Nigerian Journal of Engineering Science Research (NIJESR).
Department of Mechanical Engineering, Gen. Abdusalami Abubakar College of Engineering, Igbinedion University, Okada, Edo State, Nigeria.

ISSN: 2636-7114
Journal Homepage: https://nijesr.iuokada.edu.ng/

Data-Driven and Expert-Based Comparative Feature Analysis for Polyurethane Foam Formulation

[1*]Ese Emmanuel Enakireru, [2]Mathias Ekpu, [2]Samuel O. Sada, [2]Friday E. Ukrakpor, [3]Henry O. Orugba, [4]Francis Amadhe, [5]Rilwan Anjorin

[1] Department of Mechanical Engineering, Petroleum Training Institute, Effurun, Nigeria
[2]Department of Mechanical Engineering, Delta State University, Oleh, Nigeria
[3]Department of Chemical Engineering, Delta State University, Oleh, Nigeria
[4]Department of Production Engineering, University of Benin, Benin City, Nigeria
[5]Institute of Engineering, Technology and Innovation Management (METI), University of Port Harcourt, Nigeria

**Corresponding Author: Ese Emmanuel Enakirerua: enakireru_ee@pti.edu.ng*

Manuscript History
Received: 08/12/2024
Revised: 02/02/2025
Accepted: 28/03/2025
Published: 29/04/2025
https://doi.org/10.5281/zenodo.15310854

Abstract: *Accurately identifying the most influential chemical constituents is essential for optimizing the quality of polyurethane foams (PUFs), particularly in terms of density, cell morphology, and overall performance. This study integrates statistical, machine learning, and expert-informed techniques to evaluate feature importance for PU foam density prediction. A commercial dataset containing over 20,000 observations from a batch-slab stock PU foam facility was analyzed across seven chemical input variables. Feature selection methods included variance inflation factor (VIF) to assess multicollinearity, recursive feature elimination (RFE) integrated with support vector regression (SVR) using a radial basis function (RBF) kernel to rank features based on predictive power, and opinion discriminative analysis (ODA) to incorporate expert judgment. Fifty (50) industry professionals evaluated the features using a 5-point Likert scale, and their responses were analyzed to compute discriminative power (Dp) scores. Spearman's rank correlation coefficient (ρ) was employed to assess alignment among the models. RFE+SVR and ODA, RFE+SVR and VIF, all showed weak agreement (ρ = 0.178), while VIF strongly correlated positively with ODA (ρ = 1). This study highlights the value of combining statistical methods, algorithm-driven techniques, and expert insights to achieve more dependable feature selection in the production of polyurethane (PU) foam.*

Keywords: *Polyurethane Foams, Variance Inflation Factor, Recursive Feature Elimination, Support Vector Regression, Opinion Discriminative Analysis*

INTRODUCTION

Polyurethane foams (PUFs) are remarkably adaptable materials widely utilized across various industries, thanks to their superior mechanical strength, thermal insulation, and sound-absorbing capabilities. Typically, PUFs are produced through an exothermic chemical reaction between isocyanates and polyols, with catalysts, surfactants, and other additives often incorporated to fine-tune the material's properties (Mustafov & Seydibeyoglu, 2020). Their applications are diverse, as they range from flexible foams used in upholstered furniture, to rigid foams for insulation in construction and appliances, and thermoplastic variants found in medical equipment and footwear. PUFs also serve as key components in coatings, adhesives, sealants, and elastomers for flooring and automotive interiors (Das & Mahanwar, 2020). However, developing PUF formulations remains a complex, multivariable process, as the interactions between raw materials and processing parameters significantly influence the foam's structure and performance. Despite widespread application, scorch formation of PUFs compromises foam quality by reducing resilience and increasing the risk of failure during usage (Reed *et al.,* 2020). Numerous techniques for producing polyurethane foam have been developed and adopted worldwide (Kiss *et al.,* 2021). Polymeric foams are nearly ubiquitous today, largely due to their superior properties compared to alternative materials. Among these, polyurethane foams (PUFs) stand out as the most significant class. Their low density and thermal conductivity, coupled with favorable mechanical characteristics, make them highly effective as thermal and acoustic insulators, as well as ideal materials for structural applications and comfort-focused products (Sklenickova *et al.,* 2022).

In recent years, a variety of techniques have been explored and implemented to improve the processing of polyurethane foams (PUFs), highlighting both economic and environmental benefits. However, PUFs derived from renewable resource-based polyols often exhibit certain limitations, including lower thermal stability, greater flammability, increased sensitivity to high temperatures during fire exposure, and generally weaker mechanical properties (Yadav *et al.,* 2022). At the same time, rapid urbanization and industrial expansion linked to global economic growth have intensified concerns around ergonomics and orthopaedics needs relating to work-related tasks, rest, and sleep quality (Seyed *et al.,* 2023). In response to these challenges, Omoruwou *et al.* (2024) employed thermodynamic data with data-driven approach involving machine learning (ML) algorithms such as XGBoost, Random Forest, Support Vector Machines, and Logistic Regression to predict scorch risk. Their findings underscore the potential of ML in enhancing real-time monitoring and decision-making in PUFs production. Oyejide *et al.* (2020a) developed a specialized mixing and mold system designed for flexible polyurethane foam production, to optimize processing conditions. In a related study, the same authors utilized computational fluid dynamics to simulate a mixing mold unit within a small-scale foam batch production setup (Oyejide *et al.,* 2020b). Efforts to enhance sustainability in foam formulation have also gained momentum. Kirpluks *et al.* (2020) and Coman *et al.* (2021) successfully produced foams using tall oil- and olive oil-based polyols, respectively, demonstrating mechanical and thermal properties comparable to those of traditional petrochemical-based foams. Likewise, Leng *et al.* (2022) confirmed the viability of coconut oil-derived polyols for foam production. The reuse of polyurethane waste and the incorporation of recycled materials have also received growing attention. For instance, Kiss *et al.* (2020) showed that recycled polyols can be used to fine-tune properties such as tensile strength and airflow. Dhaliwal *et al.* (2021) enhanced the flame retardancy of soy-based foams by adding nanoclays, which help form protective char layers during combustion. Other innovative approaches include the integration of industrial by-products: Kuznia *et al.* (2021) explored the use of fly ash and microspheres in rigid PU foams, while Leszczyńska *et al.* (2020) and Husainie *et al.* (2021) investigated how additives like ground eggshells, rapeseed-derived polyols, cellulose, and chitin affect foam structure and performance. Khaleel *et al.* (2021) found that incorporating turkey feather fibers significantly enhanced the thermal and acoustic properties of rigid PU foams. Similarly, Mohammadpour & Sadeghi (2020) utilized liquefied lignin to develop foams with improved oil-absorption capabilities, pointing to new functional applications. Recent advances in computational science have also driven major shifts in polyurethane research. Task *et al.* (2023) employed machine learning for predictive modeling in slabstock foam processes, while Pugar (2023) demonstrated that hierarchical models, when informed by domain-specific knowledge, can achieve high accuracy even with limited datasets. Tools like SHAP values and Partial Dependence Plots have increased the interpretability of deep learning models, as highlighted by Rodríguez-Sánchez *et al.* (2024). Tasdemir *et al.* (2024) applied neural networks to accurately model compressive behavior under varying temperature and strain-rate conditions. In another study, Oyejide *et al.* (2023) illustrated how combining statistical methods with machine learning can help decipher the complex interactions between formulation variables, offering practical insights for industrial-scale foam production. Meanwhile, Admasu *et al.* (2022) leveraged image-based regression and generative adversarial networks (GANs) to analyze SEM images, enabling predictive optimization of foam microstructure and performance. Optimization of bio-based formulations has also benefited from hybrid methodologies, such as Zhang & Xu's (2022) integration of Taguchi experimental design with machine learning to fine-tune starch-based/EVA foams. Physics-informed machine learning approaches, such as the data-driven finite element method introduced by Korzeniowski & Weinberg (2022), bypass traditional empirical constitutive models to provide more accurate simulations of open-cell foam behavior. At the atomic level, machine learning frameworks have accelerated the discovery of structure-property relationships, enabling more effective optimization of soft materials (Leem *et al.,* 2023). These advancements mark a significant shift from traditional, empirically based methods toward more intelligent, data-driven approaches in polyurethane foam research. As the focus on sustainability, interpretability, and performance optimization continues to grow, it underscores the promising future of polyurethane materials in both academic research and industrial applications. The current study used a combination of data and expert-driven methods to assess the relative impact of various chemical inputs on polyurethane foam characteristics, including density, cell stability, and overall product quality. The key contributions of this work include: (i) the use of Variance Inflation Factor (VIF), a statistical technique employed to detect multicollinearity among the features; (ii) the application of Recursive Feature Elimination (RFE) integrated with Support Vector Regression (SVR) as an algorithmic model to identify the most relevant variables; (iii) the use of Opinion Discriminative Analysis (ODA) to validate the results from the statistical and algorithmic models, incorporating expert feedback from 50 industry professionals rated on a 5-point Likert scale; and (iv) the application of Spearman's rank correlation coefficients (ρ) to assess the degree of alignment between feature importance rankings derived from VIF, RFE+SVR, and ODA.

MATERIALS AND METHODS

In data-driven modelling, particularly for large-scale industrial processes like polyurethane (PU) foam production, the presence of many interrelated variables can lead to issues such as multicollinearity, redundant features, and increased model complexity. Identifying the most relevant variables is crucial for improving the accuracy, interpretability, and computational efficiency of the models. Feature selection plays a key role in this process by reducing dimensionality, preventing overfitting, and enhancing the generalizability of predictive models. In ensuring robust and interpretable modelling of PUF properties, the present study employed statistical methods (VIF), machine learning-based feature ranking techniques (RFE-SVR), and expert-driven (ODA) feature ranking from industry professionals in chemical constituents' importance.

2.1 Data Source and Experimental Context

The dataset used in this study was sourced from a commercial polyurethane (PU) foam manufacturing plant employing a batch-slab stock production process, located along the Upper Mission Extension axis in Benin City, Nigeria. Over an operational period of eight years, from March 2017 to May 2024, approximately 20,000 observations were systematically collected, providing a comprehensive dataset for modelling and analysis. The focus of the dataset is primarily on the chemical components used in the production process, while key mechanical parameters, such as reaction temperature, mixing speed, and curing time, were kept constant to isolate their effects on foam density. Specifically, the reaction temperature was maintained between 20°C and 25°C to ensure ambient curing conditions and minimize the formation of trapped air bubbles. The mixing speed was controlled between 2500 and

3500 rpm to ensure uniform blending of the reactants. The curing period ranged from 12 to 24 hours, allowing for complete cross-linking and stable foam structure development. Each data entry recorded the volumetric and mass measurements of essential chemical components, using flow meters, weight scales, and automated sensor systems before they were introduced into the mixing unit. This pre-mixing monitoring ensured consistent input quality and repeatable processes, laying the groundwork for controlled PU foam block production. The batch-slab stock method used by the company is a hybrid production approach, blending the efficiency of slab-stock foaming with the flexibility of batch processing. This technique offers the advantage of customizable foam properties, allowing for formulations tailored to specific application needs. The formulation process begins with the precise measurement of the chemical components. Polyol acts as the main polymer backbone in PU foam production, while methyl chloride serves as the physical blowing agent responsible for expanding the foam cells. Toluene diisocyanate (TDI) is essential for the polymerization reaction with polyol, while a tertiary amine catalyst accelerates the reaction rate. Water is added to generate carbon dioxide (CO_2), which contributes to foam expansion through chemical blowing. Stannous octoate acts as a co-catalyst, further speeding up the polymerization process, and silicon oil is used as a surfactant, improving cell structure formation and stabilizing the foam's morphology. These chemicals are mixed thoroughly in a closed system, triggering both polymerization and foaming reactions under tightly controlled conditions. Table-1 presents a summary of the chemical constituents used in polyurethane foam production at the commercial facility.

Table-1 Lists of chemical constituents (features) in puf dataset

Output Feature	Units	Lower Limits	Upper Limits
Density	Kgm^{-3}	1200.3	1220.8
Input Features	Units	Lower Limits	Upper Limits
Polyol	L	98.1	100
Methyl chloride (MC)	L	9.8	12.9
Toluene diisocyanate (TDI)	L	40	45
Tertiary amine (amine)	L	4.8	8.1
Water	L	4.6	7.9
Stannous octoate (STAN_OCT)	L	3	4.8
Silicon oil	L	18	25

2.2 Models Employed

2.2.1 Variance Inflation Factor (VIF) for Feature Selection

The Variance Inflation Factor (VIF) is a statistical measure used to assess the degree of multicollinearity within a dataset. It quantifies how much the variance of an estimated regression coefficient is increased due to collinearity with other predictor variables. Feature selection is the process of identifying and selecting the most important features from a larger set of variables. VIF is specifically designed to detect multicollinearity among independent variables in a regression model. Multicollinearity occurs when two or more predictors are highly correlated, which can cause instability in coefficient estimation, reduce model interpretability, and inflate standard errors. VIF quantifies this inflation by assessing how much the variance of a coefficient is increased due to collinearity with other predictors. According to , each predictor *Xi* in the model, a regression is run where *Xi* is the dependent variable, and all other predictors *X-i* are the independent variables.

For predictor *Xi*:

$$X_i = \beta_0 + \sum_{j \neq i} \beta_j X_j + \epsilon \tag{1}$$

The coefficient of determination Ri^2 for each regression model is calculated to assess the proportion of variance in *Xi* explained by other predictors.

Mathematically, for each predictor variable *Xi*, VIF is defined as,

$$\text{VIF}\ (\mathrm{X_i}) = \frac{1}{1 - \mathrm{R_i^2}} \tag{2}$$

where *Ri* is the coefficient of determination of a regression of *Xi* on all other predictor variables. A high Ri^2 (close to 1) means that *Xi* can be predicted well from other predictors, indicating high collinearity. Commonly accepted VIF thresholds are: (i) VIF = 1: No collinearity (ii) VIF between 1 and 5: Moderate collinearity (iii) VIF > 5 or 10: High collinearity, suggesting that the variable may need to be removed or adjusted.

2.2.2 Recursive Feature Elimination with Support Vector Regression (RFE-SVR)

RFE–SVR is a feature selection method that evaluates the relevance of features based on predictive performance. RFE (Recursive Feature Elimination) is a backward selection technique that iteratively removes the least important features, based on the weight of the machine learning model, retraining the model at each step to identify the optimal subset of features. In this study, Support Vector Regression (SVR) was chosen as the core model within the RFE framework due to its effectiveness in handling high-dimensional data and capturing nonlinear relationships. RFE works by progressively eliminating the least significant features based on model performance, retraining the model after each removal.

Support Vector Regression (SVR) aims to find a function *f*(*x*) that minimizes the deviation from the actual target values, ensuring that the difference remains within a specified threshold, ϵ, while also keeping the model's complexity as low as possible.

SVR optimization problem is formulated as:

$$\min_{w,b,\breve{\varsigma},\breve{\varsigma}^*} \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}(\breve{\varsigma}_i + \breve{\varsigma}_i^*) \tag{3}$$

Subject to:

$$\begin{cases} y_i - (w^T x_i + b) \leq \epsilon + \breve{\varsigma}_i \\ (w^T x_i + b) - y_i \leq \epsilon + \breve{\varsigma}_i^* \\ \breve{\varsigma}_i, \breve{\varsigma}_i^* \geq 0 \end{cases}$$

Where, w denotes the weight vector, b is the bias term, C represents the regularization parameter, and ξ_i and $\xi_i^*$are the slack variables introduced to handle margin violations. The parameter ε defines the precision tolerance, while x_i and y_i correspond to the input features and target values of the training data, respectively. In the RFE, the SVR is trained on the dataset, and feature importance is assessed based on the absolute values of the model coefficients, $|w_j|$. At each iteration, the feature associated with the smallest coefficient magnitude is removed from the feature set.

$$Rank(x_j) = |w_j| \tag{4}$$

This iterative procedure is repeated until the predefined number of features is retained. The model is evaluated at each iteration using cross-validated R^2 scores, and the subset with the highest score is selected. The choice of the SVR kernel as the base estimator in the RFE framework is justified by its strong generalization capabilities and robustness in capturing nonlinear relationships within the data. The RFE process involves training the SVR model, ranking features by their importance (based on coefficient weights or model impact), removing the least important feature iteratively until a defined number of features is reached or the model performance deteriorates.

2.2.3 Opinion Discriminative Analysis (ODA)

Opinion Discriminative Analysis (ODA) is a human-centric analytical approach designed to synthesize expert judgement into a ranked or scored framework for decision-making. In this study, ODA was applied to evaluate the relevance of polyurethane foam input parameters (chemical constituents) based on the domain expertise of professionals with over 10 years of industrial experience in flexible PU foam production. Features are ranked based on their ODS values in descending order. Opinion Discriminative Analysis (ODA) is often used in sentiment classification or subjective evaluation tasks. It relies on assigning discriminative weights to terms or variables based on their ability to differentiate classes. Features with the highest discriminative scores are considered most opinionated or impactful for classification. In the ODA framework, experts assign scores to each feature $x_j \in X$ based on perceived importance using a likert questionnaire scale (from 1 to 5). Normalization of Expert Opinions is mathematically shown in Eqn. (5)

$$\widetilde{O_{ij}} = \frac{O_{ij} - min(O_j)}{max(O_j) - min(O_j)} \tag{5}$$

Where O_{ij} is the original score given by expert I for feature j, $\tilde{O}_{ij}$ is the normalized opinion score. Opinion Discriminative Score (ODS) for each feature is computed by averaging normalized scores across all n experts using Eqn. (6).

$$ODS_j = \frac{1}{n}\sum_{i=1}^{n} \widetilde{O_{ij}} \tag{6}$$

ODA offers a subjective yet quantifiable mechanism to validate or contrast data-driven findings from RFE-SVR by leveraging expert intuition, especially when features show marginal variance in statistical or ML-based importance.

2.3 Development and Administration of Survey Questionnaire

In complementing the data-driven feature selection techniques and validating the practical relevance of the PUF constituents, a structured 5-point Likert-scale questionnaire was designed and administered to polymer experts within a commercial PU foam production company. The survey aimed to capture experts' judgment on the relative importance of each chemical input in influencing PU foam density, cell stability, and overall product quality. Each constituent was evaluated by the experts using the 5-point Likert scale in Table 2. The questionnaire focused on the following seven chemical constituents commonly used in PU foam formulation, (i) Polyol (ii) Methyl chloride (MC) (iii) Toluene diisocyanate (TDI) (iv) Tertiary amine (amine) (v) Water (vi) Stannous octoate (vii) Silicon oil.

Table-2 Five (5) point rensis likert scale

Scale	Interpretation
1	Not Important
2	Slightly Important
3	Moderately Important
4	Very Important
5	Extremely Important / Critical

The experts were instructed to willingly and strictly rate each chemical based on its perceived impact on foam density and performance, according to their knowledge and experience, without personal rights being intruded. The questionnaire was administered to a purposively selected sample of 30 – 50 polymer and process experts to ensure robustness and statistical reliability. This population size is justified based on common practice in expert opinion studies, where more than thirty (30) responses are typically sufficient to achieve stable mean scores, low standard errors, and reliable inter-rater consistency.

2.3.1 Survey Administration

The questionnaire was distributed through a combination of in-person administration during technical review meetings and digital forms (Google Forms) emailed to internal staff. Experts were given brief instructions and definitions to standardize understanding of each chemical's role in foam synthesis. Anonymity and confidentiality were maintained to ensure unbiased responses. The responses were compiled and scored, with mean Likert values computed for each chemical. Rankings were determined based on their average score across all respondents. Statistical measures such as standard deviation and interquartile range (IQR) were used to assess variability and consensus among experts.

2.4 Spearman's Rank Correlation Coefficient (ρ)

Spearman's Rank Correlation Coefficient (ρ) is an excellent and valid statistical method for comparing the consistency among the rankings generated. After determining feature importance using both RFE-SVR and ODA, the rankings were compared using Spearman's rank correlation coefficient, ρ, to assess the consistency between machine-learning-based selection and human expert evaluation. The interpretation is shown in Table 3.

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} \tag{7}$$

Where d_i is the difference between the ranks of the i^{th} feature from three methods, n is the total number of features.

Table-3 Spearman's rank correlation coefficient and their interpretations

Spearman's rank correlation coefficient	Interpretation
ρ = 1	Perfect agreement
ρ = 0	No correlation
ρ = -1	Perfect disagreement

The justification for using ρ is that it is non-parametric as it does not assume a normal distribution of the features or their rankings. However, it is rank-based as it compares the order or ranking of features, and it captures monotonic relationships.

RESULTS AND DISCUSSION

3.1 Recursive Feature Elimination (RFE) with Support Vector Regression (SVR)

RFE integrated with SVR and Non-linear RBF kernel seeks to find a function that approximates the relationship between the input features and the polyurethane density, which is the target. For nonlinear SVR, the model is transformed into a higher-dimensional space through a kernel function such as given in Table-4.

Table-4 Ranked attributes of PUF dataset using RFE+SVR_{RBF}

Rank	Features No.	Features
1	7	SILICON_OIL
1	3	TDI
1	5	WATER
3	6	STAN_OCT
1	4	AMINE
2	1	POLYOL
1	2	MC

Table-4 shows that the features 'MC', AMINE', WATER, SILICON_OIL, and 'TDI' have a high rank, while the features 'POLYOL' and 'STAN_OCT' are minimum ranked. Therefore, from the above, the present study excluded the commonly least two ranked features of the PUF dataset and trained machine learning algorithms with the remaining five (5) optimal features for prediction performance metrics.

Enakireru *et al.*, (2025). Data-Driven and Expert-Based Comparative Feature Analysis for Polyurethane Foam Formulation. *Nigeria Journal of Engineering Science Research (NIJESR)*, 8(1): 26-39. https://doi.org/10.5281/zenodo.15310854

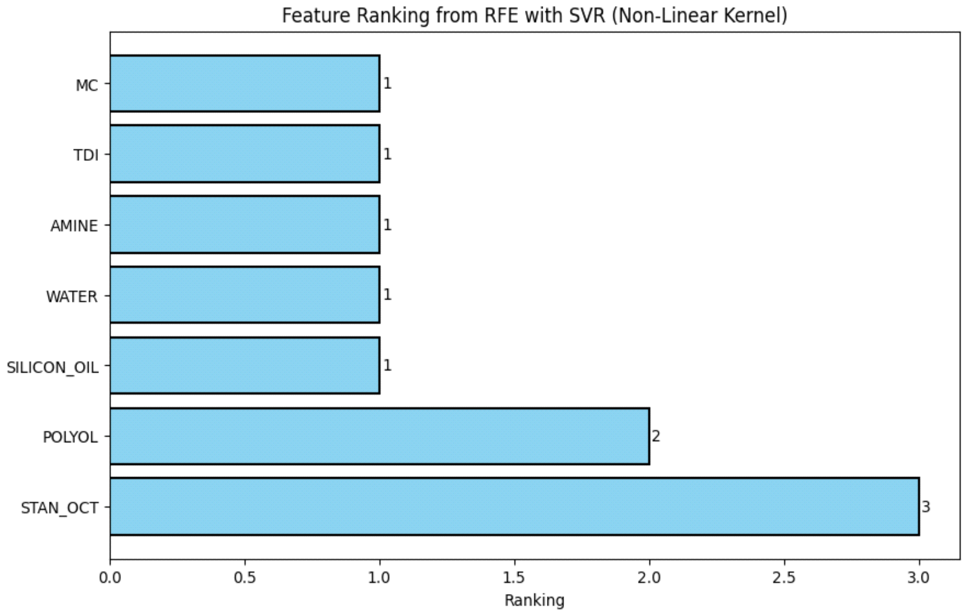

Fig. 1 Feature ranking from RFE with SVR (non-linear kernel)

RFE with SVR using a non-linear kernel (RBF) produces rankings for each feature at threshold values of zero, indicating that a ranking of 1 means the most important feature. A threshold of zero would exclude features ranked as least important (those with important scores close to or below zero). Features with positive rankings would be selected.

3.2 Variance Inflation Factor (VIF) for Feature Selection

In evaluating the degree of multicollinearity among the input variables in the combined polyurethane foam (PUF) dataset, the Variance Inflation Factor (VIF) is calculated for each feature. VIF serves as a statistical measure that quantifies how much the variance of an estimated regression coefficient increases due to collinearity with other predictors. A VIF value of 1 suggests no multicollinearity, values between 5 and 10 indicate moderate multicollinearity, and values significantly above 10 are typically considered indicative of severe multicollinearity. Table-5 presents the VIF values and corresponding feature rankings. The results demonstrate a pervasive issue of multicollinearity, as all examined variables exhibit exceedingly high VIF values. Specifically, the features ‘Amine’ and ‘Water’ display VIF values of infinity (VIF = ∞), indicating perfect multicollinearity. This implies that these variables are exact linear combinations of one or more other predictors in the dataset, which undermines their statistical independence. Additional features such as ‘Silicon oil’ (VIF = 10,662.59), ‘TDI’ (VIF = 7,435.04), ‘Methyl chloride’ (VIF = 3,291.12), ‘Polyol’ (VIF = 1,360.17), and ‘Stannous octoate’ (VIF = 1,207.20) also show alarmingly high VIF values. Fig. 2 describes the VIF result using horizontal bars.

Table-5 Ranked attributes of combined dataset using VIF

Rank	Feature	VIF
3	Amine	∞
4	Water	∞
6	Silicon oil	10662.59
2	TDI	7435.04
1	Methyl chloride	3291.12
0	Polyol	1360.17
5	Stannous octoate	1207.20

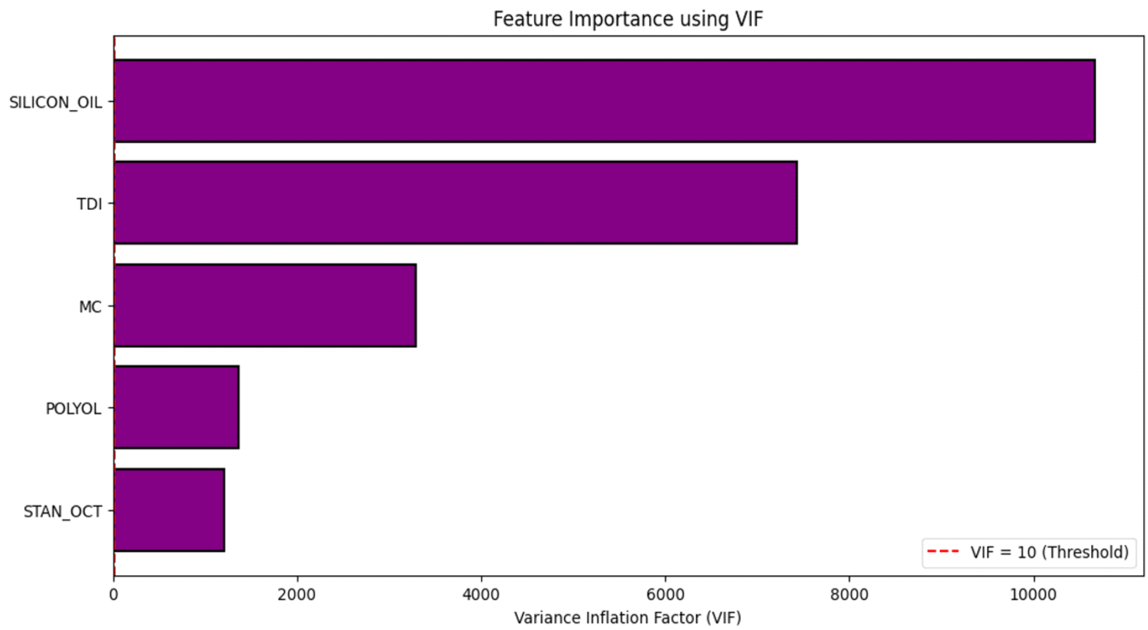

Fig. 2. Feature importance using VIF technique

Results were later compared with machine learning-based feature importance (RFE–SVR and VIF) using Spearman's rank correlation to assess consistency between human and algorithmic evaluation. The mixed-method strategy combines expert insight with computational analysis, providing deeper validation of critical features and helping to ensure the interpretability and real-world relevance of the modelling results.

3.6.2 Feature Importance Validation (ODA)

In validating and complementing the outcomes of the statistical and machine learning-based feature selection techniques, an expert-informed decision support approach using Opinion Discriminative Analysis (ODA) was employed. A structured questionnaire was designed and administered to a purposive sample of fifty (50) polymer experts and process engineers working within a commercial PU foam production facility. The experts were selected based on their active roles in formulation development, quality assurance, and process optimization, and each had at least three years of industry experience. The questionnaire tasked experts with evaluating seven key chemical constituents using a 5-point Likert scale. Each constituent was rated from 1 (Not Important) to 5 (Extremely Important), based on the respondent's assessment of its influence on polyurethane (PU) foam density, foam morphology, and curing behaviour. The resulting data were organized into a (50×7) matrix, where each row represented an individual expert, and each column corresponded to one of the seven (7) chemical features under consideration. The matrix was analyzed using the Optimized Discriminant Analysis (ODA) model, which calculated the Discriminative Power (Dp) score for each feature, as presented in Table 6. The Dp score is a quantitative measure that indicates the relative importance of a given variable based on the consistency and strength of expert opinions across the sample.

Table-6 Ranked attributes of the combined dataset using ODA

Rank	Feature	VIF
1	Silicon oil	0.91
2	TDI	0.87
3	Methyl chloride	0.83
4	Polyol	0.78

5	Stannous octoate	0.74
6	Amine	0.70
7	Water	0.65

In Table-6, the Features with higher Dp scores were considered to have stronger discriminative capability and therefore greater significance from a domain perspective. The scores are visualized comparatively in a bar chart shown in Fig. 3.

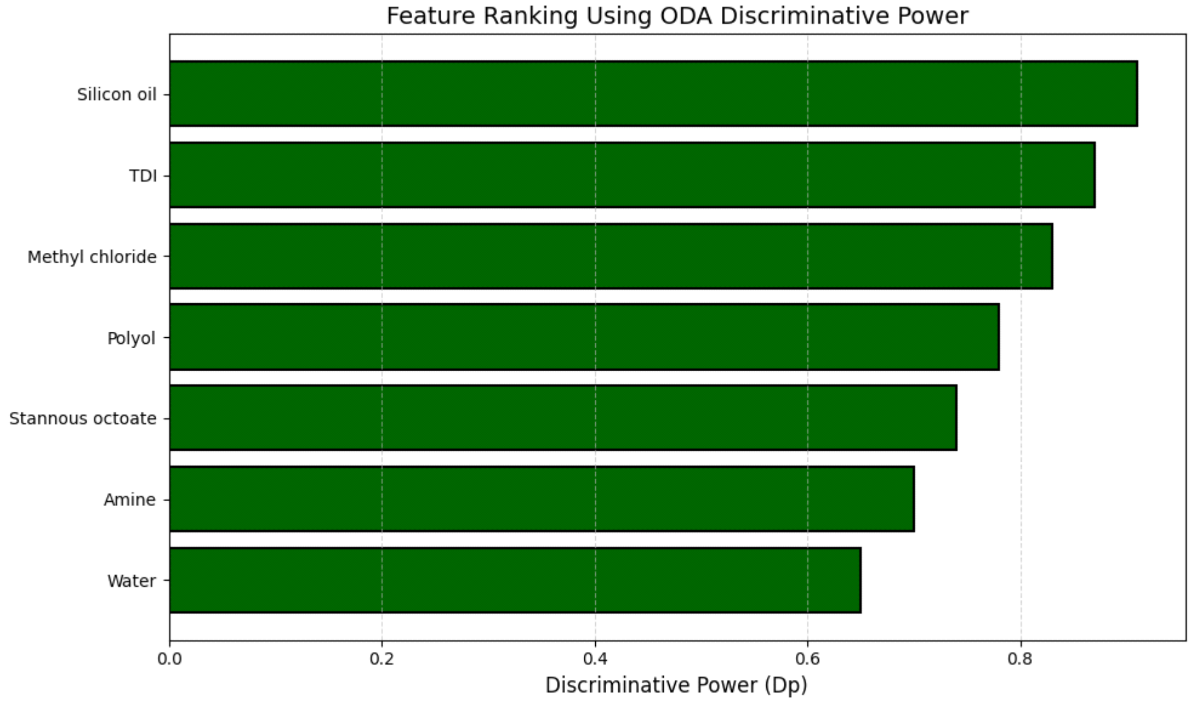

Fig. 3. Ranking of the Dp values of the PUF features

3.7 Comparison of the Feature Selection Models

The Spearman's rank correlation coefficients (ρ) indicate the level of agreement between the different feature ranking methods. In computing the Spearman's rank correlation coefficient (ρ) for the RFE+SVR, ODA and VIF, it is imperative to convert the data into comparable ranking formats. Based on the ranking output from RFE+SVRRBF (Table 4), the numeric ordinal ranks with ties averaged are computed. Rank 1 occurs 5 times (SILICON OIL, TDI, WATER, AMINE, MC), Rank 2 occurs once (POLYOL), and Rank 3 occurs once (STAN_OCT). For Rank 1 Position: (1, 2, 3, 4, 5), Average = (1+2+3+4+5) / 5 = 3.0, Position for Rank 2 = 6, Position for Rank 3 = 7. The assigned ordinal ranks for the RFE+SVR, the VIF ranking (Table 4) and the ODA ranking (Table 5) are shown in Table-7. Computing the Spearman's Rank Correlation Coefficient (ρ) using Eqn. (7) is shown in Fig. 4.

Table -7 The assigned ordinal ranks for RFE+SVR, VIF and the ODA ranking

Feature	RFE Rank	VIF Rank	ODA Rank
Silicon oil	3.0	1	1
TDI	3.0	2	2
Methyl chloride	3.0	3	3
Polyol	6.0	4	4
Stannous octoate	7.0	5	5
Amine	3.0	6	6
Water	3.0	7	7

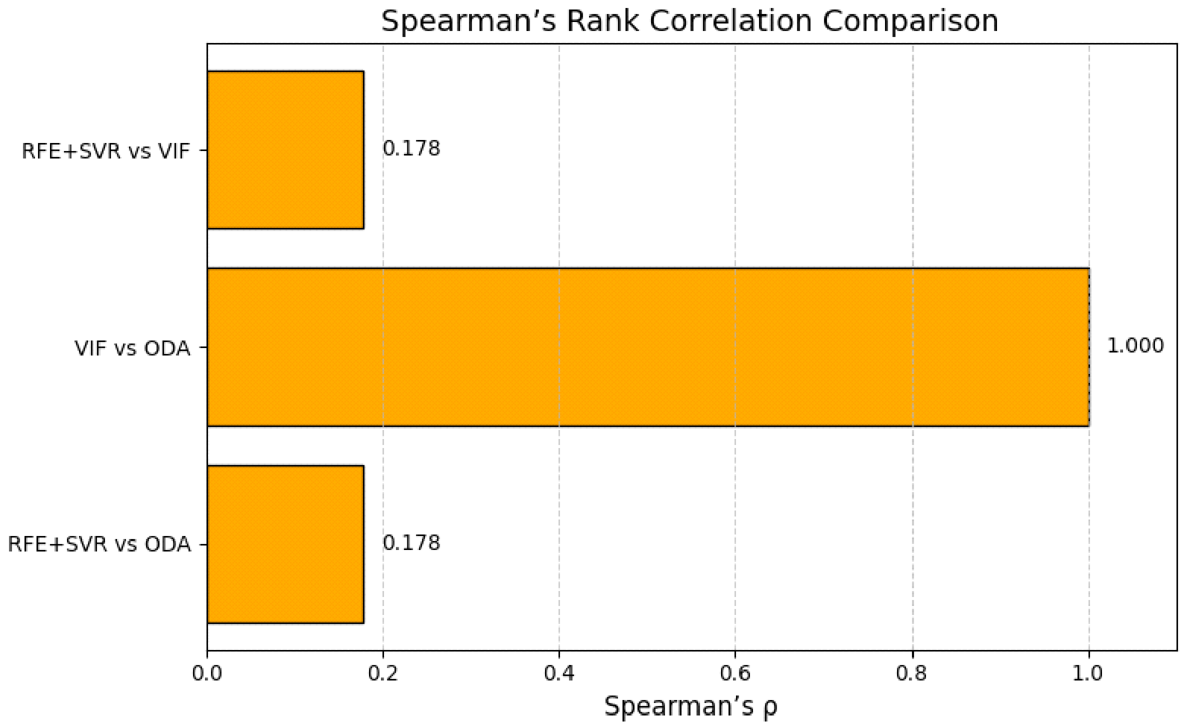

Fig. 4. Spearman's Rank Correlation Comparison

The following interpretations can be deduced from Fig. 4.

(i) RFE+SVR vs ODA (Expert Opinion): $\rho = 0.178$

This suggests a low positive correlation, maintaining that the feature importance ranking from RFE+SVR is slightly consistent with expert evaluations.

(ii) RFE+SVR vs VIF: $\rho = 0.178$

This suggests a low positive correlation, maintaining that the feature importance ranking from RFE+SVR is slightly consistent with the VIF.

(iii) VIF vs ODA (Expert Opinion): $\rho = 1$

This suggests a strong positive correlation, maintaining that the feature importance ranking from VIF is largely consistent with expert evaluations. Spearman's rank correlation coefficient (ρ) assesses the strength and direction of the association between two ranked variables. A value of $\rho = 1.000$ indicates a perfect positive correlation, meaning the two rankings are identical in order. This was observed between VIF and ODA rankings, suggesting that the expert-driven ODA closely aligns with the statistical ranking produced by the VIF method. In contrast, the correlation between RFE+SVR and both ODA and VIF is $\rho \approx 0.178$. This moderate-to-low correlation indicates that the machine learning model (RFE+SVR) assigns feature importance rankings that differ from both the expert-based ODA and the statistically driven VIF. The results indicate that, although the VIF and ODA methods exhibit a high degree of consistency, the RFE-SVR model may reveal distinct patterns of feature relevance by capturing nonlinear dependencies inherent in the data.

CONCLUSION

The integration of Optimized Discriminant Analysis (ODA) offered a robust validation framework for evaluating the consistency and relevance of feature rankings derived from machine learning models. The primary contributions of this work are summarized as follows. Variance Inflation Factor (VIF) analysis was employed as a statistical measure to detect multicollinearity among the input features. The presence of significant multicollinearity was evidenced by excessively high VIF values, most notably for the features 'Amine' and 'Water', both of which yielded infinite values. This outcome indicates the existence of perfect or near-perfect linear dependencies between these variables and other predictors in the dataset. Recursive Feature Elimination (RFE), combined with Support Vector Regression (SVR) utilizing a nonlinear radial basis function (RBF) kernel, was employed to identify the most influential variables for predicting polyurethane foam (PUF) density. Features including 'Silicon Oil', 'TDI', 'Water', 'Amine', and 'Methyl Chloride' consistently received high importance rankings. Conversely, the exclusion of lower-ranked variables such as 'Polyol' and 'Stannous Octoate' contributed to a reduction in model complexity and an improvement in predictive performance. In validating the outcomes obtained from both statistical and algorithmic feature selection methods, Opinion Discriminative Analysis (ODA) was employed. This approach incorporated expert evaluations from 50 industry professionals, who rated the importance of each feature using a 5-point Likert scale. The resulting Discriminative Power (Dp) values derived from ODA closely corresponded with the rankings produced by the RFE-SVR model, thereby reinforcing the robustness and credibility of the selected features from a domain-specific perspective. Spearman's rank correlation coefficients (ρ) were calculated to assess the degree of concordance between feature importance rankings generated by different selection methods. The results indicate a strong positive correlation, suggesting that the rankings derived from the Variance Inflation Factor (VIF) method are largely consistent with expert evaluations. Further studies should integrate data-driven methodology using expert-driven evaluations with statistical and algorithmic analyses in improving the density, cell morphology, and overall performance of PUFs.

CONFLICT OF INTEREST

The authors declare no conflict of interests.

REFERENCES

Admasu, A., Upadhyay, D., Couvreur, R., Tardiff, J., Kiziltas, A., Blanchard, P. (2022). *Deep Learning Image-based Design of Graphene-reinforced Polyurethane Foams.*

Ates, M., Karadag, S., Eker, A. A., Eker, B. (2022). Polyurethane foam materials and their industrial applications. *Polymer International*, 71(10), 1157-1163.

Coman, A.E., Peyrton, J., Hubca, G., Sarbu, A., Gabor, A.R., Nicolae, C.A., Iordache, T.V., Averous, L. (2021). Synthesis and characterization of renewable polyurethane foams using different biobased polyols from olive oil. *European Polymer Journal*, 149, 110363.

Das, A., Mahanwar, P. (2020). A brief discussion on advances in polyurethane applications. *Advanced Industrial and Engineering Polymer Research*, *3*(3), 93-101.

De Luca Bossa, F., Santillo, C., Verdolotti, L., Campaner, P., Minigher, A., Boggioni, L., Losio, S., Coccia, F., Iannace, S., Lama, G.C. (2020). Greener nanocomposite polyurethane foam based on sustainable polyol and natural fillers: Investigation of chemico-physical and mechanical properties. *Materials*, 13(1), p.211.

Dhaliwal, G. S., Bajwa, D. S., Bajwa, S. (2021). Fabrication and testing of soy-based polyurethane foam with flame retardant properties. *Journal of Polymers and the Environment*, 29, 1153-1161.

Husainie, S. M., Deng, X., Ghalia, M. A., Robinson, J., Naguib, H. E. (2021). Natural fillers as reinforcement for closed-molded polyurethane foam plaques: Mechanical, morphological, and thermal properties. *Materials Today Communications*, 27, 102187.

Khaleel, M., Soykan, U., Çetin, S. (2021). Influences of turkey feather fiber loading on significant characteristics of rigid polyurethane foam: Thermal degradation, heat insulation, acoustic performance, air permeability and cellular structure. *Construction and Building Materials*, 308, 125014.

Kirpluks, M., Vanags, E., Abolins, A., Michalowski, S., Fridrihsone, A., Cabulis, U. (2020). High functionality bio-polyols from tall oil and rigid polyurethane foams formulated solely using bio-polyols. *Materials*, 13(8), 1985.

Kiss, G., Rusu, G., Bandur, G., Hulka, I., Romecki, D. (2021). Advances in Low-Density Flexible Polyurethane Foams by Optimized Incorporation of High Amount of Recycled Polyol. *Polymers*, Vol. 13, Number 1736.

Kiss, G., Rusu, G., Peter, F., Tănase, I., Bandur, G. (2020). Recovery of flexible polyurethane foam waste for efficient reuse in industrial formulations. *Polymers*, 12(7), 1533.

Korzeniowski, T. F., Weinberg, K. (2022). Data-driven finite element computation of open-cell foam structures. *Computer Methods in Applied Mechanics and Engineering*, 400, 115487.

Kuranska, M., Pinto, J. A., Salach, K., Barreiro, M. F., Prociak, A. (2020). Synthesis of thermal insulating polyurethane foams from lignin and rapeseed based polyols: A comparative study. *Industrial Crops and Products*, 143, 111882.

Kuźnia, M., Magiera, A., Zygmunt-Kowalska, B., Kaczorek-Chrobak, K., Pielichowska, K., Szatkowski, P., Benko, A., Ziąbka, M., Jerzak, W., 2021. Fly ash as an eco-friendly filler for rigid polyurethane foams modification. *Materials*, 14(21), p.6604.

Lee, J. H., Kim, S. H., Oh, K. W. (2021). Bio-based polyurethane foams with castor oil based multifunctional polyols for improved compressive properties. *Polymers*, 13(4), 576.

Leem, J., Jiang, Y., Robinson, A., Xia, Y., Zheng, X. (2023). Data-Driven Approach to Tailoring Mechanical Properties of a Soft Material. *Advanced Functional Materials*, 33(38), 2304451.

Leng, X., Li, C., Cai, X., Yang, Z., Zhang, F., Liu, Y., Yang, G., Wang, Q., Fang, G., Zhang, X. (2022). A study on coconut fatty acid diethanolamide-based polyurethane foams. *RSC advances*, 12(21), pp.13548-13556.

Leszczyńska, M., Ryszkowska, J., Szczepkowski, L., Kurańska, M., Prociak, A., Leszczyński, M.K., Gloc, M., Antos-Bielska, M. and Mizera, K. (2020). Cooperative effect of rapeseed oil-based polyol and egg shells on the structure and properties of rigid polyurethane foams. *Polymer Testing*, 90, p.106696.

Mockeviciene, B., Vedlūga, T. (2024). Furniture Industries: Challenges of Regionalisation, Customisation and New Paradigm of Pricing. In Participation Based Intelligent Manufacturing: Customisation, Costs, and Engagement. *Emerald Publishing Limited*. pp. 47-100.

Mohammadpour, R., Mir Mohamad S. G. (2020). Effect of liquefied lignin content on synthesis of bio-based polyurethane foam for oil adsorption application. *Journal of Polymers and the Environment*, 28(3), 892-905.

Mustafov, S. D., Sen, F., Seydibeyoglu, M. O. (2020). Preparation and characterization of diatomite and hydroxyapatite reinforced porous polyurethane foam biocomposites. *Scientific reports*, 10(1), 13308.

Omoruwou, F., Ojugo, A. A., Ilodigwe, S. E. (2024). Strategic Feature Selection for Enhanced Scorch Prediction in Flexible Polyurethane Form Manufacturing. *Journal of Computing Theories and Applications*, 1(3), 346-357.

Oyejide O.J., Achebo J.I., Ozigagun A. (2023). A Factorial Study of Variables affecting the Formulation and Production of Polyurethane Foam'. *FUPRE Journal of Scientific and Industrial Research*. 6(1): 8-19.

Oyejide, O. J., Achebo, J. I., Okwu, M. O., Tartibu, L. (2020a). Development of a mixing and moulding unit for mini foam batch production plant. *Proceedings of the 2nd African International Conference on Industrial Engineering and Operations Management*, IEOM 2020 (pp. 652-663). IEOM Society.

Oyejide, O. J., Achebo, J. I., Okwu, M. O., Tartibu, L. K. (2020b). Computational Fluid Dynamics and Simulation of a Mixing Mould Unit of a Mini Foam Batch Production Plant. *Proceedings of the 2nd African International Conference on Industrial Engineering and Operations Management* (IEOM Society International) Harare, Zimbabwe, December 7-10, 2020.

Pugar, J. A. (2023). Multiscale Modeling of Polyurethane Properties via Latent Variables with Hierarchical Machine Learning (Doctoral dissertation, Carnegie Mellon University).

Reed, N., Huynh, N. U., Rosenow, B., Manlulu, K., Youssef, G. (2020). Synthesis and characterization of elastomeric polyurea foam. *Journal of Applied Polymer Science*, 137(26), 48839.

Rodríguez-Sánchez, A. E., Plascencia-Mora, H., Acevedo-Alvarado, M. (2024). Neural network-driven interpretability analysis for evaluating compressive stress in polymer foams. *Journal of Cellular Plastics*, 60(4), 237-258.

Sklenickova, K., Abbrent, S., Halecky, M., Koci, V., Benes, H. (2022). Biodegradability and ecotoxicity of polyurethane foams: A review. *Critical Reviews in Environmental Science and Technology*, *52*(2), 157-202.

Tasdemir, B., Tagarielli, V. L., Pellegrino, A. (2024). A data-driven rate and temperature dependent constitutive model of the compression response of a syntactic foam. Materials Today Communications, 39, 108790.

Task, K., McBride, M., Smiecinski, T., Wesley, W., Kang, Y., McCraw, A., Chiandussi, L. and Wandernoth, S. (2023). Digitalization and Predictive Modeling of Polyurethane Data.

Yadav, A., de Souza, F. M., Dawsey, T., Gupta, R. K. (2022). Recent advancements in flame-retardant polyurethane foams: A review. *Industrial & Engineering Chemistry Research*, *61*(41), 15046-15065.

Zhang, Y., Xu, X. (2022). Predicting mechanical performance of starch-based foam materials. *Journal of Cellular Plastics*, 58(3), 505-514.