables transition from reactive quality control to predictive process optimization with clear understanding of which variables drive quality outcomes.

4.5 Limitations and Future Research Directions

Several limitations constrain the generalizability of these findings. The limited dataset size necessitates validation with larger experimental datasets to confirm both the robustness of the feature engineering approach and the stability of the SHAP importance rankings across diverse printing conditions and material systems.

Future research should investigate whether the identified feature importance hierarchy remains consistent across different material types, printer configurations, and target properties. The development of adaptive feature engineering frameworks that can automatically adjust feature importance weighting based on specific manufacturing contexts represents a promising research direction.

The integration of real-time SHAP analysis with process monitoring systems could enable dynamic feature importance tracking, allowing manufacturing systems to adapt control strategies as equipment characteristics change over time. This approach could bridge the gap between static prediction models and adaptive manufacturing control systems.

4.6 Validation Against Previous Studies

The achieved improvements align with recent advances in physics-informed ML for AM. Faegh et al. (2025) reported 15-40% improvements when incorporating thermomechanical principles in process-structure-property modeling, consistent with the 36-51% improvements herein this study. The dominance of layer height and nozzle temperature in SHAP rankings confirms findings from multiple studies on FDM parameter optimization.

Comparison with Recent AM-ML Studies:
- Current study (Huber-Engineered): R^2 = 0.817, n=50
- Conventional ML approaches: R^2 = 0.4-0.7 (typical range)
- Physics-informed approaches: R^2 = 0.6-0.8 (Faegh 2025)

The 51.3% improvement demonstrates superior performance particularly in small-dataset scenarios (n=50), addressing a critical gap where most studies require hundreds of samples.

5 CONCLUSIONS

This study proposes a framework to systematically evaluate the effects of process parameters and their interactions on the properties of FDM-printed specimens. The results show that, in addition to conventional standalone printing parameters that are traditionally used, combinations of these parameters into engineered features with physical significance can lead to higher accuracy in ML predictions. This is evident from the feature importance ranking, where three engineered features are ranked just behind the top two original FDM printing parameters. It is also worth noting that the feature importance analysis using SHAP supports the common understanding that layer height, nozzle temperature, and the interplay among wall thickness, layer thickness, and total printed volume are key determinants of surface roughness and tensile properties in FDM-printed samples. Regarding polynomial features, the findings suggest that although feature engineering can generate many derived features, this does not necessarily guarantee that the hidden thermo-mechanical behavior of the printed material can be captured. Subsequently, some generated features may have limited utility in accurately predicting the properties of printed samples using ML models.

ACKNOWLEDGMENTS

This paper was funded by the University of Finance – Marketing. Additionally, it was completed in connection with the project "Innovative and additive technologies for sustainable energy industry", registration no. CZ.02.01.01/00/23_021/0010117 financed by the Structural Funds of European Union project. It was also realized in connection with project Students Grant Competition SP2025/062 "Specific research on progressive and sustainable production technologies", and SP2025/063 "Specific research into innovative and progressive production technologies" financed by the Ministry of Education, Youth and Sports and Faculty of Mechanical Engineering VSB-TUO.

REFERENCES

[Beşliu-Bancescu 2023] Beşliu-Bancescu, I., I. Tamaşag, & L. Slatineanu. Influence of 3D Printing Conditions on Some Physical–Mechanical and Technological Properties of PCL Wood-Based Polymer Parts Manufactured by FDM. *Polymers*, May 2023, Vol.15, pp 2305, ISSN 2073-4360. https://doi.org/10.3390/polym15102305

[Faegh 2025] Faegh, M., S. et al. A review on physics-informed machine learning for process-structure-property modeling in additive manufacturing. *Journal of Manufacturing Processes*, January 2025, Vol.133, pp 524–555, ISSN 1526-6125. https://doi.org/10.1016/j.jmapro.2024.11.066

[Gillispie 2023] Gillispie, G.J., et al. The correlation between rheological properties and extrusion-based printability in bioink artifact quantification. *Materials & Design*, September 2023, Vol.233, pp 112237, ISSN 0264-1275. https://doi.org/10.1016/j.matdes.2023.112237

[Jaganathan 2024] Jaganathan, S., et al. Advances in Optimizing Mechanical Performance of 3D-Printed Polymer Composites: A Microstructural and Processing Enhancements Review. *Advances in Polymer Technology*, January 2024, Vol.2024, pp 3168252, ISSN 0730-6679, 1098-2329. https://doi.org/10.1155/2024/3168252

[Karniadakis 2021] Karniadakis, G.E., et al. Physics-informed machine learning. *Nature Reviews Physics*, May 2021, Vol.3, pp 422–440, ISSN 2522-5820. https://doi.org/10.1038/s42254-021-00314-5

[Ng 2024] Ng, W.L., et al. Progress and Opportunities for Machine Learning in Materials and Processes of Additive Manufacturing. *Advanced Materials*, August 2024, Vol.36, ISSN 0935-9648, 1521-4095. https://doi.org/10.1002/adma.202310006

[Ranjbar 2021] Ranjbar, N., et al. Rheological characterization of 3D printable geopolymers. *Cement and Concrete Research*, September 2021, Vol.147, pp 106498, ISSN 0008-8846. https://doi.org/10.1016/j.cemconres.2021.106498

[Siddiqui 2024] Siddiqui, M.A.S., et al. Biodegradable natural polymers and fibers for 3D printing: A holistic perspective on processing, characterization, and advanced applications. *Cleaner Materials*, December 2024, Vol.14, pp 100275, ISSN 2772-3976. https://doi.org/10.1016/j.clema.2024.100275

[Trovato 2025] Trovato, M., et al. Machine learning in design for additive manufacturing: A state-of-the-art discussion for a support tool in product design lifecycle. *The International Journal of Advanced Manufacturing Technology*, March 2025, Vol.137, pp 2157–2180, ISSN 0268-3768, 1433-3015. https://doi.org/10.1007/s00170-025-15273-9

[Truong 2025] Truong, T.C., et al. A SYSTEMATIC LITERATURE REVIEW ON THE IMPACT OF GENERATIVE AI ON ADDITIVE MANUFACTURING: CURRENT STATE AND FUTURE DIRECTIONS. *MM Science Journal*, March 2025, Vol.2025, ISSN 1803-1269, 1805-0476. https://doi.org/10.17973/MMSJ.2025_03_2025010

[Zhu 2024] Zhu, Y., et al. Thermal prediction of additive friction stir deposition through Bayesian learning- enabled explainable artificial intelligence. *Journal of Manufacturing Systems*, February 2024, Vol.72, pp 1–15, ISSN 0278-6125. https://doi.org/10.1016/j.jmsy.2023.10.015

CONTACTS:

Thanh-Cong Truong, Phd.

Faculty of Data Science, University of Finance - Marketing

No. 778 Nguyen Kiem, Duc Nhuan, Ho Chi Minh, 70000, Viet Nam
ttcong@ufm.edu.vn