Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Measurement of multiple rheological properties of oil–water emulsions using a rotating device and machine learning

Kazuhiro Itoh

itoh@eng.u-hyogo.ac.jp

University of Hyogo

Sayu Hashimoto

University of Hyogo

Mehak Jain

University of Hyogo

Masakazu Morimoto

University of Hyogo

Article

Keywords: vision-based measurement, rheological parameter, non-Newtonian fluid, machine learning, small data analysis

Posted Date: November 12th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11423935/v1

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

Additional Declarations: No competing interests reported.

Measurement of multiple rheological properties of oil–water emulsions using a rotating device and machine learning

Kazuhiro Itoh[1*], Sayu Hashimoto[1], Mehak Jain[1], Masakazu Morimoto[1]

[1] University of Hyogo, 2167 Shosha, Himeji, Hyogo, 671-2280 Japan

[*] Corresponding author: itoh@eng.u-hyogo.ac.jp

Abstract

Rapid and non-intrusive measurement of the rheological properties of non-Newtonian fluids is crucial for applications such as the development of autonomous experimental systems. Herein, we propose a technique that utilizes machine learning to sequentially predict surface tension, contact angle, and viscosity characteristics based on flow images. Flow images were obtained by rotating oil–water emulsion samples in a cuvette by 90°. The extracted interface positions were used as features to train a random forest algorithm. The 1 mL emulsion sample remained entirely unconsumed, and the measurement was completed in 6 s. We demonstrated that the surface tension and contact angles of eight emulsion samples could be predicted with a NRMSE=27.7% or less using training data from only two samples.

Keywords

vision-based measurement, rheological parameter, non-Newtonian fluid, machine learning, small data analysis

1. Introduction

The integration of materials informatics and autonomous experimentation has contributed to a dramatic increase in the efficiency of materials discovery [1]. Automated experimentation using machine learning is actively being developed across various fields, including drug discovery [2] and biotechnology [3]. Artificial intelligence (AI)-driven robotic systems perform tasks such as protein synthesis in unmanned laboratories [4]. Furthermore, “metaverse laboratories” are being developed to optimize automated experiments by using deep learning to analyze correlations in material properties, ranging from the atomic scale to the bulk level, and predict future system configurations [5].

However, in many research and development scenarios, manual operations, such as transferring liquids via a pipette or test tube, are essential. Connecting different pieces of equipment and seamlessly integrating AI and robots still require a high level of expertise and technical skills. In such cases, partial or “hybrid” automated experiments that incorporate manual tasks are recommended. This approach involves a system where humans and autonomous robots work collaboratively in the same laboratory [6, 7]. Here, robots utilize the same instruments and perform the same actions as humans, and they make decisions based on features extracted through manual processes [8].

Transparent vessels, commonly used for reaction, mixing, and classification operations in biology and chemistry, present challenges for this type of hybrid automation [9, 10]. Conventional robotic operations that rely on pressure sensors have certain limitations. Consequently, vision-based

44 control systems have been proposed to enhance operational safety and versatility while achieving high
45 throughput. Robots execute tasks such as recognizing transparent vessels, performing pouring
46 operations, and screening for solubility by analyzing captured images [11-13]. Vision systems
47 contribute to higher throughput by eliminating the need for certain sensors, as well as reducing the
48 weight and space requirements of equipment. Thus, the development of such image-based
49 measurement techniques is valuable not only for automated experiments but also for experiments
50 conducted in extreme environments, such as X-ray facilities [14, 15] or microgravity conditions [16].
51 However, the amount of available data, including multiple rheological parameters, for the
52 development of new materials is limited. Applying machine learning to solve complex tasks is
53 challenging and, in some cases, infeasible, particularly for deep learning, which requires large datasets.
54 Consequently, rather than mass-producing image-based deep learning models, it is more adaptable to
55 construct a prediction system that reduces flow images to low-dimensional parameters, such as
56 viscosity and surface tension, and incorporates physical models [17].
57 Several studies on machine learning with limited data have focused on the effective use of
58 data [18-20] and the comparison of algorithms [21-23]. In small-dataset machine learning, the random
59 forest (RF) algorithm offers robust analysis because it can utilize every record in a dataset without the
60 risk of overfitting. Consequently, it is employed in fields where it is difficult to acquire large volumes
61 of data. Examples of its applications include education [20], biomedicine [24], prognosis of breast

cancer [25], marine biology [26], forest fires [22], river flow rates [23], forest management [27], screening for small-molecule additives [28], and the prediction of thermophysical properties and heat transfer in nanofluids [29].

In this study, we attempted to decompose the flow images of oil–water (O/W) emulsions as representative non-Newtonian fluids into rheological parameters using a RF algorithm. Basic experimental procedures such as stirring and pouring are related to the rheological parameters of fluids [30]. Consequently, we focused on surface tension, contact angle, and viscosity. Currently, these physical parameters are measured using separate devices and methods. When factors such as temperature control and pre- and post-cleaning are considered, the measurement of these quantities requires significant effort, making it hard to automate.

Traditional methods, such as electrochemical, optical, and spectroscopic methods for the measurement of surface tension, are unsuitable for automated systems because they require careful cleaning and laborious setup. The pendant drop method is a well-known image-based technique for measuring surface tension. Soori et al. [31] reported that by inputting images of suspension droplets into a convolutional neural network, surface tension can be predicted with high accuracy, independent of the droplet size or liquid properties. Because the surface tension of surfactant solutions depends on their concentration and relaxation time, multiple measurement techniques have been developed using machine learning [32, 33]. However, the requirement for special equipment to generate and image the

80 droplets presents challenges for implementing in-line measurements in autonomous experimentation
81 systems. Moreover, databases for deep learning only exist for surfactants [34-36] and ionic liquids
82 [37]. However, no databases exist for most other non-Newtonian fluids. Handling such fluids requires
83 time and effort to build a database.

84 Viscosity is an important rheological parameter for a wide range of fluids, including foods,
85 paints, solvents, and detergents. However, there are only a few examples of non-contact measurements.
86 Deshmukh et al. [38] enabled the high-throughput viscosity measurement of complex fluids based on
87 flow in a microchip. Methods utilizing compliant containers and pressure sensors in conjunction with
88 machine learning have also been proposed [39]; however, these methods require a substantial sample
89 volume and involve large-scale equipment. Although high-precision measurements using microfluidic
90 channels [40, 41] and machine learning have been achieved, slurry fluids present clogging problems.
91 Although a high-throughput method using pipetting actions [42] has been proposed, it is limited to
92 Newtonian fluids. Walker et al. [43] attempted to predict viscosity using a vision-based approach by
93 capturing video of a robot pouring liquids. This approach is similar in concept to that used in the
94 present study. However, it was limited to Newtonian fluids and focused on classification rather than
95 predicting specific viscosity values.

96 This study aimed to provide developers, such as those working in drug discovery and the
97 food and cosmetic industries who manually handle small samples, with a rapid and non-intrusive

rheological measurement method. It attempted to predict surface tension and viscosity simultaneously based on flow images captured while tilting a small cuvette by 90°. Because the proposed method fundamentally targets small-dataset evaluations, the RF algorithm was employed for machine learning. The proposed method enables the measurement of liquid properties without physical contact or sample consumption. Because the only required action is to tilt the container, the proposed technique has potential for future applications in in-line measurements in autonomous experimental systems.

2. Methods

2.1 Preparation of emulsion

Ten samples with different physical properties were prepared. Table 1 lists the compositions of the samples. The dispersed phase comprised linseed oil, oleic acid, stearic acid, and palmitic acid. The continuous phase consisted of a mixture of distilled water and ethanol containing sodium benzoate and Polysorbate 60. Sample weighing was performed at room temperature (20–25 °C).

The continuous and dispersed phases were placed in separate 50-mL beakers and heated to 60 °C with stirring using a magnetic stirrer. After the two phases were mixed, the mixed solution was cooled to 30 °C and stirred for 10 min. Subsequently, it was transferred to a conical tube and homogenized for 5 min using a homogenizer (Model 3000, Dremel Co.) while being maintained at 30 °C using a mantle heater. The rotational speed of the homogenizer was set to 17,000 rpm.

116 2.2 Rheological properties

117 The measurements of surface tension (T), advancing and receding contact angles (θ_F and
118 θ_R, respectively), density (ρ), and viscosity (η) were conducted in air at 30 °C. Surface tension was
119 measured by the drop weight method [44]. Nozzles with inner diameters ranging from 0.13 to 4.01
120 mm were attached to the syringe tip. A total weight of 20 drops was transferred to surface tension
121 according to Tate's law. The error in surface tension for these measurements was 5% or less.

122 The advancing and receding contact angles were measured from the images of tilting sessile
123 drops [45]. A drop placed on the surface was photographed, and the plate was tilted at an angle of 7°.
124 The plate was made of polystyrene and coated with a silicone-based water repellent. The images were
125 analyzed by spline approximation using the ImageJ software to determine the advancing and receding
126 contact angles. The reading error for these angles was less than 1%. All images of the drops are shown
127 in Supplementary Fig. S1. The density was determined by filling a 3D-printed 10-mm cube container
128 and measuring its weight.

129 The shear dependence of the viscosity, the so-called flow curve, was measured using a
130 rheometer (AR-2000EX, TA Instruments, Waters LLC). Measurements were conducted in steady-state
131 flow mode over a shear rate range $\dot{\gamma}$ from 0.001 to 1000 s^{-1}. In this study, the maximum viscosity
132 value observed at low shear rates close to $\dot{\gamma}$ = 0.001 s^{-1} was defined as the zero-shear viscosity, η_0.
133 To characterize the properties of the flow curve, the measured viscosity values in the $\dot{\gamma}$ range from

134 0.010 to 1.000 s^{-1} were fitted using the following power-law equation,

135

$$\eta = K \cdot \dot{\gamma}^{n-1} . \quad (1)$$

136

137 Here, n is the power-law index, and K [Pa·s^n] is the consistency.

138

139 2.3 Rotation and image acquisition

140 One milliliter of the prepared emulsion sample was placed in a transparent cuvette, mounted
141 on a rotating device, and spun. The test apparatus is shown in Fig. 1a. The cuvette was made of
142 polystyrene and had a 10 × 10 mm^2 cross-section. The same water-repellent agent used for the contact
143 angle measurements was pre-applied to the inner surface of the cuvette.

144 An automated rotation stage (OSMS-60YAW, Sigma Koki Co., Ltd.; stage diameter: 60 mm)
145 was employed. The motion began with the cuvette in a vertical position, followed by a 90° rotation
146 over 4 s. From 4 to 6 s, the horizontal cuvette remained stationary while the internal flow was observed.
147 Fig. 1b shows the input signal for the angular change. To prevent disturbances in the motion of the
148 interface, the angular velocity was set to be slow during the early and final stages of the rotation.

149 A video of the flow inside the cuvette was captured using two cameras, one positioned at
150 the top and the other at the front. Lightweight board-type cameras (ELP-USBFDH01M-L21, ELP Co.,

151 Ltd.) were used to perform tracking shots while moving in tandem with the cuvette. Images were
152 acquired at a resolution of 1920 × 1080 pixels and a frame rate of 30 fps.

153 The OpenCV library included in Python was used to extract still frames from the video of
154 the flowing emulsion in the MP4 format. JupyterLab with Anaconda navigator was used to execute
155 the process. Seven JPEG images captured by the top camera at 0, 1, 2, 3, 4, 5, and 6 s were used for
156 machine learning.

157 Figure 2 shows examples of the obtained images. The interface moves from left to right with
158 increasing time. The shape of the interface was not straight; instead, it assumed a complex shape
159 because of contact with the corner of the cuvette. These interface shapes were also incorporated into
160 the machine-learning process. Specifically, the interface displacement distances along the two lines
161 depicted in Fig. 2b were obtained as features L_1 and L_2 in the machine learning model. Thus, the
162 feature set L (comprising L_1 and L_2) representing the interface movement consisted of 14 data
163 points regarding the interface positions along these two lines over the interval from 0 to 6 s. All of the
164 image of interface was presented in supplementary Fig. S2.

165 The RF algorithm, implemented using the Python Scikit-learn library, was used for machine
166 learning. RF is an ensemble learning method that constructs multiple decision trees during training
167 and outputs the mode of the classes for classification tasks or the mean prediction for regression tasks
168 [46]. It is characterized by resistance to overfitting. Predictions were made using RF regression. The

target variables were the rheological parameters of the emulsion, i.e., surface tension, advancing and receding contact angles, the power-law index and consistency. The training data consisted of one-dimensional arrays of the measured values, including the interface positions L.

The prediction error for each measurement value was evaluated using relative error (RE),

$$\mathrm{RE}\ [\%] = \frac{|y_i - \hat{y}_i|}{y_i} \times 100\,. \tag{2}$$

Here, n is the number of data points, y_i is the experimental value, and $\hat{y}_i$ is the predicted value. In addition, the normalized root mean square error (NRMSE) was used as an index for evaluating prediction performance.

$$\mathrm{NRMSE}\ [\%] = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}}{\frac{1}{n}\sum_{i=1}^{n} y_i} \times 100 \tag{3}$$

3. Results

3.1 Rheological property

The results of the maximum displacement of the interface L_{max} and other rheological properties are summarized in the lower half of Table 1. The flow curves for all samples are shown in Supplementary Fig. S3. The O/W emulsion exhibited shear-thinning behavior. K and n were

185 evaluated based on the linear approximation shown in Supplementary Fig. S3.

186 Because the cuvette height was 45 mm, the interface of Sample No. 4 (L_{max} = 40.5 mm)
187 reached a point near the top of the cuvette (Table. 1 and supplementary Fig. S2). The average values
188 for surface tension, advancing and receding contact angles, n, K, and zero-shear viscosity were
189 T_{ave}= 37.5 mN/m, $\theta_{F,ave}$= 44.4°, $\theta_{R,ave}$ = 41.8°, n_{ave} = 0.258, K_{ave} = 4.744 Pa·s^n, and $\eta_{0,ave}$ =
190 626.2 Pa·s, respectively. Fig. 3 shows the relationship between L_{max}, η_0, and T. Generally, the
191 displacement of the interface increased as the viscosity decreased. The correlation coefficient between
192 L_{max} and η_0 for all measurements was −0.57. However, excluding Sample No. 10, which had an
193 exceptionally high viscosity, the coefficient became more negative, reaching −0.75, indicating a
194 moderate correlation. In contrast, the correlation coefficient between L_{max} and T was −0.56,
195 indicating no correlation.

196 3.2 Cross validation test

197 The results of leave-one-out cross-validation (LOOCV) tests are listed in Table 2. In this
198 study, we denoted, for example, trial A for the prediction of T based on the training data consisting
199 of the time-series variation of the interface position $L(t)$ and T as A [T: L, T]. In the subsequent
200 trials, the acquired image data remained the same. The LOOCV tests were performed by systematically
201 swapping out a single data point from the dataset to act as the validation set.

In Trial A, the maximum RE in surface tension evaluation was 39.1% and the RMSE was 25.3%. These errors are well within an acceptable range for practical applications, such as improving the precision of experimental operations performed by robot arms. Therefore, the flow images and procedure for extracting the positions of interfaces were deemed suitable for predicting surface tension. As shown in Fig. 3, there was no correlation between the L_{max} and T. It can thus be inferred that surface tension is related to the process or history leading up to the final distance reached by the interface.

The second row of Table 2 lists the results for Trial B [T: L, T, η_0], in which η_0 was added as a feature, with $L(t)$ and T being the same for trials A and B. Changing the training data conditions from Trial A to B led to an increase in the cross-validation prediction error. Therefore, including viscosity as a feature when predicting the surface tension resulted in overfitting. A preliminary analysis indicated that the simplest condition, that of Trial A, yielded the lowest cross-validation error.

The results of Trial C [θ_F: L, θ_F] for predicting θ_F are shown in the third row of Table 2. Advancing and receding contact angles are mutually related through the mechanical equilibrium of the tilted sessile drop. Therefore, Trial D [θ_F: L, θ_F, θ_R], which incorporates θ_R as a feature, was also conducted. In Trial C, the maximum RE was 48.9%, and the RMSE was 24.3%. In contrast, Trial D yielded a maximum RE of 32.5% and an RMSE of 15.5%. These results indicated that Trial D,

220 which utilized both θ_F and θ_R as features, was more effective for the prediction of θ_F.

221 The results of trial E [θ_R: L, θ_R] and F [θ_R: L, θ_F, θ_R] for predicting θ_R are given in the
222 fifth and sixth rows of Table 2, respectively. While Trial E showed a maximum RE of 65.6% and an
223 RMSE of 33.3%, Trial F demonstrated improvements, achieving a maximum RE of 60.1% and an
224 RMSE of 20.9%. Thus, trials D and F that used both θ_F and θ_R as features were more advantageous
225 for predicting contact angles. This was due to the mechanical correlation between the two contact
226 angles.

227 The results for trials G [n: L, n] and H [n: L, n, K] predicting n are shown in the seventh
228 and eighth rows of Table 2, respectively. As both n and K are coefficients in the power-law model
229 defined by Eq. (1), we anticipated that they would influence the flow while interacting. Therefore, we
230 anticipated that the error for Trial H would be smaller than that of Trial G. However, the RMSE of
231 trials G and H was 41.3% and 41.5%, respectively, showing negligible difference between the two.
232 Although adding feature K reduced the error in some cases, such as Sample Nos. 2, 3, 4, and 6, it
233 increased the error in other cases (Sample Nos. 9 and 10). These results indicated that including K as
234 a feature for predicting n offered minimal benefit. The results for trials I [K: L, K] and J [K: L, n,
235 K] for predicting K are shown in the ninth and tenth rows of Table 2, respectively. The RMSE for
236 trials I and J was 62.7% and 64.9%, respectively. This reveals that including n as a feature when
237 predicting K led to overfitting.

Finally, the results of Trial K [η_0: L, η_0] for predicting η_0 are shown in the last row of Table 2. A very small error (0.88%) was observed for Sample No. 4, whereas considerably larger errors were observed for Sample Nos. 1, 7, 8, and 10. Sample No. 4 is characterized by a small K and a minimum n. However, the relation between $L(t)$ and the good prediction of η_0 is currently unknown. The small displacement of the interface with L_{max} < 7 mm for Sample Nos. 7, 8, and 9 may have increased the error in predicting η_0. For Sample No. 1, the shape of the interface was asymmetric with respect to the central axis of the cuvette, as shown in Fig. 2a. It is possible that the characteristics of the interface shape were not reflected in L_1or L_2 because of this asymmetry. Indeed, the errors for predicting T, θ, n, and K were not small for Sample No. 1. It was therefore concluded that images with good interface symmetry contributed to improved prediction accuracy.

Preliminary analysis confirmed that an increase in the number of features withheld the improvement of the cross-validation error of η_0. Although η_0 is a commonly used index, it represents viscosity at low shear rates ($\dot{\gamma} \approx 0.001$ s^{-1}). The advantage of image-based methods lies in their ability to evaluate the dynamic rheological properties at practical shear rates rather than in the static region. Similarly, the range ($0.010 \leq \dot{\gamma} \leq 1.000$ s^{-1}) of approximation used for fitting to the power law would affect the prediction of K. Approximating with a viscosity model that covers a wider range of shear rates may improve the error. However, introducing a complex viscosity model increases the number of prediction targets, potentially leading to a shortage of training data.

4. Demonstration and discussion

To verify the applicability of this method in product development, a trial was conducted to predict the rheological properties of the remaining samples using two training data points. Sample Nos. 1 and 10 were selected as training data. Sample No. 1 represents the lowest values for θ_R, K, and η_0, while Sample No. 10 exhibits the highest values for K and η_0. Thus, they were selected to encompass the entire dataset.

Based on the results of the previous section, the model was trained using the following conditions: Trial A for surface tension, Trial D for the advancing contact angle, Trial F for the receding contact angle, Trial G for the power-law index, and Trial I for consistency. The zero-shear viscosity was excluded from the prediction owing to its high error rate. First, the parameters L, T, θ, K, and n were extracted from the images of Sample Nos. 1 and 10. A RF model was trained to predict the surface tension T under the Trial A condition. Next, the L data extracted from the images of Sample No. 2 was input into the RF model trained on Sample Nos. 1 and 10 to predict T for Sample No.2. After recreating an instance of the RF model with the same hyperparameters, similar predictions were performed for Sample Nos. 3–9. Using the same method, θ, K, and n were also predicted. Although L was extracted from the images only once, the machine learning processes were repeated for each target variable. This approach was adopted to minimize the risk of overfitting caused by the limited

amount of training data. If the current manual machine learning and prediction process were automated for continuous execution, the analysis time required would be negligible.

The results are summarized in Table 3. The surface tension and advancing and receding contact angles could be predicted with an accuracy of approximately 30% or lower, even with only two training samples. Although the NRMSE of the power-law index (38.4%) and consistency (72.0%) were higher than those for surface tension and contact angle, we have demonstrated the viability of using image-based machine learning for the rapid, noncontact measurement of multiple rheological properties within a practical error range.

In addition, Fig. 4 shows the relationship between the RE of various parameters and L_{max}. It was confirmed that the RE was not significantly affected by the maximum distance travelled by the interface. This suggests that the measurement of these rheological quantities requires an interface displacement of only 6.3–40.5 mm within the present error. One approach to improve accuracy would be to approximate the flow curve, including the high-shear-rate viscosity. Furthermore, extracting interface positions while accounting for asymmetry in the width direction remains a challenge for future study.

5. Conclusion

An O/W emulsion sample was placed in a transparent cuvette with a 10 × 10 mm^2 cross-

292 section, and flow images were captured while the cuvette was tilted by 90° over 4 s. Time-series
293 positions of the interface were extracted from these images and used for machine learning with the RF
294 algorithm.

295 The surface tension, contact angle, power-law index and consistency of eight samples were
296 satisfactorily predicted based on training samples. Based on cross-validation results, the training
297 features were limited to the interface position and the corresponding target quantity.

298 Surface tension and contact angle were predicted with an NRMSE of less than 27.7%.
299 Although errors were higher for the power-law index (38.4%) and consistency (72.0%), this study
300 demonstrates the viability of a rapid measurement method for multiple physical properties with
301 practical accuracy. However, the prediction error for the zero-shear viscosity, which describes the state
302 at rest, was significant.

303

304 References

305 [1] Hwang, J. et. al.; Self-driving laboratories in Korea: a new era of autonomous discovery, Digital
306 Discovery, 5(5), 1968-1980 (2026). https://doi.org/10.1039/d6dd00024j

307 [2] Chen, Y. and Xue, W.; Machine learning for de novo molecular generation: A Comprehensive
308 Review, ACS Chem. Neuroscience, 17(4), 666-680 (2026).
309 https://doi.org/10.1021/acschemneuro.5c00861

[3] Smith, A. A. et al.; Using a GPT-5-driven autonomous lab to optimize the cost and titer of cell-free protein synthesis, bioRxiv 2026.02.05.703998 (2026). https://doi.org/10.64898/2026.02.05.703998

[4] Callaway E.; 'Set it and forget it': automated lab uses AI and robotics to improve proteins, Nature, 625(7995), 436 (2024). https://doi.org/10.1038/d41586-024-00093-w

[5] Gao, Y., Lin, H. and Zhu, X.; General Aqueous System Simulation through an AI-Embedded Metaverse Chemistry Laboratory, J. Phys. Chem. Letters, 15(22), 5978-5984 (2024). https://doi.org/10.1021/acs.jpclett.4c01111

[6] Tom, G. et al.; Self-Driving Laboratories for Chemistry and Materials Science, Chem. Reviews, 124(16), 9633-9732 (2024). https://doi.org/10.1021/acs.chemrev.4c00055

[7] Veeramani, S. et al.; Human-Aware Robot Behaviour in Self-Driving Labs, HRI Companion '26: Companion Proc. 21st ACM/IEEE Int. Conf. on Human-Robot Interaction, 292 - 296 (2026). https://doi.org/10.1145/3776734.3794402

[8] Jeon, M., Yu, G., Choi, H., Kim, G. and Hwang, H.; Real-Time Automated Solubility Screening Method Using Deep Neural Networks with Handcrafted Features, Sensors 23(12), 5525 (2023). https://doi.org/10.3390/s23125525

[9] Tiong, L.C.O. et al.; Machine vision-based detections of transparent chemical vessels toward the safe automation of material synthesis, npj Comput. Mater., 10, 42 (2024). https://doi.org/10.1038/s41524-024-01216-7

[10] Salazar-Villacis P. and Benyahia, B.; The ADePT framework for assessing autonomous laboratory robotics. Commun Chem., 9, 99 (2026). https://doi.org/10.1038/s42004-026-01932-9

[11] Eppel, S., Xu, H., Bismuth, M. and Aspuru-Guzik, A.; Computer Vision for Recognition of Materials and Vessels in Chemistry Lab Settings and the Vector-LabPics Data Set, ACS Cent. Sci., 6, 1743-1752 (2020). https://dx.doi.org/10.1021/acscentsci.0c00460

[12] Schober, D., Güldenring, R., Love, J. and Nalpantidis, L.; Vision-based robot manipulation of transparent liquid containers in a laboratory setting, arXiv 2404.16529v1 (2024). https://doi.org/10.48550/arXiv.2404.16529

[13] Shiri, P. et al.; Automated solubility screening platform using computer vision, iScience 24, 102176 (2021). https://doi.org/10.1016/j.isci.2021.102176

[14] Ozgulbas, D.Y. et al.; Robotic pendant drop: containerless liquid for μs-resolved, AI-executable XPCS. Light Sci Appl 12, 196 (2023). https://doi.org/10.1038/s41377-023-01233-z

[15] Peter A. B. and Tyler B. M.; The Autonomous Formulation Laboratory: An Open Liquid Handling Platform for Formulation Discovery Using X-ray and Neutron Scattering. Chem. Mater. 35(3), 846–852 (2023). https://doi.org/10.1021/acs.chemmater.2c03118

[16] Seehanam, S. et al.; Universal and economical experimental platform for colloidal mixing lab-on-chip in parabolic flight. Sci. Rep. 15, 19971 (2025). https://doi.org/10.1038/s41598-025-04368-8

[17] Swischuk, R. Mainini, L., Peherstorfer, B. and Willcox, K.; Projection-based model reduction:

346 Formulations for physics-based machine learning, Computers & Fluids, 179, 704-717 (2019).
347 https://doi.org/10.1016/j.compfluid.2018.07.021.
348 [18] Xu, P., Ji, X., Li, M. and Lu, W.; Small data machine learning in materials science. npj Comput.
349 Mater., 9, 42 (2023). https://doi.org/10.1038/s41524-023-01000-z
350 [19] Kraljevski I., Ju, Y.C., Ivanov, D., Tschöpe, C. and Wolff, M.; How to Do Machine Learning
351 with Small Data? - A Review from an Industrial Perspective, arXiv, 2311.07126 (2023).
352 https://doi.org/10.48550/arXiv.2311.07126
353 [20] Steinert S. et al.; A refined approach for evaluating small datasets via binary classification using
354 machine learning. PLoS ONE, 19(5): e0301276 (2024). https://doi.org/10.1371/journal.pone.0301276
355 [21] Wantoro, A., Zulkifli, Bintoro, P., Andika, T.H. and Ardhy, F. et al.; Performance Evaluation of
356 Classification Multi Algorithms on Small Dataset: A Comparative-Based Analysis, 2025 Tenth Int.
357 Conf. on Informatics and Computing (ICIC), Lampung, Indonesia, 2025, pp. 1-6,
358 https://doi.org/10.1109/ICIC68054.2025.11309491.
359 [22] Zhou, S.; An Analysis of The Small Sample Datasets Based on Machine Learning, EITCE'22;
360 Proc. 2022 6th Int. Conf. Electronic Inform. Technol. Comp. Enging., pp. 1654-1658 (2022).
361 https://doi.org/10.1145/3573428.3573720
362 [23] Fang, W. et al.; An evaluation of random forest based input variable selection methods for one
363 month ahead streamflow forecasting, Sci Rep. 14, 29766 (2024). https://doi.org/10.1038/s41598-024-

81502-y

[24] Sunwoo Han, S., Williamson, B. D. and Fong, Y.; Improving random forest predictions in small datasets from two-phase sampling designs, BMC Med. Inform. Decis. Mak., 21, 322 (2021). https://doi.org/10.1186/s12911-021-01688-3

[25] Zhang, H. and Wang, M.; Search for the smallest random forest, Stat Interface, 2(3):381 (2009). https://10.4310/sii.2009.v2.n3.a11

[26] Luan, J., Zhang, C., Xu, B., Xue, Y. and Ren, Y.; The predictive performances of random forest models with limited sample size and different species traits, Fisheries Research, 227, 105534 (2020). https://doi.org/10.1016/j.fishres.2020.105534.

[27] Diamantopoulou, M.J., Özçelik, R. and Genç, S.K.; Evaluation of the random forest regression machine learning technique as an alternative to ecoregional based regression taper modelling, Computers and Electronics in Agriculture, 239 (A), 110964, (2025). https://doi.org/10.1016/j.compag.2025.110964.

[28] Liu, A. L. et al.; Small Data Machine Learning: Classification and Prediction of Poly (ethylene terephthalate) Stabilizers Using Molecular Descriptors. ACS Appl. Polym. Mater. 2 (12), 5592–5601 (2020). https://doi.org/10.1021/acsapm.0c00921

[29] Sharma, P. et al.; Recent Advances in Machine Learning Research for Nanofluid-Based Heat Transfer in Renewable Energy System. Energy Fuels, 36 (13), 6626–6658 (2022).

https://doi.org/10.1021/acs.energyfuels.2c01006

[30] Miyazawa, C., Sakagami, K., Konno, N. and Nonomura, Y.; In-Situ Observation of Pouring a High-Viscosity Liquid. Technologies, 8(4), 69 (2020). https://doi.org/10.3390/technologies8040069

[31] Soori, T., Rassoulinejad-Mousavi, S.M., Zhang, L., Rokoni, A. and Sun, Y.; A machine learning approach for estimating surface tension based on pendant drop images, Fluid Phase Equilibria, 538, 113012 (2021). https://doi.org/10.1016/j.fluid.2021.113012.

[32] Saddiqi, H. A. et al.; Predictive modeling of emulsion stability and drop characteristics using machine learning: A study on surfactant influence and time dynamics, Food and Bioproducts Processing, 149, 190-198 (2025). https://doi.org/10.1016/j.fbp.2024.11.019.

[33] Dankloff, P. F. J. et al.; An autonomous robotic module for efficient surface tension measurements of formulations. npj Comput Mater., 11, 358 (2025). https://doi.org/10.1038/s41524-025-01842-9

[34] Seddon, D., Müller, E.A. and Cabral, J.T.; Machine learning hybrid approach for the prediction of surface tension profiles of hydrocarbon surfactants in aqueous solution, J. Colloid and Interface Science, 625, 328-339 (2022). https://doi.org/10.1016/j.jcis.2022.06.034.

[35] Chen J. et al.; Prediction of critical micelle concentration (CMC) of surfactants based on structural differentiation using machine learning, Colloids and Surfaces A: Physicochemical and Engineering Aspects, 703(1), 135276 (2024). https://doi.org/10.1016/j.colsurfa.2024.135276.

[36] Hödl, S. L. et al.; SurfPro - a curated database and predictive model of experimental properties

400 of surfactants. Digital Discovery, 4 (5), 1176–1187 (2025). https://doi.org/10.1039/d4dd00393d

401 [37] Gugulothu, N. Mohan, M., Marland, M.R., Smith, J. C., and Kidder M. K.; Deep Learning

402 Approaches for Predicting the Surface Tension of Ionic Liquids. J. Chem. Inf. Model, 65(12), 5856–

403 5867 23 (2025). https://doi.org/10.1021/acs.jcim.5c00158

404 [38] Deshmukh, K., Mitra, K. and Bit, A.; Influence of Non-Newtonian Viscosity on Flow Structures

405 and Wall Deformation in Compliant Serpentine Microchannels: A Numerical Study, Micromachines,

406 14, 1661 (2023). https://doi.org/10.3390/mi14091661

407 [39] Satpute, N.V. et al.; Data-Driven and Machine-Learning-Based Real-Time Viscosity

408 Measurement Using a Compliant Mechanism, Appl. Sci., 14, 10992 (2024).

409 https://doi.org/10.3390/app142310992

410 [40] Carnicer, V., Alcázar, C., Orts, M.J., Sánchez, E. and Moreno, R.; Microfluidic rheology: A new

411 approach to measure viscosity of ceramic suspensions at extremely high shear rates, Open Ceramics,

412 5, 100052 (2021). https://doi.org/10.1016/j.oceram.2020.100052.

413 [41] Giudice, F. D. and Barnes, C.; Rapid Temperature-Dependent Rheological Measurements of Non-

414 Newtonian Solutions Using a Machine-Learning Aided Microfluidic Rheometer, Anal. Chem., 94 (8),

415 pp.3617–3628 (2022). https://doi.org/10.1021/acs.analchem.1c05208

416 [42] Soh, B. W., Chitre, A., Yang Lee, W., Bash, D. Kumar, J. N. and Hippalgaonkar, K.; Automated

417 pipetting robot for proxy high-throughput viscometry of Newtonian fluids. Digital Discovery, 2(2),

481-488 (2023). https://doi.org/10.1039/d2dd00126h

[43] Walker, M., Pizzuto, G., Fakhruldeen, H and Cooper, A. I.; Go with the flow: deep learning methods for autonomous viscosity estimations, Digital Discovery, 2(5), pp.1540–1547 (2023). https://doi.org/10.1039/d3dd00109a

[44] Lee, B. B, Ravindra P. and Chan, E. S.; A Critical Review: Surface and Interfacial Tension Measurement by the Drop Weight Method. Chem. Eng. Communications, 195(8), 889–924 (2008). https://doi.org/10.1080/00986440801905056

[45] Hodgson, G., Passmore, M., Skarysz, M., Garmory, A. and Paolillo, F.; Contact angle measurements for automotive exterior water management. Exp Fluids 62, 119 (2021). https://doi.org/10.1007/s00348-021-03219-2

[46] Mohanty, A. and Gao, G.; A Survey of Machine Learning Techniques for Improving Global Navigation Satellite Systems, arXiv, 2406.16873 (2024). https://doi.org/10.48550/arXiv.2406.16873

Data availability

Data supporting the findings of this study are available from the corresponding author upon request.

Acknowledgements

We express our gratitude to Mr. Koto Tanaka, a graduate student at the University of Hyogo, for

meticulously conducting our experiments.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Author contributions

Kazuhiro Itoh: Conceptualization, Validation, Writing − Review and editing

Sayu Hashimoto: Data curation, Investigation, Visualization

Mehak Jain: Resources

Masakazu Morimoto: Methodology, Supervision

Competing interests

This research was funded by a joint research effort conducted by K. Itoh and Asada Chemical Industry Co., Ltd. Asada Chemical Industry was not involved in the content of the paper in any way.

Figure legends

Fig. 1. Experimental apparatus: (a) rotating device and (b) input signal representing the angle from the vertical.

454 **Fig. 2.** Time evolution of the interface of Sample Nos. 1 and 4.

455 **Fig. 3.** Relation between L_{max}, η_0, and T.

456 **Fig. 4.** Dependence of RE on L_{max}.

457 **Table 1.** Composition and rheological parameters of the emulsions.

458 **Table 2.** LOOCV test errors [%].

459 **Table 3.** Demonstration of estimation errors [%].

460

461 **Table 1.** Composition and rheological parameters of the emulsions.

Sample	1	2	3	4	5	6	7	8	9	10
Linseed oil [mL]	12.2	12.2	13.7	13.6	13.7	13.1	13.1	13.6	13.6	13.1
Oleic acid [mL]	6.1	6.1	6.9	6.9	6.9	6.9	6.6	6.9	6.9	6.9
Stearic acid [g]	1.3	1.3	1.5	1.5	1.5	2.0	1.4	1.5	1.5	2.0
Palmitic acid [g]	0.4	0.4	0.5	0.5	0.5	0.5	0.4	0.5	0.5	0.5
Distilled water [mL]	9.5	8.4	7.0	8.4	7.8	8.5	9.5	8.4	8.4	7.0
Ethanol [mL]	3.8	3.9	3.8	3.4	3.8	3.8	3.8	3.9	3.9	3.8
Polysorbate 60 [g]	1.7	0.2	1.7	0.2	1.9	0.2	0.2	0.2	0.2	1.7
Sodium benzoate [mg]	-	1.0	-	0.5	-	0.5	-	1.0	0.5	-
ρ [kg/m^3]	887	843	896	851	892	893	856	865	980	914
T [m·N/m]	27.2	24.9	33.2	44.1	26.2	42.7	50.4	42.7	47.8	35.9
θ_F [deg.]	34.7	39.5	59.2	44.2	42.6	39.2	47.9	44.0	33.7	59.1
θ_R [deg.]	23.3	33.7	58.8	38.8	39.8	39.8	38.0	48.6	38.8	58.7
n [-]	0.343	0.236	0.240	0.156	0.353	0.174	0.363	0.362	0.195	0.157
K [Pa·s^n]	1.132	1.970	3.954	1.676	6.692	3.436	4.320	6.805	5.776	11.68
η_0 [Pa·s]	64.20	253.3	379.6	196.9	461.9	624.0	479.2	531.8	860.2	2411
L_{max} [mm]	37.5	32.2	15.4	40.5	32.4	22.7	6.90	6.30	7.00	6.90

462

463

Table 2. LOOCV test errors [%].

Sample	1	2	3	4	5	6	7	8	9	10	NRMSE
$A\,[T\colon L, T]$	34.9	25.3	26.9	39.1	28.2	21.3	21.8	0.1	11.9+	28.1	25.3
$B\,[T\colon L, T, \eta_0]$	38.6	39.0	18.1	38.9	39.5	16.7	25.5	0.5	11.1	28.8	26.6
$C[\theta_F\colon L, \theta_F]$	22.2	1.3	29.6	15.2	13.5	9.6	6.4	17.0	48.9	32.5	24.3
$D[\theta_F\colon L, \theta_F, \theta_R]$	18.1	1.7	23.3	11.9	1.5	8.8	8.4	1.0	32.5	13.7	15.5
$E\,[\theta_R\colon L, \theta_R]$	65.6	1.3	33.2	26.7	14.4	6.2	2.8	34.5	63.8	39.3	33.3
$F[\theta_R\colon L, \theta_F, \theta_R]$	60.1	13.7	21.5	9.9	0.4	10.7	17.1	11.6	26.1	22.3	20.9
$G[n\colon L, n]$	45.6	16.6	26.9	96.4	22.4	37.6	23.0	41.6	55.9	59.1	41.3
$H[n\colon L, n, K]$	45.9	6.79	23.6	82.2	27.6	30.7	25.8	37.7	61.1	79.9	41.5
$I[K\colon L, K]$	119	39.2	42.5	11.6	67.6	30.3	73.4	0.71	42.8	57.6	62.6
$J[K\colon L, n, K]$	157	39.7	36.8	81.1	66.8	41.3	72.3	4.69	43.2	59.8	64.9
$K[\eta_0\colon L, \eta_0]$	414	47.2	44.6	0.88	44.0	34.8	144	128	51.2	72.9	106

Table 3. Demonstration of estimation errors [%].

Sample	2	3	4	5	6	7	8	9	NRMSE
T [m·N/m]	17.7	1.14	33.5	14.4	23.7	33.6	21.6	30.0	27.7
θ_F [deg.]	2.66	15.3	8.15	0.52	26.4	9.23	18.7	55.4	20.4
θ_R [deg.]	3.26	18.5	15.8	11.6	14.7	4.18	27.9	63.4	22.8
n [-]	26.3	7.15	91.0	19.6	31.1	42.5	42.3	7.03	38.4
K [Pa·s^n]	87.0	101	120	32.8	123	103	28.6	51.5	72.0

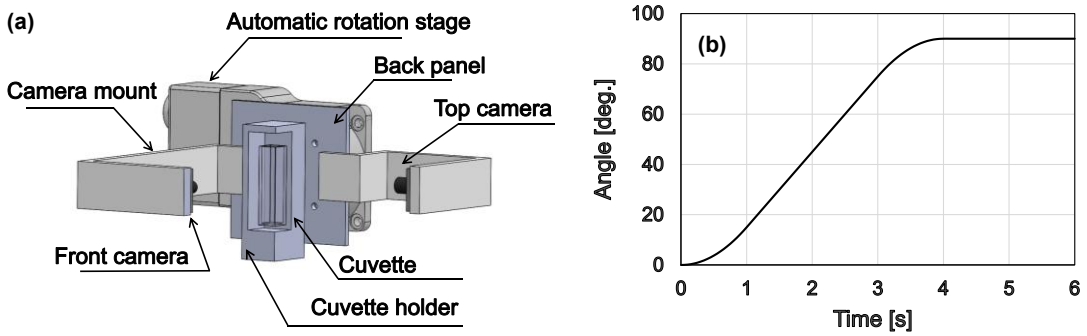


Fig. 1. Experimental apparatus: (a) rotating device and (b) input signal representing the angle from the vertical.

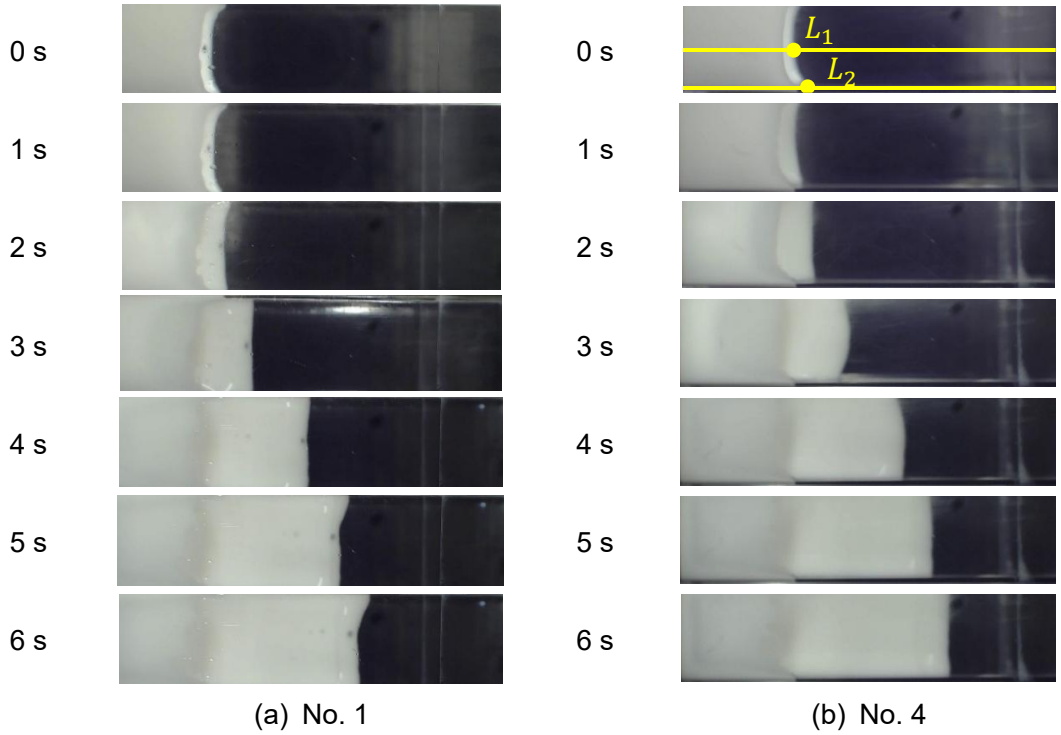
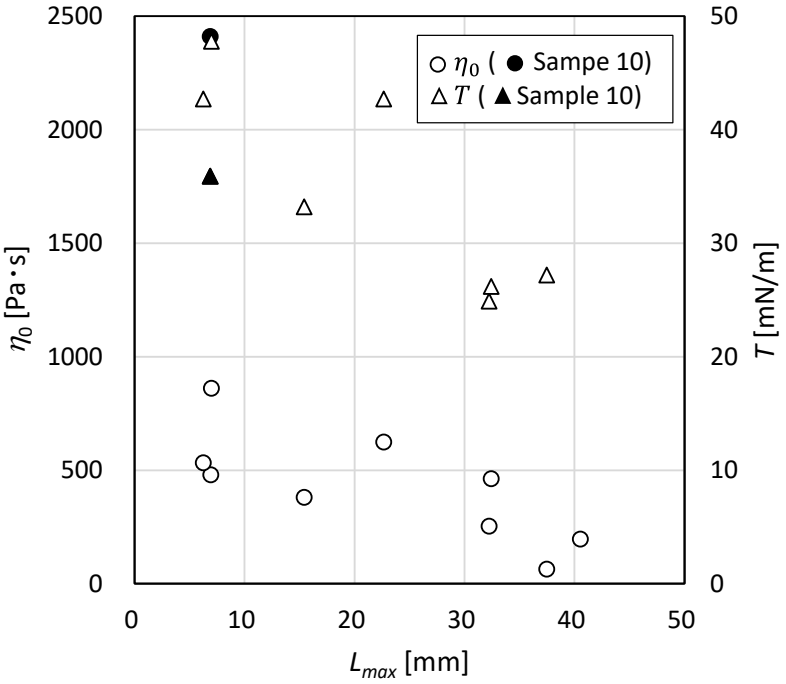


Fig. 2. Time evolution of the interface of Sample Nos. 1 and 4.

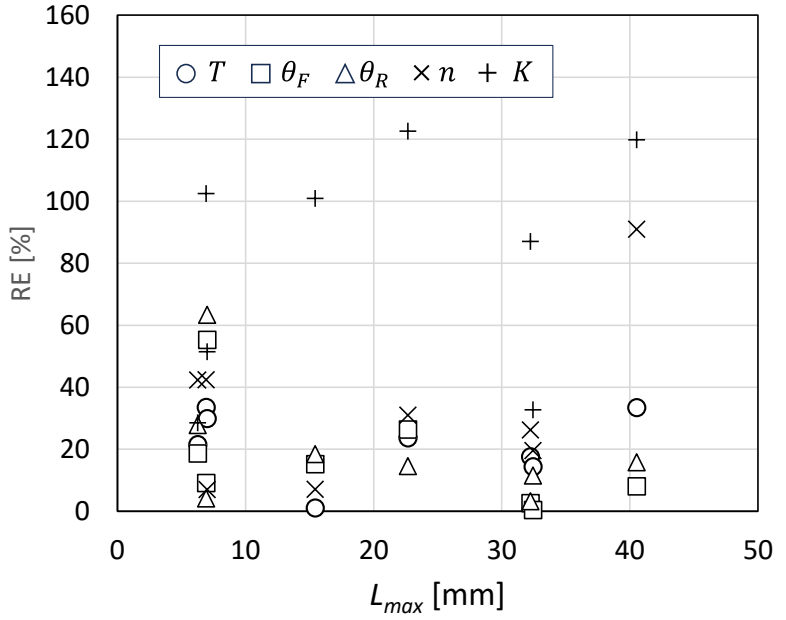


488

489 **Fig. 3.** Relation between L_{max}, η_0, and T.

490

491



492

493 **Fig. 4.** Dependence of ER on L_{max}.

494

495

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- 261019supplementarysheet.docx