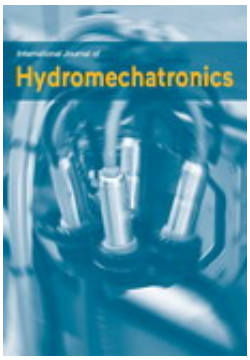


International Journal of Hydromechatronics

ISSN online: 2515-0472 - ISSN print: 2515-0464
https://www.inderscience.com/ijhm

Artificial intelligence-based viscosity prediction of polyalphaolefin-boron nitride nanofluids

Omer A. Alawi, Haslinda Mohamed Kamar, Mustafa Mudhafar Shawkat, Mohammed M. Al-Ani, Hussein A. Mohammed, Raad Z. Homod, Mazlan A. Wahid

DOI: 10.1504/IJHM.2024.10063148

Article History:

Received:	01 August 2023
Last revised:	30 October 2023
Accepted:	03 November 2023
Published online:	30 April 2024

Artificial intelligence-based viscosity prediction of polyalphaolefin-boron nitride nanofluids

Omer A. Alawi and
Haslinda Mohamed Kamar

Department of Thermofluids,
School of Mechanical Engineering,
Universiti Teknologi Malaysia,
81310 UTM Skudai, Johor Bahru, Malaysia
Email: omeralawi@utm.my
Email: haslinda@utm.my

Mustafa Mudhafar Shawkat

Department of Petroleum Engineering,
School of Chemical and Energy Engineering,
Faculty of Engineering,
Universiti Teknologi Malaysia,
81310, Johor Bahru, Malaysia
Email: mudhafar@graduate.utm.my

Mohammed M. Al-Ani

Center for Artificial Intelligence Technology (CAIT),
Faculty of Information Science and Technology,
Universiti Kebangsaan Malaysia (UKM),
Bangi, Selangor, Malaysia
Email: mohmed_alanni@yahoo.com

Hussein A. Mohammed*

School of Engineering,
Edith Cowan University,
270 Joondalup Drive, WA 6027, Australia
Email: hussein.mohammed@ecu.edu.au
*Corresponding author

Raad Z. Homod

Department of Oil and Gas Engineering,
Basrah University for Oil and Gas,
Aljamhoryah, Alzahraa District, Altijari Street, Basra, Iraq
Email: raadahmood@yahoo.com

Mazlan A. Wahid

Department of Thermofluids,
School of Mechanical Engineering,
Universiti Teknologi Malaysia,
81310 UTM Skudai, Johor Bahru, Malaysia
Email: mazlan@utm.my

Abstract: Predicting viscosity's nanofluids can benefit all domains, including energy, thermofluids, power systems, energy storage, materials, cooling, heating, and lubrication. The objective of this study to predict the dynamic viscosity of polyalphaolefin-hexagonal boron nitride (PAO/hBN) nanofluids using four main parameters: shear rate, shear stress, nanomaterials mass fraction, and temperature. Moreover, three hybrid ensemble learning models (Bayesian ridge-random forest, Bayesian ridge-MLP regressor and Bayesian ridge-AdaBoost regressor) were developed for the current task. The forward sequential feature selector (FSFS) created four input combinations (models). Model 4 showed the best prediction accuracy, followed by models 2, 3 and 1. The computational findings showed that ensemble learner 1 was slightly outperformed by ensemble learner 3. Meanwhile, among the predictive models, ensemble learner 2 consistently placed third. Besides, the research results demonstrated that creating predictive models based on all input parameters can produce a precise prediction matrix. Overall, the study recommended exciting conclusions on predicting a nanolubricant's viscosity for use in heat transfer applicants.

Keywords: nanofluids; viscosity; polyalphaolefin; PAO; machine learning; ensemble learning; boron nitride.

Reference to this paper should be made as follows: Alawi, O.A., Kamar, H.M., Shawkat, M.M., Al-Ani, M.M., Mohammed, H.A., Homod, R.Z. and Wahid, M.A. (2024) 'Artificial intelligence-based viscosity prediction of polyalphaolefin-boron nitride nanofluids', *Int. J. Hydromechatronics*, Vol. 7, No. 2, pp.89–112.

Biographical notes: Omer A. Alawi received his PhD in Mechanical Engineering (Nanostructured Materials) from Universiti Teknologi Malaysia, an MEng in Mechanical Engineering (Metal Oxides) from Universiti Tenaga Nasional, and Bachelor's in Mechanical Engineering (Heat Transfer) from University of Anbar. His current research activities are including nanomaterials preparation and characterisation, energy technologies and AI for climate change and atmospheric science.

Haslinda Mohamed Kamar is a Lecturer at the School of Mechanical Engineering, Faculty of Engineering, Universiti Teknologi Malaysia (UTM) in Johor Bahru, Johor. She has been a faculty member since 1993. She teach heat transfer and thermodynamics courses. Her areas of interest include automotive air-conditioning, thermal comfort, indoor air quality (IAQ), natural ventilation as a passive cooling strategy in buildings, and computational fluid dynamics (CFD) modelling and simulation.

Mustafa Mudhafar Shawkat is a Master's degree holder in Chemical and Energy Engineering from Universiti Teknologi Malaysia. His research interests are on petroleum exploration, petroleum geology, enhance oil recovery and machine learning application on the naturally fractured reservoir.

Mohammed M. Al-Ani is a Master's degree holder in Artificial Intelligence Technology from Universiti Kebangsaan Malaysia. His research interests are on machine learning, classification, deep learning, natural language processing and sentiment analysis.

Hussein A. Mohammed is a Senior Lecturer in Mechanical Engineering at the School of Engineering of Edith Cowan University, Australia. His research interests are on renewable energy technologies, nanotechnology, thermofluids of various energy systems, CFD numerical simulations and aerodynamics studies. He has published more than 192 refereed papers in prestigious peer-reviewed international journals with high impact factors. He has recently been featured among the World's Top 2% Scientists List from 2020 to 2023, according to a Stanford University study. He has been listed in Australia's Top 40 Superstars Researchers who are lifetime achievers in physics and mathematics/thermal sciences.

Raad Z. Homod received his BSc degree from Basra University in 1991, and MSc degree from UM, Kuala Lumpur, Malaysia, both in Mechanical Engineering (with distinction). He received his PhD in Mechatronics at UNITEN, Malaysia. He was working in HVAC system engineering. His research interests include nonlinear systems analysis, optimal control, process control, neural networks, HVAC, buildings model and modelling sensors with fuzzy logic.

Mazlan A. Wahid is a Professor in Department of Thermofluid, Faculty of Engineering, School of Mechanical Engineering. He received his Bachelor degree from Embry-Riddle Aeronautical University, Daytona Beach, Florida, Master's degree from Universiti of Leeds, Yorkshire, England, UK, while he obtained his PhD Degree from State University of New York, Buffalo, New York, USA. His current research focus is on combustion, energy and heat transfer, sustainable low emission combustion, high-speed reacting flow, sustainable energy conversion and heat transfer enhancement.

1 Introduction

1.1 Research background and motivation

Nanofluids, a fluid with outstanding thermophysical properties, have gotten a lot of interest in recent years because of their potential uses in various industrial fields (Wang et al., 2023; Lingala, 2023). It is critical to accurately determine the thermophysical parameters of nanofluids, such as dynamic viscosity and thermal conductivity, to assess their heat transfer effectiveness in these applications (Ali and Salam, 2020; Rashidi et al., 2021; Rashidi et al., 2021). However, due to the high cost of nanoparticles, experimental evaluation of nanofluids' complete spectrum of thermophysical properties is not economically possible. As a result, building predictive models for estimating nanofluid characteristics has become a critical area of research (Asadi et al., 2019; Zendehboudi et al., 2019; Zhou et al., 2023b). Researchers have used various strategies, including experimental, numerical, and intelligent methods, to analyse, predict, and optimise the thermophysical properties of nanofluids. Among these qualities, dynamic viscosity and thermal conductivity are important in evaluating heat transfer in industrial applications

such as pumping power and convective heat transfer (Gholizadeh et al., 2020). Despite the complexity of molecular interactions and the hydrodynamics of nanofluids, it is crucial to accurately determine the dynamic viscosity of these materials as a rheological parameter. In conclusion, research into the thermophysical properties of nanofluids is an area of science that is quickly developing, and the creation of prediction models for calculating the properties of nanofluids is essential for ensuring their practical usage in various industrial applications.

1.2 Literature review

Numerous theoretical and experimental studies have been carried out recently to understand better the effects of various variables on the improvement of dynamic viscosity and thermal conductivity of nanofluids, including temperature, solid volume fraction, nanoparticle size, base fluid type, shear rate, and nanoparticle shape (Garoosi, 2020; Amin et al., 2021). The majority of these studies have concentrated on determining the dynamic viscosity of nanofluids using different types of nanoparticles, such as aluminum oxide (Al_2O_3), titanium dioxide (TiO_2), silicon dioxide (SiO_2), copper oxide (CuO), silicon carbide (SiC) and single-walled carbon nanotubes (SWCNTs), in different volume fraction ranges suspended in water (DW), deionised water (DIW), ethylene glycol (EG), and mixtures of these base fluids (Porgar et al., 2023; Younes et al., 2022; Alshuhail et al., 2023; Yang et al., 2017; Kalsi et al., 2023). However, it can be used to measure nanofluids' relative viscosity in a lab, so some researchers have created theoretical models to calculate the dynamic viscosity to cut down on time-consuming laboratory experiments (Hemmati-Sarapardeh et al., 2018). Additionally, empirical models based on conventional mathematical correlations have been developed (Wang et al., 2016). An experimental study measured the thermo-physical properties of polyalphaolefin oil modified with nano additives like PAO6 as a base fluid mixed with zirconium oxide (ZrO_2) and boron nitride (BN) nanoparticles (NPs) and graphene nanoplatelets (GNPs) (Guimarey et al., 2019). They observed that, the greatest increase in viscosity is obtained when GNPs were used as an additive. Also, the combination of polyalphaolefin (PAO) oil with hexagonal boron nitride (hBN) was prepared for thermal management and lubrication applications (Sleiti, 2020). The viscosity as a function of temperature (from –20 to 70°C) and volume fraction (0.25–1%) decreased with temperature for both base fluid and PAO/hNB nanofluids and increased with concentration. In another study, hexagonal boron nitride (h-BN) nanoparticles were dispersed in (PAO6) for comprehensive experimental examination (Jiang et al., 2021). Rheological characterisation showed that the h-BN nano-lubricants exhibited non-Newtonian behaviour at low shear rates (<36.69 s^{-1}) and Newtonian behaviour at higher shear rates.

Artificial neural network (ANN) models have been extensively employed in experimental studies in recent years to assess the relative thermal conductivity and dynamic viscosity of diverse nanofluids. However, for many nanofluids, the hypothesised correlations are inadequately precise and effective (Esfe et al., 2016; Karimipour et al., 2018). Advanced machine learning methods, such as the adaptive neuro-fuzzy inference system (ANFIS), multilayer perceptron artificial neural network (MLP-ANN), least square support vector machine (LLSVM), and others, are now being used by researchers to address this problem. For instance, Mehrabi et al. (2013) developed an FCM-ANFIS model to estimate the effective viscosity of nanofluids, which is restricted to a particular

kind of nanofluid and employs volume fraction, temperature, and nanoparticle size as input variables to predict the effective viscosity of nanofluids. A hybrid ANN-based group method of data handling (GMDH) technique was utilised by Atashrouz et al. (2014) to calculate the dynamic viscosity of nine nanofluids containing Al_2O_3, SiO_2, TiO_2, and CuO nanomaterials. MLP-ANN and radial basis function neural network (RBF-ANN) models were presented by Heidari et al. (2016) and colleagues, and they demonstrated good accuracy in predicting the viscosity of nanofluids using the same dataset. The viscosity of Newtonian nanofluids with various nanoparticle kinds and base fluids can be estimated using soft computing approaches, such as ANFIS, gene expression programming (GEP), MLP-ANN, and LLSVM, according to Chen et al. (2007). The MLP model was developed by Ansari and his team (2018) utilising various training algorithms and other transfer functions. Input factors for this study include temperature, shear rate values, nanoparticle size, density, and nanoparticle concentration. To estimate the relative viscosity of nanofluids, Baghban et al. (2019) developed a hybrid ANFIS and particle swarm strategy with as input parameters temperature, nanoparticle diameter, nanofluid density, volumetric fraction, and base fluid viscosity. However, the dataset that was gathered did not contain any carbon nanotubes.

Recently, some robust machine learning models were developed to estimate the thermophysical properties of nanofluids; two modern novel machine learning approaches, a Bayesian optimised support vector machine and a wide neural network, were used to model-predict the thermophysical properties of Al_2O_3-GO nanofluids with a robust predictive efficiency of 97.15–99.91% (Kanti et al., 2023b). ANFIS models with different types of clustering techniques, including grid partitioning (GP), subtractive clustering (SC), and fuzzy c-means (FCM), were utilised to estimate the thermophysical properties of water-based oxide-MWCNT hybrid nanofluids (Zhang et al., 2023). SC-based ANFIS approach presented the highest precision model for the nanofluids viscosity ($R = 0.99887$ and MAPE = 0.4206%). Application of novel machine-learning techniques was applied to examine the effects of pH on the stability and thermal properties of copper oxide (CuO), graphene oxide (GO), and their hybrid nanofluid (HNF) at different mixing ratios (Kanti et al., 2023a). Their ML results revealed that the Bayesian optimised support vector machine (BoA-SVM) was superior to Bayesian optimised boosted regression trees (BoA-BRT) for the viscosity model. An explainable artificial intelligence (XAI) technique called the Bayesian approach optimised Gaussian process regression was employed to develop a predictive model for the thermophysical properties of polydisperse SiO_2/aqueous glycerol nanofluids (Sharma et al., 2023). The test XAI approaches were shown as robust because of the high correlation values, which ranged from 99.68% to 99.99%, along with minimal modelling errors.

1.3 Research objectives

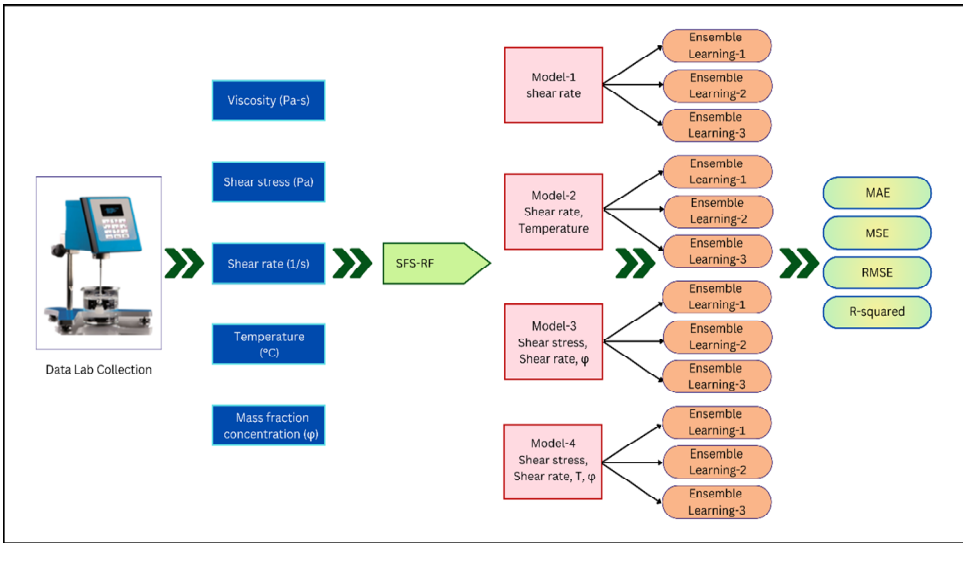
This research is significant because it fills a knowledge gap in optimising nanolubricant viscosity. The thermophysical properties of nanofluids have been extensively studied, but the best way to maximise the viscosity of nanolubricants has received less attention. Furthermore, the use of machine learning in this field of study has been limited. The study aims to use experimental data to predict the dynamic viscosity of a nanolubricant made of a mixture of polyalphaolefin (PAO) and hexagonal boron nitride (hBN) nanomaterials. In order to examine the impact of many parameters on viscosity, including

temperature, shear stress, mass fractions of nanoparticles, and shear rate, the study separated the computational models into four sections. This is crucial because, by recognising how these factors affect the viscosity of the nanolubricant, the viscosity may be optimised. The study used three different hybrid machine learning techniques – Bayesian ridge + random forest, Bayesian ridge + MLP regressor and Bayesian ridge + AdaBoost regressor – to assess the overall performance of the computational models. After a thorough statistical analysis, these machine-learning models were utilised to examine the experimental data and make predictions regarding the viscosity of the nanolubricant. This is crucial because it shows how machine learning can help optimise nanolubricant viscosity. Overall, this study contributes to the field by filling a gap in the literature on nanolubricant viscosity optimisation, exploring the effect of various parameters on viscosity, and proving the promise of machine learning in this area of research.

2 Methodology

As per Figure 1, the current research design is divided into several phases, such as data collection from lab measurements, features engineering and selection using sequential forward selection (SFS), development of four combinations, application of three different hybrid AI-driven models of each input combination and evaluate the performance of each hybrid model through four metrics.

Figure 1 Flowchart of the proposed methodology (see online version for colours)



2.1 Data collection and statistical analysis

The raw data shown here relates to the articles previously published in Sleiti (2020, 2021). The experimental measurements were conducted using the AR-G2 rheometer from TA Instruments to assess the viscosity values for pure PAO and the PAO/hBN

nanolubricants. Different data were taken, including viscosity (Pa-s), shear stress (Pa), shear rate (1/s), and temperature (°C). Moreover, the weight concentration of the nanomaterials was in three different values: (φ = 0% for pure PAO) and (φ = 0.6% and 1.0% for PAO/hBN nanolubricants).

Table 1 Main parameters of nanofluids data analysis

		Shear stress (Pa)	*Shear rate (1/s)*	*Temperature (°C)*	*Mass fraction (φ)*	*Viscosity (Pa-s)*
N	Valid	537	537	537	537	537
	Missing	0	0	0	0	0
Mean		9.9005	259.7481	24.9991	0.5307	0.1701
Std. error of mean		0.00406	10.62038	1.11632	0.01774	0.01097
Median		9.9340	173.3000	25.0000	0.6000	0.0573
Mode		9.99	41.40	–19.70	0.00	0.02
Std. deviation		0.09419	246.10889	25.86887	0.41100	0.25420
Variance		0.009	60569.587	669.198	0.169	0.065
Skewness		-0.982	0.986	0.000	–.231	2.547
Std. error of skewness		0.105	0.105	0.105	0.105	0.105
Kurtosis		0.030	0.036	–1.199	–1.505	7.279
Std. error of kurtosis		0.210	0.210	0.210	.210	0.210
Range		0.38	988.08	89.70	1.00	1.55
Minimum		9.62	6.42	–19.70	0.00	0.01
Maximum		10.00	994.50	70.00	1.00	1.56
Percentiles	25	9.8390	51.6850	2.5000	0.0000	0.0231
	50	9.9340	173.3000	25.0000	0.6000	0.0573
	75	9.9800	421.4000	47.5000	1.0000	0.1931

As stated above, the dataset includes four different input parameters, such as shear rate, shear stress, temperature, and mass percentage of hBN nanomaterials, to estimate the values of dynamic viscosity, as shown in Table 1. The dataset has 537 valid points and no missing values. Furthermore, Table 1 displays descriptive statistics of input and output properties using metrics such as mean, standard deviation, variance, skewness, and kurtosis to provide a complete picture of the relationships between independent and dependent variables. The Pearson correlation coefficient (PCC) between input and output variables are described in Figure 2. Temperature input showed the strongest correlation with dynamic viscosity, with a negative coefficient correlation of 0.741. At the same time, shear rate and share stress indicated almost the exact correlation of ±0.55. Meanwhile, the mass fraction of BN nanoparticles showed a lower impact of 0.172.

2.2 Sequential forward selection

Sequential forward selection (SFS) is a wrapper feature selection algorithm that improves model performance and interpretability by adding one feature at a time to the current set of selected features (Zhou et al., 2023a) (see Figure 3). The process stops when a certain number of features have been added, or no further improvement can be made. SFS is

computationally expensive and can lead to sub-optimal solutions if the relationship between the features and target variable is nonlinear. A new method that combines the SFS-based feature selection strategy with the random forest (RF) machine learning algorithm is proposed to solve these weaknesses. This method accurately identifies the importance of each feature, removing unneeded and duplicate features. The present study investigated the influence of different input parameters on predicting the dynamic viscosity of nanolubricants using four feature combinations (models), as reported in Table 2. The hyperparameters tuning of the SFS algorithm are as follows (random forest regressor(), k_features = 4, forward = true, scoring = 'R^2', verbose = 2, cv = 3).

Figure 2 Correlation coefficient between input and output variables (see online version for colours)

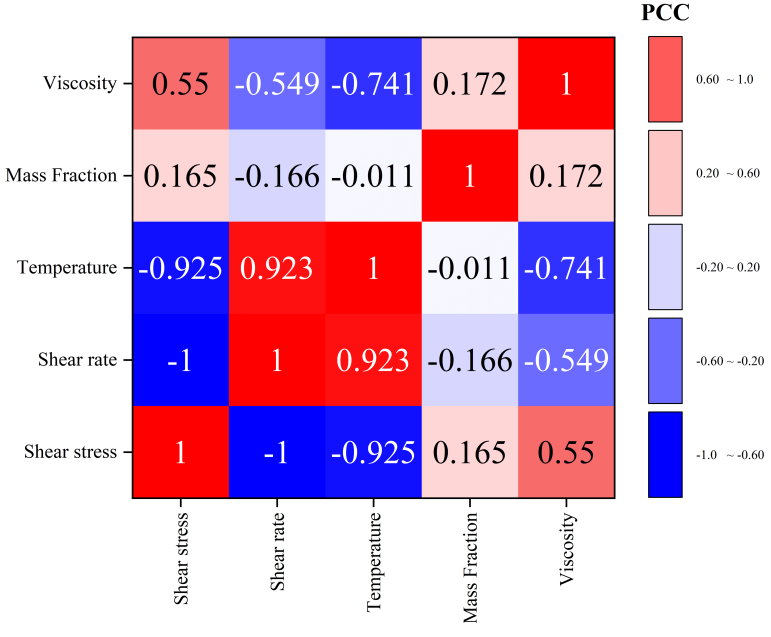
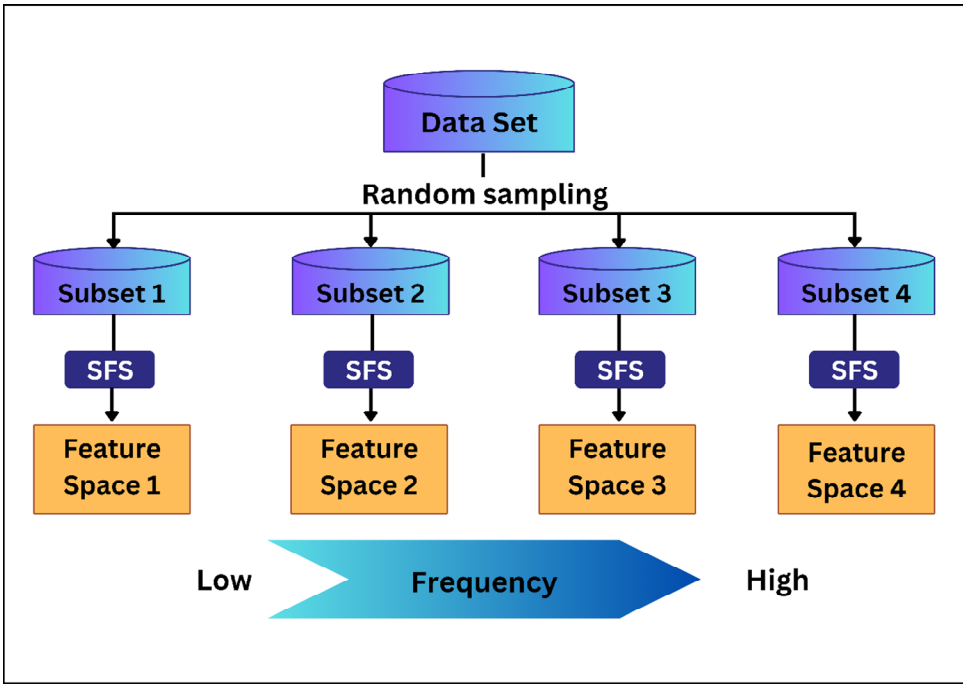


Table 2 Features combinations according to forward sequential feature selector

#	*Feature names*	*cv_scores*	*avg_score*	*ci_bound*	*std_dev*	*std_err*
Model 1	Shear rate	[0.99969451 0.99336954 0.9990032]	0.99736	0.00637	0.00283	0.002
Model 2	Shear rate, temperature	[0.99785936 0.99886052 0.9978461]	0.99819	0.00107	0.00048	0.00034
Model 3	Shear stress, shear rate, mass fraction	[0.99640226 0.99568815 0.99935141]	0.99715	0.00357	0.00159	0.00112
Model 4	Shear stress, shear rate, temperature, mass fraction	[0.99218607 0.99612162 0.99876728]	0.99569	0.00608	0.00270	0.00191

Figure 3 Sequential forward selection using random forest process (see online version for colours)



2.3 Machine learning regressors

The Bayesian ridge algorithm, random forest algorithm, MLP regressor algorithm, and adaptive boost regressor are the four machine learning approaches developed in this study to estimate the values of nanolubricant viscosity. The feature selection process was utilised to determine the relationship between input features and the target variable to determine the most informative features for predicting viscosity. After selecting the best features, each algorithm was trained on the data using these features. Finally, three hybrid models were developed using ensemble voting regressor implementation with the Bayesian ridge algorithm as the base algorithm, and the other three models combined to create three predictors.

2.3.1 Bayesian ridge algorithm

The Bayesian ridge algorithm is a linear regression method that utilises Bayesian methods to regularise the model (Cruz et al., 2021). This probabilistic algorithm estimates the weight coefficients of the linear model through maximum a posteriori (MAP) estimation. One of the key advantages of this approach is that it can efficiently manage scenarios in which the number of features in the dataset exceeds the number of samples, which frequently leads to overfitting when using typical linear regression models. The Bayesian ridge algorithm incorporates past information and presumptions about the model's parameters into the estimate process, which helps prevent overfitting. The algorithm provides estimates of the model's uncertainty, which can be advantageous

in specific applications such as decision-making processes. These estimates can be helpful when the model's accuracy is critical and understanding the confidence level in the model's predictions is required (Salem et al., 2022). The objective function to be minimised is as follows:

$$\min_{w,\alpha} \frac{1}{2n_{samples}} \|y - Xw\|_2^2 + \frac{\alpha}{2}\|w\|_2^2 + \frac{1}{2}\lambda\|w\|_2^2 \quad (1)$$

where w is the weight vector, α is the precision of the noise, X is the input parameters, y is the nanolubricant viscosity, $2n_{samples}$ is the dataset size, and λ is the hyperparameter for the ridge (L2) regularisation.

2.3.2 Random forest algorithm

The random forest algorithm is an effective technique for solving regression problems (Zekić-Sušac et al., 2021). An ensemble method creates multiple decision trees to predict continuous output variables. The random forest algorithm employs the bagging approach, randomly selecting samples from the input data and making a decision tree for each sample. Then, the algorithm calculates the prediction of each tree, and the final output is the average of all the trees' predictions. The random forest approach uses several trees and can capture complex correlations, such as nonlinearity and feature interaction, between input and output variables. Moreover, the algorithm includes a feature selection mechanism called feature randomness, which randomly selects a subset of features for each tree and finds the best split at each node. This method results in decorrelation between the trees and reduces the correlation between the features and output variables, lowering the prediction variance. The random forest algorithm is a versatile and robust method that can handle large datasets with many features. It is widely used in various regression applications such as time series prediction, high dimensional data analysis, and spatial data modelling (Parzinger et al., 2022; Wang et al., 2022). This study implements the following hyperparameters with the random forest algorithm: (max_depth = 5, oob_score = true, bootstrap = true).

2.3.3 MLP regressor algorithm

A supervised learning approach called a multilayer perceptron (MLP) uses artificial neural networks for regression tasks (Esfe et al., 2022; Ghritlahre and Prasad, 2018). The design comprises several interconnected layers of artificial neurons, where the input layer receives input data, one or more hidden layers process the data, and the output layer generates the prediction. The weights linking the neurons in each layer are changed during training to improve the network's performance. A labelled dataset is necessary for MLP training, which involves presenting the network with input-output pairs and adjusting the weights to reduce the discrepancy between the predictions and actual outputs. The network can be used to predict new, unexplored data after training. MLP is helpful in various applications, including time series forecasting, natural language processing, and image identification, because it can simulate nonlinear interactions between inputs and outputs. However, MLP is susceptible to overfitting, mainly when many hidden layers or neurons exist. Different regularisation strategies, including dropout, weight decay, and early halting, can be used to address this problem (Ghazvini

et al., 2022). In this model, two hidden layers (200,200) with activation = 'relu' were implemented. The equation for a single neuron's output is given by:

$$f(x) = \sigma\left(\sum_{i=1}^{n} w_i \cdot x_i + b\right) \tag{2}$$

where $f(x)$ is the output of the neuron, x_i are the input parameters, w_i are the weights associated with the inputs, b is the bias term, σ is the activation function.

2.3.4 Adaptive boost regressor

AdaBoost is a widely used ensemble method in regression tasks, particularly thermal engineering (Adun et al., 2022; Chen et al., 2022; Zhou et al., 2020). The central concept of AdaBoost is to combine several weak models, such as decision trees, to form a robust ensemble model. The weak models are trained iteratively, with each iteration focusing on the examples previously misclassified by the ensemble. This iterative process helps improve the ensemble's overall performance by reducing the individual models' bias and variance. AdaBoost's key advantage is handling complex and nonlinear relationships between inputs and outputs. Moreover, it is robust to outliers and can be applied to various applications such as predicting temperature, pressure, and other thermodynamic properties in thermal engineering. This study used AdaBoost parameters: (base_estimator = RF, random_state = 0, n_estimators = 10). The weights are determined based on the performance of each weak learner. The formula for the prediction is:

$$F(x) = \sum_{t=1}^{T} \alpha_t h_t(x) \tag{3}$$

where $F(x)$ is the final prediction, T is the number of weak learners, α_t are the weights for each weak learner's prediction, $h_t(x)$ is the prediction of the t^{th} weak learner.

2.4 Ensemble voting regressor

Ensemble voting regressor is a powerful technique that combines multiple base models' predictions to form a more robust prediction (Phyo et al., 2022). This method employs the principle of majority voting to make a final prediction. Numerous base models are initially trained on the same dataset, making predictions on new, unseen data. Then, the base models' predictions are combined by taking the majority vote on the predictions. In a scenario with three base models, two of which predict a particular value, that value is selected as the final prediction. Ensemble voting can be applied to both classification and regression tasks. One of the primary advantages of ensemble voting is its ability to enhance the model's overall performance by reducing the variance and bias of the individual base models. Also, it improves the model's generalisation, making it more robust for new unseen data. This method can be applied in various applications, including thermal engineering, weather forecasting, and energy management.

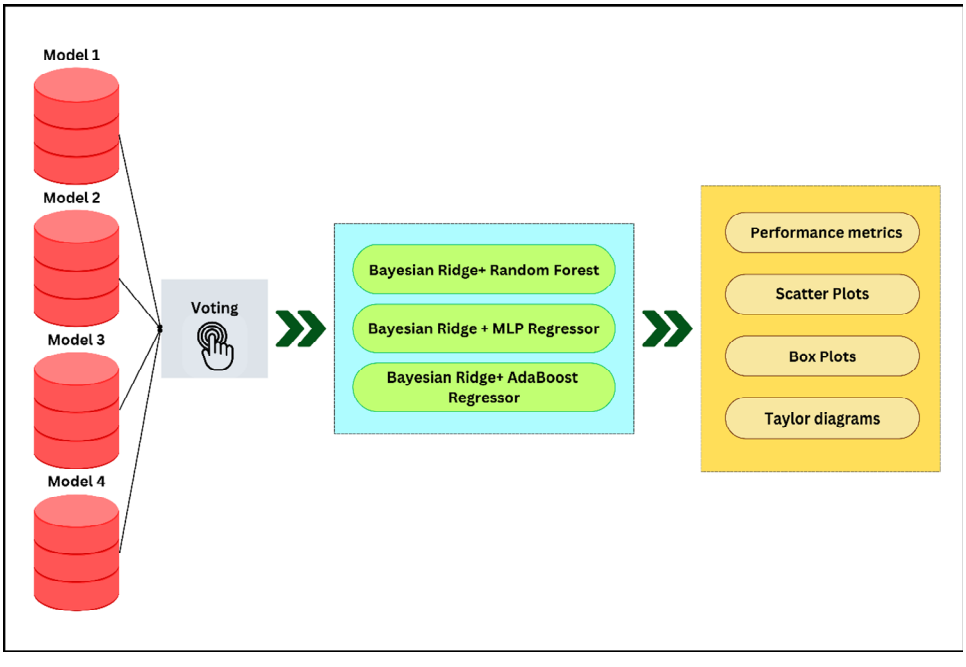
After the performance assessment of the diverse traditionally used ML regressions, the two best regressions have been recognised. At that point, these two main regressions will be used for the next step to obtain the best output from the dataset. Then, the voting regressor will be utilised. As per Figure 4, three ensemble learning (hybrid) were developed as ensemble learning 1 (Bayesian ridge + random forest), ensemble learning 2 (Bayesian ridge + MLP regressor) and ensemble learning 3 (Bayesian ridge+ AdaBoost

regressor). It is worth noting that the Bayesian ridge is consistent with all machine learning algorithms due to its linear correlations between the input parameters and the dynamic viscosity. The equation for the ensemble voting regressor's prediction can be expressed as follows:

$$F(x) = \sum_{i=1}^{n} w_i \cdot f_i(x) \tag{4}$$

where $F(x)$ is the final prediction made by the ensemble, n is the number of individual regressors (base models) in the ensemble, w_i is the weight associated with the i^{th}, regressor's prediction, $f_i(x)$ is the prediction made by the i^{th} regressor for the input sample x.

Figure 4 Schematic diagram of ensemble voting regressor (see online version for colours)



2.5 Performance metrics

In the current study, four performance metrics, mean absolute error (MAE), mean squared error (MSE), root mean squared error (RMSE), and R-squared (R^2), were utilised to evaluate the performance of machine learning models in predicting the dynamic viscosity of nanolubricants. Generally, a model with a lower MAE, MSE, and RMSE and a higher R^2 better fits the data. However, it is essential to remember that the choice of metric depends on the specific problem and the goals of the analysis.

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|\mu_i - \hat{\mu}\right| \tag{5}$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left|\mu_i - \hat{\mu}\right|^2 \tag{6}$$

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left|\mu_i - \hat{\mu}\right|^2} \tag{7}$$

$$R^2 = 1 - \frac{\sum\left(\mu_i - \hat{\mu}\right)^2}{\sum\left(\mu_i - \hat{\mu}\right)^2} \tag{8}$$

where μ_i is the predicted value of viscosity's nanofluid and $\hat{\mu}$ is the mean value of viscosity's nanofluid.

3 Modelling results and analysis

Table 3 presents the performance metrics of four different combinations (models 1, 2, 3 and 4) during their training and testing phases using three different hybrid ensemble learning approaches (ensemble learnings 1, 2 and 3). As stated above, the metrics include mean squared error (MSE), root mean squared error (RMSE), R-squared (R^2), and mean absolute error (MAE).

Overall, the results demonstrate that the model itself (prediction approach) and the ensemble learning technique (number of input features) that are used both affect machine learning performance. For example, model 2 (two input parameters) and model 4 (four parameters) outperform model 1 (one input parameter) and model 3 (three input parameters) throughout both the training and testing stages. This suggests that models 2 and 4 are the most accurate models in predicting nanolubricant viscosity with two and four input parameters. In contrast, models 1 and 3 show lower accuracy than models 2 and 4. This conclusion reveals that models 1 and 3 implementation is more cost-effective in employing one and three features to predict the viscosity value rather than four features. In specific, ensemble learnings 1 and 3 consistently outperform ensemble learning 2. This finding can be attributed to the development of tree-based algorithms that show higher accuracy than neural networks for the current task.

Finally, the ensemble learning approach selected can considerably impact the models' accuracy and generalisability, and the models' performance can vary based on their complexity and other parameters. As a result, it is critical to carefully select and evaluate various ensemble learning strategies in order to improve model performance and produce more accurate predictions.

This part shows the scatter plots of the coefficient of determination (R^2) using four different models (combinations) during their training and testing scenarios using three different ensemble machine learning techniques are presented in Figure 5. For model 1, the R^2 values during the testing phase for the three hybrid methods are close, ranging from 0.74953 to 0.82509. However, the R^2 value during the training phase is higher, ranging from 0.75008 to 0.82799. For model 2, the R^2 value during the testing phase for the three ensemble learning methods ranges from 0.75460 to 0.910885, with ensemble learning 1 and ensemble learning 3 showing slightly higher values. The R^2 value during the training phase is higher than the testing phase for all the ensemble learning methods,

ranging from 0.78923 to 0.92025. For model 3, the R^2 values during the testing phase for all three ensemble learning methods range from 0.79440 to 0.82786, with ensemble learning 3 showing the highest value. The R^2 value during the training phase is higher than the testing phase for all the ensemble learning methods, ranging from 0.79548 to 0.82786. For model 4, the R^2 value during the testing phase for all three ensemble learning methods ranges from 0.91800 to 0.93902, with ensemble learning 3 showing the highest value. The R^2 value during the training phase is higher than the testing phase for all the ensemble learning methods, ranging from 0.93965 to 0.94714.

Table 3 Quantitative evaluation of models 1, 2, 3 and 4

Model 1	*Training phase*			*Testing phase*		
	Ensemble learning 1	*Ensemble learning 2*	*Ensemble learning 3*	*Ensemble learning 1*	*Ensemble learning 2*	*Ensemble learning 3*
MSE	0.01072	0.01558	0.01072	0.01281	0.01831	0.01279
RMSE	0.10354	0.12481	0.10355	0.11318	0.13532	0.11308
R^2	0.82799	0.75008	0.82797	0.82478	0.74953	0.82509
MAE	0.07122	0.09285	0.07105	0.07873	0.09985	0.07848
Model 2	*Training phase*			*Testing phase*		
	Ensemble learning 1	*Ensemble learning 2*	*Ensemble learning 3*	*Ensemble learning 1*	*Ensemble learning 2*	*Ensemble learning 3*
MSE	0.00501	0.01314	0.00497	0.00658	0.01794	0.006515
RMSE	0.07075	0.11462	0.07050	0.08109	0.13394	0.080714
R^2	0.91969	0.78923	0.92025	0.91005	0.75460	0.910885
MAE	0.04500	0.08652	0.04507	0.05323	0.09420	0.053199
Model 3	*Training phase*			*Testing phase*		
	Ensemble learning 1	*Ensemble learning 2*	*Ensemble learning 3*	*Ensemble learning 1*	*Ensemble learning 2*	*Ensemble learning 3*
MSE	0.01082	0.01275	0.01073	0.01284	0.01503	0.01293
RMSE	0.10402	0.11291	0.10358	0.11331	0.12260	0.11371
R^2	0.82640	0.79548	0.82786	0.82437	0.79440	0.82314
MAE	0.07134	0.08274	0.07104	0.07866	0.09106	0.07858
Model 4	*Training phase*			*Testing phase*		
	Ensemble learning 1	*Ensemble learning 2*	*Ensemble learning 3*	*Ensemble learning 1*	*Ensemble learning 2*	*Ensemble learning 3*
MSE	0.00333	0.00376	0.00329	0.00446	0.00599	0.00450
RMSE	0.05775	0.06133	0.05740	0.06677	0.07742	0.06705
R^2	0.94650	0.93965	0.94714	0.93902	0.91800	0.93850
MAE	0.03934	0.04306	0.03947	0.04269	0.04878	0.04290

All the models show higher R^2 values during the training phase than the testing phase, showing machine learning algorithms' ability to generalise well to unseen data in predicting the value of dynamic viscosity. The performance of the models during the testing phase is better when ensemble learning 3 is used. However, the choice of the hybrid machine learning approach and the input combination's performance varies

depending on the model complexity, number of features (data size), interpretability, training time, and hyperparameters tuning.

Figure 5 Scatter plots of models 1, 2, 3 and 4 using different ensemble learners versus predicted viscosity (see online version for colours)

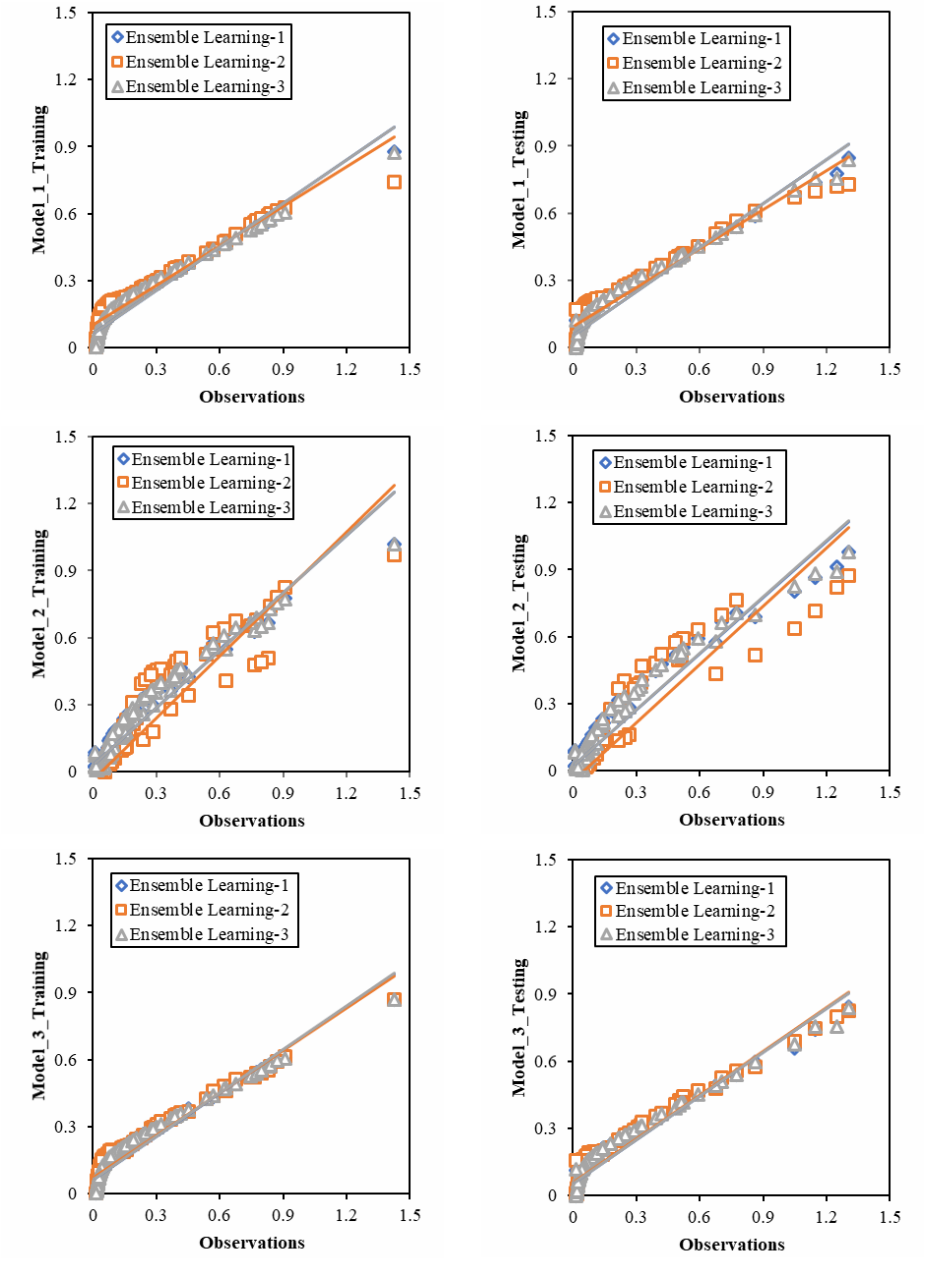
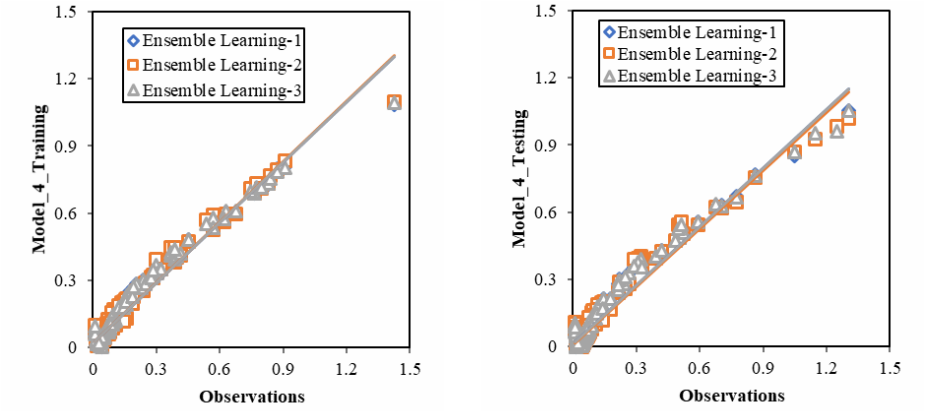


Figure 5 Scatter plots of models 1, 2, 3 and 4 using different ensemble learners versus predicted viscosity (continued) (see online version for colours)



The boxplots of the developed models during the training/testing scenarios are presented in Figure 6. Box plots illustrate the distribution and probability density between the observed and predicted values of dynamic viscosity. Models 1, 2, 3 and 4 testing scenarios were evaluated using three different ensemble learning methods: ensemble learnings 1, 2 and 3.

The results indicate that, in the testing scenario of model 1, the median and standard deviation of the ensemble learnings 1, 2 and 3 were (0.130, 0.179, and 0.130) and (0.189, 0.173, and 0.189), respectively. In the testing scenario of model 2, the median and standard deviation were (0.083, 0.030, and 0.083) and (0.237, 0.260, and 0.237) for ensemble learnings 1, 2 and 3, respectively. In the testing scenario of model 3, the median and standard deviation were (0.130, 0.158, and 0.130) and (0.188, 0.193, and 0.188) for ensemble learnings 1, 2 and 3, respectively. Finally, in the testing scenario of model 4, the median and standard deviation were (0.056, 0.055, and 0.056) and (0.243, 0.244, and 0.243) for ensemble learnings 1, 2 and 3, respectively.

Overall, while median performance values suggest that hybrid model-3 consistently performs well, followed by hybrid model 1 and hybrid model 2, it is essential to consider the standard deviation, which indicates variability in performance. Hybrid model 2 exhibits higher variability, making it less stable, while hybrid models 1 and 3 display more consistent results. Combination 2 consistently yields the lowest median performance, while combination 4 tends to perform the best.

In addition to the graphical comparison, a Taylor diagram (Taylor, 2001; Ali et al., 2022) was also employed to evaluate the developed hybrid models based on their correlation and standard deviation. Figure 7 presents the Taylor diagram for the three hybrid learners with different input combinations for training and testing data.

Figure 6 Boxplots presentation of developed models for the training/testing phases, (a) model-1 training (b) model-1 testing (c) model-2 training (d) model-2 testing (e) model-3 t raining (f) model-3 testing (g) model-4 training (h) model-4 testing (see online version for colours)

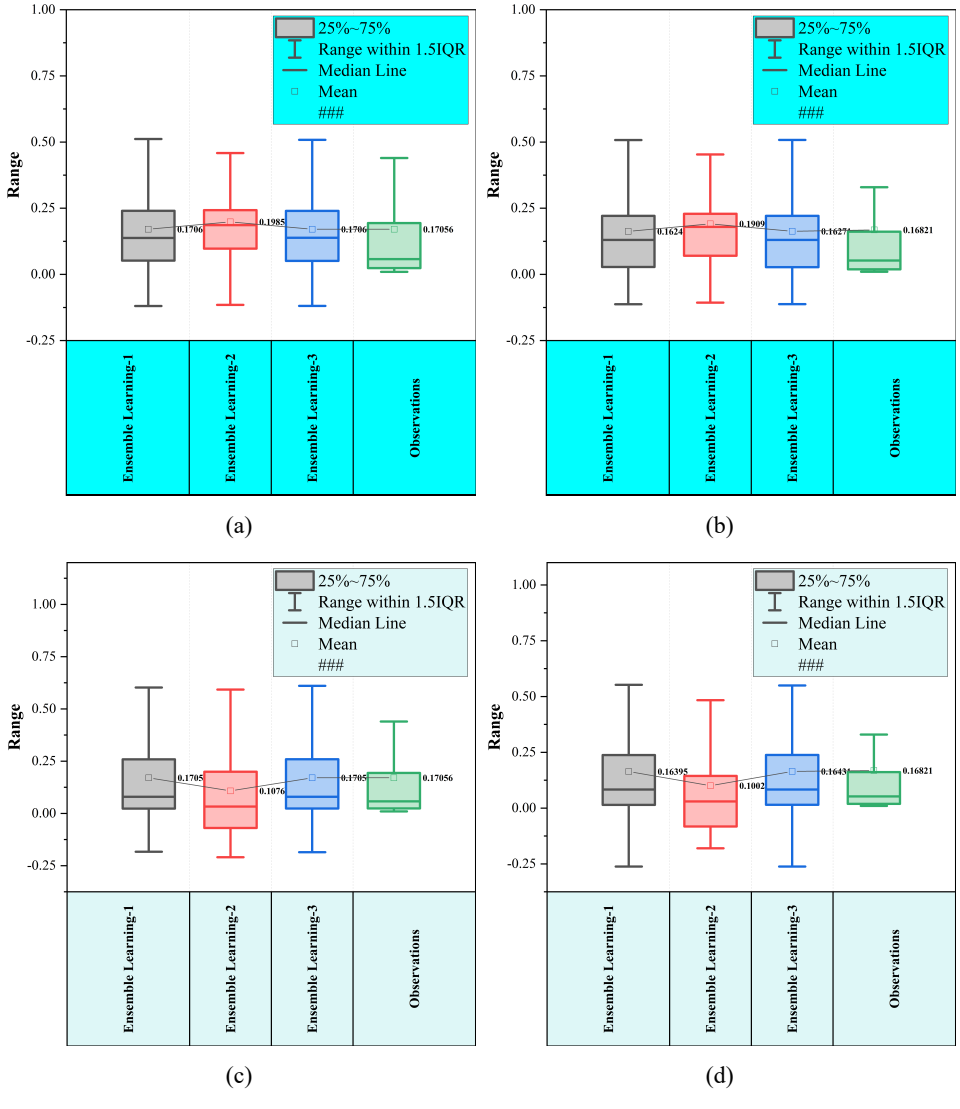
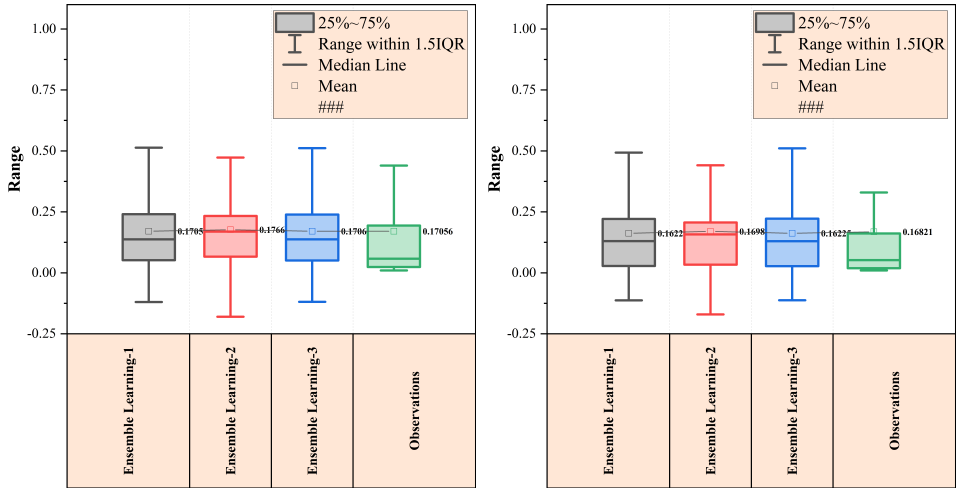
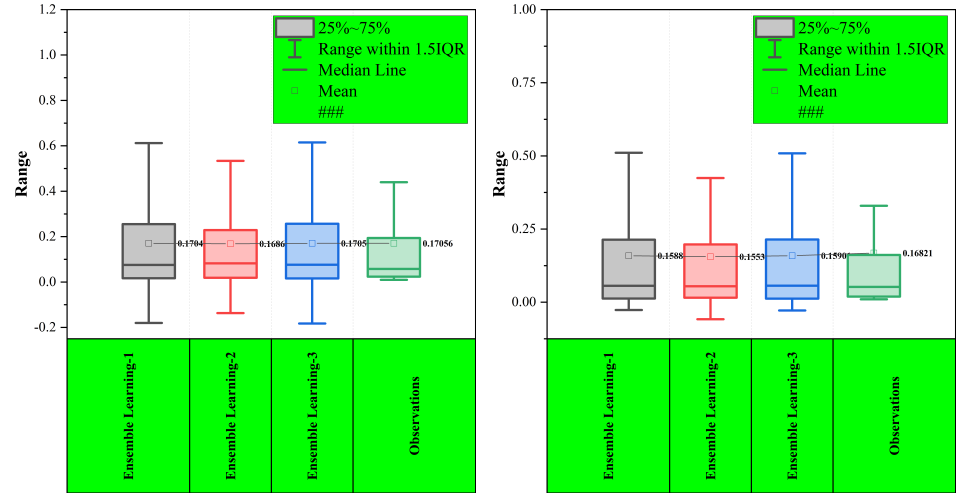


Figure 6 Boxplots presentation of developed models for the training/testing phases, (a) model-1 training (b) model-1 testing (c) model-2 training (d) model-2 testing (e) model-3 training (f) model-3 testing (g) model-4 training (h) model-4 testing (continued) (see online version for colours)

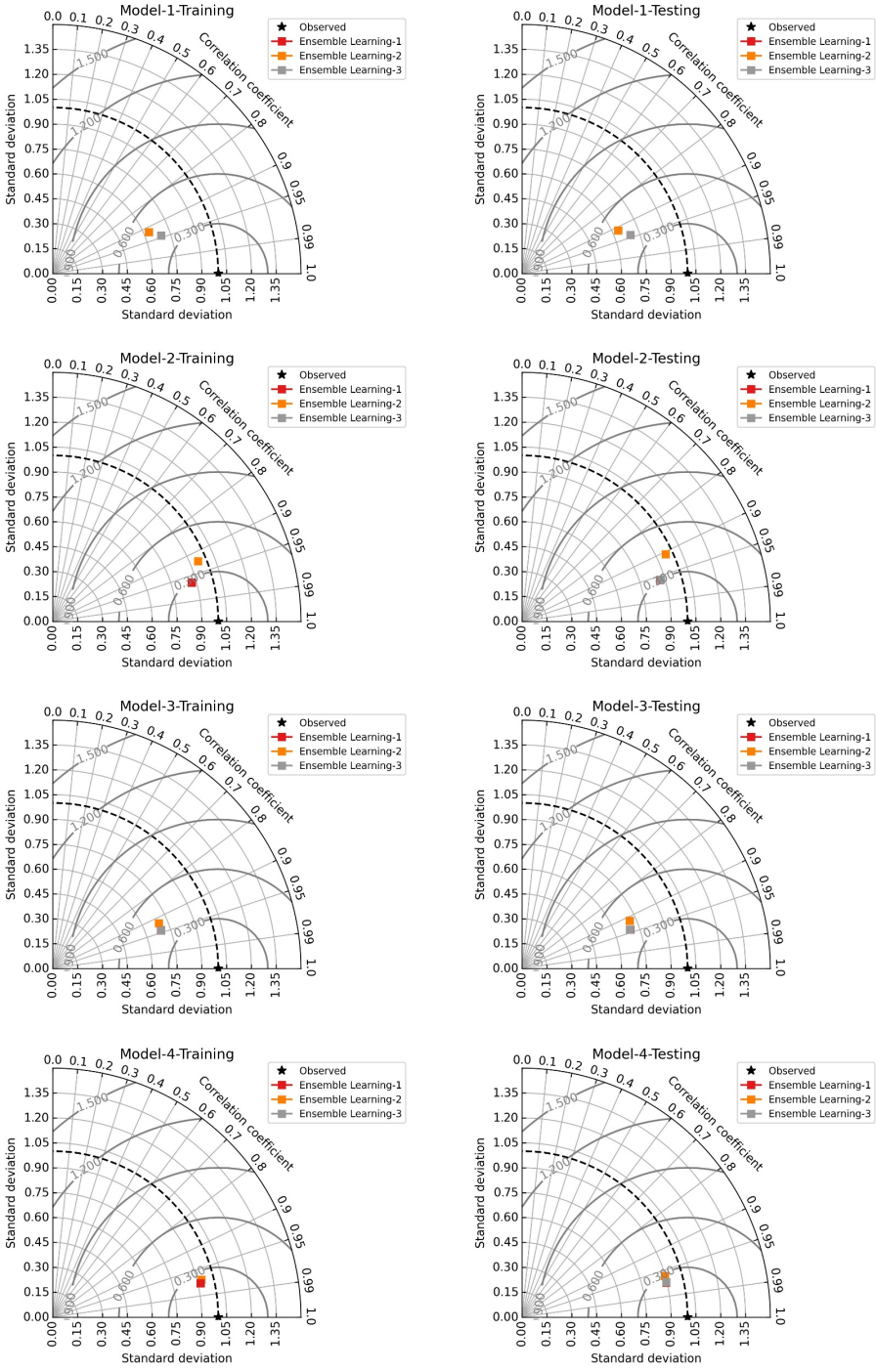


(e) (f)



(g) (h)

Figure 7 Taylor plots of performance metrics for the four predictive models (see online version for colours)



According to the testing Taylor diagram findings, combination 4 performs best in estimate accuracy, followed by combinations 2, 1 and 3. Moreover, it is observed that the hybrid model utilising Bayesian ridge along with random forest generally performs slightly better than the model using Bayesian ridge with AdaBoost regressor, which in turn serves better than the model with Bayesian ridge and MLP regressor in all combinations. These findings confirm the high accuracy and reliability of the proposed ensemble learning models in predicting the dynamic viscosity of Newtonian oil-based nanofluids.

It should be noted that the Taylor diagram provides a valuable visualisation for comparing multiple statistical measures, such as correlation and standard deviation, of different models. The diagram identifies the optimal model in terms of these measures and facilitates the comparison of other models in a single plot. Therefore, using the Taylor diagram in addition to boxplots provides a more comprehensive evaluation of the developed ensemble learning models.

4 Algorithms reviews

The current research employed three different hybrid machine learning algorithms to predict the viscosity of polyalphaolefin-boron nitride (PAO-BN) nanofluids due to their significant challenges in materials science and engineering. Bayesian ridge was the base model in the hybrid algorithms due to its linear regression technique that maintains the linear regression coefficients using a probabilistic framework. The Bayesian ridge was combined with random forest, MLP, and AdaBoost in the first, second, and third models, respectively. In addition, the primary purpose of combining two algorithms is to compensate for the weaknesses of the other, aiming for a more robust and accurate prediction of PAO-BN nanofluid viscosity. Random forest is an ensemble method based on decision trees. It is known for capturing complex relationships in the data and utilising nonlinear patterns. The MLP is an ANNs capable of learning complex patterns and relationships. It's often used for nonlinear regression tasks. AdaBoost is an ensemble method that combines multiple weak learners into strong learners, giving more weight to misclassified samples. It is often used to improve the performance of decision trees. As per the above explanations, it is clear that, hybrid models 1 and 3 outperformed model 2 when four input parameters were included. Moreover, second combination can replace combination 4 to reduce the computational cost and have two input parameters rather than four.

5 Conclusions and future recommendations

The current study aimed to address the challenge of forecasting the viscosity of nanofluids, which is a complex and time-consuming task. To achieve this goal, several machine learning (ML) models were employed, including Bayesian ridge+ random forest, MLP regressor + Bayesian ridge and Bayesian ridge + AdaBoost regressor, which were implemented with three ensemble learners. The modelling process considered several relevant nanofluid characteristics, such as shear stress, shear rate, temperature, and mass fractions.

The main findings of the study are:

1 The proposed machine learning models demonstrated high accuracy and reliability in predicting the dynamic viscosity of nanofluids utilising temperature, shear rate, mass fraction, and shear stress.

2 The Bayesian ridge + AdaBoost regressor model was the most effective in predicting the dynamic viscosity relative to the three hybrid learners.

3 Model 4 (combination with four parameters) was the most accurate in the testing phase, with $R^2 > 94\%$, followed by Models 2, 1, and 3 (combination with 2, 1 and 3 parameters).

4 This study was the first to present the dynamic viscosity of nanofluids through these four combinations, providing a reliable and robust technology for dynamic viscosity prediction.

In conclusion, the results of this study demonstrate the potential of the developed ML models for accurately predicting the dynamic viscosity of nanofluids. However, future research could explore additional essential issues, such as using optimisation models with multiple models with the same parameters or the same applied model with different scales of input data.

Acknowledgements

This work would not have been possible without the support of Universiti Teknologi Malaysia (UTM), operated by the Research Management Center (RMC) under the Research University Grant number (06E10).

References

Adun, H. et al. (2022) 'Impact of data processing and robust machine learning process on accurate estimation of specific heat capacity property in energy storage applications', *Journal of Energy Storage*, 1 November, Vol. 55, Part A, p.105359.

Ali, A.R.I. and Salam, B. (2020) 'A review on nanofluid: preparation, stability, thermophysical properties, heat transfer characteristics and application', *SN Applied Sciences*, Vol. 2, No. 10, p.1636.

Ali, Z.H. et al. (2022) 'Developing an integrative data intelligence model for construction cost estimation', *Complexity*, Vol. 2022, Article ID 4285328, 18pp.

Alshuhail, L.A., Shaik, F. and Sundar, L.S. (2023) 'Thermal efficiency enhancement of mono and hybrid nanofluids in solar thermal applications – a review', *Alexandria Engineering Journal*, 1 April, Vol. 68, pp.365–404.

Amin, A-T.M., Hamzah, W.A.W. and Oumer, A.N. (2021) 'Thermal conductivity and dynamic viscosity of mono and hybrid organic-and synthetic-based nanofluids: a critical review', *Nanotechnology Reviews*, Vol. 10, No. 1, pp.1624–1661.

Ansari, H. et al. (2018) 'A new comprehensive model for relative viscosity of various nanofluids using feed-forward back-propagation MLP neural networks', *International Communications in Heat and Mass Transfer*, February, Vol. 91, pp.158–164.

Asadi, A. et al. (2019) 'Effect of sonication characteristics on stability, thermophysical properties, and heat transfer of nanofluids: a comprehensive review', *Ultrasonics Sonochemistry*, November, Vol. 58, p.104701.

Atashrouz, S., Pazuki, G. and Alimoradi, Y. (2014) 'Estimation of the viscosity of nine nanofluids using a hybrid GMDH-type neural network system', *Fluid Phase Equilibria*, 25 June, Vol. 372, pp.43–48.

Baghban, A. et al. (2019) 'Developing an ANFIS-based swarm concept model for estimating the relative viscosity of nanofluids', *Engineering Applications of Computational Fluid Mechanics*, Vol. 13, No. 1, pp.26–39.

Chen, H. et al. (2007) 'Rheological behaviour of ethylene glycol based titania nanofluids', *Chemical Physics Letters*, Vol. 444, Nos. 4–6, pp.333–337.

Chen, L. et al. (2022) 'Optimization and comparison of machine learning methods in estimation of carbon dioxide loading in chemical solvents for environmental applications', *Journal of Molecular Liquids*, 1 March, Vol. 349, p.118513.

Cruz, J. et al. (2021) 'Selection of characteristics by hybrid method: RFE, ridge, lasso, and Bayesian for the power forecast for a photovoltaic system', *SN Computer Science*, Vol. 2, No. 3, p.202.

Esfe, M.H. et al. (2016) 'Designing an artificial neural network to predict dynamic viscosity of aqueous nanofluid of TiO2 using experimental data', *International Communications in Heat and Mass Transfer*, July, Vol. 75, pp.192–196.

Esfe, M.H., Kamyab, M.H. and Toghraie, D. (2022) 'Statistical review of studies on the estimation of thermophysical properties of nanofluids using artificial neural network (ANN)', *Powder Technology*, March, Vol. 400, p.117210.

Garoosi, F., (2020) 'Presenting two new empirical models for calculating the effective dynamic viscosity and thermal conductivity of nanofluids', *Powder Technology*, 15 April, Vol. 366, pp.788–820.

Ghazvini, M. et al. (2022) 'Optimization of MLP neural network for modeling flow boiling performance of Al_2O_3/water nanofluids in a horizontal tube', *Engineering Analysis with Boundary Elements*, December, Vol. 145, pp.363–395.

Gholizadeh, M. et al. (2020) 'Prediction of nanofluids viscosity using random forest (RF) approach', *Chemometrics and Intelligent Laboratory Systems*, 15 June, Vol. 201, p.104010.

Ghritlahre, H.K. and Prasad, R.K. (2018) 'Application of ANN technique to predict the performance of solar collector systems – a review', *Renewable and Sustainable Energy Reviews*, March, Vol. 84, pp.75–88.

Guimarey, M.J. et al. (2019) 'Thermophysical properties of polyalphaolefin oil modified with nanoadditives', *The Journal of Chemical Thermodynamics*, April, Vol. 131, pp.192–205.

Heidari, E., Sobati, M.A. and Movahedirad, S. (2016) 'Accurate prediction of nanofluid viscosity using a multilayer perceptron artificial neural network (MLP-ANN)', *Chemometrics and Intelligent Laboratory Systems*, 15 July, Vol. 155, pp.73–85.

Hemmati-Sarapardeh, A. et al. (2018) 'On the evaluation of the viscosity of nanofluid systems: modeling and data assessment', *Renewable and Sustainable Energy Reviews*, January, Vol. 81, Part 1, pp.313–329.

Jiang, H. et al. (2021) Elucidation of the thermophysical mechanism of hexagonal boron nitride as nanofluids additives', *Advanced Powder Technology*, Vol. 32, No. 8, pp.2816–2827.

Kalsi, S. et al. (2023) 'Thermophysical properties of nanofluids and their potential applications in heat transfer enhancement: a review', *Arabian Journal of Chemistry*, November, Vol. 16, No. 11, p.105272.

Kanti, P.K. et al. (2023a) 'The effect of pH on stability and thermal performance of graphene oxide and copper oxide hybrid nanofluids for heat transfer applications: application of novel machine learning technique', *Journal of Energy Chemistry*, July, Vol. 82, pp.359–374.

Kanti, P.K. et al. (2023b) 'The stability and thermophysical properties of Al_2O_3-graphene oxide hybrid nanofluids for solar energy applications: application of robust autoregressive modern machine learning technique', *Solar Energy Materials and Solar Cells*, May, Vol. 253, p.112207.

Karimipour, A. et al. (2018) 'A new correlation for estimating the thermal conductivity and dynamic viscosity of CuO/liquid paraffin nanofluid using neural network method', *International Communications in Heat and Mass Transfer*, March, Vol. 92, pp.90–99.

Lingala, S.S. (2023) 'Ionic-liquid-based nanofluids and their heat-transfer applications: a comprehensive review', *Chem. Phys. Chem.*, 16 November, Vol. 24, No. 22, p.e202300191.

Mehrabi, M., Sharifpur, M. and Meyer, J.P. (2013) 'Viscosity of nanofluids based on an artificial intelligence model', *International Communications in Heat and Mass Transfer*, April, Vol. 43, pp.16–21.

Mehta, B. et al. (2022) 'Synthesis, stability, thermophysical properties and heat transfer applications of nanofluid – a review', *Journal of Molecular Liquids*, 15 October, Vol. 364, p.120034.

Parzinger, M. et al. (2022) 'Comparison of different training data sets from simulation and experimental measurement with artificial users for occupancy detection – using machine learning methods random forest and LASSO', *Building and Environment*, September, Vol. 223, p.109313.

Phyo, P-P., Byun, Y-C. and Park, N. (2022) 'Short-term energy forecasting using machine-learning-based ensemble voting regression', *Symmetry*, Vol. 14, No. 1, p.160.

Porgar, S., Oztop, H.F. and Salehfekr, S. (2023) 'A comprehensive review on thermal conductivity and viscosity of nanofluids and their application in heat exchangers', *Journal of Molecular Liquids*, 15 September, Vol. 386, p.122213.

Rashidi, M.M. et al. (2021) 'Thermophysical properties of hybrid nanofluids and the proposed models: an updated comprehensive study', *Nanomaterials*, Vol. 11, No. 11, p.3084.

Salem, H. et al. (2022) 'Predictive modelling for solar power-driven hybrid desalination system using artificial neural network regression with Adam optimization', *Desalination*, 15 January, Vol. 522, p.115411.

Sharma, K. et al. (2023) 'Prognostic modeling of polydisperse SiO2/Aqueous glycerol nanofluids' thermophysical profile using an explainable artificial intelligence (XAI) approach', *Engineering Applications of Artificial Intelligence*, November, Vol. 126, Part B, p.106967.

Sleiti, A.K. (2020) 'Heat transfer measurements of polyalpha-Olefin-boron nitride nanofluids for thermal management and lubrication applications', *Case Studies in Thermal Engineering*, December, Vol. 22, p.100776.

Sleiti, A.K., (2021) 'Dataset for measured viscosity of polyalpha-olefin-boron nitride nanofluids', *Data in Brief*, April, Vol. 35, p.106881.

Taylor, K.E. (2001) 'Summarizing multiple aspects of model performance in a single diagram', *Journal of Geophysical Research: Atmospheres*, Vol. 106, No. D7, pp.7183–7192.

Wang, J. et al. (2023) 'A review on nanofluid stability: preparation and application', *Renewable and Sustainable Energy Reviews*, December, Vol. 188, p.113854.

Wang, L. et al. (2016) 'Investigation on viscosity of Fe_3O_4 nanofluid under magnetic field', *International Communications in Heat and Mass Transfer*, March, Vol. 72, pp.23–28.

Wang, R., Li, K. and Su, Y-X. (2022) 'Prediction of ameloblastoma recurrence using random forest – a machine learning algorithm', *International Journal of Oral and Maxillofacial Surgery*, Vol. 51, No. 7, pp.886–891.

Yang, L. et al. (2017) 'Recent developments on viscosity and thermal conductivity of nanofluids', *Powder Technology*, 15 July, Vol. 317, pp.348–369.

Younes, H. et al. (2022) 'Nanofluids: key parameters to enhance thermal conductivity and its applications', *Applied Thermal Engineering*, 5 May, Vol. 207, p.118202.

Zekić-Sušac, M., A. Has, and M. Knežević, (2021) 'Predicting energy cost of public buildings by artificial neural networks, CART, and random forest', *Neurocomputing*, 7 June, Vol. 439, pp.223–233.

Zendehboudi, A. et al. (2019) 'Data-driven methods for estimating the effective thermal conductivity of nanofluids: a comprehensive review', *International Journal of Heat and Mass Transfer*, March, Vol. 131, pp.1211–1231.

Zhang, Z. et al. (2023) 'Optimized ANFIS models based on grid partitioning, subtractive clustering, and fuzzy C-means to precise prediction of thermophysical properties of hybrid nanofluids', *Chemical Engineering Journal*, 1 September, Vol. 471, p.144362.

Zhou, B. et al. (2023a) 'XGBoost-SFS and double nested stacking ensemble model for photovoltaic power forecasting under variable weather conditions', *Sustainability*, Vol. 15, No. 17, p.13146.

Zhou, L. et al. (2020) 'Machine learning algorithms to predict flow condensation heat transfer coefficient in mini/micro-channel utilizing universal data', *International Journal of Heat and Mass Transfer*, Vol. 162, p.120351.

Zhou, T. et al. (2023b) 'Electrokinetic transport of nanoparticles in functional group modified nanopores', *Chinese Chemical Letters*, Vol. 34, No. 6, p.107667.