

Machine Learning-Powered Predictive Analytics for Evaluating the Performance of Energized Foams

Yousif Almumen

University of Adelaide

Rakan Alhindi

University of Adelaide

Ammar Almishari

University of Adelaide

Maria Gonzalez Perdomo

`maria.gonzalez@adelaide.edu.au`

University of Adelaide

Afshin Tatar

University of Adelaide

Sardar Salar Najib

University of Adelaide

Fuyuan Yang

University of Adelaide

Ceridwen Nguyen

University of Adelaide

Research Article

Keywords: Energized foams, Hydraulic fracturing, High-pressure, high temperature (HPHT), Machine learning, Foam rheology, Apparent viscosity prediction

Posted Date: October 16th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9798716/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Machine Learning-Powered Predictive Analytics for Evaluating the Performance of Energized Foams

Yousif Almumen^A, Rakan Alhindi^A, Ammar Almishari^A, Maria Gonzalez Perdomo^{A*}, Afshin Tatar^A,
Sardar Salar Najib^A, Fuyuan Yang^A, Ceridwen Nguyen^A

School of Chemical Engineering, Adelaide University, Adelaide, South Australia, Australia

*Correspondence to: Email: maria.gonzalez@adelaide.edu.au

Abstract

Hydraulic fracturing of tight and ultra-low-permeability reservoirs increasingly relies on energized foam-based fluids to reduce water use while maintaining proppant transport. Predicting foam rheology under high-pressure, high-temperature (HPHT) reservoir conditions remains challenging because pressure, temperature, shear, and foam degradation act simultaneously, while laboratory replication is costly, time-consuming, hazardous, and limited in scope.

This study presents a coupled static–dynamic machine-learning framework for predicting apparent viscosity of nanoparticle- and surfactant-stabilized CO₂ and N₂ foams under reservoir-representative elevated pressure and temperature conditions. Experimental rheology data from 11 laboratory tests (3,963 datapoints), spanning pressures up to 3,040 psi and temperatures up to 122 °C, were used to train and evaluate models. Static models used thermodynamic and operational parameters, whereas dynamic models incorporated time-dependent features to represent viscosity evolution.

Among the evaluated approaches, a multilayer perceptron (MLP) provided the strongest static performance among the deployment-oriented non-tree models ($R^2 = 0.88$), while Lasso regression produced the most stable dynamic forecasts (mean root mean square error, RMSE = 0.06 in logarithmic space). Feature-importance analysis identified temperature-related variables as dominant in static predictions, whereas lagged viscosity governed dynamic behaviour.

The novelty lies in integrating reservoir-representative foam rheology, nanoparticle–surfactant-stabilized CO₂/N₂ systems, and complementary static and time-aware machine-learning models within one predictive framework. The approach provides a physically grounded basis for forecasting viscosity evolution, formulation screening, and operational decision-making, while remaining constrained by the range and variables represented in the experimental dataset.

Keywords: Energized foams, Hydraulic fracturing, High-pressure, high temperature (HPHT), Machine learning, Foam rheology, Apparent viscosity prediction.

1. Introduction

Rheological behavior of fracturing fluids is a critical factor governing the efficiency and success of hydraulic fracturing operations in unconventional reservoirs. Apparent viscosity directly controls proppant transport, fracture geometry, and the ability to maintain conductive flow paths after treatment. Under high-pressure, high-temperature (HPHT) reservoir conditions, accurately characterizing and controlling fluid rheology becomes increasingly challenging, yet essential for designing effective injection strategies and ensuring operational reliability.

The growing dependence on unconventional hydrocarbon resources has intensified the application of hydraulic fracturing in formations with extremely low permeability. In such environments, fracturing fluids must possess sufficient viscosity and stability to suspend and transport proppants—such as coated sands and ceramic materials—deep into complex fracture networks, thereby preventing premature settling and loss of fracture conductivity (Tran et al. 2020; Tran et al. 2025). Conventional water-based fracturing fluids are economically attractive; however, they often suffer from excessive water use, potential formation damage, and limited proppant-transport efficiency under demanding reservoir conditions. Energized foam-based fracturing fluids, particularly CO₂- and N₂-based foams, have therefore attracted increasing attention because they reduce liquid requirements while providing useful proppant-carrying capacity and cleanup characteristics.

Foam stability and rheology are strongly governed by the surfactant system and its interaction with the gas–liquid interface. At elevated temperature and salinity, surfactant performance can deteriorate, promoting liquid drainage, foam coalescence, viscosity loss, and reduced fracturing performance. Experimental studies have shown that incorporating silica nanoparticles (SNPs) alongside surfactants can enhance foam stability, foamability, apparent viscosity, and proppant suspension (Gonzalez Perdomo and Wan Madihi 2022; Li et al. 2022; Tran et al. 2023a; Bai et al. 2025). In particular, nanoparticle–surfactant interactions can strengthen interfacial films and improve resistance to thermal and mechanical degradation, although the magnitude of the benefit depends on surfactant chemistry, concentration, salinity, and temperature (Tran et al. 2023a, b).

Despite these advances, experimental evaluation of foam rheology under reservoir-representative elevated pressure and temperature conditions remains costly, time-consuming, and operationally complex. Laboratory measurements require specialized equipment and are often limited in their ability to explore wide parameter spaces involving pressure, temperature, shear rate, foam quality, and chemical composition. This creates an engineering need for reliable and scalable methods that can extend the value of experimentally measured rheological data without implying validity beyond the domain supported by those measurements.

Machine learning (ML) offers one route to this challenge by enabling data-driven representation of nonlinear relationships between fluid composition, operating conditions, and rheological response. Across porous-media and hydraulic-fracturing applications, ML has increasingly been used to support prediction, optimization, and interpretation of complex multivariable systems (Delpisheh et al. 2024; Karami et al. 2025). For foam-based fracturing fluids, ML-driven viscosity prediction could complement laboratory testing by supporting formulation screening and scenario evaluation within the range of conditions represented by the available data.

This study addresses these challenges by developing an integrated experimental–computational framework for predicting the apparent viscosity of nanoparticle–surfactant–stabilized CO₂ and N₂ foams under reservoir-representative elevated temperature and pressure conditions. Specifically, this paper contributes by: (i) utilizing reservoir-representative elevated pressure and temperature experimental rheological data to characterize foam behavior across a wide range of pressures, temperatures, shear rates, foam qualities, and chemical compositions; (ii) explicitly incorporating nanoparticle–surfactant interaction effects within multi-gas (CO₂ and N₂) foam systems to capture stabilization mechanisms relevant to field-scale applications; and (iii) implementing and systematically comparing both static regression-based and dynamic, time-aware machine learning frameworks for foam apparent viscosity prediction. By unifying rheological experimentation, nanoparticle–surfactant chemistry, and static–dynamic machine learning within a single predictive workflow, this study advances beyond existing approaches and provides a robust, data-driven tool for optimizing foam-based fracturing fluid design and injection strategies in challenging unconventional reservoirs.

Background and related work

Energized foams have gained attention as alternatives to conventional water-based fracturing fluids, particularly in tight and ultra-tight reservoirs where water sensitivity, proppant transport, and fracture conductivity are important constraints. Compared with slickwater, foam-based systems can reduce liquid use and formation damage while providing sufficient apparent viscosity for proppant transport (Gonzalez Perdomo and Wan Madihi 2022; Tran et al. 2020; Verma et al. 2025; Wanniarachchi et al. 2017). These attributes make CO₂- and N₂-based foams attractive for water-sensitive formations and other applications where rapid cleanup is desirable.

Foam rheology is central to proppant transport and fracture conductivity. Energized foams commonly exhibit non-Newtonian and shear-thinning behaviour, allowing viscosity to decrease during high-shear pumping while retaining greater carrying capacity at lower shear rates (Gonzalez Perdomo and Wan Madihi 2022; Verma et al. 2025). Experimental and simulation studies further show that foam rheology influences proppant settling, spatial distribution, and ultimately the effectiveness of proppant placement within hydraulic fractures (Tran et al. 2020; Tran et al. 2025).

Considerable research has therefore focused on improving foam stability through nanoparticle–surfactant systems. Surfactants adsorb at gas–liquid interfaces to stabilize foam films, while nanoparticles can anchor at or near the interface and retard liquid drainage and coalescence (Li et al. 2022; Zhu et al. 2022). The combined interfacial and colloidal effects can increase apparent viscosity and thermal stability, although performance depends strongly on nanoparticle type, surfactant chemistry, and concentration (Tran et al. 2023a; Bai et al. 2025). Recent work on nanoparticle-enhanced surfactant fracturing fluids similarly emphasizes the importance of nanoparticle–surfactant interactions for rheological stability and proppant transport under demanding reservoir conditions (Bai et al. 2025).

Foam performance remains sensitive to environmental and compositional conditions. Increasing temperature generally accelerates drainage and destabilization, while salinity can alter electrostatic interactions and promote nanoparticle aggregation; both effects can reduce foam viscosity and proppant-suspension capacity (Tran et al. 2023a, b). Pressure and gas type also influence foam texture and rheology through gas compressibility, solubility, and phase behaviour (Gonzalez Perdomo and Wan Madihi 2022; Al-Darweesh et al. 2024). A recent JPEPT review likewise identifies temperature, pressure, salinity, foam quality, and gas type as key controls on foam stability and rheology under downhole conditions, while noting growing interest in data-driven predictive methods (Gowida et al. 2026).

Parallel to these experimental developments, data-driven methods have increasingly been applied to foam rheology and petroleum-engineering prediction problems. Machine-learning models can represent nonlinear relationships among temperature, pressure, salinity, shear rate, foam quality, nanoparticle concentration, and surfactant concentration (Al-Darweesh et al. 2024; Olukoga et al. 2022). Broader reviews also identify growing use of ML for prediction and hybrid modelling in porous-media systems (Delpisheh et al. 2024). Within petroleum engineering, recent work has applied ML to hydraulic-fracturing efficiency and optimization (Karami et al. 2025), while time-series and deep-learning models have been evaluated for forecasting unconventional-reservoir production (Laalam et al. 2023). These studies illustrate the value of data-driven approaches when nonlinear or sequential information is important, although their target variables differ from foam viscosity.

Existing studies draw on several broad families of machine-learning models, each offering different balances of transparency, flexibility, and predictive capacity (Géron 2022; Olukoga et al. 2022; Swamynathan 2017). Linear and polynomial regression provide parsimonious baselines for identifying first-order relationships. Support vector regression (SVR), including radial basis function (RBF) kernels, can represent nonlinear

dependencies and have previously been applied to foam-viscosity prediction under varying operating conditions (Gonzalez Perdomo and Wan Madihi 2022).

Tree-based models, including classification and regression trees (CART) and ensemble extensions, can capture threshold effects and variable interactions, while k-nearest neighbors (KNN) provide flexible local interpolation in experimental data spaces. Neural-network models such as multilayer perceptrons (MLPs) can represent strongly nonlinear, coupled dependencies. These model families therefore provide complementary perspectives on foam behaviour, but their suitability depends on dataset size, validation strategy, and the intended use of the prediction.

Interpretability methods are useful for assessing whether complex ML models rely on physically plausible variables. Model-agnostic approaches such as SHapley Additive exPlanations (SHAP) and permutation feature importance (PFI) attribute changes in model output or performance to individual predictors and allow comparison of variable influence across different model classes (Fisher et al. 2019; Lundberg and Lee 2017). Tree-based models also provide intrinsic importance measures, although these measures should be interpreted as model diagnostics rather than direct evidence of causal physical mechanisms.

Model performance for continuous targets is commonly evaluated using the coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE). R^2 describes the proportion of variance represented by a model, whereas RMSE penalizes larger errors and MAE summarizes the average absolute deviation (Géron 2022; Swamynathan 2017). Error metrics are expressed in the units of the modelled target; in the present study, models trained on log-transformed viscosity are therefore evaluated in logarithmic space before predictions are transformed back to real viscosity for engineering interpretation.

Boosting is an ensemble strategy in which weak learners are trained sequentially so that each new learner focuses on reducing residual error from the preceding ensemble (Géron 2022; Swamynathan 2017). Gradient-boosting implementations, including Extreme Gradient Boosting (XGBoost) and the histogram-based gradient boosting regressor (HistGBR), can perform strongly in nonlinear regression but require appropriate control of model complexity and validation to limit overfitting.

Despite these advances, experimental and data-driven studies retain important limitations. Laboratory foam-rheology datasets relevant to fracturing are commonly generated within equipment-constrained ranges, often below about 120 °C and 3000 psi, and may vary substantially in formulation, gas type, salinity, and measurement protocol (Gonzalez Perdomo and Wan Madihi 2022; Tran et al. 2020; Tran et al. 2023a; Olukoga et al. 2022). Recent reviews indicate that deeper and more demanding subsurface applications can impose more severe thermal, pressure, and salinity stresses, reinforcing the need to state clearly the experimental domain represented by any predictive model (Bai et al. 2025; Gowida et al. 2026).

Collectively, recent research has advanced three related areas: experimentally grounded understanding of foam rheology and stability, nanoparticle–surfactant stabilization of fracturing foams, and data-driven prediction of rheological or hydraulic-fracturing behaviour (Al-Darweesh et al. 2024; Bai et al. 2025; Karami et al. 2025; Olukoga et al. 2022; Tran et al. 2023a, b; Tran et al. 2025; Verma et al. 2025). However, these strands are usually addressed separately, and most ML studies of foam rheology remain predominantly static rather than explicitly modelling temporal viscosity evolution.

Important limitations therefore remain. Experimental studies provide mechanistic and formulation-specific insight but are necessarily restricted to finite laboratory operating ranges, while high-accuracy ML studies may use relatively limited datasets that do not explicitly represent the temporal history of foam degradation. The combined consideration of reservoir-representative elevated pressure and temperature rheology, nanoparticle–surfactant-stabilized CO₂/N₂ systems, and both static and time-aware ML within a single workflow remains uncommon in the literature reviewed here.

In particular, integration of experimentally generated elevated-pressure and elevated-temperature rheology data with time-aware modelling of viscosity evolution remains limited. This constrains the interpretability and transferability of existing predictive tools when rheological history is important. The present study therefore develops an integrated experimental–computational framework that combines foam-rheology measurements, nanoparticle–surfactant chemistry, and complementary static and dynamic ML models within a physically grounded predictive methodology.

2. Methodology

This section describes the experimental dataset, static and dynamic prediction tasks, model-development procedures, validation strategy, diagnostics, feature-importance analysis, and the integrated prediction utility used to estimate energized-foam apparent viscosity under reservoir-representative elevated pressure and temperature conditions.

Experimental data and problem definition

The study incorporated 11 laboratory experiments and 3,963 datapoints involving CO₂ and N₂ foams. Experimental conditions spanned pressures from 947 psi to 3,040 psi and temperatures from 25 °C to 122 °C. The target variable for the machine-learning models was the natural logarithm of apparent viscosity, which improved numerical stability, reduced skewness, and facilitated learning of temperature- and shear-dependent trends. The time-resolved measurements enabled both instantaneous and temporal analyses of foam behaviour.

Two prediction tasks were defined. The static task estimated apparent viscosity from instantaneous thermodynamic, mechanical, and compositional conditions, including pressure, temperature, shear rate, and formulation descriptors, under a quasi-equilibrium assumption. The dynamic task represented time-dependent viscosity evolution and incorporated prior rheological history, reflecting the hypothesis that foam response depends not only on the current state but also on preceding thermal and mechanical exposure.

Accordingly, the static model was used to initialize baseline predictions that feed into the dynamic model, which refines the response over time and provides a more physically representative description of viscosity evolution under reservoir-representative elevated pressure and temperature conditions. This flow is presented in Figure 1.

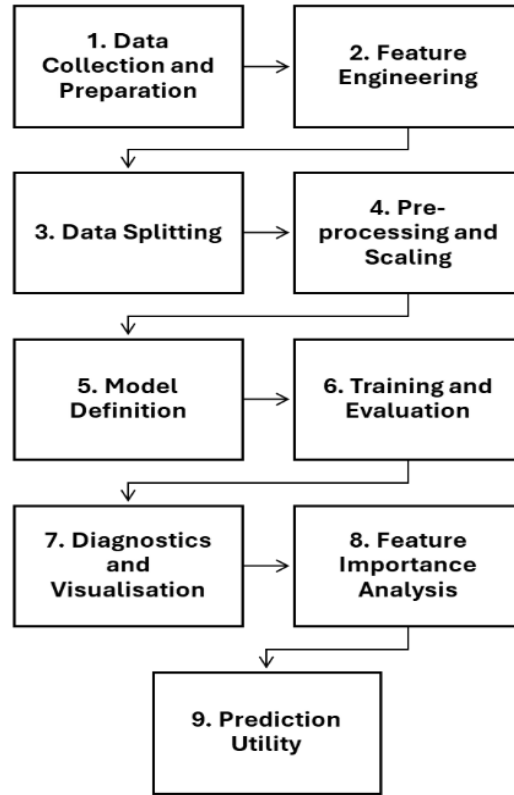


Fig. 1. Methodology outline

Figure 2 summarizes the overall data and model development pipeline. All analysis was conducted in Python using libraries from scikit-learn, XGBoost, and SHAP, following reproducible and statistically robust procedures.

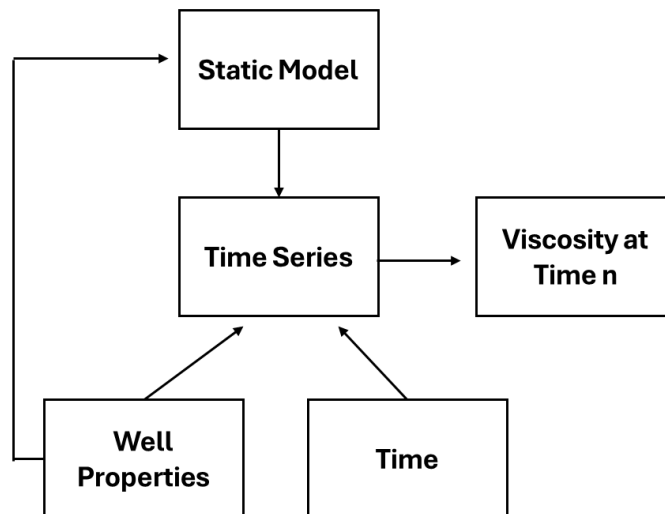


Fig. 2. Model development workflow for static and dynamic viscosity prediction.

Static modelling approach

The static modelling approach predicted energized-foam apparent viscosity from instantaneous system conditions. Input variables represented the thermodynamic and rheological state of each measurement and included temperature (K), inverse temperature (1/K), pressure (psi), shear rate (1/s), surfactant concentration (mL/100 mL), reduced temperature (T_r), reduced pressure (P_r), and time (s). The target variable was the natural logarithm of apparent viscosity, $\log(\mu)$, to reduce skewness and improve numerical stability. Table 1 summarizes the variables included in the static modelling dataset and their descriptive statistics.

Table 1. Description of input variables used in the static modelling framework.

Statistic	Time (s)	T (K)	P (psi)	Shear rate (s^{-1})	μ (cP)	T_r	P_r	Surfactant (mL/100 mL)	1/T (K^{-1})	$\log(\mu)$
count	3963	3963	3963	3963	3963	3963	3963	3963	3963	3963
mean	2913.838	352.3422	1635.9601	225.2615	11.3881	2.1076	2.5216	1.8724	0.0029	2.029
min	10	298.25	947	1	0.0851	0.9802	0.9287	1.4	0.0025	-2.4639
25%	1420	329.35	1010	170	3.82	1.1966	0.9623	2	0.0026	1.3402
50%	2840	363.65	1040	170	7.41	2.3633	2.1125	2	0.0027	2.0028
75%	4260	383.15	3000	206	18.4	2.8839	2.803	2	0.003	2.9124
max	6600	395.15	3040	1000	233	3.1311	6.1548	2	0.0034	5.451
std	1747.1867	33.0577	917.9433	185.9201	10.9713	0.8297	1.6517	0.2456	0.0003	0.9459

As listed in Table 2, the evaluated approaches covered linear, kernel-based nonlinear, neural-network, tree-based, and boosting model families. All models were wrapped within scikit-learn Pipeline objects to ensure identical preprocessing during cross-validation and prediction.

Table 2. Machine learning models used in the study.

Model Type	Algorithms
Linear	Ridge, Lasso
Smooth nonlinear and neural networks	SVR, MLP
Tree-Based	Decision Tree, Random Forest, Histogram Gradient Boosting, XGBoost

Grid search was used for models with relatively small hyperparameter spaces (e.g., Ridge and Lasso), whereas randomized search was used for models with larger search spaces (e.g., MLP, SVR, and XGBoost), as summarized in Table 3. The optimization objective was minimization of RMSE for log-transformed viscosity.

Table 3. Summary of hyperparameter tuning settings.

Model Type	Hyperparameters	Search Range / Values
Ridge / Lasso	Regularization α	[0.001, 0.01, 0.1, 1, 10]
SVR (RBF & Polynomial)	C γ (Gamma)	[0.5, 1, 5, 10] ['scale', 'auto']

	Degree (Polynomial only)	[2, 3]
MLP	Hidden layer sizes	[(20,), (40,), (20, 10), (40, 20, 10)]
	α (Regularization)	[1e-4, 1e-3, 1e-2]
	Learning rate	[1e-3, 5e-4]
	Max iterations	3000
Decision Tree	Max depth	[3, 5, 10, None]
Random Forest	Number of estimators	[200, 400]
	Max depth	[None, 10, 20]
HistGBR	Max iterations	[200, 300, 400]
XGBoost	Number of estimators	[300, 500]
	Max depth	[3, 5, 7]

To evaluate generalization across independent experimental runs, the dataset was split by experiment rather than by individual measurement. Nine of the 11 experiments were allocated to model development and two were retained as unseen experiments. This cross-validation (CV) strategy preserved independence between subsets and reduced leakage from repeated measurements within the same experiment. High-pressure CO₂ and N₂ trials (~3000 psi) were included in model development so that the training data represented the upper-pressure regime available in the experimental dataset.

Figure 3 shows the number of samples contributed by each experiment and identifies the experiments assigned to model development and unseen testing.

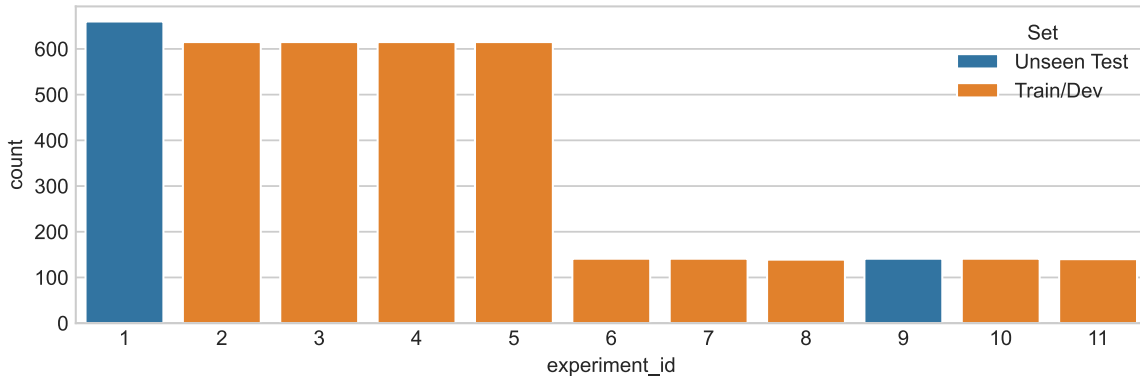


Fig. 3. Distribution of experiments across training and testing datasets.

Figure 4 shows the pressure distribution of the model-development and unseen-test samples. The training/validation data include both the ~1000 psi and ~3000 psi regimes, whereas the unseen-test samples are concentrated near ~1000 psi. This difference in pressure coverage is considered when interpreting unseen-experiment performance.

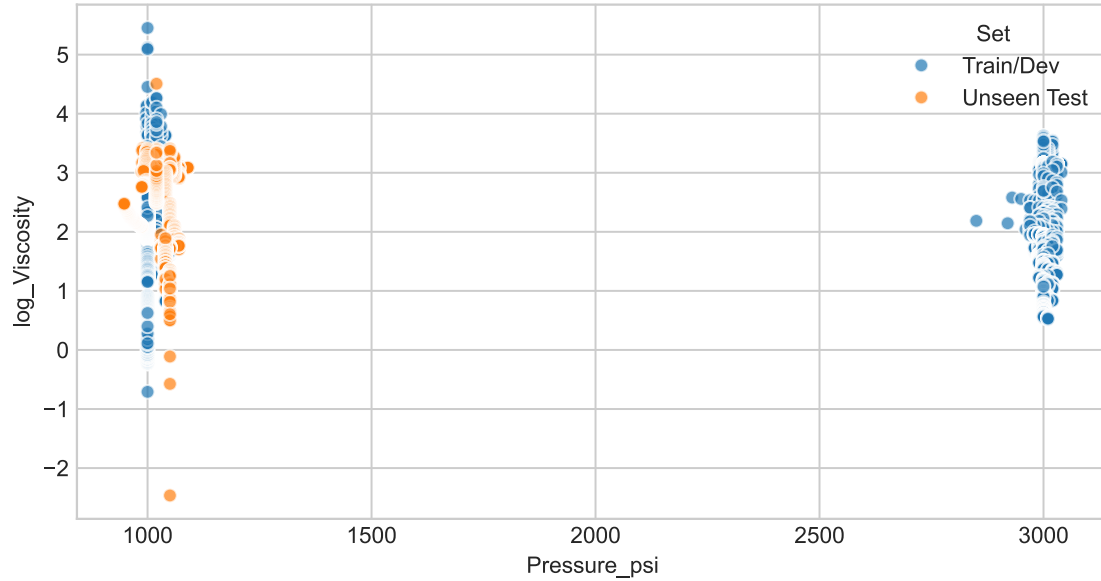


Fig. 4. Distribution of samples across the operational pressure range.

The final static model was selected by considering predictive accuracy, stability, physical plausibility, and suitability for continuous prediction. Tree-based models produced the highest test-set accuracy, but their piecewise structure provides limited basis for extrapolation outside the sampled training domain. Among the non-tree candidates, the multilayer perceptron (MLP) achieved the highest test-set R^2 and lowest errors and was therefore selected for the static branch. Diagnostic checks also showed limited systematic bias, while feature-importance analysis indicated strong sensitivity to temperature-related variables and secondary influence from shear rate and reduced pressure.

Dynamic (time-dependent) modelling

A dynamic modelling framework was developed to account for the time dependence of apparent viscosity observed during the foam-rheology experiments. The objective was to predict viscosity using both instantaneous operating conditions and temporal information.

Dynamic features extended the static inputs with time-dependent descriptors that represent how the foam responds to changing operating conditions. For each experiment, the training features included:

- Rolling means and standard deviations (windows = 3, 5, 10 steps) for T, P, and $\log(\mu)$
- Lagged values of $\log(\mu)$ (lags = 1, 2, 3)
- Normalized delta features (ΔT , ΔP , $\Delta \gamma$) and normalized time.

This formulation preserved causality and allowed the model to learn temporal patterns associated with viscosity evolution. Rows lacking valid lag data were discarded to prevent temporal leakage, yielding a chronologically consistent dataset for sequence-aware learning. Foam quality was not available as a continuous measurement during dynamic testing and was therefore not explicitly represented in the dynamic model. Surfactant concentration remained constant within each experiment and was retained as a compositional descriptor. Gas phase was represented through the critical temperature and critical pressure of the injected gas rather than categorical indicators, allowing the model to respond continuously to thermodynamic differences between CO_2 and N_2 systems. Other potentially influential variables, including brine salinity and nanoparticle loading, were unavailable as time-varying model inputs and were not included.

For the dynamic modelling branch, data within each experiment were divided chronologically to preserve time ordering. As illustrated in Figure 5, the earlier portion of each experiment was used for model development, while the final portion was retained for internal temporal testing.

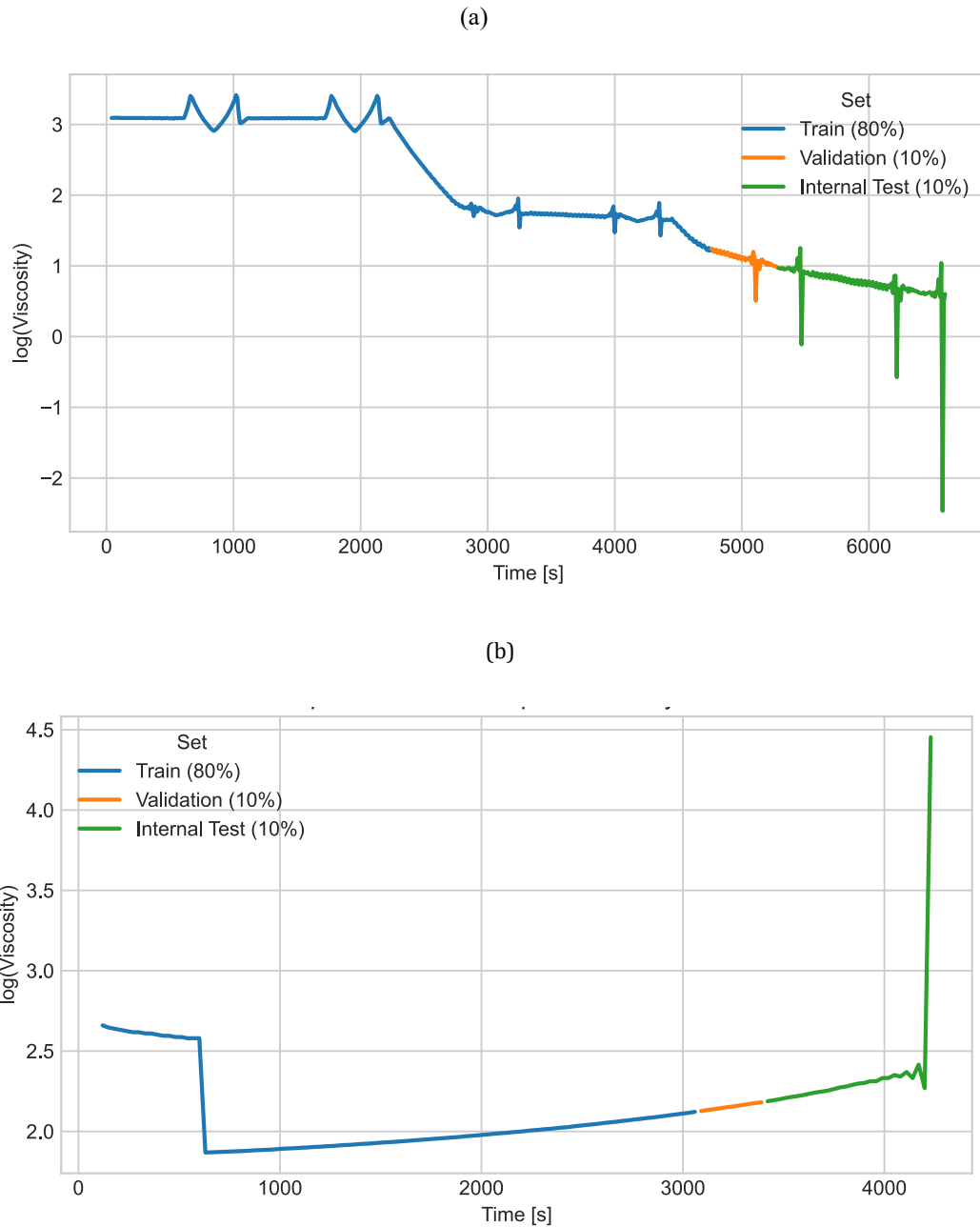


Fig. 5. Temporal data split within representative training experiments.

For the dynamic branch, Lasso regression was selected because it produced the lowest RMSE and MAE within the recursive forecasting framework while maintaining stable and physically plausible viscosity trajectories. Lasso regularization constrained coefficient growth and reduced amplification of prediction error during sequential forecasting.

Model training and evaluation

Model performance was evaluated in logarithmic and real-viscosity space using the coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE). Real-space evaluation was obtained by exponentiating predictions from $\log(\mu)$, enabling viscosity comparisons in field units. In addition to pointwise accuracy, temporal stability was evaluated by examining prediction trajectories and the variability of validation metrics across experiments. Static models were trained using five-fold cross-validation within the model-development subset to support consistent model comparison.

Dynamic models were trained using a walk-forward (recursive forecasting) strategy. At each prediction step, the model-predicted log-viscosity was fed back as an input for subsequent time steps, and rolling statistics were updated sequentially to reflect the evolving prediction history. Training was performed incrementally, beginning with 20% of each experiment for model fitting and 10% for validation, and progressively expanding the training window until 80% of the experiment was included. This approach intentionally allowed prediction errors to propagate over time, providing a realistic assessment of model stability and suitability for sequential viscosity forecasting. For the dynamic modelling framework, predictions were generated recursively and reintroduced as lagged inputs, while rolling means and standard deviations were recalculated dynamically to emulate sequential experimental forecasting. RMSE and MAE were emphasized as the primary evaluation metrics, as they directly quantify the magnitude of sequential prediction errors and penalize large deviations that could distort the physical realism of predicted viscosity trajectories.

Because sequential prediction can accumulate error, RMSE and MAE were emphasized for dynamic model selection alongside inspection of trajectory shape and physical plausibility. Models intended for the final coupled utility were required to remain numerically stable under recursive forecasting. Tree-based models were retained as diagnostic benchmarks but were not used in the final coupled static–dynamic prediction framework.

Model diagnostics

Model diagnostics were conducted to verify data integrity, assess model calibration, and evaluate temporal forecasting behavior. Prior to training, the distribution of log-transformed viscosity was examined to confirm suitability for modelling, and all input features were screened to ensure completeness and valid physical ranges. Training and testing subsets were compared to confirm overlapping thermodynamic domains. For dynamic models, time ordering within each experiment was verified to prevent data leakage.

Following training, predicted–measured scatter plots and residual distributions were used to assess calibration, bias, and error structure. For the dynamic framework, time-series comparisons of predicted and measured viscosity were examined to assess whether the forecast trajectories remained physically plausible under changing shear and temperature conditions. These diagnostics complemented the numerical performance metrics and informed final model selection.

Feature-importance analysis

To interpret model behavior and quantify variable influence, feature-importance analyses were performed using three complementary approaches. SHAP (SHapley Additive exPlanations) was applied to decompose predictions into feature contributions using a representative subset of the training data. Permutation Feature Importance (PFI) was evaluated on the test set to assess sensitivity of model performance to individual variables. For tree-based models included for benchmarking, intrinsic impurity-based feature importance was computed to provide a reference comparison. Feature-importance rankings were then generated for the final static and dynamic models to enable consistent comparison of variable influence across instantaneous and time-dependent formulations.

Integrated prediction utility

The integrated prediction utility couples the selected static and dynamic models in a single simulation framework. It accepts user-defined pressure, temperature, gas phase, surfactant concentration, shear, and time histories and reconstructs the engineered features required by the models. Initial apparent viscosity is estimated using the static model, after which dynamic predictions are propagated recursively to represent viscosity evolution through the prescribed operating sequence. The utility outputs predicted viscosity trajectories and derived power-law parameters for scenario-based evaluation. Its use is intended for screening and interpretation within or near the experimental domain represented by the training data rather than as a substitute for independent field validation.

3. Results

In this section, the performance of both static and dynamic regression models for predicting foam viscosity is presented and analyzed. Static models are first evaluated to assess their ability to capture instantaneous physical dependencies, followed by dynamic models that forecast temporal behavior using lagged and rolling inputs. Feature importance analyses and physical interpretations highlight the key variables driving viscosity, while illustrative predictions demonstrate the practical utility of the developed framework. Finally, the engineering implications of the model's predictive capability under HPHT and dynamic shear conditions are discussed.

Static model performance

A variety of regression models were developed and tested, including Ridge, Lasso, Support Vector Regression (SVR), Random Forest, Multilayer Perceptron, and XGBoost. Their performance was assessed using standard evaluation measures such as R^2 , RMSE, and MAE, as illustrated in Table 4.

Table 4. Static model performance in logarithmic space.

Model	Type	R^2 (Train)	CV (log)	RMSE	R^2 (Test)	Test RMSE	Test MAE
<i>Logarithmic Space</i>							
HistGBR	Tree	0.9983	0.084		0.9268	0.239	0.145
Random Forest	Tree	0.9987	0.090		0.9197	0.250	0.163
Decision Tree	Tree	1.0000	0.121		0.9139	0.259	0.164
XGBoost	Tree	0.9997	0.064		0.9116	0.263	0.170
MLP	Non-Tree	0.9788	0.152		0.8844	0.300	0.209
Lasso	Non-Tree	0.9676	0.178		0.8218	0.373	0.237
SVR_RBF	Non-Tree	0.9828	0.129		0.8178	0.377	0.226
Ridge	Non-Tree	0.9765	0.153		0.6186	0.546	0.307
SVR_Poly	Non-Tree	0.9804	0.135		0.4548	0.652	0.383
<i>Real Space</i>							
HistGBR	Tree				0.9107	2.578	1.555
DecisionTree	Tree				0.9082	2.614	1.839
RandomForest	Tree				0.8483	3.361	1.922
XGBoost	Tree				0.7761	4.083	2.319
MLP	Non-Tree				0.7705	4.134	2.510
Lasso	Non-Tree				0.4194	6.575	3.544
SVR_RBF	Non-Tree				0.3448	6.984	3.424

Ridge	Non-Tree	-3.4951	18.294	7.058
SVR Poly	Non-Tree	-5.2107	21.504	9.041

Among these models, Tree-based models achieved the highest overall accuracy in log space, but they were not considered suitable for deployment because they cannot extrapolate beyond the training domain. Among the extrapolative models, the Multilayer Perceptron (MLP) achieved the highest test-set R^2 and lowest error values, making it the best candidate for capturing the nonlinear dependence of viscosity on temperature, pressure, and shear rate.

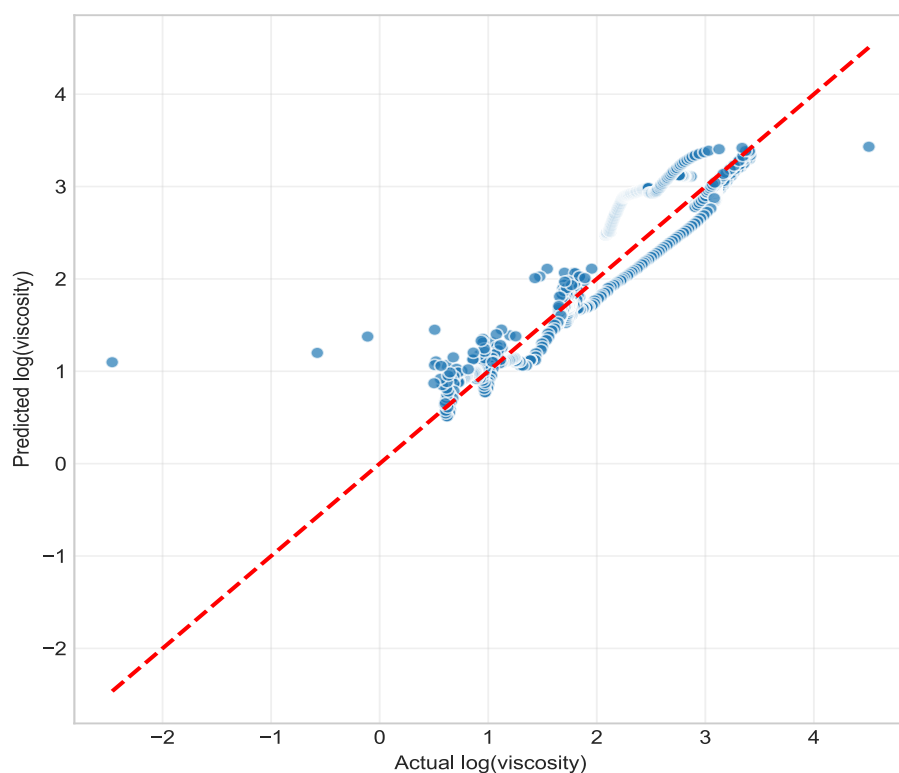


Fig. 6. Actual versus predicted viscosity in logarithmic space using the MLP model.

Figure 6 compares predicted and measured $\log(\text{viscosity})$ for the selected MLP model. Most points lie close to the 1:1 line across the central viscosity range, although larger deviations occur at sparsely sampled extremes. The pattern indicates good overall calibration within the evaluated test data while also identifying regions in which prediction uncertainty is greater.

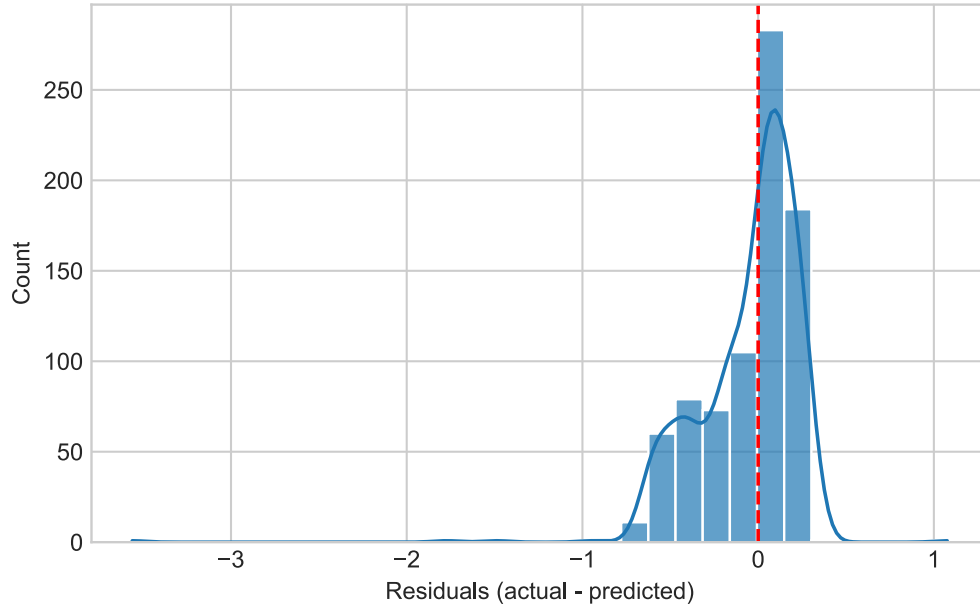


Fig. 7. Residual distribution for actual and predicted viscosity in logarithmic space.

The residual distribution in Figure 7 is concentrated near zero, indicating limited overall bias. Because residuals are defined as actual minus predicted viscosity, the small negative tail corresponds to isolated overpredictions rather than underpredictions. The asymmetric tail nevertheless indicates that a small number of observations are predicted less accurately than the main body of the data.

Overall, the static results show that instantaneous physical inputs provide useful viscosity estimates and that the MLP offers the strongest performance among the non-tree candidates considered for the coupled framework. The residual outliers also indicate that the model should not be treated as uniformly accurate across all portions of the experimental domain.

Dynamic prediction performance

The dynamic models were evaluated with additional caution because recursive forecasts use lagged and rolling inputs derived from preceding predictions. Small step-wise deviations can therefore accumulate through a sequence. Table 5 summarizes validation performance across the evaluated dynamic models, while Table 6 reports Lasso performance for internal temporal testing and completely unseen experiments. In this setting, RMSE and MAE are emphasized alongside R^2 because the magnitude and accumulation of sequential error are central to forecast stability.

Table 5. Summary of validation metrics for dynamic model performance.

Model	Mean R^2	Mean RMSE	Mean MAE
Lasso	0.1621	0.060	0.050
SVR_Poly	-82.3638	0.308	0.278
Ridge	-292.2470	0.574	0.414
SVR_RBF	-945.8880	0.396	0.359
MLP	-2082.7400	2.438	1.937

As shown in the validation summary in Table 5, the Lasso model achieved the most stable performance among the evaluated approaches, with the lowest RMSE and MAE values. In contrast, models such as the MLP and SVR exhibited larger accumulated errors during recursive prediction.

Table 6. Summary of testing metrics for dynamic model performance.

Test	R^2	RMSE	MAE
Internal 80-100%	0.436	0.185	0.138
Unseen experiments	-0.190	0.473	0.449

Table 6 shows that Lasso performance declined when moving from internal temporal testing to completely unseen experiments. For the internal test segment, $R^2 = 0.436$, $RMSE = 0.185$, and $MAE = 0.138$, whereas the unseen experiments yielded $R^2 = -0.190$, $RMSE = 0.473$, and $MAE = 0.449$. The negative unseen-experiment R^2 indicates limited pointwise generalization across independent experiments, even though the model retained aspects of the broader temporal trend. This distinction is important when interpreting the dynamic model as a forecasting aid rather than a universally transferable predictor.

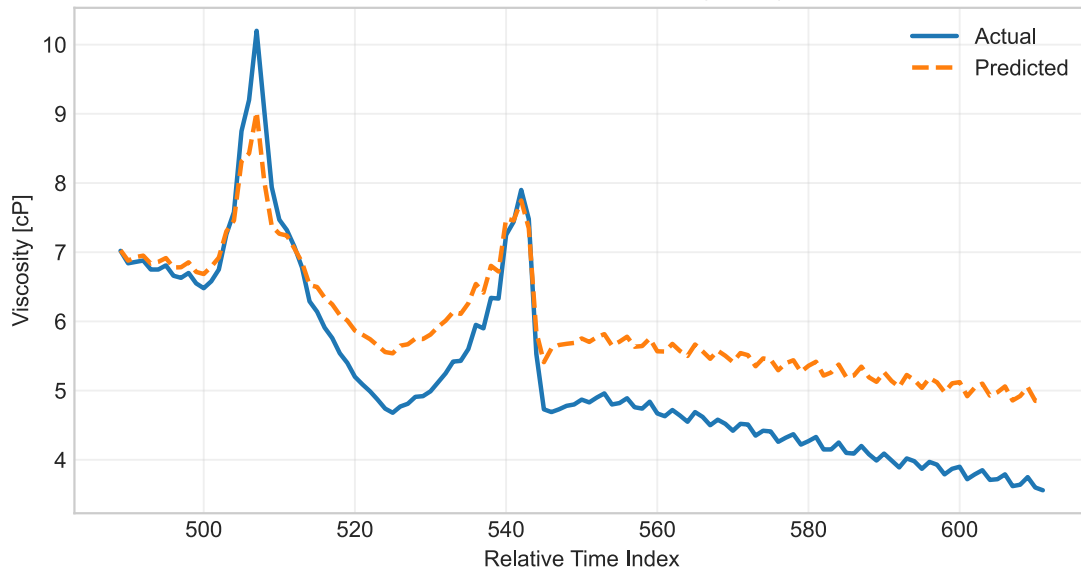


Fig. 8. Internal dynamic forecast using the Lasso regression model.

Figure 8 illustrates an internal temporal forecast for experiment 3. The predicted viscosity follows the broad trajectory of the measured series but smooths some of the sharper peaks. This behaviour is consistent with the moderate internal-test R^2 and relatively low RMSE and MAE values and indicates stable recursive forecasting within the represented experiment.

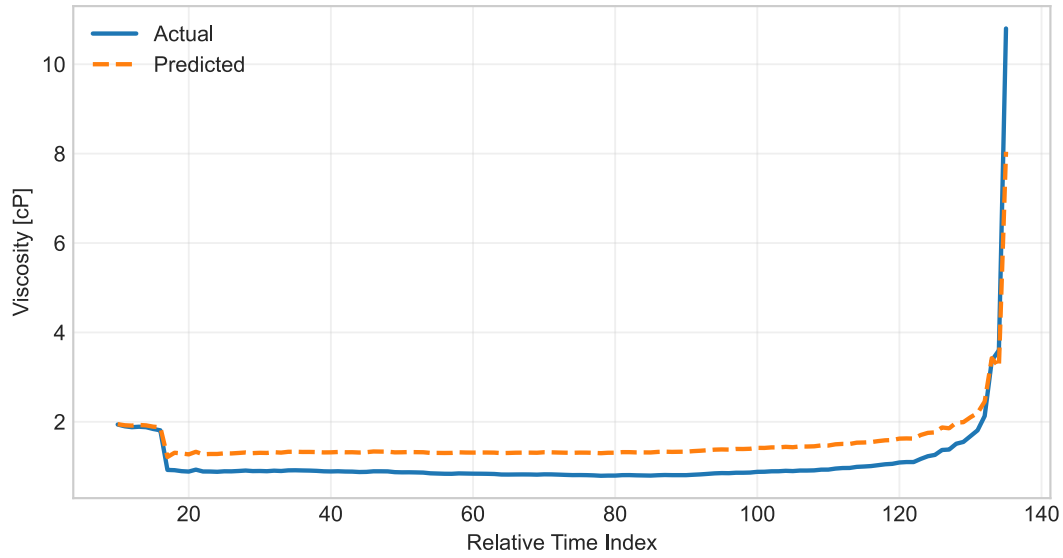


Fig. 9. Unseen dynamic forecast using the Lasso regression model.

Figure 9 shows the forecast for experiment 8, which was not used for model development. The Lasso prediction follows the broad shape of the viscosity trajectory but exhibits systematic offsets over much of the sequence and underestimates the final peak. Accordingly, the figure supports qualitative trend retention but also illustrates the reduced pointwise accuracy quantified for unseen experiments in Table 6.

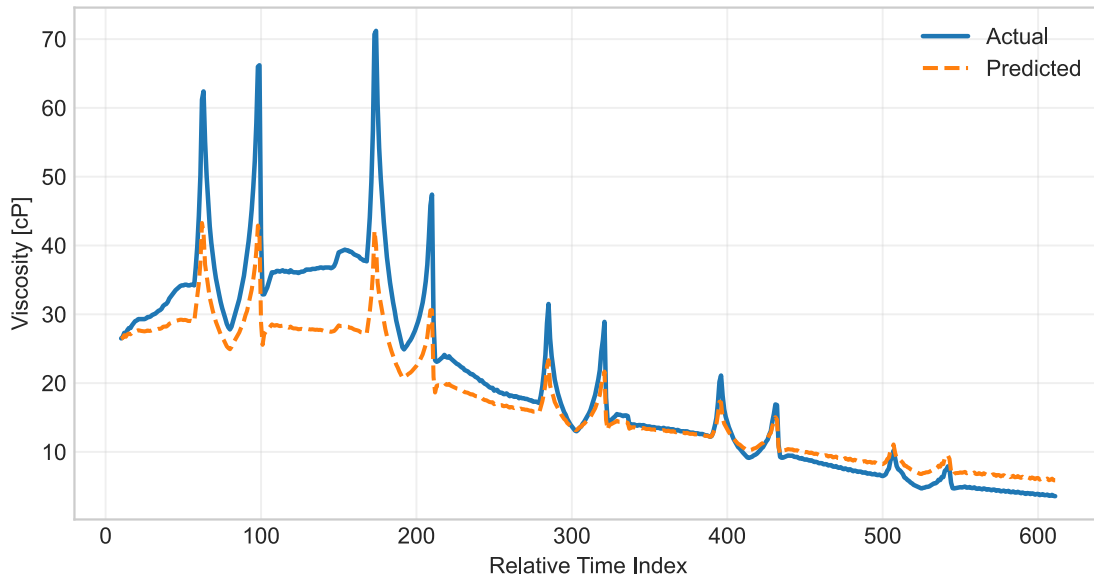


Fig. 10. Full-sequence viscosity forecast for a representative experiment.

Figure 10 presents a full-sequence forecast for experiment 3. The Lasso model reproduces the overall decline in viscosity and several local changes, while smoothing or underestimating some high-viscosity peaks. This representative sequence demonstrates stable recursive behaviour within a well-represented experiment, but it should not be interpreted as evidence of equivalent accuracy for all unseen operating conditions.

Feature importance and physical interpretation

Three complementary approaches were used to interpret model behaviour: permutation feature importance, tree-based intrinsic importance, and SHAP analysis. Figures 11 and 12 summarize the static-model importance results, while Figure 13 presents the corresponding dynamic-model analyses and the comparison of leading static and dynamic features.

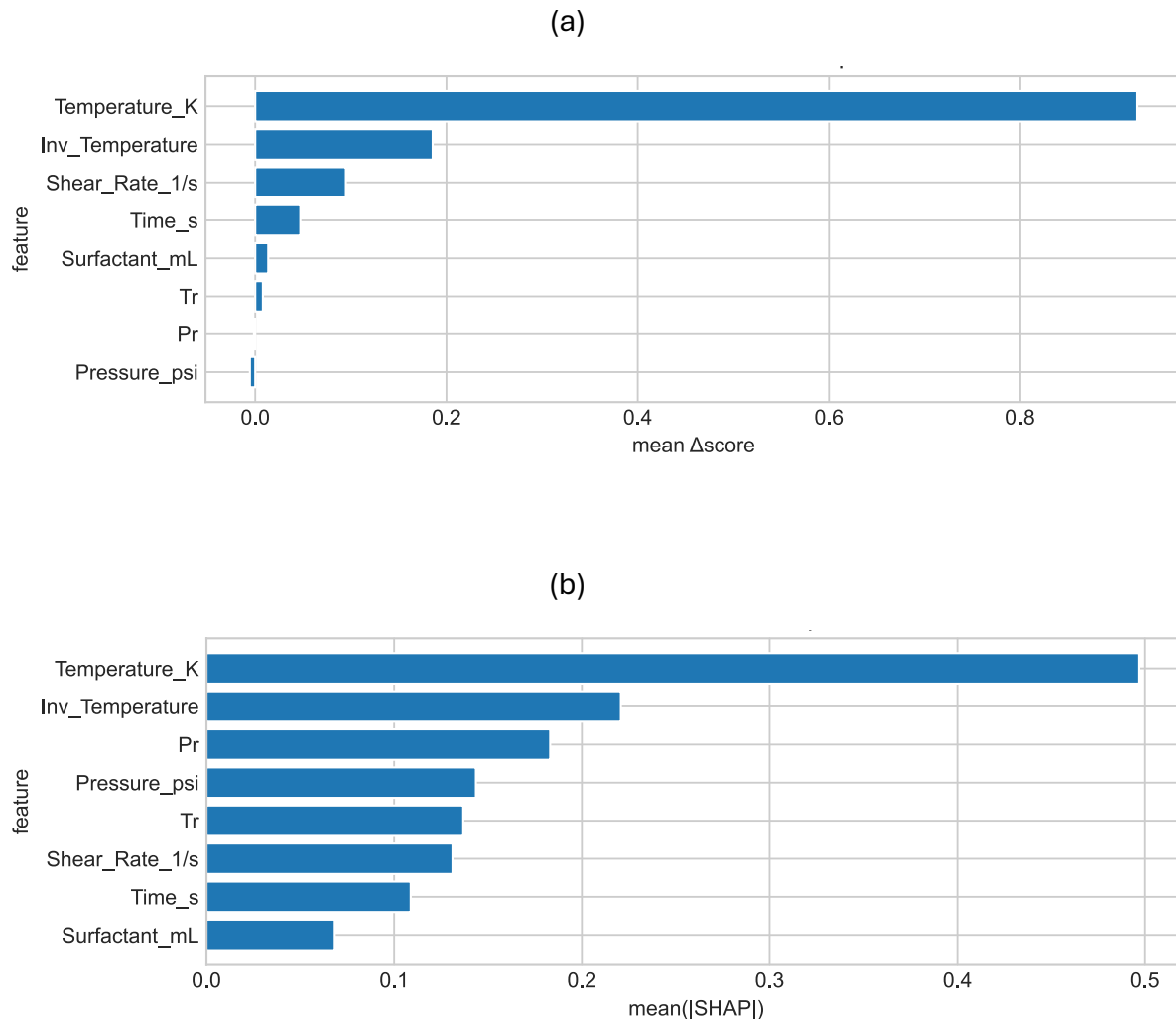


Fig. 11. Feature-importance analysis for the static MLP model: (a) Permutation importance, (b) SHAP importance

For the static MLP, both SHAP and permutation analyses identify temperature and inverse temperature as the dominant predictors, with shear rate and reduced pressure providing smaller contributions (Figure 11). This ordering is consistent with strong thermal sensitivity of foam viscosity within the experimental dataset. The random-forest intrinsic importance in Figure 12 emphasizes a different combination of variables, illustrating that importance rankings are model dependent and should be interpreted as diagnostics rather than universal physical rankings.

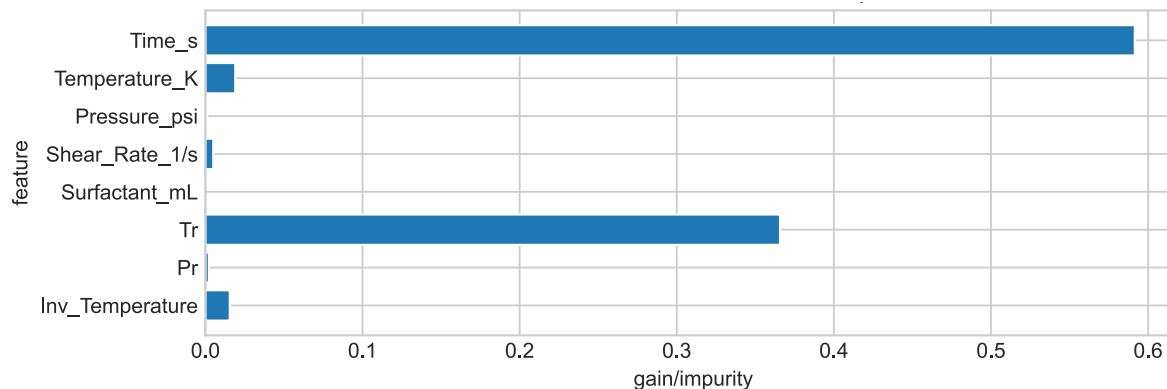


Fig. 12. Random forest intrinsic feature importance for the static modelling framework.

Figure 13 shows that lagged viscosity and rolling viscosity descriptors dominate the dynamic-model importance rankings. This result is consistent with the sequence-aware formulation: the model relies strongly on recent rheological history to predict subsequent viscosity. The importance rankings support the role of temporal information in the dynamic model but do not, by themselves, establish causal degradation mechanisms.

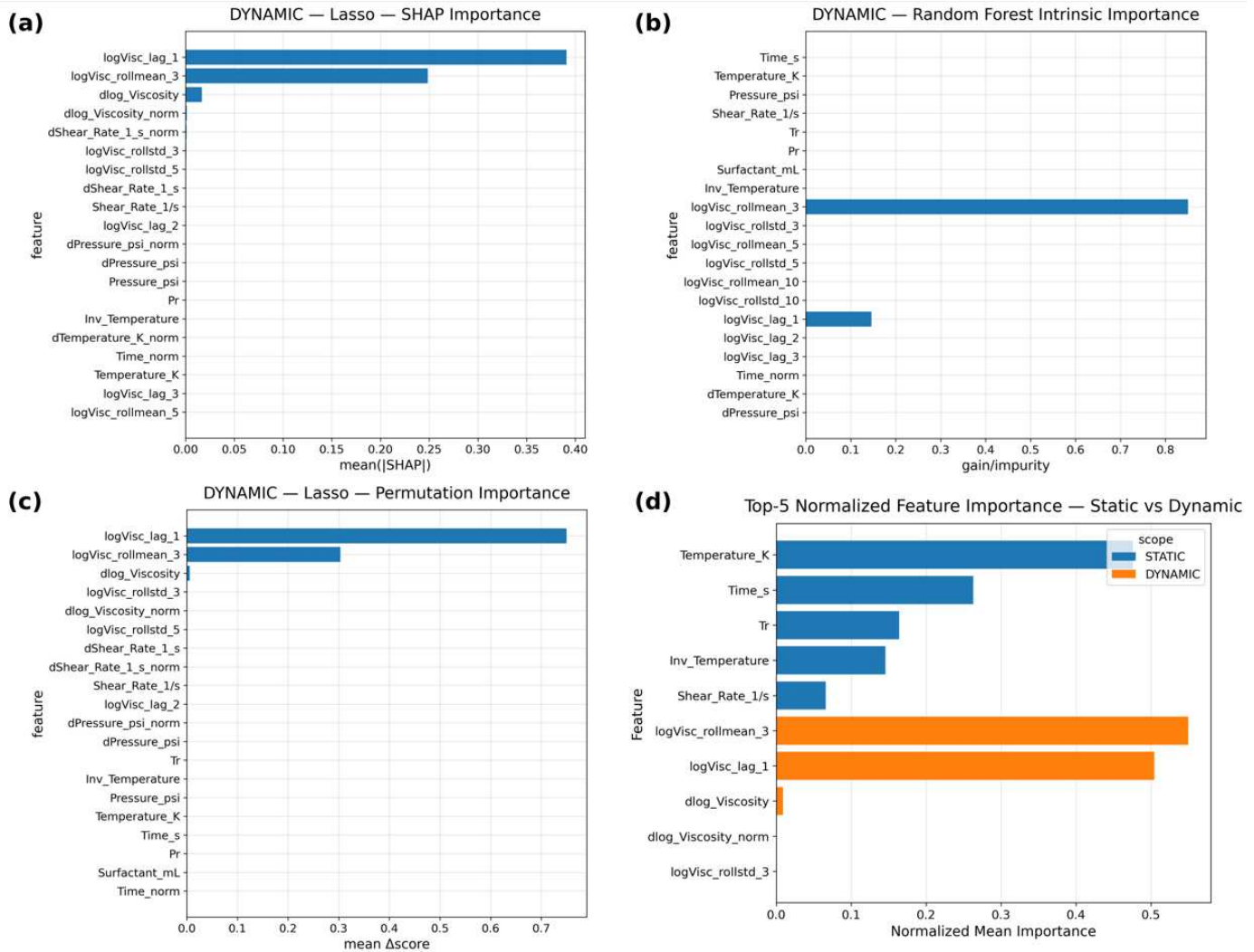


Fig. 13. Dynamic model feature-importance analysis: (a) SHAP importance, (b) random forest intrinsic importance, (c) permutation importance, and (d) comparison of the top five static and dynamic features.

Prediction utility

The following examples illustrate how the coupled static–dynamic framework can be used to explore viscosity response under prescribed temperature and shear histories. These plots are scenario-based demonstrations of the prediction utility; they are not independent field validations of the model.

Figure 14 shows the experimental N₂-foam viscosity profile used as a reference for comparison. The profile captures the temporal evolution of viscosity, including transient peaks and an overall decline associated with the imposed test sequence.

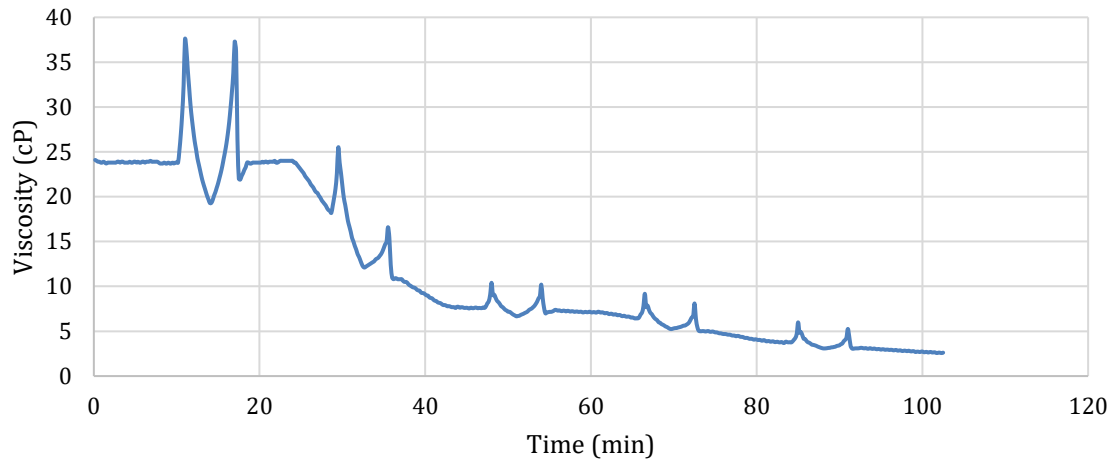


Fig. 14. Experimental viscosity profile of N₂ foam at 3000 psi.

Figure 15 shows the synthetic temperature profile used for the prediction-utility example, increasing from approximately 30 °C to 120 °C.

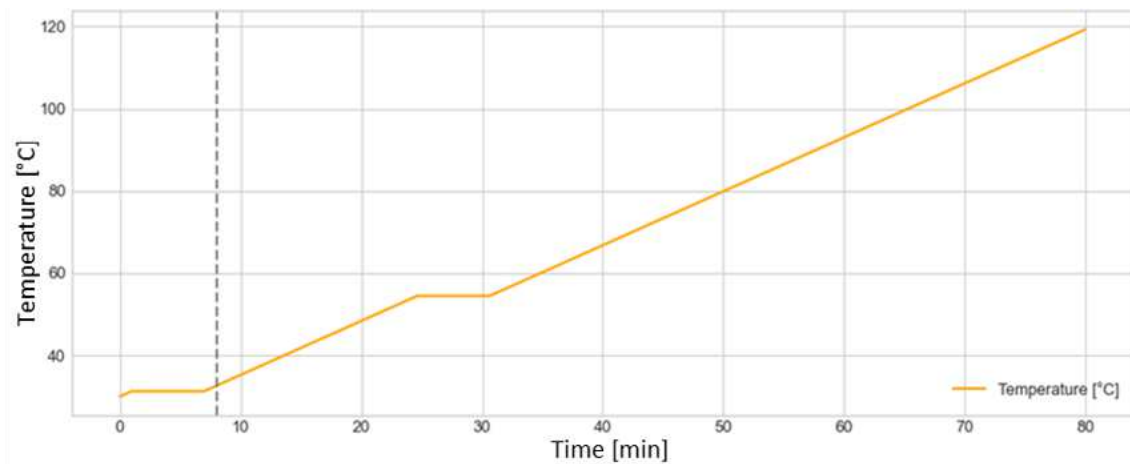


Fig. 15. Simulated temperature increase from surface to reservoir conditions.

Figure 16 shows the two synthetic shear-rate ramps used to test model response during changing and quasi-steady shear conditions.

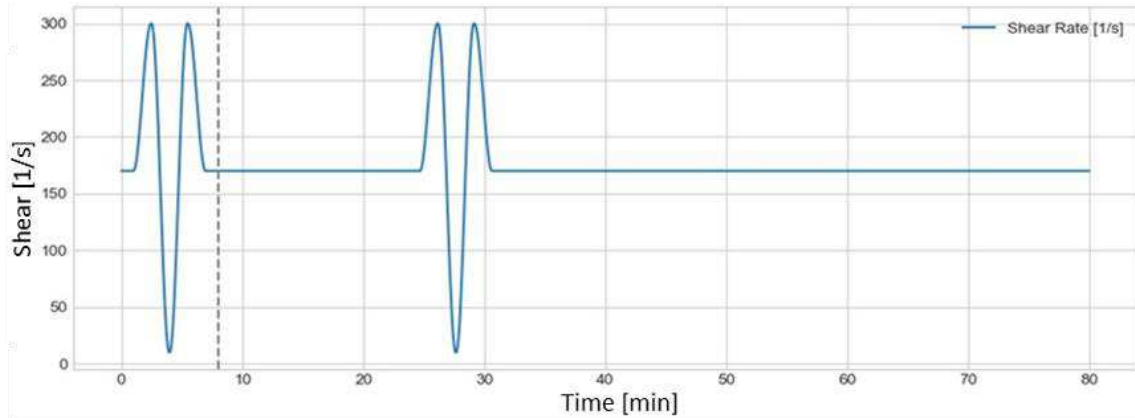


Fig. 16. Synthetic operational ramp profiles used for predictive evaluation.

Figure 17 compares static–dynamic predictions using Lasso and Ridge in the dynamic branch. The Lasso trajectory shows a smooth decrease in viscosity that is consistent with the expected thermal trend and the experimental reference profile, whereas Ridge develops a divergent trajectory under the same recursive sequence.

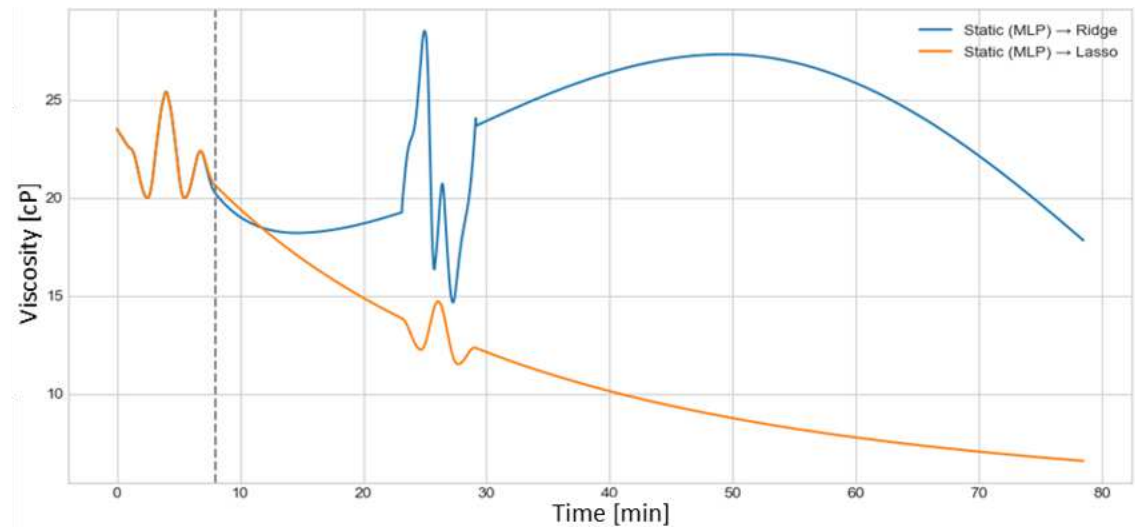


Fig. 17. Static–dynamic viscosity prediction for N₂ foam at 3000 psi.

The Ridge forecast in Figure 18 departs from the expected viscosity trend during recursive prediction, illustrating poorer sequential stability than Lasso for this example. This behaviour supports the selection of Lasso for the dynamic branch of the coupled utility.

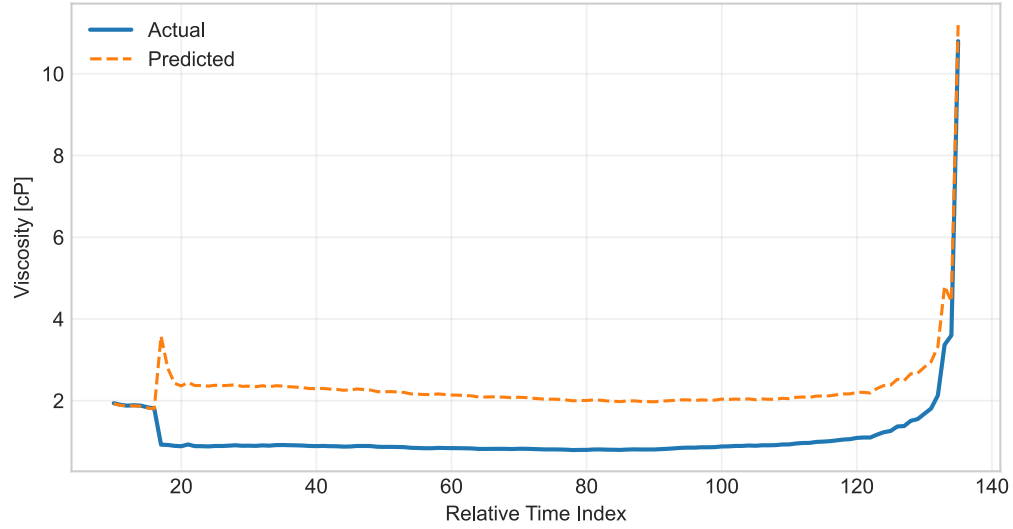


Fig. 18. Ridge regression forecast for the full experimental sequence.

Figure 19 presents shear stress versus shear rate together with the fitted power-law parameters n and k for the two ramp segments. These parameters provide compact descriptors of the predicted shear-dependent rheological response generated by the utility.

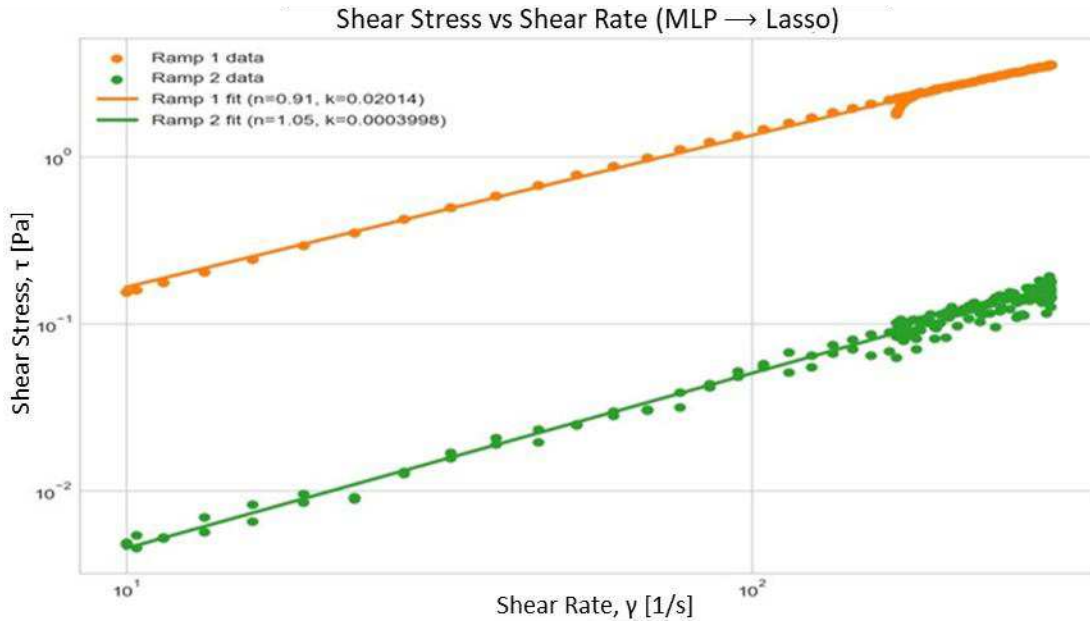


Fig. 19. Shear stress versus shear rate and corresponding rheological parameters (n and k).

Taken together, the scenario plots show that the coupled framework can generate stable and physically interpretable viscosity trajectories when the selected Lasso dynamic model is used. The results support the utility of combining static initialization with temporal forecasting for scenario screening, while the unseen-experiment metrics show that broader validation is required before field-scale predictive use.

4. Engineering implications

The coupled framework enables forward evaluation of foam apparent viscosity under prescribed temperature and shear histories, moving the analysis beyond isolated steady-state measurements. Its most defensible use is as a screening and decision-support tool within, or close to, the experimental domain represented in the training data rather than as a validated predictor for arbitrary field conditions.

For formulation development, the framework can compare predicted viscosity decay under common operating scenarios and help identify formulations or operating windows that warrant targeted laboratory verification. Because the model preserves key dependencies on temperature, shear, and recent viscosity history, it can be used to examine how rheological response evolves through a sequence rather than at a single operating point.

For elevated-temperature and elevated-pressure design, the prediction utility can support sensitivity analysis of prescribed downhole-like temperature and shear profiles. However, the unseen-experiment results demonstrate that quantitative accuracy decreases when experimental conditions differ from those used for model development. Consequently, model output should be interpreted together with the represented training domain and should not replace laboratory or field validation for new formulations or operating regimes.

A further benefit is the potential to reduce the number of exploratory experiments required during early-stage formulation screening. Machine-learning predictions can be used to prioritize conditions for confirmatory testing, while experimental measurements remain essential for calibration, validation, and extension of the model domain.

Overall, the static–dynamic approach provides an interpretable framework for integrating experimental rheology with sequential prediction. Its value lies in organizing and extending information from the available tests, while its current limitations; particularly the small number of independent experiments and absence of continuously measured foam quality, salinity, and nanoparticle loading—define clear priorities for further validation.

5. Discussion

The results show that complementary static and dynamic machine-learning models can represent different aspects of energized-foam rheology within the available experimental dataset. The static model captures nonlinear relationships among instantaneous operating variables, whereas the dynamic model introduces recent viscosity history and rolling descriptors to represent sequential behaviour. The results also show that model stability and physical plausibility must be considered alongside pointwise accuracy, particularly when predictions are extended to independent experiments.

Static versus dynamic modelling performance

The static results indicate strong nonlinear dependence of apparent viscosity on the measured operating conditions. Tree-based models achieved the highest test-set accuracy, but the MLP provided the best performance among the non-tree models selected for deployment and offered a smooth functional mapping for the coupled utility. This selection therefore reflects a balance between predictive performance and intended model use rather than an assertion that the MLP is universally superior to tree-based methods.

Dynamic forecasting produced a different performance pattern. Recursive prediction allows error to accumulate because each predicted viscosity contributes to later lagged and rolling features. Lasso provided the most stable recursive behaviour among the evaluated models, but its performance declined on completely unseen experiments ($R^2 = -0.190$; $RMSE = 0.473$; $MAE = 0.449$). Thus, the dynamic model is better

interpreted as a stable trend-forecasting component for scenario analysis than as a high-accuracy predictor across arbitrary independent experiments.

Physical interpretation of feature importance

Feature-importance analysis provides additional insight into model behaviour. Temperature-related variables dominated the static MLP rankings, while shear rate and reduced pressure made smaller contributions. In the dynamic model, lagged viscosity and rolling viscosity descriptors were dominant, confirming that recent rheological history is central to the model formulation. These rankings are physically plausible, but they describe how the fitted models use the available predictors and should not be interpreted as direct causal evidence of individual foam-degradation mechanisms.

The strong influence of temporal descriptors is consistent with the observed path dependence of foam response, in which drainage, coalescence, and structural rearrangement can evolve through time. However, the present dataset does not isolate these mechanisms individually, and variables such as foam quality, salinity, and nanoparticle loading were not available as continuous dynamic inputs. Mechanistic interpretation should therefore remain linked to the experimental evidence rather than inferred solely from feature-importance rankings.

Implications for predictive modelling under reservoir-representative elevated pressure and temperature conditions

A key contribution of this work is the separation and subsequent coupling of static and dynamic prediction tasks. The static model provides an initial viscosity estimate from instantaneous thermodynamic and compositional descriptors, while the dynamic model updates the trajectory using temporal history. This architecture offers a practical way to distinguish state-dependent and history-dependent information within one workflow and supports scenario evaluation without claiming that the model replaces full-physics or experimental validation.

The results also highlight important limitations. Performance degradation on unseen experiments indicates sensitivity to experimental conditions not fully represented in the training data, particularly variables such as foam quality, salinity, and nanoparticle concentration that were unavailable for dynamic measurement. These limitations are not failures of the modelling approach, but rather reflections of data constraints. Incorporating additional experimentally measurable features would likely improve both static accuracy and dynamic robustness.

Engineering relevance and deployment considerations

From an engineering perspective, the integrated utility provides a structured method for comparing viscosity evolution across prescribed operating scenarios and prioritizing conditions for further testing. Its use should remain within or near the represented experimental domain unless additional validation demonstrates transferability to more extreme or compositionally different conditions. This limitation is particularly important because the current dataset comprises only 11 independent experiments.

Importantly, the results caution against indiscriminate use of complex ML models in operational settings. Models intended for deployment must be evaluated not only on accuracy, but also on stability, interpretability, and physical consistency. In this regard, the final selection of MLP for static prediction and Lasso for dynamic forecasting reflects a deliberate balance between model flexibility and robustness.

Positioning within existing literature

The present findings extend a literature base that has largely developed along separate experimental and data-driven directions. Recent studies have advanced nanoparticle–surfactant fracturing-fluid design and foam

stability (Bai et al. 2025; Tran et al. 2023a, b; Verma et al. 2025), quantified the effect of foam rheology on proppant placement (Tran et al. 2025), and applied ML to foam-rheology prediction or hydraulic-fracturing optimization (Al-Darweesh et al. 2024; Olukoga et al. 2022; Karami et al. 2025). Broader reviews also identify increasing use of hybrid and data-driven models in porous-media systems (Delpisheh et al. 2024), while recent JPEPT work highlights both ML-assisted hydraulic-fracturing design and continuing stability challenges in foam systems (Gowida et al. 2026; Karami et al. 2025). Against this background, the main contribution of the present work is the explicit coupling of static and time-aware viscosity models using experimentally generated CO₂/N₂ foam rheology data, together with an interpretation of both model utility and generalization limits.

6. Future work and recommendations

While this study demonstrates the viability of machine learning for predicting the apparent viscosity of energized foams under reservoir-representative elevated pressure and temperature conditions, several opportunities exist to further enhance model capability and robustness.

Future research should explore advanced sequence-based deep learning architectures, such as recurrent neural networks and transformer models, to better capture complex temporal dependencies in viscosity evolution during multi-stage or long-duration experiments. Expanding and diversifying the experimental dataset is also recommended, particularly through the inclusion of additional elevated temperature and pressure measurements across a wider range of gas systems, salinity levels, and nanoparticle concentrations, which would improve model generalizability and enable more rigorous validation.

Incorporating physics-informed modelling approaches represents another promising direction. Embedding rheological and thermodynamic constraints, such as power-law or Herschel–Bulkley formulations, within the machine learning framework could improve extrapolation reliability under untested reservoir-representative elevated pressure and temperature conditions while enhancing physical interpretability. In addition, the integration of uncertainty quantification techniques, such as bootstrapping, Bayesian inference, or Monte Carlo dropout, would allow confidence bounds to be placed on viscosity predictions, supporting risk-informed engineering decisions.

Finally, further analysis of feature interactions and explainability is warranted. Building on the current SHAP and permutation-based analyses, future work could investigate coupled effects between temporal descriptors and thermodynamic variables to refine feature selection, reduce model complexity, and improve predictive efficiency.

7. Conclusions

This study developed and evaluated an integrated static–dynamic machine-learning framework for predicting the apparent viscosity of energized foams under reservoir-representative high-pressure, high-temperature (HPHT) conditions. The principal conclusions are:

- The experimental dataset provided a robust basis for complementary static and dynamic modelling, comprising 11 CO₂ and N₂ foam experiments and 3,963 datapoints spanning 947–3,040 psi and 25–122 °C.
- For static viscosity prediction, the multilayer perceptron (MLP) provided the strongest performance among the deployment-oriented non-tree models, achieving a test coefficient of determination (R²) of 0.8844 and a root mean square error (RMSE) of 0.300 in logarithmic space. Tree-based models

607 achieved higher in-domain numerical accuracy but were retained as benchmarks rather than used in
 608 the final coupled utility.

- 609 • For dynamic forecasting, Lasso regression provided the most stable recursive predictions, with a
 610 mean RMSE of 0.060 and mean absolute error (MAE) of 0.050 in logarithmic space during
 611 validation. Performance decreased on completely unseen experiments ($R^2 = -0.190$; RMSE = 0.473;
 612 MAE = 0.449), indicating that the dynamic model retains broad temporal trends more reliably than
 613 pointwise accuracy across independent experiments.
- 614 • Feature-importance analysis identified distinct controls on static and dynamic predictions.
 615 Temperature-related variables dominated the static model, with shear rate and reduced pressure
 616 providing secondary contributions, whereas lagged viscosity and rolling viscosity statistics
 617 dominated dynamic predictions, confirming the importance of rheological history in time-dependent
 618 foam behavior.
- 619 • The principal innovation of this work is the integration of experimentally generated reservoir-
 620 representative foam rheology, nanoparticle–surfactant-stabilized CO₂ and N₂ systems, and
 621 complementary static and time-aware machine-learning models within a unified predictive
 622 framework. Within the literature reviewed, this combination is uncommon and provides a structured
 623 way to couple instantaneous and history-dependent rheological information.
- 624 • The framework provides a basis for screening foam formulations and evaluating viscosity evolution
 625 under prescribed operating conditions, but its predictions remain constrained by the experimental
 626 domain and the limited number of independent tests. Continuous measurements of variables such
 627 as foam quality, salinity, and nanoparticle loading, together with broader independent validation,
 628 are required before field-scale predictive deployment.

629 **List of Acronyms**

630 CART – Classification and regression tree

631 CV – Cross-validation

632 HistGBR – Histogram-based gradient boosting regressor

633 HPHT – High-pressure, high-temperature

634 KNN – K-nearest neighbours

635 MAE – Mean absolute error

636 ML – Machine learning

637 MLP – Multilayer perceptron

638 PFI – Permutation feature importance

639 RBF – Radial basis function

640 RMSE – Root mean square error

641 SHAP – SHapley Additive exPlanations

642 SNP – Silica nanoparticle

643 SVR – Support vector regression

644 XGBoost – Extreme Gradient Boosting

645 **Statements and Declarations**

646 **Funding**

647 The current study and the authors received no financial support or specific grant from any funding agency in
648 the public, commercial, or not-for-profit sectors.

649 **Author contributions**

650 Y. A., R. A., and A. A. developed the machine learning models, performed the computational analysis, and
651 contributed to manuscript preparation. A. T. provided machine learning advisory support and reviewed the
652 technical modelling framework. S. S. N., C. N., and F. Y. contributed to manuscript drafting, restructuring,
653 literature integration, and revision of the journal article. M. E. G. P. conceived and led the scientific research,
654 supervised the project, interpreted the results, and conducted the final editing and revision of the manuscript.
655 All authors reviewed and approved the final manuscript.

656 **Competing interests**

657 The authors declare no competing interests.

658 **Data availability**

659 The data used to support the findings of this study are available from the corresponding author upon request.

660 **8. References**

- 661 Al-Darweesh J, Aljawad MS, Tariq Z, Alajmei S, Yan B, Kamal MS (2024) Prediction of foam rheology
662 models parameters utilizing machine learning tools. *ACS Omega* 9:20397–20409.
663 <https://doi.org/10.1021/acsomega.4c00965>
- 664 Bai Y, Chen Y, Wu L, Luo P, Zhang X, Yang Z, Chen X (2025) Research progress on nanoparticle-enhanced
665 surfactant-based fracturing fluids: A comprehensive review. *Chem Eng Res Des* 220:29–48.
666 <https://doi.org/10.1016/j.cherd.2025.06.040>
- 667 Delpisheh M, Ebrahimpour B, Fattahi A, Siavashi M, Mir H, Mashhadimoslem H, Abdol MA, Ghorbani M,
668 Shokri J, Niblett D, Khosravi K, Rahimi S, Alirahmi SM, Yu H, Elkamel A, Niasar V, Mamlouk M (2024)
669 Leveraging machine learning in porous media. *J Mater Chem A* 12:20717–20782.
670 <https://doi.org/10.1039/D4TA00251B>
- 671 Fisher A, Rudin C, Dominici F (2019) All models are wrong, but many are useful: Learning a variable's
672 importance by studying an entire class of prediction models simultaneously. *J Mach Learn Res* 20:1–81
- 673 Géron A (2022) Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow, 3rd edn. O'Reilly
674 Media, Sebastopol
- 675 Gonzalez Perdomo ME, Wan Madihi S (2022) Foam based fracturing fluid characterization for an optimized
676 application in HPHT reservoir conditions. *Fluids* 7:156. <https://doi.org/10.3390/fluids7050156>
- 677 Gowida A, Shah JP, Elkatatny S (2026) Foam systems for underbalanced and geothermal drilling: a critical
678 review of stability challenges and research frontiers. *J Petrol Explor Prod Technol* 16:10.
679 <https://doi.org/10.1007/s13202-025-02114-4>
- 680 Karami A, Akbari A, Kazemzadeh Y, Nikraves H (2025) Enhancing hydraulic fracturing efficiency through
681 machine learning. *J Petrol Explor Prod Technol* 15:5. <https://doi.org/10.1007/s13202-024-01914-4>

- Laalam A, Tomomewo OS, Khalifa H, Bouabdallah N, Ouadi H, Tran TH, Perdomo ME (2023) Comparative analysis between empirical correlations and time series models for the prediction and forecasting of unconventional Bakken wells production. In: SPE Asia Pacific Unconventional Resources Symposium. Society of Petroleum Engineers. <https://doi.org/10.2118/217297-MS>
- Li S, Li S, Du K, Zhu J, Shang L, Zhang K (2022) Synthesis and stability of switchable CO₂-responsive foaming coupled with nanoparticles. *iScience* 25:105091. <https://doi.org/10.1016/j.isci.2022.105091>
- Lundberg SM, Lee SI (2017) A unified approach to interpreting model predictions. In: *Advances in Neural Information Processing Systems* 30:4765–4774
- Olukoga T, Totaro M, Feng Y (2022) Ensemble machine learning for predicting viscosity of nanoparticle-surfactant-stabilized CO₂ foam. *SPE Reserv Eval Eng* 25:794–814. <https://doi.org/10.2118/209377-PA>
- Swamynathan M (2017) *Mastering machine learning with Python in six steps: A practical implementation guide to predictive data analytics using Python*. Apress, Berkeley. <https://doi.org/10.1007/978-1-4842-3063-4>
- Tran T, Gonzalez Perdomo ME, Wilk K, Kasza P, Amrouch K (2020) Performance evaluation of synthetic and natural polymers in nitrogen foam-based fracturing fluids in the Cooper Basin, South Australia. *APPEA J* 60:227–241. <https://doi.org/10.1071/AJ19062>
- Tran T, Gonzalez Perdomo ME, Haghighi M, Amrouch K (2023a) Effects of cationic and anionic surfactants on the stability, rheology and proppant suspension of nanoparticle-stabilized fracturing foams at elevated temperature. *Geoenergy Sci Eng* 228:212041. <https://doi.org/10.1016/j.geoen.2023.212041>
- Tran T, Gonzalez Perdomo ME, Haghighi M, Amrouch K (2023b) Experimental study of the effects of salinity on nanoparticle-surfactant foams for fracture stimulation application. *Gas Sci Eng* 115:205007. <https://doi.org/10.1016/j.jgsce.2023.205007>
- Tran T, Nguyen GH, Gonzalez Perdomo ME, Haghighi M, Amrouch K (2025) Simulation study of the effects of foam rheology on hydraulic fracture proppant placement. *Processes* 13:378. <https://doi.org/10.3390/pr13020378>
- Verma A, Pal N, Kumar N, Al-Yaseri A, Ali M, Ojha K (2025) Nanoparticles stabilized foam fluid for hydraulic fracturing application of unconventional gas reservoirs: A review of the properties, progress and future prospects. *Pet Res* 10:149–165. <https://doi.org/10.1016/j.ptlrs.2024.08.004>
- Wanniarachchi WAM, Ranjith PG, Perera MSA (2017) Shale gas fracturing using foam-based fracturing fluid: a review. *Environ Earth Sci* 76:91. <https://doi.org/10.1007/s12665-017-6399-x>
- Zhu Y, Chen T, Cui Z (2022) Multiple Pickering emulsions stabilized by the same particles with different extent of hydrophobization in situ. *Front Chem* 10:950932. <https://doi.org/10.3389/fchem.2022.950932>